FT2019 PROCEEDINGS

AEROSPACE TECHNOLOGY CONGRESS 2019

SUSTAINABLE AEROSPACE INNOVATION IN A GLOBALISED WORLD

STOCKHOLM WATERFRONT CONGRESS CENTRE, OCT 8-9, 2019

Proceedings of the 10th Aerospace Technology Congress

October 8-9, 2019, Stockholm, Sweden

Organized by:

Flyg- och Rymdtekniska Förening (FTF)
The Swedish Society of Aeronautics and Astronautics

www.ftfsweden.se

In Cooperation with:

INNOVAIR
The Swedish Strategic Innovation Programme for Aeronautics

www.innovair.org

Conference Location:

Stockholm Waterfront Congress Centre
Nils Ericsons Plan 4
111 64 Stockholm
Sweden

Conference webpage: www.ft2019.se

PREFACE

Since the start of the congress in 1992 as a national aerospace congress, the Aerospace Technology Congress became an international institution with the CEAS Conference 2013 hosted by the Swedish Society for Aeronautics and Astronautics. The Aerospace Technology Congress represents nowadays the largest national aerospace conference series in Sweden with a significant contribution from Europe and Worldwide, as it has become a Swedish forum for international collaboration. However, it has not lost its unique focus to both academia and industry. This becomes evidence in the selection of keynote speakers as well in the presenter and participants statistics where countries with industrial and research connection are dominating.

Highlights of the Conference:

• 143 oral presentations and 7 poster presentations in up to 12 parallel technical sessions
• 31 Reviewed Technical Papers
• 13 Keynote speeches
• 1 Panel discussion on synergies between aeronautics and astronautics
• 9 Sponsors & Exhibitors

WELCOME

On behalf of The Swedish Society for Aeronautics and Astronautics (Flyg- och rymdtekniska Föreningen, FTF) and The Swedish Strategic Innovation Programme for Aeronautics (Innovair), it is our pleasure to warmly welcome you to Stockholm to join our Congress FT2019, Sustainable Aerospace Innovation in a Globalised World.

The Congress is the tenth in our series of Aerospace congresses in Sweden. Every third year since 1992, FTF is hosting a national aerospace congress devoted to the progress of aeronautics and astronautics in Sweden. The Congresses have become the most important forum for aerospace companies and engineers in Sweden. Top-level managers from the leading European and international aerospace consortiums are invited for plenary lectures on the most important topics of today. Congress attendees will also have unique opportunities for networking and shall benefit from the scientific exchange during both formal and informal meetings with colleagues from Sweden and abroad.

So welcome again to the 10th Aerospace Congress!
Väl mött!

Dr. Anders Blom

Program Director Innovair
The Strategic Innovation
Programme for Aeronautics

Dr. Roland Karlsson

Chairman FTF
The Swedish Society of
Aeronautics and Astronautics

ORGANISED BY

The Swedish Society of Aeronautics and Astronautics (FTF) in cooperation with the Strategic Innovation Programme for Aeronautics (Innovair)

PARTNERS

We are proud to present the Aerospace Technology Congress 2019 partners.

Gold   Silver   Bronze

SUSTAINABLE AEROSPACE INNOVATION IN A GLOBALISED WORLD

MAJOR TOPICS 2019

- Major cooperative projects
- Environmentally friendly technology
- Aircraft and spacecraft system analysis
- Engines
- Structures, materials, and processes
- Aircraft and spacecraft technology
- Sub-system and system technology
- Operational availability, maintenance and support

www.ft2019.se

ORGANIZING COMITTEES

CONFERENCE CHAIR
Dr. Anders Blom
Dr. Roland Karlsson

PROGRAM COMMITTEE

Ordinary:

Kaj Lundahl, chairman
Björn Jonsson, secretary
Roland Karlsson
Christina Ahremark
Anders Blom
Oscar Hag
Tomas Ireman
Lasse Karlsen
Petter Krus
Lars Lundberg
Robert Lundberg
Mats-Olof Olsson
Odd Romell
Henrik Runnemalm
Vijay Sharan
Emil Vinterhav

Associates:

Eva Bernhardsdotter
Tomas Grönstedt
Dan Henningsson
Per Ingvarson
Christopher Jouannet
Olle Norberg
Olle Persson
Ingo Staack
Sytze Veldman
Mats Åbom

ORGANIZING COMMITTEE

Roland Karlsson, chairman
Björn Jonsson, secretary
Anders Blom
Olle Bååthe
Anders Gustafsson

Kaj Lundahl
Bengt Moberg
Louise Fischer, student representative
Anna Linderoth, Meetagain
Caroline Knies, Meetagain

PROCEEDINGS COMMITTEE

Ingo Staack
Petter Krus
Ragnar Larsson
Ardeshir Hanifi
Victor Juliano De Negri
Emilia Villani
Tomas Grönstedt

Associates:

Malin Åkermo
Roger Larsson
Kristian Sandahl
Kerstin Johansen
Mats Björkman
Magnus Sethson
Roland Gårdhagen
Alessandro Dell'AmicoLuis Ribeiro
Robert Braun
Alejandro Sobrón
Martin Hochwallner
David Lundström
Jonas Lundberg

EDITOR´S NOTE

The Aerospace Technology Congress 2019 represents the 10th iteration of aerospace technologies advancements over a period of almost 30 years.

The congress title **Sustainable Aerospace Innovation in a Globalised World** indicates the current challenges. At times of Friday for Future demonstrations and discussions about flight shame, urban air mobility and sustainable society, aerospace is suddenly no longer the high-tech showcase that it had been in the past. The congress title points out that these are no longer national problems and it is now more important than ever that the worldwide aerospace community – including academia, industry, politics and the society – has to work together to overcome today´s mobility and transportation problems that are not only limited to global warming (carbon dioxide emissions).

Currently it seems there is no single solution to these problems. Consequently, any possible technological progress must be applied to achieve the advanced stated goals of emission reduction. Research topics to be included may extensively enlarge classical aeronautical fields; operation, holistic lifecycle-based product development, intermodal freight/passenger transportation and so on are only a few keywords of a globalized world constantly in motion. This is reflected in the diversity of the 31 technical papers that are collected in this proceeding out of the 156 oral presentations at FT2019, ranging from low-TRL material research to production, maintenance, failure mitigation and operational aspects, software and avionics engineering to domain-specific or vehicle-holistic design analyses. It is and has always been the strength of the FTF Technology Congress to attract researches, technicians and experts from academia, research institutes and industry alike. This broad background of our contributors and participants enables an intense and broad knowledge exchange and networking which may lead to tomorrows solutions!

At this point we also want to thank our reviewers. Without their support this proceedings would not have been possible.

Dr. Ingo Staack   **Prof. Petter Krus**

Division for Fluid- and Mechatronic Systems
Department for Management and Engineering
Linköping University, Sweden

CONTENTS

I. Aerospace System Analysis

A Requirements Engineering-based Approach for Defining a System Modification Process during Aircraft Operation
Washington Carvalho Tricote Resende, Luís Gonzaga Trabasso and Marina Mendonça Natalino Zenun

Modeling, Simulation and Control of an Aircraft with Morphing Wing
Leonardo Barros da Luz, Wilcker Neuwald Schinestzki, Carlos Eduardo de Souza and Pedro Paglione

Modelling of Transonic and Supersonic Aerodynamics for Conceptual Design and Flight Simulation
Petter Krus and Alvaro Abdallah

An Ontological Approach to System-of-Systems Engineering in Product Development
Ludvig Knöös Franzén, Ingo Staack, Christopher Jouannet and Petter Krus

A study for an MDO Process Applied to Conceptual Design of a Remotely Piloted Aircraft
Maria Luiza Cassão Gatelli, Carlos Eduardo de Souza and Marcos Daniel de Freitas Awruch

A Requirements Engineering-based Approach for Defining a System Modification Process during Aircraft Operation

Washington Carvalho Tricote Resende, Luís Gonzaga Trabasso, and Marina Mendonça Natalino Zenun

Department of Mechanical Engineering, Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, São Paulo/Brazil
E-mail: washingtontricote@gmail.com, gonzaga@ita.br, marina.natalino@hotmail.com

Abstract

From the systems engineering standpoint, an aircraft may be considered a complex system consisting of multiple subsystems operating in an integrated manner, within several operational environments. During the operation, deviations from the design may arise due to the interaction among the aircraft subsystems and their environments. Such deviations require the aircraft manufacturer to redesign a subsystem in order to obtain a solution of high effectiveness and avoid additional costs of a second rework. This paper presents a requirements engineering-based approach to define or modify a process in which an aircraft subsystem is redesigned. The approach is comprised of seven steps, which include the stakeholder, functional and physical analyses previously presented in the total view framework. The proposed activities also include defining the *measures of effectiveness* (MoEs), as well as a *validation and verification* (V&V) strategy for both the system and the requirements. The present work highlights the importance of requirements definition and management during a modification that occurs at an advanced life cycle stage. The steps presented may be seen as guidelines either to modify or define a process. Therefore, the way each step is deployed depends upon the complexity of the system and organizational issues. A case study has been carried out to evidence the performance difference between a system modification process currently used at an aerospace company and the process modified with the approach presented herein. The outcome shows that the new process would potentially reduce the previously found problems by around 60%, and it is believed that this figure would be even greater if the process were actually implemented.

Keywords: systems modification, requirements engineering, complex systems

1 Introduction

A system can be defined as an integrated set of elements, subsystems, or assemblies that accomplish a defined objective. The term *elements* refers to hardware, software, firmware, processes, people, information, techniques, facilities, services, and other support elements. Systems are organized into a hierarchy, consisting of systems elements, which can be broken down into new system elements until an atomic level [1]. A system is considered complex when the interaction between its components can give rise to large events that cannot be located back in the properties or micro-level behaviour of the components. These emergent effects are really difficult to foresee [2]. An aircraft fits the definition of a complex system and, in this paper, it was considered the higher-level system.

Even though the aircraft life cycle presents its stages in a chronological sequence, as shown in fig. 1, several operational environments and systems interactions may lead to the necessity of modifying some subsystems at an advanced life cycle stage. Thus, despite all development effort, the redesign of aircraft subsystems may be needed during aircraft operation.

Modifying a system at an advanced life cycle stage implies a higher financial impact [1]. Therefore, it is extremely important to have a process that enables the development of successful solutions, avoiding additional costs of a second rework. Besides the financial aspect, a possible second rework could also adversely affect customer satisfaction and, consequently, the corporate image.
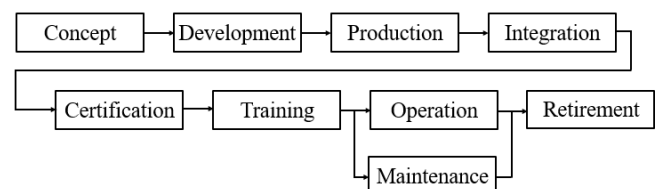

Figure 1: Aircraft life cycle stages. Adapted from [3].

When the operator, e.g. airlines for commercial aviation, reports a system failure to the aircraft manufacturer, the company needs to internalize the problem by involving a

11

multidisciplinary team to support the customer and to design a solution of high effectiveness. Although the way in which the issue is internalized depends on the organizational structure of the company, the engineers are the process stakeholders responsible for the system redesign. Thus, the activities to be performed by the engineers shall be well structured and focus on problem solving. This paper uses the requirements engineering, which is a subset of systems engineering, to define a system modification process to be applied during aircraft operation.

2 Requirements engineering-based approach

The requirements engineering-based approach presented herein makes use of two frameworks: the total view framework and the framework for a *sufficient set of requirements* (SCoRe) [3] [4]. The former was built by combining the concepts of both systems and concurrent engineering. The latter is based on the former, but it emphasizes the definition of a sufficient set of requirements. Instead of covering the whole life cycle of the system, the scope of this work focus on an advanced life cycle stage. Thus, the three-dimensional analysis presented in the total view framework was tailored to fit the modification of a system at the operation stage. The stakeholder, functional and physical analysis, i.e. the three-dimensional analysis, is performed focusing on the problems reported by the customers in order to obtain a sufficient set of requirements. Some authors defended that the set of requirements must be complete [5]. As it may be difficult to identify completeness, the term sufficient was presented in the framework SCoRe and it is used here.

The system modification process proposed is shown in fig. 2. Each step contains several activities that guide the development of a solution of high effectiveness. The first five steps are adapted from the total view framework. The process also includes activities involving the organization of the requirements into a hierarchy and the definition of a verification and validation strategy for both the system and the requirements.
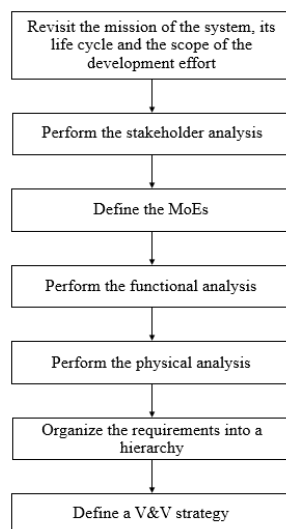
Figure 2: The system modification process

The process presented in fig. 2 can be applied to modify systems of different levels of complexity. When the aircraft is analysed as the higher-level system, it can be broken down into flight controls, avionics, propulsion, environmental control, interior systems, among others. These systems are aircraft subsystems and all of them are subjected to modifications in an advanced life cycle stage. The complexity of the modification will depend on the complexity of the system, thus the seven steps of the system modification process must be adapted according to this complexity. It is equally important to adapt the process taking into account the stakeholders' background and/or company standards. Each step of the system modification process is detailed hereafter.

2.1 First step of the system modification process

This step consists of revisiting the mission of the system, its life cycle and the scope of the development effort. The mission of the system states its purpose or reason of being and guides the integrated process [6]. Although the mission of the system is often defined at the beginning of the life cycle, the process proposes revisiting the mission during a system redesign so that the requirements that guide the modification of the system are aligned with the mission. The scope of the development effort evidences other life cycle stages in which the development organization has partial or total responsibility [6].

In this step, the abstraction level of system hierarchy is defined and all subsequent analyses shall be performed for all layers of the hierarchy. When defining the mission and life cycle of an aircraft subsystem, they are within the mission and life cycle of higher-level systems. Therefore, it may be more interesting to define these concepts for the higher-level system.

Even though this paper focus on the operational stage of the life cycle, this is not the only stage to be considered during analysis. Some problems that occur in operation can be related to production issues and may also occur during the maintenance of the system. Hence, it is important to define the stages to be analysed, as well as their different scenarios, which are either the decomposition of a stage or the alternatives in each stage - e.g. preventive and corrective maintenance [6]. The activities within the first step are presented in fig. 3.

Figure 3: Inputs, activities and outputs of the first step

2.2 Second step of the system modification process

In this step, the stakeholder analysis is performed, in which the system and organization stakeholders are identified for each level of system hierarchy and for each stage of the life

cycle. Stakeholder concerns are used to define the stakeholder requirements [6]. All requirements must be written following a standard so that they are easily and equally understood by all process stakeholders. One way of standardising the language used for requirements is using boilerplates throughout the process [7]. All activities contained in the second step are shown in fig. 4.
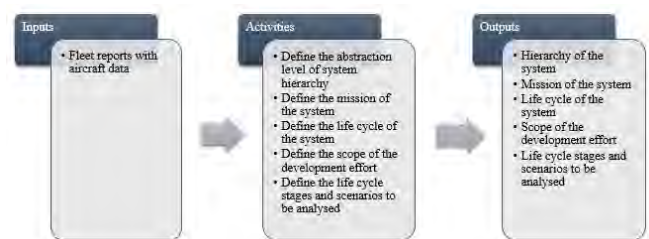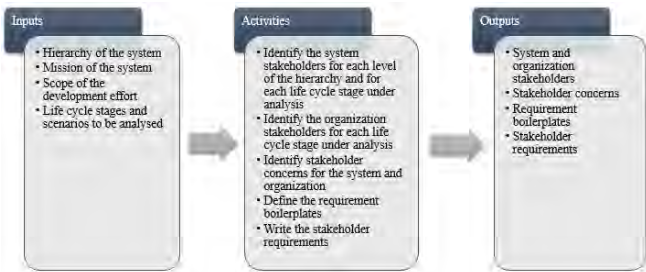

Figure 4: Inputs, activities and outputs of the second step

The definition of stakeholder requirements is within the problem domain, in which any reference to a particular solution is avoided in order to allow the engineers to devise the best solution without preconceived ideas. Functional and physical characteristics of the system are defined afterwards in the solution domain and are stated through system requirements and architectural design [7].

2.3 Third step of the system modification process

From the system and organization stakeholders and their concerns, the MoEs and their acceptance criteria are defined, as shown in fig. 5. The MoEs measure how the system meets the stakeholder requirements [6].

Figure 5: Inputs, activities and outputs of the third step

2.4 Fourth step of the system modification process

The functional analysis is within the solution domain and it is performed in this step. The environment elements that interact with the system are identified, as well as *information* (INF), *material* (MAT), and *energy* (ENR) exchanged between these elements and the system. The functional context diagrams are made for the system and organizations for each life cycle stage under analysis. For the system, each hierarchy level must be considered [6].

To completely represent the system functioning, the systems engineers' activities also include doing the list of events, the functional structure diagram, and the state transition diagram. This way, at the end of this step, the engineers will be able to know the environment elements that interacts with the system, the way in which the interactions occur, as well as the functions and states of the system. With all this information, the first subset of system requirements is written. In this step, the interface requirements among the system and

environment elements are captured. All activities contained in this step are listed in fig. 6.

Figure 6: Inputs, activities and outputs of the fourth step

2.5 Fifth step of the system modification process

The physical analysis consists of making the architecture context diagram by identifying the physical components responsible for the exchange of INF, MAT and ENR between the system and the environment elements [6]. The second subset of system requirements is defined. The activities to be performed in this step are presented in fig. 7.

Figure 7: Inputs, activities and outputs of the fifth step

2.6 Sixth step of the system modification process

At this point, all stakeholders and systems requirements will have been defined. Thus, it is necessary to organize all these requirements into a unique hierarchy, allowing correct requirements traceability, as shown in fig. 8. Stakeholder requirements are at the highest level and there is no hierarchical difference among them. On the other hand, system requirements are organized according to the system hierarchy. The highest level system requirements are linked to stakeholder requirements.
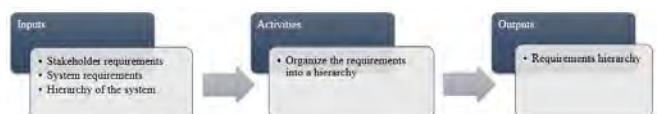
Figure 8: Inputs, activities and outputs of the sixth step

If the system being modified was developed in accordance with a requirements baseline at the development stage, a new baseline shall be created to add the new set of requirements obtained in this step. If a new requirement is equivalent to an existing one and the reported problems show that the system did not meet this requirement, it indicates that the system and/or requirements were not correctly verified and validated, thus an adequate V&V strategy must be defined. A traceability analysis shall also be performed to identify the links between the new requirements and existing ones. If an existing requirement is affected, the need to modify, revalidate and reverify that requirement must be assessed.

2.7 Seventh step of the system modification process

In this step, a V&V strategy is defined following the activities shown in fig. 9. Validation of a requirement consists of ensuring that the requirement is justified and relevant to the stakeholders needs. Validating a system means to demonstrate that it meets the requirements, mission or business profile of the stakeholders, as well as their operational scenarios. The stakeholders shall be involved in the validation process. Verifying a requirement is to check the way it is written and the application of its desired characteristics, such as unambiguity, consistency, completeness, uniqueness, traceability, verifiability, among others. Verifying a system is the activity of checking whether its characteristics or properties are in accordance with requirements, architecture and design [1].
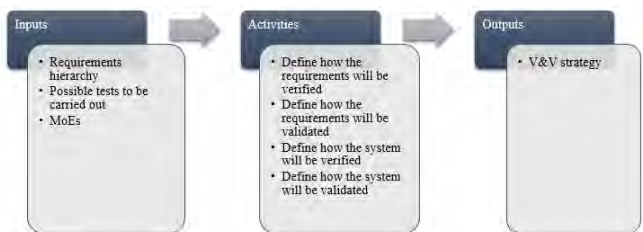
Figure 9: Inputs, activities and outputs of the seventh steps

3 Case study

The process of modifying an interior system is different from other systems in terms of measuring customers' perception. Stakeholder concerns about interior systems, e.g. concerns involving appearance and comfort aspects, are often qualitative, which makes measuring the effectiveness of the solution a difficult and time consuming task.

This particularity of interior systems can make the standardization of the system modification process more difficult, and the process becomes dependent on the expertise of its stakeholders. In order to avoid this dependence and to exemplify the approach presented in this paper, a folding table installed on an executive aircraft was chosen to be analysed in the case study. Some customers have reported problems in the folding table of an executive aircraft after the modification of the system through the AS-IS process. The problems were mostly related to damage to the finishing material and to issues involving the functionality of the mechanism.

The current system modification process, i.e. the AS-IS process, used to modify the folding table has been adapted by one of the authors to include the seven steps previously presented in fig. 2. The AS-IS process was modelled according to a methodology that divide process modelling into three phases: data collection, definition of the AS-IS model, and validation of the model. The first phase consists of defining the goal and the detail level of the model, as well as analysing the process. The second phase involves selecting the modelling tools and defining the AS-IS process design. In the third phase, the model is tested and, if needed, some modifications are made [8]. The AS-IS process deficiencies were identified from the information obtained from the

interviews with the process stakeholders and from a value analysis of each activity. The deficiencies were prioritize in a *gravity, urgency and tendency* (GUT) matrix and the activities related to them were replaced by the seven requirements-driven steps. The methodology used herein to model a system modification process integrates process modelling with requirements engineering in the TO-BE process definition. This methodology is shown in fig. 10.
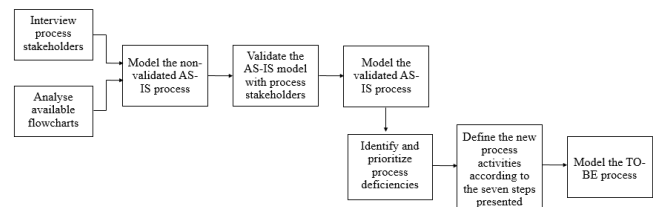
Figure 10: Methodology used to model the AS-IS and TO-BE processes

The activities of the TO-BE process have been carried out by one of the authors to simulate the activities to be performed by the engineers if the process were actually implemented. After that, the set of requirements obtained was compared against the reported problems, showing that the TO-BE process addressed several problems that have not been identified through the AS-IS process.

The outcomes of each step for the folding table in operation are presented hereafter. According to the system hierarchy, the folding table can be considered a subsystem of the interior, which in turn is a subsystem of the aircraft, as shown in fig. 11.
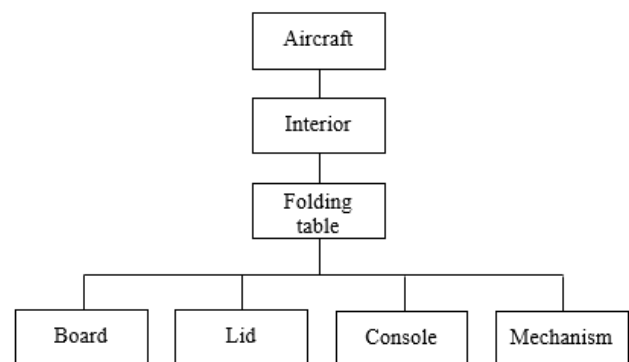
Figure 11: System hierarchy

3.1 System mission, life cycle and scope of development effort

In this case study, the mission has been defined in the second level of the hierarchy, because the folding table mission is within the interior mission, which is *to provide comfort, practicality, entertainment and safety to passengers from start to finish of the trip*. As the life cycle of an aircraft includes all its subsystems, the engineer can use the stages previously presented in fig. 1. The responsibility for the development of interior systems is shared between the aircraft manufacturer and the interior supplier. Thus, both companies are considered development organizations. The aircraft

manufacturer also participates solely or partially in the production, certification, training and maintenance stages, characterizing its broad scope of the development effort.

3.2 Stakeholder analysis

Based on boilerplates available in the literature [7] [9], the requirement boilerplates shown in Table 1 have been defined and used in writing requirements. The placeholders are the parts that vary from requirement to requirement.

Table 1: Requirement boilerplates for the case study

Domain	Boilerplate
Problem	The *<stakeholder type>* shall be able to *<capability>*
	While *<operational condition>*, the *<stakeholder type>* shall be able to *<capability>*…
	… within *<performance>* of *<event>*
	… at least *<quantity>* times per *<unit time>*
	… for a period of at least *<quantity>* *<unit time>*
Solution	The *<system>* shall (not) *<system response>*
	When *<optional precondition>* *<trigger>*, the *<system>* shall (not) *<system response>*
	If *<optional preconditions/trigger>*, then the *<system>* shall (not) *<system response>*
	While *<in an specific state>*, the *<system>* shall (not) *<system response>*
	Where *<feature in included>*, the *<system>* shall (not) *<system response>*

Figures 12 and 13 show the context diagram of the aircraft interior and the folding table in operation, respectively. The elements of the diagram are the stakeholders of the system in operation and their concerns are listed in bullet points.
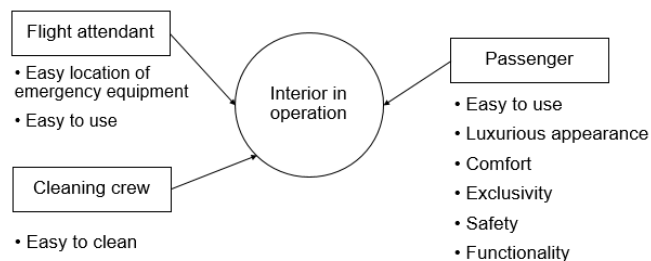
Figure 12: Stakeholders of the aircraft interior in operation and their concerns
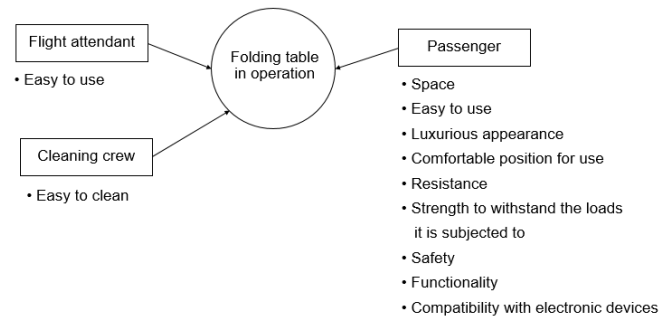
Figure 13: Stakeholders of the folding table in operation and their concerns

Through the diagrams of fig. 12 and 13, two requirements were obtained from the stakeholder *passenger* and his/her concern about the *functionality* and *luxurious appearance* of the system. As these concerns are the same for both the aircraft interior and the folding table, the requirements were stated for the higher level system, so any other interior subsystem should also meet these requirements. Therefore, no differentiation was made between hierarchy levels for stakeholder requirements. The requirements were stated as follows:

- The passenger shall be able to use all aircraft interior features on all flights *(type of requirement: capability)*.
- The passenger shall be able to feel satisfied with the appearance of the aircraft interior *(type of requirement: capability)*.

3.3 MoEs

Table 2 shows the MoEs defined from the passenger's concerns about the *resistance* of the folding table. These MoEs illustrates the particularity of interior systems in quantifying customers' perception. The concept of no permanent damage may vary from customer to customer and the number of cycles can be defined based on flight data and laboratory tests, which must consider the contamination of the finishing material by substances contained in drinks and foods. The normal conditions of use defined in the criteria must be well defined and documented.

Table 2: MoEs for the folding table

MoE	Criterion
Resistance to scratching	No permanent scratches on the finishing material after X cycles, under normal conditions of use
Resistance to pressure marks	No permanent dimples in the finishing material after X cycles, under normal conditions of use
Stain resistance	No permanent stains on the finishing material after X cycles under normal conditions of use

Abrasion resistance	No noise indicating abrasion of the components of the mechanism after X cycles, under normal conditions of use

3.4 Functional analysis

Figures 14 and 15 show the functional context diagram for the interior and folding table in operation, respectively. The analysis has been performed for both levels of hierarchy to properly capture the interface requirements between the folding table and the systems that interact and interfere in its operation. In the diagrams, the direction of the arrows indicates the direction in which ENR and MAT are exchanged between the system under analysis and other systems or stakeholders.
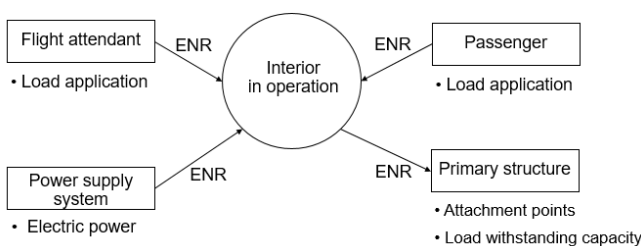
Figure 14: Functional context diagram for the aircraft interior in operation
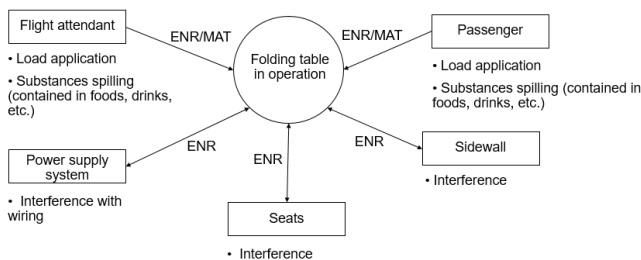
Figure 15: Functional context diagram for the folding table in operation

Table 3 presents the list of events to identify the system functions. The functional structure diagram related to the list of events is shown in fig. 16, whereas the state transition diagram is shown in fig. 17.

Table 3: Event list to identify system functions

System in operation	
System stimulus	System response
1. User press the lid	1. The system opens the lid
2. User press the table board	2. The system partially lifts the table board
3. User pull up the table board	3. The system locks the table board when it is completely released
4. User moves the table board down	4. The system locks the table board horizontally
5. User unfolds the table board	5. The system locks the outer part of the table board horizontally
6. User makes use of the table	6. The system withstands user-imposed loads
7. User folds the outer part of the table board	7. The system retracts the outer part over the inner part of the table board
8. User raises the table board to the upright position	8. The system opens the lid
9. User stows the table board	9. The system locks the table board inside the console
10. User closes the lid	10. The system is not in use
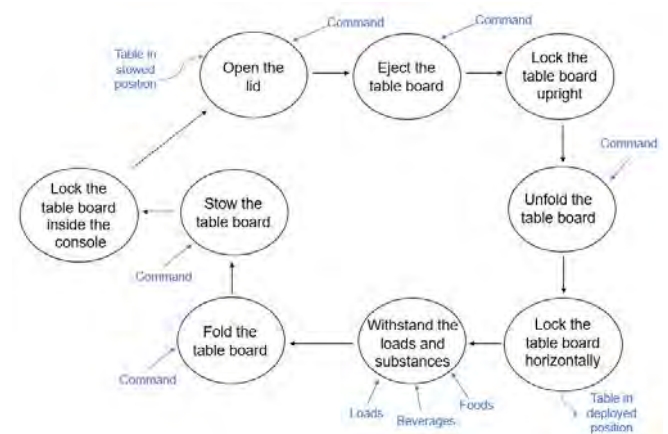
Figure 16: Functional structure diagram of the folding table
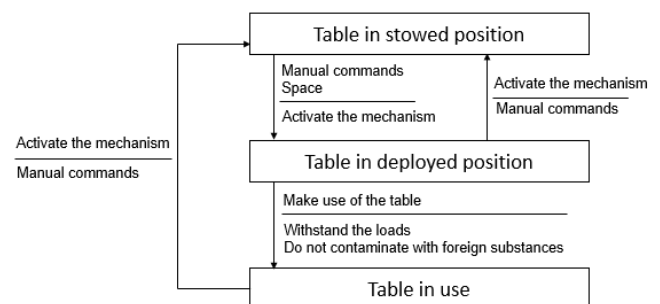
Figure 17: State transition diagram of the folding table

The following requirements have been identified through the functional analysis:

- While under normal conditions of use, the aircraft interior shall operate for X cycles without damage to the finishing material *(type of requirement: functional)*.
- While under normal conditions of use, the aircraft interior shall operate for X cycles without damaging its functionality *(type of requirement: functional)*.

- While being stowed by the user, the table shall not have its movement blocked by contact between any parts *(type of requirement: interface/constraint)*.
- When the table board reaches a position of $(X \pm Y)°$, the lid shall be in vertical position (90°) *(type of requirement: interface)*.
- While the table is stowed, the table shall not have contact between parts with finishing material *(type of requirement: constraint)*.
- While the table is stowed, the outer part of the table board shall not have contact with the inner part of the table board *(type of requirement: interface/constraint)*.
- The table shall not have contact between its metal parts and finishing material *(type of requirement: interface/constraint)*.
- While the table is stowed, the table board shall not have contact with the mechanism *(type of requirement: interface/constraint)*.

Regarding the type of requirement, the requirements above were classified as interface requirements because they state the functional relationships that exist among the folding table elements. Requirements whose text contains *shall not* have also been classified as constraints.

3.5 Physical analysis

Figures 18 and 19 show, respectively, the architecture context of the interior and the folding table in operation, in which the physical components responsible for the exchange of ENR and/or MAT were identified.
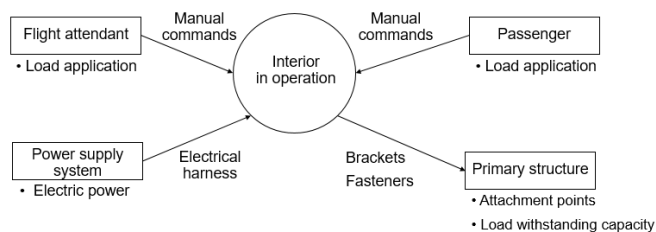
Figure 18: Architecture context for the aircraft interior in operation
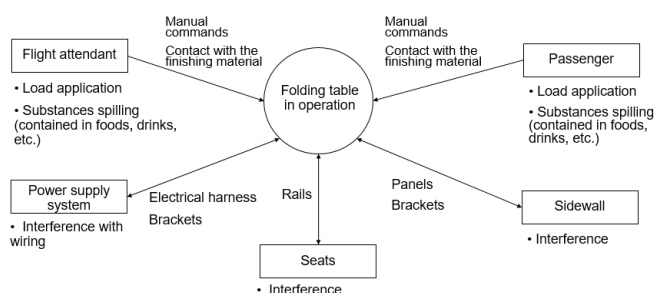
Figure 19: Architecture context for the folding table in operation

The following requirements have been identified through the physical analysis.

- The table shall absorb the impacts coming from contact with the sidewall panel *(type of requirement: interface)*.
- The table shall not have abrasion damage after X cycles (type of requirement*: constraint)*.

These requirements were classified as interface or constraint requirements, similar to the classification used in the functional analysis. The difference here is that the interface requirement states a physical relationship between the folding table and the sidewall, which is an environment element.

3.6 Requirements hierarchy

The stakeholder requirements and the set of system requirements were organized into a hierarchy, as shown in fig. 20. An identification number was assigned to each requirement, in which the first digit corresponds to the hierarchy level as follows: 1) first level – stakeholder requirements; 2) second level – interior requirements; 3) third level – folding table requirements; 4) fourth – table elements requirements. The second digit is used to differentiate the requirements within the same level.
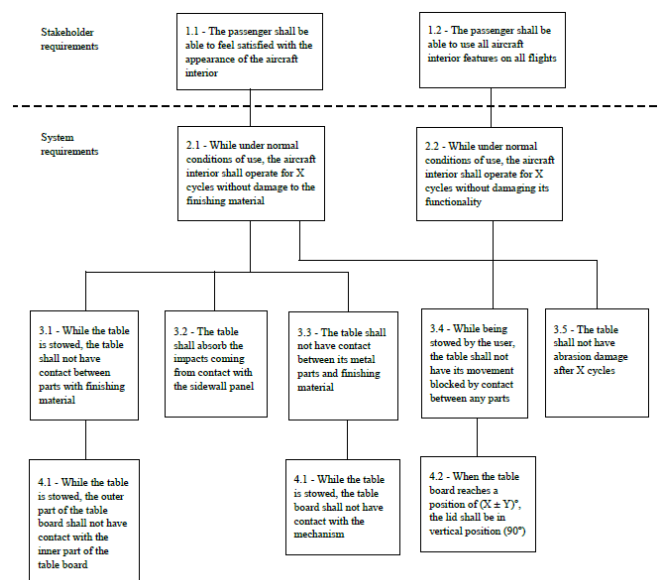
Figure 20: Requirements hierarchy

The requirements hierarchy focus on finishing materials and mechanism interfaces issues to solve the problems reported from customers. However, many other requirements could be identified through the stakeholders, functional and physical analyses. Thus, for each case, the engineers will guide the direction of the analyses with the aim of addressing the reported problems. If some reported problem is related to any organizational issue, all analyses performed for the system must also be carried out for the organizations.

In case the aircraft subsystem is developed and manufactured by a supplier, both the aircraft manufacturer and the supplier will be involved in the system modification process. In this case study, for instance, the folding table supplier should

break down the folding table requirements into lower-level requirements until the level of components.

3.7 V&V strategy

The system and the requirements could be verified and validated as follows:

Requirements verification: the requirements have the characteristics of a good requirement (e.g. unique, traceable, unambiguous, etc. [5]) and have been written using the boilerplates of Table 1.

Requirements validation: the requirements are relevant to the needs of the stakeholders in the hierarchy shown in fig. 20, in which all system requirements are linked to stakeholder requirements.

System verification: this process must begin at the lowest level of the hierarchy and consists of verifying that the system meets the system requirements. The requirements could be verified through analyses, tests, simulations, traceability, among other methods.

System validation: the system meets the stakeholder requirements by complying with their children requirements, as well as by achieving the effectiveness expected by the customer, defined through the MoEs.

3.8 Outcome

A set of requirements was obtained after executing the activities of the proposed system modification process. The reported system problems were compared against these requirements. In case the problem would be avoided by meeting one or more requirements, it was considered that the problem would be solved. The outcome shows that the new process would potentially avoid 60% of the problems. It is believed that this figure would be even greater if the process were actually implemented. This is because the stakeholder, functional and physical analyses would not be carried out by only one person, but rather by a multidisciplinary team with the support of all process stakeholders to obtain a sufficient set of requirements.

4 Conclusion

The approach presented in this paper can be used by an organization to define or improve its system modification process in order to obtain a solution of high effectiveness. By achieving the effectiveness desired by customers, additional costs of a second rework are avoided and customer satisfaction is increased. The steps of the system modification process are comprehensive and can be tailored to fit different types of systems and organizations.

In addition to contributing to the industry by presenting a system modification process, this paper also contributes to the literature by presenting the use of requirements engineering to modify systems at an advanced life cycle stage. Instead of covering the whole life cycle, this work showed the importance of defining new requirements to modify an aircraft subsystem at the operation stage.

The steps of the system modification process proposed herein proved viable with the case study outcome, which showed a potential reduction of the problems reported during the system operation.

References

[1] INCOSE, Systems engineering handbook: a guide for system life cycle processes and activities, 4th red., San Diego: John Wiley & Sons, 2015.

[2] S. Dekker, Drift into failure: from hunting broken components to understanding complex systems, Boca Raton: Taylor & Francis, 2011.

[3] M. M. N. Zenun, *Modelo para obtenção do conjunto suficiente de requisitos no desenvolvimento de sistemas de aeronaves,* São José dos Campos, SP: Instituto Tecnológico de Aeronáutica (ITA), 2015.

[4] G. Loureiro, *A systems engineering and concurrent engineering framework for the integrated development of complex products,* Loughborough: Loughborough University, 1999, p. 362.

[5] A. T. Bahill och A. M. Madni, Tradeoff Decisions in System Design, Los Angeles, CA: Springer, 2017.

[6] G. Loureiro, A. G. Adinolfi, C. E. V. Ribeiro och R. C. B. d. Andrade, "System concurrent engineering for the development of an aeronautical navigation system," *Product: Management & Development,* vol. 8, nr 2, pp. 107-122, 2010.

[7] E. Hull, K. Jackson och J. Dick, Requirements engineering, London: Springer, 2011.

[8] C. S. T. Amarala, H. Rozenfeldb, J. M. H. Costab, M. d. F. d. A. Magon och Y. M. Mascarenhas, "Improvement of radiology services based on the process management approach," *European Journal of Radiology,* nr 78, p. 377–383, 2011.

[9] J. Dick och J. Llorens, "Using statement-level templates to improve the quality of requirements," i *24th International Conference on Software & Systems Engineering and their Applications*, Paris, 2012.

Modeling, Simulation and Control of an Aircraft with Morphing Wing

Leonardo Barros da Luz[*], Wilcker Neuwald Schinestzki[*], Carlos Eduardo de Souza[**], and Pedro Paglione[†]

[*]Aerospace Engineering Graduation Course, Federal University of Santa Maria, Santa Maria, Brazil
[**]Mechanical Engineering Department, Federal University of Santa Maria, Santa Maria, Brazil
[†]Visiting Professor, Federal University of Santa Maria, Santa Maria, Brazil
[1]E-mail:leobdl.formula@gmail.com, wilcker.formula@gmail.com, carlos.souza@ufsm.br, pedro.paglione@gmail.com

Abstract

The technological advances, mainly in the development of new materials, recovered the interest in the application of morphing wings in aircraft. Due to the potential of replacing conventional control surfaces by morphing surfaces, the present work presents the modeling of aerodynamics, dynamics and control design of an aircraft with morphing wings. The morphing concept is given by changing the camber of the trailing edge along the wingspan. For aerodynamics modeling, it was adopted unsteady strip theory and, for dynamics modeling, it was used rigid body mechanics, considering the displacement of the center of mass and the time-varying inertia tensor. Finally, control design is performed using Exponential Mapping Controller (EMC) method. The results showed that, for the adopted variable geometry configuration, the influence of the center of mass displacement and the inertia variation on the aircraft behavior were insignificant, whereas the influences of the unsteady aerodynamics were significant. Consideration of the unsteady aerodynamic effects increases the magnitude of the aircraft movements, necessitating a greater control action.

Keywords: Morphing wings, Modeling, Aerodynamic, Dynamics, Control

1 Introduction

The development of aircraft over the years led to its use in a variety of applications. The research about flight performance improvement is inserted in this perspective, due to the spread in vehicular electronic computation capacity, what allows, today, for fast computations at a very low cost, and consequently, use in any kind of aircraft.

One of the alternatives for performance improvement relies on the development of more efficient aerodynamic flight controls, specially the variable geometry mechanisms concept. Those are the mechanisms that are designed to adapt itself to changes in the mission environment [1]. The development of such mechanisms seeks inspiration from nature, specially in the birds flight, which achieved a very efficient condition because of biological evolution [2].

The inspiration in nature was already present in the early years of aviation development. In the XIX century, many visionaries developed bird inspired mechanisms that allowed geometry changes aiming performance improvement or flight control [3]. However, the demand for aircraft made of stiffer materials in the place of flexible ones, prevented their further development, and those mechanisms could not be found anymore in most aircraft [2].

Nowadays, the use of variable geometry mechanisms is possible again because of technological development in different fields, specially in the materials and structures area, what brought the possibility of designing flexible yet fail safe structures. The work of [4] described many aircraft designs that consider variable geometry concepts. On example is the F-14 fighter, which alters its sweep in flight to achieve better performance on different flight phases. The more recent work of [5] describes the development of an aircraft that has all its geometry variable using an innovating structural concept.

1.1 The variable geometry devices

The variable geometry devices can be divided into two categories: discrete and continuous [2]. The discrete ones are those found in conventional aircraft, which have single functionalities operating in some flight phases, such as the flaps, slats and rectractable landing gears. The continuous ones usually have multiple functionalities, operating in different flight phases. One example is the bird wing.

1.2 The Categories of Continuous Changes in Wing Geometry

In [4], the different ways of changing wing geometry are divided between changing shape in wing plan, out of wing plane and in profile. Changes in wing plan shape include changes in wingspan, chord and fluff, while out-of-plane shape changes include variations in torsion, dihedral and flexion. For profile changes, the parameters that vary are arching and thickness.

For the profile change, [6] presents two distinctions for variations in arching, leading edge variation and trailing edge variation.

The work of [7] presents a very interesting concept of a variable geometry wing, which can be characterized in the category of change in trailing edge bending, as [6]. The concept is given the Spanwise Morphing Trailing Edge (SMTE), which consists of changing the incidence of the trailing edge smoothly over the entire wingspan. To make this smooth transition, active and passive surfaces are used, active surfaces are responsible for the trailing edge movement, while passive surfaces transition to the other active surface.

To carry out the studies of this work, we opted for the SMTE concept of [7], because it has a high application potential, directly impacting the autonomy of the aircraft.

1.3 The Advantages of Using Variable Geometry Wing

The use of wings with variable geometry can mainly affect the aerodynamics and its control, seeking to increase its performance. In the work of [6] the advantages of each of the variable geometry wing categories are presented. The advantages for different types of variable geometry are listed in the items below.

- **Variable camber**: It is capable of changing the lift distribution, having advantages in the takeoff and landing phases and can be applied at the trailing or leading edge. Lead edge application can be a lower noise and drag alternative to conventional slats. Application at the trailing edge can reduce drag, making control surfaces more efficient.

- **Variable thickness**: You can change the drag of the profile by changing its thickness, directly impacting the location of the transition point from laminar to turbulent regime.

- **Variable wingspan**: Aircraft with high elongation have low maneuverability and high aerodynamic efficiency, however those with low aspect ratios have good maneuverability and low aerodynamic efficiency. The variable wingspan can enjoy the advantages of low aspect ratios and high aspect ratios.

- **Variable sweep**: Variable sweep can combine the advantages of a non-sweep wing at low speed, take-off and landing stages with the advantages of sweep wings at high speed speeds.

- **Variable twist**: Variable twisting can relieve maneuvering and bursting loads. In addition, the twisting of the wing may alter the lift distribution along the wingspan and may have the function of a control surface.

In the experimental work of [8] data were obtained that demonstrated an improvement in the aerodynamic efficiency of the aircraft with the use of a variable wingspan, showing a 17 % increase in autonomy. In addition, FlexSys, founded in 2000 by Dr. Sridhar Kota, indicates that using its FlexFoil technology can reduce drag by a range of 5 % to 12 % for long-range fixed wing aircraft, representing a huge fuel economy [9]

1.4 Modelings

To fulfill the objectives of the work, the modeling of kinematics and dynamics, aerodynamics and control is required. Therefore, in the subsections below, the literature review of the kinematics and dynamics modeling applied to aircraft with variable geometry in subsection 1.4.1, aerodynamics in subsection 1.4.2 and the control application in subsection 1.4.3.

1.4.1 Kinematics and Dynamics Modeling

To perform the dynamics modeling of a variable geometry wing aircraft, one can choose two different techniques, adopt the aircraft with a single rigid body or as the union of several rigid bodies.

In the works of [10] and [11] are presented modeling of the dynamics of a wing aircraft with variable geometry using rigid body mechanics. However, some effects that arise when adopting a change in geometry should be taken into account. The main effect would be a significant change in the center of gravity and moments of inertia, so the inertia tensor is a function in time and has a dynamics, which may be close to the frequency ranges of rigid body dynamics.

Due to the considerations provided by [11], it is clear that the equations of standard rigid body motion cannot be applied. An alternative would be to use multibody dynamics methods, but depending on the choice of method one can find large systems of equations, resulting in a significantly higher computational cost. So in [12] another approach to modeling is provided which consists of continuing to treat the aircraft as a single body but utilizing the displaced center of mass of the origin and relaxing the stiffness condition, making the inertia tensor an explicit function in time. This approach may be a more efficient alternative, because its computational cost is lower compared to the dynamics of multiple bodies, and due to these factors, this approach was chosen for this work.

In [13] aircraft dynamics modeling is performed using a tool called SimMechanics, present in Matlab software, which uses the standard Newtonian dynamics of forces and torques. SimMechanics is a block diagram modeling environment where a body can be modeled by joining multiple rigid bodies through joints.

1.4.2 Aerodynamic Modeling

Aerodynamic methods applied to wings with variable geometry are divided into two broad categories, stationary and non-stationary. Moreover, within these categories one can have a division between linear and nonlinear methods.

The linear methods, both stationary and non-stationary, are based on potential flow theory, having a restricted application to thin airfoils and small angles of attack. To overcome these constraints of linear methods, nonlinear aerodynamic methods such as Computational Fluid Dynamics (CFD) are used.

These, however, require a higher computational cost.

In [6] a review of the different aerodynamic methods used in wing modeling with variable geometry was performed. For the finite wing modeling within the stationary method category, the most commonly used are the Vortex Lattice Method (VLM) and the nonlinear vortex lattice method. For the non-stationary method category, the use of Unsteady Vortex Lattice Method (UVLM) and Doublet Lattice Method (DLM) is highlighted. It is also noted the large use of CFD, which may be non-stationary and stationary, the tendency of increasingly using CFD in these studies is due to its ability to obtain results very close to experimental, especially in the nonlinear region. Finally, [6] concludes that the UVLM method has great potential for modeling involving bending, thickness, torsion and wingspan alterations.

In the aerodynamic modeling performed by [14], we highlight the use of CFD with a Spalart-Allmaras turbulence model, a one-equation model, which has a lower computational cost than other models. turbulence. In [15], the turbulence model used was $k - omega$, a two equation model, which requires a higher computational cost than Spalart-Allmaras, but the results obtained were very close to those obtained. experimental results.

In the work of [16] the aerodynamic method used is strip theory, with the generalized non-stationary theory of Theodorsen applied to an airfoil. The results for the theory's predicted flutter velocity and frequency were very close to the experimental ones.

In the work of [17] a comparative study is carried out, evaluating loads in flight, between quasi-stationary band theory, Vortex Lattice Method (VLM), non-stationary band theory and Doublet Lattice Method (DLM). The author concludes that non-stationary band theory is the best candidate for use in initial analysis, considering the relationship between computational cost and correct prediction of aerodynamic effects. So, in this work, it was decided to use the non-stationary band theory, with modifications to include inflection, dihedral and profile characteristics.

1.4.3 Control Application

There are different control techniques that can be applied to control aircraft with variable geometry, but some complications arise due to the dependence of dynamics on changing geometry. There are two ways to handle control of aircraft with variable geometry, as suggested by [11]. Variable geometry can be considered as a configuration change requiring different controllers in each configuration, or variable geometry can be studied as the control method. Using variable geometry as control effectors, problems of non-unique solutions arise as the number of control variables. Due to the problem of non-unique solutions, the optimal control allocation method may be an alternative.

In [11] some complications are presented in the control design phase of a variable geometry wing aircraft, as in one of the primordial stages of the control design, the system linearization. In the linearization step a linear system of a dynamics

described as

$$\dot{x} = f(x, u) \tag{1}$$

However, for an aircraft with variable geometry, dynamics also become a function that depends on the μ geometry change setting and the $\dot{\mu}$ surface rate of change. Therefore, the dynamics should be described as:

$$\dot{x} = f(x, u, \mu, \dot{\mu}) \tag{2}$$

In the work of [6] a review of the control techniques used in aircraft with variable geometry is presented, highlighting the use of the pseudo-inverse allocation method and quadratic programming allocation within the actuator constraints. In [10] an optimal control technique is used for different flight conditions, with small performance losses in off-project conditions.

For this work, we opted for the Exponential Mapping Controller (EMC) method, developed by [18]. This method is based on the Sliding Mode Control (SMC) and Neuro-Fuzzy Control (NFN) methods, combining some advantages of both methods, demonstrating an excellent ability to solve dynamic control problems with terms. variations in time with some ease.

2 Mathematical models

This section presents the description of the main mathematical models used in the proposed framework. The first model presented in subsection 2.1 is the aircraft's kinematics and dynamics. Modeling the complete dynamics of the aircraft requires the aerodynamic forces, which are provided by the aerodynamic model presented in subsection 2.2. Finally, the control design model is presented in subsection 2.3, which will be responsible for commanding and stabilizing the aircraft.

2.1 Kinematics and Dynamics Model

The aircraft motion is represented by the translation and rotation of the body reference system (BRS) relative to the inertial reference system (IRS). The BRS, illustrated in Fig. 2.1, is defined with the axis X_b poiting in the direction of the aircraft's nose, axis Y_b on the right side of the aicraft, when looking in the positive direction of axis X_b, and axis Z_b points downward. Its origin is displaced from the center of mass of the aircraft. The IRS is fixed at the initial position of the aircraft and at the coincides with the BRS at $t = 0$.

The translation is characterized by the position vectors $\mathbf{R}_0$, $\mathbf{R}_{cm}$ and $\mathbf{r}_{cm}$, whereas the rotation is by the attitude of the BRS towards the IRS. The attitude is obtained using the angles of Euler ϕ, θ, ψ defined with the rotation along the x, y and z axis, respectively. The attitude matrix is given by the rotation sequence z-y-x.

The forces and moments acting on the reference systems are defined according to fig. 2, with the positive moments and
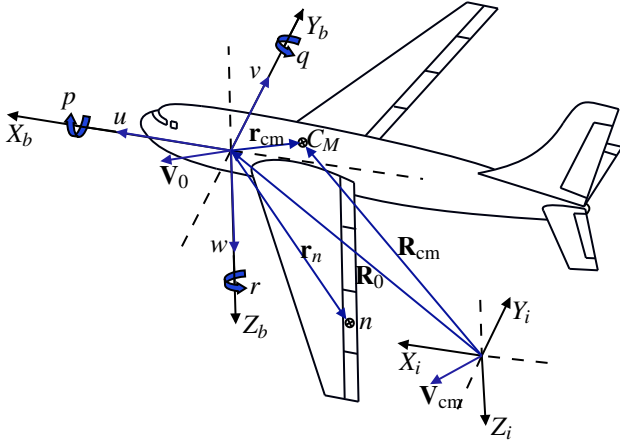
Figure 1: Definition of inertial reference system and body.

body forces in the direction of the BRS axes and the positive aerodynamic forces in the opposite direction of the SAR's axes.
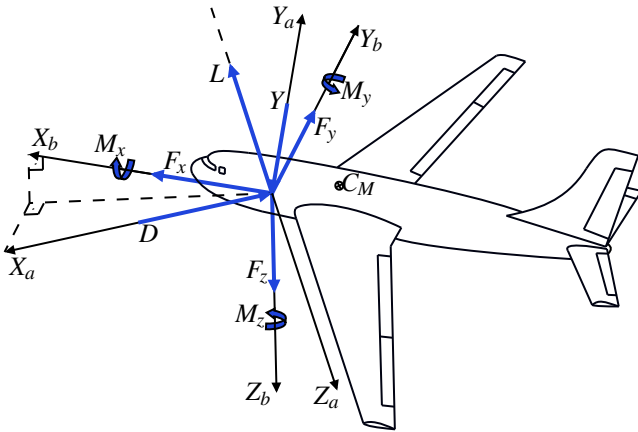

Figure 2: Representation of forces and moments acting on reference systems.

2.1.1 Dynamics

The translation dynamics considers the external forces acting on the body according to fig. 2 and the velocities u, v and w according to fig. 2.1. Furthermore, write the velocity $\mathbf{V}_{cm}$ of the center of mass, illustrated in fig. 2.1, with respect to IRS as:

$$\mathbf{V}_{cm} = \mathbf{V}_0 + \dot{\mathbf{r}}_{cm} + \boldsymbol{\omega} \times \mathbf{r}_{cm} \tag{3}$$

where $\mathbf{V}_0$ is the inertial velocity of the BRS origin with components u, v and w, $\boldsymbol{\omega}$ the angular velocity vector of the body with components p, q and r, respectively, $\mathbf{r}_{cm}$ the center of mass position vector relative to the origin of the BRS.

Using Newton's second law, assuming the aircraft has constant mass, then the resulting external force must be equal to the product of the total mass of the aircraft, m, by the time-derived center of mass velocity, $\dot{\mathbf{V}}_{cm}$, so that:

$$\mathbf{F}_{ext} = m\dot{\mathbf{V}}_{cm} \tag{4}$$

By deriving the eq. (3) in time, we obtain, after handle the terms, the translation dynamics for the origin of the BRS, in relation to the inertial system, written as:

$$m\dot{\mathbf{V}}_0 + m(\boldsymbol{\omega} \times \mathbf{V}_0) + m\ddot{\mathbf{r}}_{cm} + 2m(\boldsymbol{\omega} \times \dot{\mathbf{r}}_{cm})$$
$$+ m\dot{\boldsymbol{\omega}} \times \mathbf{r}_{cm} + m(\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r}_{cm}) = \mathbf{F}_{ext} \tag{5}$$

where $\mathbf{F}_{ext}$ is the vector of external forces with components F_x, F_y and F_z, respectively, and m is the total mass of the aircraft.

For the rotation dynamics, the rate of change of the amount of total external angular motion is obtained as presented by [12],

$$\mathbf{M}_{ext} = \dot{\mathbf{h}} + m\left(\mathbf{r}_{cm} \times \dot{\mathbf{V}}_0\right) \tag{6}$$

where $\mathbf{M}_{ext}$ is the total external moment vector with components M_x, M_y and M_z, respectively, and $\mathbf{h}$ is the angular momentum BRS, which is expressed as:

$$\mathbf{h} = \overline{\mathbf{J}} \cdot \boldsymbol{\omega} + \sum_{i=1}^{N_n} (m_n)_i \left((\mathbf{r}_n)_i \times (\mathbf{v}_n)_i\right) \tag{7}$$

In the above equation, $\overline{\mathbf{J}}$ is the aircraft's inertia matrix, $\mathbf{r}_n$ the mass element position vector n shown in fig. 2.1, $\mathbf{v}_n$ the velocity vector of the mass element, m_n the mass of the mass element, and N_n is the amount of mass elements.

Solving the derivative of eq. (7), substituting the result in eq. (6) gives the equation of rotation dynamics described as

$$\overline{\mathbf{J}} \cdot \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \overline{\mathbf{J}} \cdot \boldsymbol{\omega} + \dot{\overline{\mathbf{J}}} \cdot \boldsymbol{\omega} + m\mathbf{r}_{cm} \times (\dot{\mathbf{V}}_0 + \boldsymbol{\omega} \times \mathbf{V}_0)$$
$$+ \boldsymbol{\omega} \times \sum_{i=1}^{N_n} (m_n)_i \cdot (\mathbf{r}_n)_i \times (\mathbf{v}_n)_i + \sum_{i=1}^{N_n} (m_n)_i (\mathbf{r}_n)_i \times (\dot{\mathbf{v}}_n)_i$$
$$+ \sum_{i=1}^{N_n} (m_n)_i (\dot{\mathbf{r}}_n)_i \times (\mathbf{v}_n)_i = \mathbf{M}_{ext} \tag{8}$$

Analyzing the eqs. (5) and (8) we see a dynamic coupling, due to the second derivative of the vector $\mathbf{r}_{cm}$, giving rise to the term $\dot{\boldsymbol{\omega}}$ in the translation dynamics. This coupling is solved using the approach presented by [19], which consists in solving a coupled linear system, defined as

$$\begin{bmatrix} m\overline{\mathbf{I}} & \overline{\mathbf{S}}_R^T \\ \overline{\mathbf{S}}_R & \overline{\mathbf{J}} \end{bmatrix} \cdot \begin{bmatrix} \dot{\mathbf{V}}_0 \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_F \\ \mathbf{Q}_M \end{bmatrix} \tag{9}$$

where

$$\overline{\mathbf{S}}_R = m\tilde{\mathbf{r}}_{cm} \tag{10}$$

$$\mathbf{Q}_F = \mathbf{F}_{ext} - m(\boldsymbol{\omega} \times \mathbf{V}_0) - m\ddot{\mathbf{r}}_{cm}$$
$$-2m(\boldsymbol{\omega} \times \dot{\mathbf{r}}_{cm}) - m(\boldsymbol{\omega} \times \boldsymbol{\omega} \times \mathbf{r}_{cm}) \tag{11}$$

$$\mathbf{Q}_M = \mathbf{M}_{ext} - \boldsymbol{\omega} \times \overline{\mathbf{J}} \cdot \boldsymbol{\omega} - \dot{\overline{\mathbf{J}}} \cdot \boldsymbol{\omega} - m\mathbf{r}_{cm} \times (\boldsymbol{\omega} \times \mathbf{V}_0)$$
$$- \boldsymbol{\omega} \times \sum_{i=1}^{N_n} (m_n)_i (\mathbf{r}_n)_i \times (\mathbf{v}_n)_i - \sum_{i=1}^{N_n} (m_n)_i (\mathbf{r}_n)_i \times (\dot{\mathbf{v}}_n)_i \tag{12}$$
$$- \sum_{i=1}^{N_n} (m_n)_i (\dot{\mathbf{r}}_n)_i \times (\mathbf{v}_n)_i$$

where $\overline{\mathbf{I}}$ the identity matrix, $\overline{\mathbf{S}}_R$ the dynamic-coupling matrix, $\tilde{\mathbf{r}}_{cm}$ the anti-symmetric matrix of $\mathbf{r}_{cm}$, $\mathbf{Q}_F$ the vector of external forces and $\mathbf{Q}_M$ the vector of external moments, considering terms that do not depend on $\dot{\mathbf{V}}_0$ and $\dot{\boldsymbol{\omega}}$.

2.1.2 Kinematics

For the formulation of the translation kinematics, we consider the position vector $\mathbf{R}_0$ in fig. 2.1, with components x_0 and y_0 defining the horizontal displacements and z_0 the altitude considering the origin of IRS at sea level. These components are measured from the BRS origin relative to the IRS origin. Therefore, one can express the inertial velocity $\mathbf{V}_0$ as:

$$\mathbf{V}_0 = \begin{bmatrix} \dot{x}_0 \\ \dot{y}_0 \\ \dot{z}_0 \end{bmatrix} \tag{13}$$

Using the velocities u, v and w described in SRC, the translation kinematics can be written as:

$$\begin{bmatrix} \dot{x}_0 \\ \dot{y}_0 \\ \dot{z}_0 \end{bmatrix} = (\overline{\mathbf{C}}_b^i)^T \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{14}$$

To obtain the rotation kinematics, initially the angular velocities in the BRS are described, as

$$\Omega_b = \begin{bmatrix} \dot{\phi}_b \\ \dot{\theta}_b \\ \dot{\psi}_b \end{bmatrix} = \overline{\mathbf{C}}_1 \cdot \overline{\mathbf{C}}_2 \cdot \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} + \overline{\mathbf{C}}_1 \cdot \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix} \tag{15}$$

And assume an angular velocity $\boldsymbol{\omega} = [p, q, r]$, in order to write the equation to be solved as:

$$\Omega_b = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{16}$$

Then, solving the eq. (16) obtains the rotation kinematics described as:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & \operatorname{sen}(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\operatorname{sen}(\phi) \\ 0 & \operatorname{sen}(\phi)\sec(\theta) & \cos(\phi)\sec(\theta) \end{bmatrix} \cdot \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{17}$$

2.1.3 Morphing Modeling

Aircraft kinematics and dynamics require the first and second derivatives of the position vectors $\mathbf{r}_{cm}$ and $\mathbf{r}_n$, as well as the first derivative of the $\overline{\mathbf{J}}$ inertia matrix. The complete dynamics of the aircraft is a system of ordinary differential equations, therefore care must be taken with the derivative approximations, because when adopting approximation methods as finite differences a time-step dependent truncation error is entered. This error increases the error in the ordinary differential system solution method.

Therefore, due to the complications introduced by the finite difference approximation, the approximation provided by [12] was used. To obtain the derivatives we use a function or a numerical routine capable of obtaining the parameters of interest as a function of the deflections δ of each mass element, obtaining the derivatives according to the equations below:

$$\mathbf{r}_{cm} = f(\delta_1, \delta_2, .., \delta_{Nn}) \tag{18}$$

$$\dot{\mathbf{r}}_{cm} = \sum_{k=1}^{Nn} \frac{\partial \mathbf{r}_{cm}}{\partial \delta_k} \cdot \dot{\delta}_k \tag{19}$$

$$\ddot{\mathbf{r}}_{cm} = \dot{\delta}^T \cdot \left(\sum_{j=1}^{Nn} \sum_{k=1}^{Nn} \frac{\partial^2 \mathbf{r}_{cm}}{\partial \delta_j \partial \delta_k} \right) \cdot \dot{\delta} + \sum_{k=1}^{Nn} \frac{\partial \mathbf{r}_{cm}}{\partial \delta_k} \cdot \ddot{\delta}_k \tag{20}$$

To obtain the derivatives of $\mathbf{r}_n$ and the $\overline{\mathbf{J}}$ inertia matrix just repeat the process performed for the position vector $\mathbf{r}_{cm}$.

However, using this approximation introduces the dependence with the first and second derivative of the δ deflections, that is, a second order model of the actuator dynamics is required for each deflection, as follows:

$$\dot{\delta}_p = -2\zeta \omega_n \cdot \delta_p + \omega_n^2 \cdot (\delta_c - \delta) \tag{21}$$

$$\dot{\delta} = \delta_p \tag{22}$$

where δ_c is the commanded deflection provided by the control project, and δ_p a variable substitution for the first derivative of δ.

These dynamics must be introduced into the system solution, increasing the number of states of the complete dynamics.

To obtain the values of inertia and center of mass position required for derivative approximations, it is assumed that the panels responsible for the morphing effect are concentrated masses positioned at the center of mass of each panel. Then, the position of the aircraft's center of mass relative to the BRS is obtained by:

$$\mathbf{r}_{\text{cm}} = \frac{1}{m} \cdot \sum_{i=1}^{N_n} (m_n)_i \cdot (\mathbf{r}_n)_i \qquad (23)$$

and the resulting inertia matrix,

$$\overline{\mathbf{J}} = \overline{\mathbf{J}}_F + \sum_{i=1}^{N_n} (m_n)_i \cdot (\tilde{\mathbf{r}}_n)_i \cdot (\tilde{\mathbf{r}}_n)_i^T \qquad (24)$$

where $\overline{\mathbf{J}}_F$ is the inertia matrix of components that are not part of the morphing structure and $\tilde{\mathbf{r}}_n$ is the anti-symmetric matrix of $\mathbf{r}_n$.

2.2 Aerodynamic model

To obtain the aerodynamic forces and moments, one can make use of stationary, non-stationary or quasi-stationary aerodynamic methods. The difference between these approaches is in considering the effects of flow in time. In the scope of this work, the analysis of the variant effects in time is inserted, therefore the non-stationary approach will be adopted.

In order to understand the adopted method, the topics about the typical subsection on which the method is based, in the subsection 2.2.1, are presented the formulation of the unsteady strip theory, in the subsection 2.2.2, the contribution of the movements of rigid body, in subsection 2.2.3, finally, obtaining the derivatives of stability and control in subsection 2.2.4.

2.2.1 Typical Section

For the formulation of unsteady strip theory, two-dimensional formulations are required, such as the analytical equations developed by [20] for a typical section with 3 degrees of freedom, these being the angle θ, δ and the displacement h, represented in fig. 3. In order to obtain the complete analytical solution for non-stationary aerodynamic loading, [20] considered a thin profile and two-dimensional, incompressible, potential flow subject to a simple harmonic motion.

Given the modeling assumptions, [20] do perform overlaps potential flows, being divided into a non-circulatory part, related to the airfoil, and another circulating part, related to the vortex mat extending from the trailing edge to infinity. With these potentials, forces and moments can be calculated through the integration of pressures over chord and the relationship between circulatory and non-circulatory forces is given by Theodorsen function $C(k)$ [21].
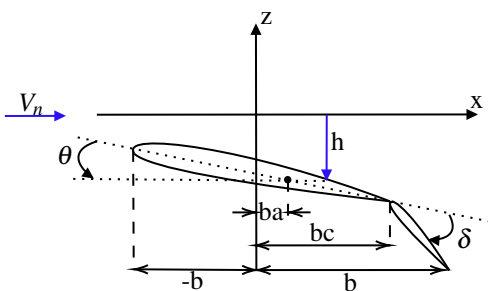
Figure 3: Typical section with three degrees of freedom.

In fig. 3, V_n represents the flow velocity, bc the distance between the control surface articulate point and the origin, ba the distance between the elastic axis and the origin, b the semi-chord.

2.2.2 Unsteady Strip Theory

Unsteady strip theory is based on the idea of representing three-dimensional aerodynamic flows by dividing the surface of interest into strips along its span, where the solutions developed for a typical section, as illustrated in the previous subsection, apply. Thus, three-dimensional effects such as the wingtip effect are neglected.

For the theory to be able to represent a wing with a sweep and dihedral angle, the modifications presented by [22] must be perform. Furthermore, in the work of [23] the modification is perform to accommodate the characteristics of the profile. Using the formulations provided by both works, the final formulation of the strip theory is called Modified Strip Theory (MST), described by the equations:

$$Q = \dot{h} + V_n \theta + V_n \sigma \tan(\Lambda_{\text{ea}}) + b\left(\frac{C_{l\alpha,n}}{2\pi} + ac_n - a\right)$$
$$\left(\dot{\theta} + V_n \tau \tan(\Lambda_{\text{ea}})\right) + \frac{V_n \delta}{\pi} T_{10} + \frac{b\dot{\delta}}{2\pi} T_{11} \qquad (25)$$

$$L = -\pi\rho b^2 \left[\ddot{h} + V_n \dot{\theta} + V_n \dot{\sigma} \tan(\Lambda_{\text{ea}}) - ba\left(\ddot{\theta} + V_n \dot{\tau} \tan(\Lambda_{\text{ea}})\right)\right]$$
$$-C_{l\alpha,n}\rho V_n bC(k)Q + \pi\rho b^2 \dot{\delta} T_4 + \rho b^3 \ddot{\delta} T_1 \qquad (26)$$

$$M_\alpha = \pi\rho b^2 V_n(\dot{h} + V_n \sigma \tan(\Lambda_{\text{ea}})) + \pi\rho b^3 a\left(\ddot{h} + V_n \dot{\sigma} \tan(\Lambda_{\text{ea}})\right)$$
$$-\rho b^3 V_n \dot{\delta}\left(T_1 - T_8 - T_4(c-a)\right) - \rho b^4 \ddot{\delta}\left(-T_7 - T_1(c-a)\right)$$
$$-\pi\rho b^2 \left(\frac{1}{8} + a^2\right)\left(\ddot{\theta} + V_n \dot{\tau} \tan(\Lambda_{\text{ea}})\right) - \rho b^2 V_n^2 \delta T_4$$
$$-2\pi\rho b^2 V_n Q\left(\frac{1}{2} - C(k)\frac{C_{l\alpha,n}}{2\pi}(a - ac_n)\right)$$
$$+\pi\rho b^2 V_n^2\left(\theta - ab\tau\tan(\Lambda_{\text{ea}})\right) \qquad (27)$$

$$C_{l\alpha,n} = \frac{C_{l\alpha}}{\cos(\Lambda_{\text{ea}})}, \quad V_n = V\cos(\Lambda_{\text{ea}}), \quad k = \frac{\omega b}{V_n} \qquad (28)$$

In the above equations Q is the strip downwash, M_α the strip pitch moment, L the strip lift, σ the strip dihedral, τ the torsion of the strip, Λ_{ea} the sweep angle relative to the elastic axis, $C_{l\alpha,n}$ angular coefficient of the curve $C_l \times \alpha$, c the dimensionless position of the profile control surface, a the dimensionless position of the elastic profile axis, ac_n the dimensionless position of the profile aerodynamic center, and ρ the flow density.

For the discretization of the surface to be in accordance with the formulation presented, the strips must be positioned perpendicularly along the elastic axis. Each of the ranges receives the system of equations presented above.

2.2.3 Contribution of Rigid Body Movements

This subsection presents the contributions of rigid body movements to enable the use of theory in a flight mechanics problem. These contributions are discussed in the work of [24], in order to find the contributions of the aircraft α and β angles and the effective angles θ_R, σ_R and Λ_R of each strip, as well as the contribution of the angular velocity of the body to the velocity $\dot{h}$ of each strip.

Given a function that relates the effective angles of each strip to the body's α and β angles, one can describe the effective angles as:

$$\theta_R = f_{\theta_R}(\alpha,\beta), \quad \Lambda_R = f_{\Lambda_R}(\alpha,\beta), \quad \sigma_R = f_{\sigma_R}(\alpha,\beta) \quad (29)$$

Therefore, the angle Λ_{ea}, θ and the displacement h from the formulation of the previous subsection should be replaced, in each strip, with:

$$
\begin{aligned}
\Lambda_{ea} &= \Lambda_{ea} + f_{\Lambda_R}(\alpha,\beta) \\
\theta &= f_{\theta_R}(\alpha,\beta) - \alpha_0 \\
h &= y' f_{\sigma_R}(\alpha,\beta)\cos(\Lambda_{ea})
\end{aligned}
\quad (30)
$$

where α_0 is the angle of attack with zero lift of the profile and y' is the y coordinate along the elastic axis.

The contribution of the angular velocities p, q and r of the body is understood as an induced velocity with respect to the distance from the elastic axis of the strip to the body system reference point. Then the $\dot{h}$, σ and τ variables are replaced by:

$$
\begin{aligned}
\sigma &= f_{\sigma_R}(\alpha,\beta)\cdot\cos(\Lambda_{ea}) \\
\tau &= \frac{\partial f_{\theta_R}(\alpha,\beta)}{\partial y'} - \frac{\alpha_0}{\partial y'} \\
\dot{h} &= L_p\, p + L_q\, q + L_r\, r + y' \cdot \frac{\partial f_{\sigma_R}(\alpha,\beta)}{\partial t}\cdot\cos(\Lambda_{ea})
\end{aligned}
\quad (31)
$$

where L_p, L_q, and L_r are the distances between the elastic axis of the strip and the origin of the body system for each angular velocity.

2.2.4 Stability and Control Derivatives

The formulation of the stability and control derivatives considers a linearized condition around an equilibrium permanent flight condition, so any state variable x_i can be described in relation to its balance plus your disturbance $x_i(t) = x_i|_{\text{eq}} + \Delta x_i(t)$. However, the force and moment expressions presented are valid for harmonic motion with frequency ω, so you must express the perturbations in a state variable as $\Delta x_i = \Delta\bar{x}_i \cdot e^{i\omega t}$.

Writing the expressions of forces and moments in relation to perturbations of state variables α, β, p, q and r gives the linearized forces and moments in relation to a movement harmonic. However, the interest is in describing the forces and moments for any movement. For this, the principle is used

that any physical response in time can be approximated by a combination of harmonic movements. Therefore, the imaginary variable $i\omega = s$ can be substituted in the formulation and the Theodorsen function $C(k)$ can be used for the frequency domain using the modified 0 and 1 Bessel functions [25].

In the frequency domain we get a state perturbation vector $\Delta\mathbf{x}(s)$, the force $L(s)$, and the momentum $M_\alpha(s)$ as:

$$\Delta\mathbf{x}(s) = \begin{bmatrix} \Delta\alpha(s) & \Delta\beta(s) & \Delta p(s) & \Delta q(s) & \Delta r(s) & \Delta\delta(s) \end{bmatrix}^T \tag{32}$$

$$L(s) = L|_{\text{eq}} + \Delta\mathbf{L}(s)\cdot\Delta\mathbf{x}(s) \tag{33}$$

$$M_\alpha(s) = M_\alpha|_{\text{eq}} + \Delta\mathbf{M}_\alpha(s)\cdot\Delta\mathbf{x}(s) \tag{34}$$

The expressions of the forces and moments in the eqs. (33) and (34) are applied to each range and can be summed vectorically, taking into account the direction of action of each force and moment. By grouping the terms, get a $\overline{\mathbf{A}}_E(s)$ matrix for forces and moments in equilibrium and a $\overline{\mathbf{A}}(s)$ matrix for perturbative terms, called the matrix of aerodynamic influence coefficients.

To leave the frequency domain and obtain the temporal response of forces and moments, the inverse Laplace transform must be used. However, the direct application of the transform cannot be performed directly, because Theodorsen $C(s)$ function is not Laplace-invertible. Therefore, one must perform the rational function approximation of the $\overline{\mathbf{A}}(s)$ matrix. The method for approximating the matrix elements used in this work is Roger's method.

With the approximation performed, one can use the inverse Laplace transform and then obtain the forces and moments in the time domain as the expression below.

$$\mathbf{F}(t) = \begin{bmatrix} F_x(t) & F_y(t) & F_z(t) & M_x(t) & M_y(t) & M_z(t) \end{bmatrix}^T \tag{35}$$

$$
\mathbf{F}(t) = \overline{\mathbf{A}}_E(t) + \overline{\mathbf{A}}_0(t)\cdot\Delta\mathbf{x}(t) + \left(\frac{b_{ref}}{V}\right)\overline{\mathbf{A}}_1(t)\cdot\Delta\dot{\mathbf{x}}(t) + \\
\overline{\mathbf{A}}_2(t)\left(\frac{b_{ref}}{V}\right)^2\cdot\Delta\ddot{\mathbf{x}}(t) + \sum_{i=1}^{n_{\text{lag}}}\overline{\mathbf{A}}_{i+2}(t)\cdot\mathbf{x}_i^{\text{lag}}(t)
\tag{36}
$$

where b_{ref} is the aircraft's reference wingspan, V the total aircraft speed and $\mathbf{x}_i^{\text{lag}}(t)$ o vector of states representing the terms of aerodynamic delay.

Finally, each of the components of the $\mathbf{F}(t)$ vector can be described in terms of the coefficients of influence, called stability derivatives, when referring to state variables, and control, when referring to to control deflections. For dimensionlessness we use the aerodynamic mean chord c_{ref} and the wing plan area S_{ref}.

2.3   Control project model

This section presents the control design method, with its formulation and characteristics. To understand the method, the necessary steps for the control design in subsection 2.3.1 and the design routine adopted in subsection 2.3.2 are presented.

2.3.1   The EMC Method

The EMC method implements an SMC-inspired approach with a heuristically defined nonlinear mapping function. The shape and boundary of the mapping function is defined by the operator. The limit is derived from basic information about the actuators and its shape is defined based on knowledge of system behavior. For EMC implementation, follow these steps:

1. Switched error calculation:

$$e_t = \frac{x_{ref} - x}{e_r} \tag{37}$$

where x_{ref} is the reference of any state x and e_r the first design parameter.

2. The parameter e_t is restricted to being between -1 and 1:

$$e_s = \begin{cases} -1, & \text{se } e_t < -1 \\ e_t, & \text{se } -1 \le e_t \le 1 \\ 1, & \text{se } e_t > 1 \end{cases} \tag{38}$$

3. Exponential Function Calculation:

$$u_e = \text{sign}(e_s)\left(1 - ||e_s| - 1|^{2^{-B}}\right) \tag{39}$$

where $\text{sign}(e_s)$ is a function that returns the e_s and B sign the second project parameter, usually inserted between -10 and 10.

4. Control action calculation:

$$u = \frac{u_{\max} - u_{\min}}{2}(u_e - 1) + u_{\max} \tag{40}$$

where $u_{\max}$ and $u_{\min}$ the maximum and minimum limits of the control action.

In practice EMC needs only two parameters and control stops, making it easy to implement and adjust.

2.3.2   Project Routine

With the specifications of the steps of the EMC method, you can use it in a design routine. The routine consists of a reference entry and an initial kick of the e_r and B parameters for each of the state variables to be traced. In Figure 4 is the system block diagram, with the state variables used in the control project.

The project can be divided into two, controller and stabilizer, as shown above. The controller is responsible for controlling the V, H and ϕ tracked variables, while the stabilizer is responsible for stabilizing the aircraft, using as input the comparison between the angles θ and α .
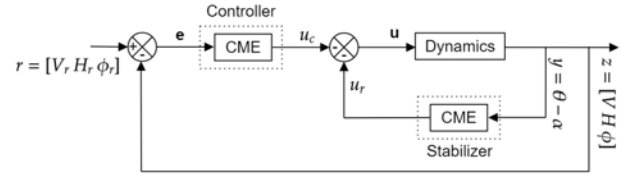
Figure 4: Block diagram for the control project.

The system is resolved with the *ode15s* function of the MATLAB © software, and thereafter the errors are stored between references and tracked states. As a penalty criterion, the time integral multiplied by the absolute value of the error $|e(t)|$ is used, a very useful criterion used to penalize transient responses [26]. The criterion can be described as:

$$f_{\text{crit}} = \int_0^\infty t|e(t)| \cdot dt \tag{41}$$

Finally, the final e_r and B parameters are obtained by minimizing the f_{crit} function, using a minimization function such as *fminsearch* function of MATLAB ©.

3   Numerical Studies

3.1   Reference Aircraft

To evaluate the proposed formulation, a simple reference aircraft was chosen, similar to the aircraft presented by [27]. It is a flying wing unmanned aerial vehicle (UAV), seen in Fig. 5, with a reference chord of 0.276 m, wingspan of 1.2 m, total mass of 0.9 kg and a sweep angle of $30°$. In addition, the flight condition is defined as a cruise flight with a speed of 12 m/s and an altitude of 100 m.

Figure 5: Reference aircraft [27].

However, the mass distribution of this aircraft is unknown. Thus, defining this distribution, becomes a challenge, directly affecting the stability of the aircraft, and making it difficult to integrate the equations system and consequently the control design.

To overcome this problem, a mesh of ten elements along the chord and twenty elements along the span is considered, and each element is considered as a point mass. The aircraft's inertia and center of mass position are computed with equations (23) and (24). Different total masses are defined for each area of the model, so that a mass of 40% of the total mass is considered to be in the central region, where the propulsive

system is located, and each half wing has 30% of total mass. The resulting center of mass is used as reference point for the calculation of moments of inertia. In fig. 6, the elements of each half wing are highlighted in blue, the central region in red, and the center of mass identified with a blue dot. Table 1 shows the values of moments of inertia and the position of the center of mass.

Figure 6: *The reference model discretized in elements for mass and inertia computation with different densities.*

Table 1: *Moment of inertia values and center of mass coordinates of reference aircraft.*

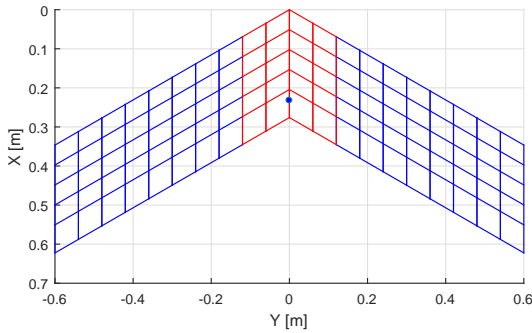| Inertial parameters	Values (S.I)
I_{xx}	0,0681
I_{yy}	0,0116
I_{zz}	0,0797
$I_{zx} = I_{yx} = I_{zy}$	0
x_{cm}	0,2317
$y_{cm} = z_{cm}$	0

The aerodynamic model is composed of 20 strips, divided into each side, all with the same area, as seen in Fig. 7. There are two control surfaces areas, one at the inner portion of a each half wing, comprising 5 strips, and another at the outer portion, comprising 3 strips. The fuselage area is modeled with a total of 4 strips, which are superposed.

The aircraft has propulsive limitations, with a maximum power of 260 W, so the maximum speed attained by the aircraft is 22.0 m/s, while the minimum speed is 9.0m/s for no loss of support. In addition, it also has limitations of control surface deflections due to actuator constraints, which are $20°$ for maximum deflection and $-20°$ for minimum deflection [27].

3.2 Response to Control Surface Disturbances

The analysis of the dynamics response to control surface disturbances is important for the understanding of aircraft behavior regarding the use of controls. To perform this analysis, we chose to only disturb the elevator command with a doublet entry of a $\Delta \delta_p = \pm 1°$ for 2 seconds. In addition, three different dynamics were considered, the first disregards the effects
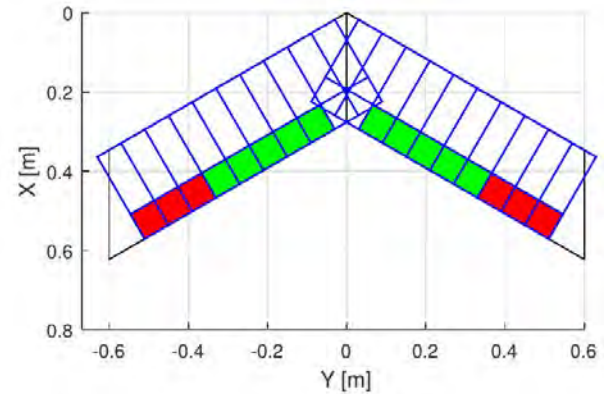
Figure 7: *The aerodynamic model, considering the wing discretized into 20 strips.*

of non-stationary aerodynamics and temporal variations of inertial parameters, the second disregards only the effects of temporal variations of inertial parameters, and the third only the effects of aerodynamics. non-stationary. In fig. 8 the black results refer to the first dynamics, blue to the second and red to the third when disturbed by the elevator control.

By analyzing the above results it can be seen that the behavior of the aircraft is consistent with the expected modeling, as a positive disturbance in the elevator decreases the angle of attack α and the angle of pitch θ, since This disturbance generates a negative pitching moment. In addition, as the angle of attack decreases, speed increases due to decreased drag, and altitude decreases due to decreased lift. The control result shows the actuator dynamics, which have relatively fast dynamics.

Performing a comparative analysis between the different dynamics, it is clear that the first and third dynamics have identical results, while the second dynamics have different results, but maintains the physical sense. This difference in magnitude of results can be explained by analyzing non-circulatory or apparent mass terms, which consider the displacement of air mass in surface motion, increasing the magnitude of forces and moments acting on motion.

In fig. 9 we find the variations of the inertial parameters when considering the third dynamic, as discussed in the previous paragraph, this dynamic does not influence the results, ie the variations are very small and consequently result in forces. and very small moments.

4 Conclusions

The work presented the modeling of a Exponential Mappping Controller applied to morphing wing. The aerodynamic model considers the modified strip theory.

By using the flight dynamics that considers the time-varying displacement of the center of mass and the resulting inertia, it is necessary to use a model able of computing these variables continously and estimate the associate derivatives. Approximations of derivatives of inertial parameters should be performed with caution. The system solution needs a variable
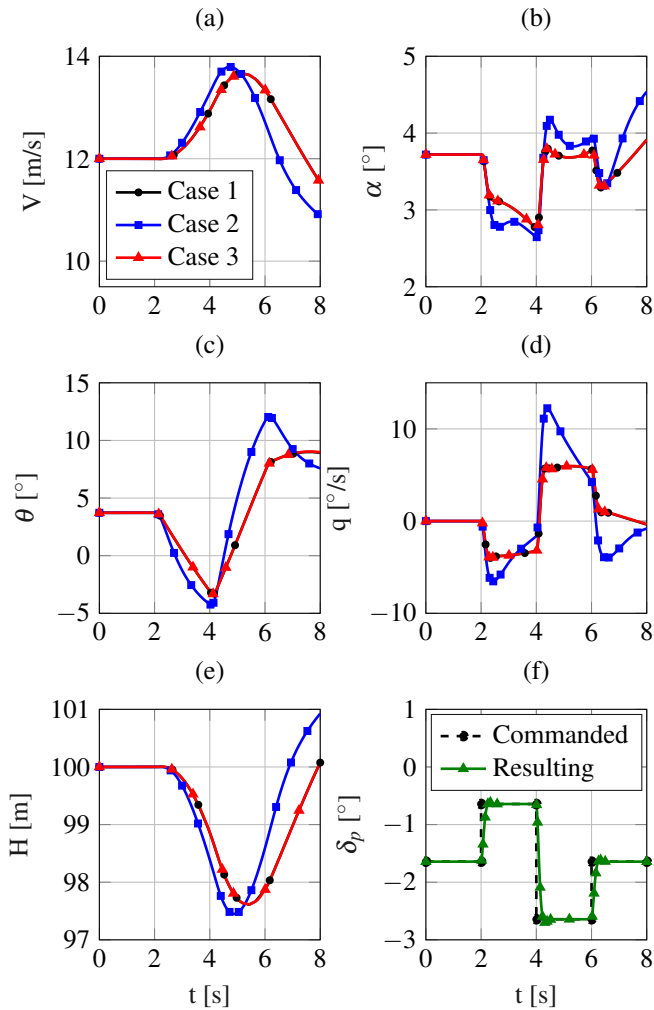
Figure 8: Response of state variables to elevator control disturbance.

Figure 9: Time variations in center of mass position and moments of inertia for a disturbance in elevator control.

time step, which makes it impossible to use more usual derivative approximations, such as the finite difference method. So, to circumvent this problem one can use linear relationships between inertial parameters and deflections, needing to know the dynamics of the actuator.

The aerodynamic method should be chosen carefully, as some non-stationary aerodynamic methods depend on the time step, such as the UVLM, requiring a more complete numerical study when it comes to the union between the equation system solution and the aerodynamics.

For the use of non-stationary strip theory, it is necessary to know the first and second derivatives of state variables, but these quantities are not known before the aerodynamic forces and moments are obtained. To work around this problem, an estimate is made using the linearized dynamics model, considering that the linearized matrices remain constant over time.

The time variations of the center of mass position and the aircraft inertia proved to be insignificant, mainly due to the category adopted for variable geometry. It is possible that a more aggressive morphing concept might affect more the flight

dynamics. This investigation is future step in the present research.

The use of non-stationary aerodynamics is required for a more coherent description of aircraft movement, as non-stationary aerodynamic effects are significant in the aircraft's ultimate behavior.

The control method used was very effective for the simulated cases adopted.

Nomenclature

Designation	Denotation	Unit
$\mathbf{F}_{ext}$, $\mathbf{Q}_F$, L	Force	N
$\mathbf{M}_{ext}$, $\mathbf{Q}_M$, M_α	Moment	N.m
$\mathbf{h}$	Angular moment	N.rad
ψ, θ, ϕ, α, β	Angle	rad
ω	Angular velocity	rad/s
$\mathbf{V}_{cm}$, $\mathbf{V}_0$, $\mathbf{v}_n$	Velocity	m/s
$\mathbf{\bar{J}}$, $\mathbf{\bar{J}}_F$	Moment of inertia	kg.m^2
$\mathbf{R}_{cm}$, $\mathbf{R}_0$, $\mathbf{r}_n$, $\mathbf{r}_{cm}$	Distance	m
m, m_n	Mass	kg
L_p, L_q, L_r	Distance	m
b_{ref}, c_{ref}	Distance	m
S_{ref}	Area	m^2

References

[1] Terrence A. Weisshaar. Morphing aircraft technology - new shapes for aircraft design. NATO Research and Technology Organisation (RTO), 2006.

[2] Rafic Ajaj, Christopher Beaverstock, and Michael Friswell. Morphing aircraft: The need for a new design philosophy. *Aerospace Science and Technology*, page 12, 2015.

[3] Charles H. Gibbs-Smith. *Aviation: An Historical Survey from Its Origins to the End of World War*, volume 1. NMSI Trading Ltd., 2 edition, 2013.

[4] Onur Bilgen Michael I. Friswell Daniel J. Inman Silvestro Barbarino, Rafic M. Ajaj. A review of morphing aircraft. *Journal Of Intelligent Material Systems And Structures*, page 12, 2011.

[5] Nicholas B. Cramer, Daniel W. Cellucci, Olivia B. Formoso, Christine E. Gregg, Benjamin E. Jenett, Joseph H. Kim, Martynas Lendraitis, Sean S. Swei, Greenfield T. Trinh, Khanh V Trinh, and Kenneth C. Cheung. Elastic shape morphing of ultralight structures by programmable assembly. *Journal Of Smart Materials and Structures*, page 14, 2019.

[6] Daochun Li, Shiwei Zhao, Andrea Da Ronch, Jinwu Xiang, Jernej Drofelnik, Yongchao Li, Lu Zhang, Yining Wu, Markus Kintscher, Hans Peter Monner, Anton Rudenko, Shijun Guo, Weilong Yin, Johannes Kirn, Stefan Storm, and Roeland De Breuker. A review of modelling and analysis of morphing wings. *Progress in Aerospace Sciences*, page 12, 2018.

[7] Alexander Pankonien and Daniel J. Inman. Experimental testing of spanwise morphing trailing edge concept. Active and Passive Structures and Integrated Systems, 2013.

[8] A. Tarabi, S. Ghasemloo, and M. Mani. Experimental investigation of a variable-span morphing wing model for an unmanned aerial vehicle. *The Brazilian Society of Mechanical Sciences and Engineering*, page 9, 2015.

[9] flexsys, 2018.

[10] Kenneth E. Boothe. Dynamic modeling and flight control of morphing air vehicles. Master of science, University of Florida, Gainesville, Florida, 2004.

[11] Thomas S. Koch. Stability and control of a morphing vehicle. Master of science in aeronautics and astronautics, University of Washington, Seattle, Washington, 2012.

[12] B. Obradovic and K. Subbarao. Modeling of flight dynamics of morphing-wing aircraft. *Journal of aircraft*, 48:11, 2011.

[13] N. Ameri, M. H. Lowenberg, and M. I. Friswell. Modelling the dynamic response of a morphing wing with active winglets. Hilton Head, South Carolina, 2007. Atmospheric Flight Mechanics Conference and Exhibit.

[14] Joaquim R.R.A. Martins David A. Burdette. Design of a transonic wing with an adaptive morphing trailing edge via aerostructural optimization. *Aerospace Science and Technology*, page 12, 2018.

[15] Eun Jung Chae, Amin Moosavian, Alexander M. Pankonien, and Daniel J. Inman. A comparative study of a morphing wing. Conference on Smart Materials, Adaptive Structures and Intelligent Systems, 2017.

[16] Ivan Wang, S. Chad Gibbs, and Earl H. Dowell. Aeroelastic model of multisegmented folding wings: Theory and experiment. *Journal of aircraft*, 49:11, 2012.

[17] Thiemo M. Kier. Comparison of unsteady aerodynamic modelling methodologies with respect to flight loads analysis. AIAA Atmospheric Flight Mechanics Conference, 2005.

[18] Hildebrando F. de Castro, Pedro Paglione, and Carlos Henrique Ribeiro. Exponential mapping controller applied to aircraft. AIAA Guidance, Navigation, and Control Conference, 2012.

[19] Ahmed A. Shabana. *Dynamics of Multibody Systems*, volume 1. Cambridge University Press, 3 edition, 2005.

[20] T. Theodorsen and I. E. Garrick. Mechanism of flutter. a theoretical and experimental investigation of the flutter problem, 1940.

[21] Theodore Theodorsen. General theory of aerodynamic instability and the mechanism of flutter, 1949.

[22] J. G. Barmby, H. J. Cunningham, and I. E. Garrick. Study of effects of sweep on the flutter of cantilever wings, 1951.

[23] Jr. E. Carson Yates. Calculation of flutter characteristics for finite-span swept or unswept wings at subsonic and supersonic speeds by a modified strip analysis, 1958.

[24] Grégori Pogorzelski. Dinâmica de aeronaves flexíveis empregando teoria das faixas não-estacionária. Master of science, Instituto Tecnológico de Aeronáutica, São Jóse dos Campos, SP-Brasil, 2010.

[25] Raymond L. Bisplinghoff, Holt Ashley, and Robert L. Halfman. *Aerolasticity*. Dover Publications, Inc., 1996.

[26] K. Mohamed Hussain, R. Allwyn Rajendran Zepherin, M. Shantha Kumar, and S.M. Giriraj Kumar. Comparison of pid controller tuning methods with genetic algorithm for foptd system. *Journal of Engineering Research and Applications*, page 7, 2014.

[27] Henrik Grankvist. Autopilot design and path planning for a uav. Defence and Security, Systems and Technology, 2006.

Modelling of Transonic and Supersonic Aerodynamics for Conceptual Design and Flight Simulation

Petter Krus and Alvaro Abdallah[*]

Department of Management and Engineering, Linköping University, Linköping, Sweden
E-mail: petter.krus@liu.se, pfalvaro@sc.usp.br
[*]Department of Aeronautical Engineering, University of Saõ Paulo, Saõ Paulo, Brazil

Abstract

In this paper a comprehensive set of models of aerodynamics for conceptual design that also can be used for flight simulation is presented. In particular these are transonic/supersonic lift coefficient, moment and induced drag. The models are based on a range of models found in literature that have been compiled and reformulated to be more practical to use. Many models in literature exhibit discontinuities or infinities e.g. around Mach one. Care has to be taken to formulate benign expressions that can be used in all parts of the flight envelope. One observation is that models are showing that induced drag have a significant influence on supersonic performance, especially at high altitude and with elevated load factor. The models are implemented in a simulation model so that performance can be evaluated at a mission level already in conceptual design.

Keywords: Supersonic, simulation, conceptual design

1   Introduction

Traditionally conceptual design has not been involving flight simulation, although this is gaining in importance. In conceptual design modelling is aimed to predict the behavior and performance of the finished product. As such they should not necessarily show the exact behaviour for the aircraft at the sketchy level of conceptual design. E.g. at the conceptual design stage aerodynamic adjustments such as fillets etc, are not defined and hence a very accurate model such as an advanced CFD model of the aerodynamics will not provide relevant result. At the conceptual design only the important configuration and dimensions of the aircraft, are laid out. Hence models at the conceptual stage are on an entirely different nature than in later part of design where the geometry has a high degree of fidelity. In this paper a comprehensive aerodynamic model with a minimum number of coefficients is established, that is suitable for simulation at the conceptual design stage, in such a way that they can give a "best guess" of the performance of the finished product. Even though the models use a minimum of parameters they can provide a model of high fidelity if they are established e.g. from wind tunnel or flight testing. Therefore the same model can be used well into the design process as better data becomes available.

2   Geometric modeling

In order to have a representative model an existing aircraft in this case the F-16, was chosen as a basis for study in this work. In addition the X-29 forward swept experimental aircraft was also used to validate the model since there are data published

in [1].

3   Modelling for mission simulation

For evaluating the performance of the aircraft in a realistic scenario a system simulation model was built that could be used in a mission simulation. The flight dynamics model is here based on a 6 degree of freedom rigid body model that is connected to an aerodynamic model. This was presented in Krus et al. [2] and in Abdalla [3]. The aircraft model can have different number of wings, with an arbitrary number of control surfaces, and a body with its inertia characteristics. The aerodynamic model is here based on a static version of the model presented in [4], although the unsteady effects can of course also be included. The control surfaces are modeled both with a linear increase of lift force with deflection and the corresponding increase in induced drag. There is also a cross coupling effect of drag for control surfaces on the same wing e.g. ailerons and flaps. In this way also the effect of trim drag on performance is automatically included, and the effect of reduced weight, as fuel is consumed.

4   Modelling transonic/supersonic characteristics

Here, a substantial part of the mission under study is in the trans-sonic and supersonic regimes. Therefore, it is important to have models that capture these characteristics in an adequate way. There are basically three effects that are modeled in the supersonic regime. The first is the *wave drag*. In the
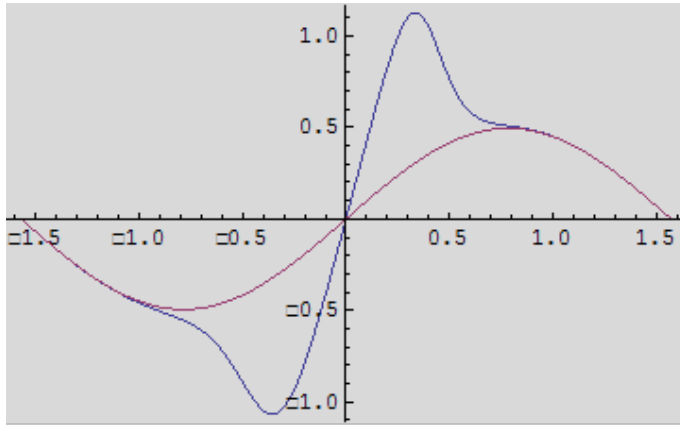
Figure 1: Non-linear aerodynamic model.

model there is a Mach number dependent coefficient that is added to the parasitic drag coefficient.

The second is that the *aerodynamic center* is moved backwards from the quarter cord position for subsonic to approximately half cord for supersonic. The third is that the lift sloop is changed. Furthermore the induced drag is also changed as a consequence of the lift sloop and due to loss of leading edge suction. The transition between subsonic and supersonic is modelled with a logistic (sigmoid) function to produce a soft transition.

4.1   Aerodynamic Drag Estimation

Aerodynamic drag in the supersonic range is composed of skin friction drag, form and interference drag, and wave drag. Skin friction drag can be basically considered to be the same for the whole range, while for the form and interference drag it are included in the wave drag in the supersonic region. The wave drag coefficient is, see [5] or [6] roughly proportional to:

$$C_{dw} \propto \frac{1}{\sqrt{M^2 - 1}} \qquad (1)$$

This expression, however, has a singularity at $M = 1$, which is clearly not realistic. In order to remove the singularity (1) is modified into:

$$C_{dw} = C_{dw0} \frac{k_{dw}}{(((M - k_{dwm})^2 - 1)^2 + k_{dw}^4)^{1/4}} \qquad (2)$$

Here C_{dw1} i the maximum wave drag and k_{dw} is a shape parameter. A low value of k_{dw} leads to a quick decay of the wave drag coefficient with Mach number. k_{dwm} is another shape parameter that can be used to move the Mach number of the maximum drag. Both these values are non-dimensional and are typically less than one. This function only applies to the supersonic region. Therefore it is multiplied with a logistic function to produce a soft step starting at the critical Mach number where the drag rise starts. This function is f_M is:

$$f_M = \frac{1}{1 + e^{-8\frac{M - (1 - \delta_M/2)}{\delta_M}}} \qquad (3)$$

Here:

$$\delta_M = 1 - M_{crit} \qquad (4)$$

where M_{crit} is the critical Mach number. Plotting this function for $\delta_M = 0.2$ yeilds Fig. 2.
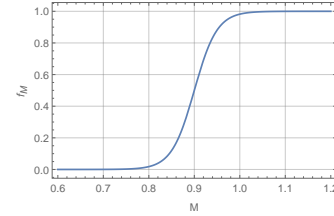
Figure 2: The logistic function for $\delta_M = 0.2$. Hence for this example $M_c rit = 0.8$.

The full expression is then

$$C_{dw} = f_M C_{dw0} \frac{k_{dw}}{(((M - k_{dwm})^2 - 1)^2 + k_{dw}^4)^{1/4}} \qquad (5)$$

With $C_{dw0} = 0.0264$, $k_{dwm} = 0.05$, $k_{dw} = 0.5$ and $\delta M = 0.2$ the function in Fig. 3 is obtained.
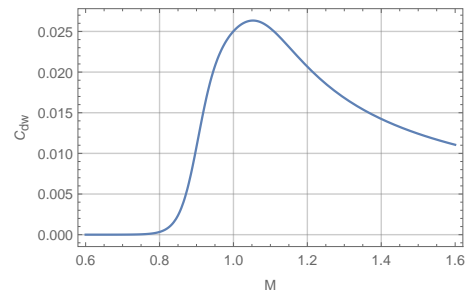
Figure 3: Wave drag coefficient as a function of Mach number, using Eq. 5

The most critical parameter for the performance prediction is the maximum wave drag contribution, i.e. C_{dw}.

Under simplified assumptions the Sears-Haack body have the lowest transonic drag. This was shown first suggested by Wolfgang Haack [7].

The theoretical wave drag coefficient for a Sears-Haack body at Mach $M = 1$ (also the maximum value according to this theory) can be calculated as (with cross-section area as reference area):

$$C_{dw,SH} = 9\pi \frac{S_{max}}{2L^2} \qquad (6)$$

For aircraft design it is more usual to use the wing area as reference area. The expression then becomes:

$$C_{dw,SH} = 9\pi \frac{S_{max}}{2L^2} \left(\frac{S_{max}}{S_{ref}} \right) \qquad (7)$$

where $S_{max} = \pi r_{max}^2$ is the maximum cross section area of the Sears-Haack body, and L is the corresponding length. S_{ref} is another area that should be used as reference are you for C_{dw}. For an aircraft it is usually the wing area and for a rocket or

missile the reference area is usually a cross section area and it could be the same as S_{max}. This value gives a lower limit to the transonic drag. For real supersonic aircraft it is a factor larger. this can be accomodated for by introducing a correction factor E_{wd} as in [6]. The wave drag can then be calculated as:

$$C_{dw} = E_{wd}C_{dw,SH} \tag{8}$$

The correction factor can be estimated by comparing to values of typical existing designs. Another way is to use a more elaborate method, based on a more detailed geometric analysis based on [8], [9] and [10]. For a more elaborate discussion on practical implementation of such metods, see e.g. in [11] and [12], [13], and in [14].

4.2 Supersonic lift

The liftsloop of a wing changes already for higher subsonic Mach number. For an infinite straight wing the lift sloop can be calculated as:

$$C_{L\alpha,sub} = \frac{C_{L\alpha,0}}{\beta} \tag{9}$$

Here $C_{L\alpha,sub}$ is the lift-sloop for low Mach numbers, which can be derived eg. from panel code. The factor β is:

$$\beta = \sqrt{1 - M^2} \tag{10}$$

Going into the supersonic region the lift curve changes dramatically. The lift curve can here be calculated from

$$C_{L\alpha,sub} = \frac{4}{\beta} \tag{11}$$

where for this case:

$$\beta = \sqrt{M^2 - 1} \tag{12}$$

However, this is only valid for a wing where the reference area is equal to the actual effective lifting area. For a whole aircraft these are not necessarily the same. therefore a nondimensional lifting area for the whole wing body combination S_0 is introduced as:

$$S_0 = S_{wb}/S_{ref} \tag{13}$$

Eq. 9 can then be rewriten as:

$$C_{L\alpha,sup} = S_0\frac{4}{\beta} \tag{14}$$

Otherwise the supersonic lift would be dependent on the chosen reference area. β is here changed into an expression that is valid in both ranges:

$$\beta = ((M^2 - 1)^2)^{1/4} \tag{15}$$

However, if (15) is used in (11) or (9) this yields a singularity at $M = 1$. In reality this does not occur, and the lift curve is smoothed out in the transonic range. Therefore the β is modified into:

$$\beta = ((M^2 - 1)^2 + \varepsilon_M^4)^{1/4} \tag{16}$$

Here ε_M is a factor that removes the singularity while the asymptotes are unaffected.

For a swept wing 14 is only valid when the Mach cone, see Fig. 4 has passed the leading edge. This means that there is prolonged transition region for this case. This Mach number can be calculated using geometry. The Mach number where the mach cone coincides with the leading edge of the wing is then.

$$M_1 = \sqrt{1 + \tan^2 \Lambda} \tag{17}$$

Note that for a complete aircraft the notion of Mach cone becomes more complex.

$M_1 = \sqrt{1 + \tan^2 \Lambda}$

45 degrees ->M1=1.2
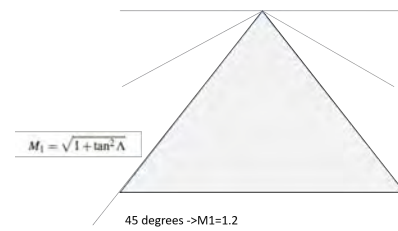
Figure 4: Mach cone touching the leading edge of a delta wing.

An expression that can be used for the whole range is then:

$$C_{L\alpha} = C_{L\alpha,sub}(1 - f_{ML}) + C_{L\alpha,sup}f_{ML} \tag{18}$$

Here

$$f_{ML} = \frac{1}{1 + e^{-4\frac{M - (1 + \delta_{ML}/2)}{\delta_{ML}}}} \tag{19}$$

This is again the logistic function, where

$$\delta_{ML} = M_1 - 1 = \sqrt{1 + \tan^2 \Lambda} - 1 \tag{20}$$
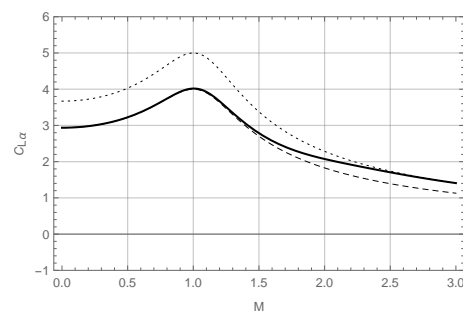
Figure 5: The lift coefficient with respect to α. The dotted line is without leading edge suction and the dashed line is with leading edge suction. The solid line is a blend of of both for a high Mach cone angle

4.3 Supersonic moment

In the supersonic regime the neutral point is moving backwards from approximately the quarter cord position of the wing to approximately the half cord position. This leads to an increased stability of the aircraft and also an increasing negative pitch moment that has to be counteracted by the elevator,

elevons and/or canard. In the transonic region the moment characteristics can be very complex and show dramatic variations. However, this requires a more detailed analysis of the aerodynamics, and a high fidelity geometric model that is not available in conceptual design, normally the aircraft should not operate in the this region so it is of less importance for performance predictions. A simple model that captures the main effect of moving the neutral point a quarter cord is:

$$C_m = C_{m0} - C_L \frac{MAC}{4} f_M \qquad (21)$$

where f_M is a factor that is typically less than one. For a straight wing $f_M = 1$ but for e.g. a cranked delta wing $f_M < 1$.

4.4 Supersonic lift dependent drag

In supersonic flight the lift dependent drag is generally small compared to the wave drag. Hence a very precise model is not needed. It can even be argued that the subsonic model can be used for simulation. However, here a simple expression is shown that at least capture the main effects. In the subsonic regime induced drag can in general be calculated from:

$$C_{Di} = \frac{C_L^2}{\pi e A} \qquad (22)$$

This can be rewritten as:

$$C_{Di} = \frac{C_{L\alpha}^2}{\pi e A} \alpha^2 = C_{Di\alpha^2,sub} \alpha^2 \qquad (23)$$

For a straight wing in supersonic flow there is no leading edge suction. This means that the lift dependent drag can be found from the lift from trigonometric relations:

$$C_{Di,sup} = C_L \tan(\alpha) \approx C_{L,sup} \alpha = C_{L\alpha^2,sup} \alpha^2 = C_{Di\alpha^2,sup} \alpha^2 \qquad (24)$$

Using Eg. (14) yields

$$C_{Di,sup} = S_0 \frac{4}{\beta} \alpha^2 \qquad (25)$$

In the firmly supersonic regime Eq. (23) can still be used if the Oswald efficiency factor e is replaced with:

$$e_{sup} = \frac{C_{L,sup}^2}{\pi A C_{Di,sup}} = \frac{C_{L\alpha,sup}^2}{\pi A S_0 \frac{4}{\beta}} \qquad (26)$$

where e_{sub} is the subsonic efficiency factor.

For a swept wing the behaviour is more complex. When the trailing edge of the wing is inside the Mach cone there can be some leading edge suction, but this leads to another level of detail that might not be available at the conceptual design stage. However, once the leading edge of the wing is outside the Mach cone, that is when the Mach number $M > M_1$, the expression derived here for the supersonic induced drag is valid. Therefore, an approximation suggested here is to assume that the lift function goes from leading edge suction gradually to a situation without leading edge suction at $M = M_1$. When there is leading edge suction the Oswald efficiency factor for the subsonic case is used and then this is gradually moved to the value for supersonic speed.

$$e = e_{sub}(1 - f_{ML}) + e_{sup} f_{ML} \qquad (27)$$

In this way Eg. 23 can be used for the whole range.
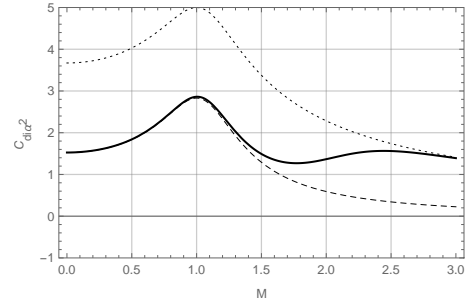
Figure 6: The coefficient for induced drag with respect to α^2. The dotted line is without leading edge suction and the dashed line is with leading edge suction. The solid line is a blend of both for a high Mach cone angle

5 Comparison of contributions to drag

In order to have an idea of the importance of the different drag components an example with coefficients chosen to loosely resemble the F-16 flying at 10000m with a weight of 12000kg was made. Calculation the corresponding drag contributions using:

$$D = C_D \rho \frac{v^2}{2} \qquad (28)$$

where $\rho = 0.4 \text{kg/m}^2$ (corresponding to an altitude of 10000 m) and the speed of sound $a = 295$ m/s. The mass of the aircraft was set to 12000 kg resulted in a required lift force of $L = 117840$ N. The lift coefficient was then calculated from

$$L = C_L \rho \frac{v^2}{2} \qquad (29)$$

which yields:

$$C_L = \frac{2L}{\rho v^2} \qquad (30)$$

The Drag coefficients are shown in Fig. 7 The contributions
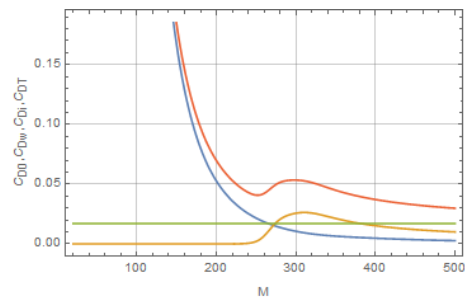
Figure 7: The contributions of parasitic, the wave drag and the induced drag as well as the total drag.

of the different drag components are shown in Fig. 8 where also the total drag is shown. It shows that the contribution of induced drag is important at supersonic speed. In e.g. a turn, a load factor for two would increase the induced drag four times, which would make it on par with the parasitic drag.

6 Conclusion

In this paper a generic model to model trans-sonic and supersonic aerodynamic characteristics is presented. Different
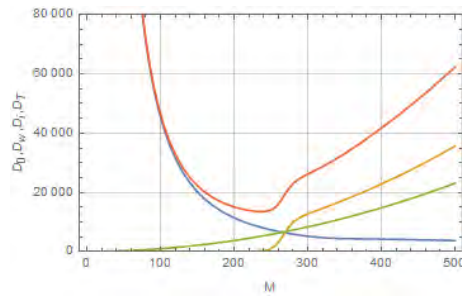
Figure 8: The contributions of parasitic, the wave drag and the induced drag as well as the total drag.

models for different parts of the envelope has been combined into continuous functions that can be used for flight simulation. One conclusion is that supersonic induced drag can be substantial and needs to be considered in conceptual design. Even though the velocity is high, the induced drag coefficient goes up when leading edge suction is lost. In addition the neutral point is moved backwards increasing the need for trim that is further increasing the induced drag.

References

[1] Edwin J. Saltzman and John W. Hicks. In-flight lift-drag characteristics for a forward-swept wing aircraft and comparisons with contemporary aircraft). Technical report, NASA Technical paper 3414, 1994.

[2] Petter Krus, Robert Braun, and Peter Nordin. Aircraft System Simulation for Preliminary Design. In *28th International Congress of the Aeronautical Sciences*, Brisbane, 2012. ICAS.

[3] Alvaro Martins Abdalla, Grönstedt Tomas Gazetta, H., and Petter Krus. 4:th CEAS European Air & Space Conference. In *The Effect of Engine Dimensions on Supersonic Aircraft Performance*, Linköping, 2013. Linköping University Electronic Press.

[4] Christopher Jouannet and Petter Krus. Unsteady aerodynamic modelling: a simple state-space approach. In *AIAA Aerospace sciences meeting and exhibit*, Reno, USA, 2005.

[5] John D. Anderson. *Introduction to Flight, Fourth Edition*. WXB/McGraw-Hill, 2000.

[6] Daniel P Raymer. *Aircraft Design: A Conceptual Approach*. AIAA, 6th edition, 2018.

[7] Wolfgang Haack. Geschossformen kleinsten Wellenwiderstandes, 1941.

[8] Richard T Whitcomb. A study of Zero-Lift-Drag-Rise Characterisics of Wing-Body Combinations Near the Speed of Sound. Technical report, NACA Technical Report 1273., 1956.

[9] Richard T Whitcomb and Thomas C Kelly. A Study of the Flow over a 45 Degree Sweptback Wing-fuselage Combination at Transonic Mach Numbers. Technical Report April, NACA-RM-L52D01, 1957.

[10] Robert T. Jones. Theory of wing-body drag at supersonic speeds. Technical report, NASA, 1953.

[11] Raghu Chaitanya Munjulury, Ingo Staack, Alvaro Martins Abdalla, Tomas Melin, Christopher Jouannet, and Petter Krus. Knowledge-based design for future combat aircraft concepts. In *29th Congress of the International Council of the AeronauticalSciences (ICAS 2014) :*. The University of São Paulo (USP), Brazil, 2014.

[12] Raghu Chaitanya Munjulury, Alvaro Martins Abdalla, Ingo Staack, and Petter Krus. Knowledge-based future combat aircraft optimization. In *30th Congress of the International Council of the AeronauticalSciences (ICAS 2016) :*, Belo Horizonte, Brazil, 2016.

[13] Raghu Chaitanya Munjulury and Alvaro Martins Abdalla. Analysis of Radar Cross Section and Wave Drag Reduction of Fighter Aircraft. In *Aerospace Technology, FT2016*, Stockholm, Sweden, 2016.

[14] Matheus Urzedo Quirino, Mathieu Calvo, Raghu Chaitanya Munjulury, and Alvaro Martins Abdalla. Analysis of the influence of the external configuration on the wave drag ofsupersonic business jets. In *31st Congress of the International Council of the Aeronautical Sciences, ICAS 2018*, Belo Horizonte, Brazil, 2018.

An Ontological Approach to System-of-Systems Engineering in Product Development

Ludvig Knöös Franzén , Ingo Staack , Christopher Jouannet[*], and Petter Krus

Department of Management and Engineering (IEI), Linköping University, Linköping, Sweden
[*]Overall Design and System Integration, Saab Aeronautics, Linköping, Sweden
E-mail: ludvig.knoos.franzen@liu.se

Abstract

This paper presents an approach to system-of-systems engineering for product development with the use of ontology. A proposed method for building as well as using ontology to generate and explore system-of-systems design spaces based on identified system-of-system needs is presented. The method is largely built to cover the first levels of related work, where a process for system of systems in the context of product development is introduced. Within this work, it is shown that scenarios for a system-of-systems can be used to identify needs and subsequently the system-of-systems capabilities that fulfils them. The allocation of capabilities to possible constituent systems is used to show the available design space. The proposed method of this paper therefore addresses these initial challenges and provides a framework for approaching the system-of-systems design space creation using ontology. A case study is used to test the method on a fictitious search and rescue scenario based on available resources and information from the Swedish Maritime Administration. The case study shows that a representation of a system-of-systems scenario can be created in an ontology using the method. The ontology provides a representation of the involved entities from the fictitious scenario and their existing relationships. Defined ontology classes containing conditions are used to represent the identified needs for the system-of-systems. The invocation of a description logic reasoner is subsequently used to classify and create an inferred ontology where the available system-of-systems solutions are represented as sub-classes and individuals of the defined classes representing the needs. Finally, several classes representing different possible system-of-systems needs are used to explore the available design space and to identify the most persistent solutions of the case study.

Keywords: Systems Engineering, System-of-Systems, Ontology, Description Logic Reasoning, Design Space Creation

1 Introduction

Interest in *system-of-systems engineering* (SoSE) for aeronautical product development has seen steady growth in recent years, and aerospace systems are becoming more and more interconnected with their operational environments [1]. This, together with rapid advancements in technology, generates a desire for systems to collaborate to achieve capabilities that are not reachable by the individual systems alone [2]. In an ever-changing world, traditional approaches to product development for aerospace systems fall short when external factors such as politics, economics, regulations, technologies and doctrines affect the initially specified requirements for the systems. This problem becomes more pronounced when both the long product development time and the expected lifetime for aerospace systems are taken into account. A *system-of-systems* (SoS) perspective for product development of aerospace systems puts a focus on SoS needs and the required SoS capabilities that fulfil them. The corresponding requirements for *Constituent Systems* (CS) are sub-

sequently generated depending on the intended SoS architecture [1]. This SoS conceptualization of aerospace product development consequently takes the process to a more abstract level, where capabilities that are not achievable by the individual systems can be generated through system collaboration. A definition of SoS presented by [3] states that a SoS is separated from a typical complex system by five characteristic properties. These are operational independence of components, managerial independence of components, geographical distribution of components, evolutionary development of components and that the system experiences emergent behaviour. This definition of SoS is used throughout this paper. A proposed holistic product development in a SoS context was presented in [1], where the process of development was divided into five main levels of interest. The proposed SoS design process can be seen in Fig. 1.

The first level of interest presented in Fig. 1 describes how the needs and boundary conditions of the SoS can generate SoS capabilities from possible scenarios. It is stated in [1]
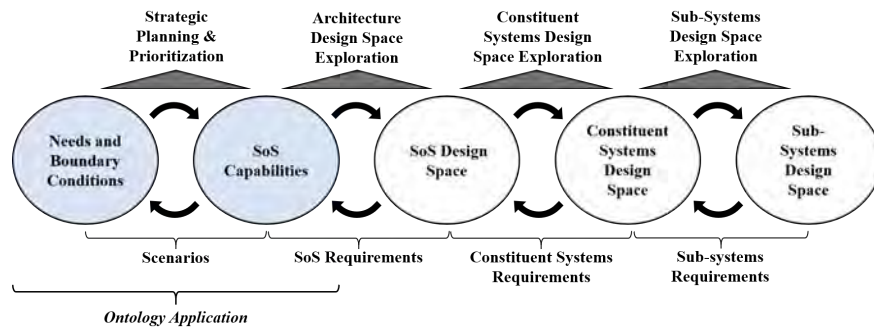
Figure 1: A representation of the holistic SoS design process from [1]. This paper covers the content under Ontology Application.

that analyses can be performed on the process by varying the initial conditions and boundaries to see how required capabilities respond to changes in the SoS needs. This consequently generates a design space of available capabilities that fulfil the SoS needs. The identified SoS capabilities from needs and scenario analyses can then be distributed to constituent systems that make up the SoS design space and architecture. The last two levels of the process describe how CS and their *sub-systems* (SS) should be chosen depending on the requirements generated in the previous levels. It is believed that ontologies are needed to connect the SoS process levels in a coherent way using common language and semantics.

The work presented in this paper therefore aims to provide a method for approaching the first two levels of the SoS-process and generating a SoS design space as presented in [1] using ontology. Furthermore, the influence of changing SoS needs on the available capabilities and SoS design space are also subject to investigation. This paper thereby contributes to the realization of the first two levels of interests shown in Fig. 1. A simple *search and rescue* (SAR) mission based on the Swedish Maritime Administration [4] is used as an implementation example to test the proposed method.

2   Frame of reference

The goal of the presented work is to provide a method for breaking down SoS needs into suitable capabilities and CS that together specify an available SoS design space. This chapter presents work that has been carried out in areas related to this paper and other approaches intended for the problem outlined in the introduction.

2.1   Capability- and system-of-systems engineering

The definition of a system according to the *International Council On Systems Engineering* (INCOSE) specifies that "*A system is a construct or collection of different elements that together produce results not obtainable by the elements alone*". [5]. The term SoS is defined by [6] as "*An interoperating collection of component systems that produce results unachievable by the individual systems alone*". As previously mentioned, [3] specifies that a Complex System or SoS is distinguished from a "conventional" system's definition by five different characteristics properties. [7] presents a

collection of views on SoS including a definition of SoS types. These include virtual, collaborative, acknowledged and directed SoS. It is furthermore argued that SoS are rarely developed as SoS. The process referred to as SoSE is rather initialized once an assessment of the SoS performance and capabilities begins. SoSE consequently involves the planning, analysis, organization and integration of CS. Capabilities performed by the CS are combined together with SoSE into SoS capabilities not achievable by the individual systems [2, 8]. This focus on capabilities for SoS can be referred to as capability engineering [9]. The process of capability engineering involves the identification of desired capabilities to be performed by the SoS, as well as investigating the possible options for attaining these capabilities [10]. Furthermore, [10] presents a method for supporting SoSE in translations of SoS capabilities into requirements. *Unified Modelling Language* (UML) object models are here used to model and increase the understanding of the involved systems with their respective functions, thus enabling explorations and trades to achieve desired capabilities. The *Systems Modelling Language* (SysML) has also been used to model SoS in various studies such as [11] and [12].

Enterprise architecture frameworks for defence industries such as the *US Department of Defence Architecture Framework* (DoDAF), the *UK Ministry of Defence Architecture Framework* (MoDAF) and the *NATO Architecture Framework* (NAF) have been used to model SoS in a *model based systems engineering* (MBSE) focused approach [13]. These architecture frameworks are intended to capture the operational, system, service, maintenance and information views of the SoS, among others. Other main actors within SoS and identified methods for approaching complex systems and SoS have been presented in [1]. [1] furthermore describes a holistic product development approach for SoS as mentioned in the introduction. The first of the proposed SoS-process levels shown in Fig. 1 describes the SoS needs and boundary conditions. The understanding of SoS needs and boundary conditions directly influences the strategic planning and prioritization, which together with the intended scenario defines the suitable SoS capabilities.

2.2 Ontology and ontological engineering for system of systems

A different approach to SoSE involves using ontologies. An ontology is an "*explicit specification of a conceptualization*" [14]. This is further explained as a formal and explicit representation of a given domain that involves knowledge of the involved entities and the relationships that exist between them [15]. Ontologies have seen a steady increase in usage for areas such as *systems engineering* (SE), SoSE and capability engineering [16–18]. An ontology can be used to enhance the interoperability aspect of a system or SoS [16], and it has been shown in [19] that scalability is also increased. Another approach to further enhance interoperability but on an ontology level is suggested in [20], where it is explained that a domain-neutral top-level ontology structure can support both the creation of new domain ontologies and the re-usability of existing ones. Top-level ontology examples can be found in [21]. Domain-specific ontologies are ontologies intended to describe individual systems or domains of interest. Such domain ontologies can for example be found in [22], which describes an ontology for aircraft design, or [23] where an ontology for information systems interoperability is presented.

An ontology can be implemented in different ontology languages. Current standards include languages such as the *Web Ontology Language* (OWL) and the *Resource Description Language* (RDF). OWL is based on RDF, but with the advantages of being better equipped for description logics and constraints checking [24]. An ontology made in OWL is composed of individuals, classes and their properties, which together are used to describe concepts of the intended domain modelled in the ontology. OWL supports the use of description logic reasoners, enabling the creation of more complex concepts out of simpler ones. Description logic reasoners can also be used to check for inconsistencies in the implemented ontology [25]. Different types of reasoners support different features [26]. Automatic reasoning over large ontologies requires large computational resources as indicated in [27, 28]. The scalability for automatic reasoning is thus hindered by computational resources. Ontologies that utilize heuristics can contribute to the efficiency of reasoning in domains containing incomplete data and consequently a large number of axioms [29]. There are also various optimization techniques which can contribute to the efficiency of reasoning [28].

Ontology-based approaches for modelling SoS are complementary to UML and SysML due to their ability to describe chosen domains from different terminologies and perspectives [18, 19]. Studies from Georgia Tech's *Aerospace Design Laboratory* (ASDL) have shown that ontologies can be used to model and prune the design space of cyber-physical systems in the context of conceptual design [30]. It has also been shown that matrix-based approaches such as an *Interactive Reconfigurable Matrix of Alternatives* (IRMA) can be used to illustrate and create the available design space from the knowledge captured in the ontology [31].

The presented frame of reference for this paper has identified methods and approaches for the modelling and usage of SoS in the context of aerospace product development. Figure 2 shows a representation of where the proposed scope of this research is situated. Based on the gathered information, a method of modelling an ontology intended for design space explorations on SoS is proposed in the next chapter.
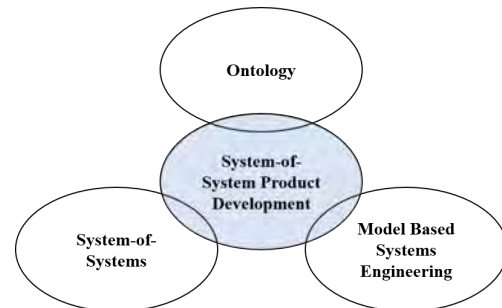
Figure 2: The positioning of the performed work compared to other areas and disciplines

3 Method

This chapter introduces a method for translating SoS needs into capabilities, and subsequently a way of generating the available design space for a SoS. It builds upon the holistic engineering approach for SoS introduced in [1], which can be seen in Fig. 1. As mentioned earlier, this holistic engineering approach suggests that the development is divided into five levels of interest, where each of the levels are recurrent and interrelated with each other. The presented levels are associated with a respective design space of solutions that should be derived by successive investigation and exploration of the previous levels.

3.1 Design space for system-of-systems capabilities

The first level of the presented process involves the description of the intended scenarios that the SoS is believed to be situated in during its lifetime. This description includes definitions for high-level frames of interest such as geopolitics, economics, customer needs, technology, laws and regulations, which together specify the overall boundary conditions and needs for the SoS. These initial conditions and needs must be varied based on different possible scenarios in order to achieve a holistic perspective on the intended SoS [1]. The influence and uncertainty of the high-level frames of interest are thereby evaluated. Suitable capabilities that can fulfil the needs should then be identified in order to generate the SoS capability design space. The capability design space is also subject to changing boundary conditions and needs, and a similar approach by varying scenarios should be used to increase the understanding of the available design space. These variations on initial conditions are used to explore the design space, consequently identifying the persistent solutions least affected by changes. These "resistant" capabilities represented in the design space are deemed to be suitable strategical choices based on the trade of boundary conditions and needs.

The distribution of the chosen capabilities can later be explored in a variety of different SoS architectures where CS and SS are assigned with the required capabilities. The process explained above is illustrated in Fig. 3.

Figure 3: A detailed view of the needs to SoS capability process.
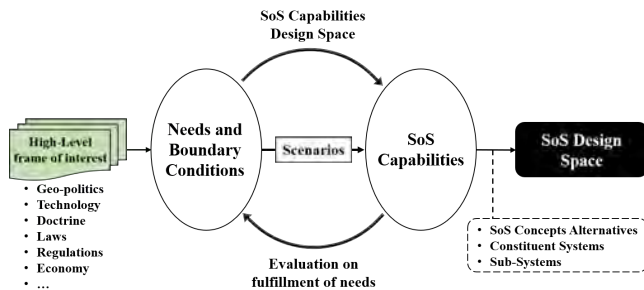
The method proposed in this paper suggests that ontology should be used to represent the main components presented in Fig. 3. The ontology provides a mapping of the involved actors and high-level frames of interest as well as the existing relationships between them. The following section describes a suggested approach for modelling an ontology intended for SoS design space generation and exploration.

3.2 Ontology design and usage

An ontology can be created in several ways, and it is even argued that there is no one correct way of modelling a domain since it depends on the application in mind [15]. The ontology creation method introduced in this paper is largely based on guidelines presented in [15] and [20], and is intended to be used with OWL. It follows eight successive steps which are illustrated in Fig. 4.

3.2.1 Step 1

The aim of the initial step is to determine the scope of the content and domain that is to be represented. Typically, this would correspond to the intended usage for the SoS to be developed, for example SAR.

Step (A) in Fig. 4 is an intermediate step which introduces the option of using a top-level ontology structure. There are several benefits with top-level ontologies that are explained in [32]. These include improved interoperability between domain ontologies.

3.2.2 Step 2

Holistic analyses of possible scenarios for the intended SoS domain and scope are used to identify entities to be implemented in the ontology. Such entities include laws, regulations, weather conditions, available assets, capabilities, needs and more. It is important to specify terms and vocabularies carefully to enable coherent formalism so that ambiguity in definitions is avoided.

Ontologies can suffer from a reinventing-the-wheel syndrome where several ontologies are created for the same domain [32]. It is therefore desirable to consider using existing ontologies for the SoS design space generation instead of creating a redundant one [15]. If a decision is made to utilize existing external ontologies, these should be modified to fit the intended formalism of the ontology under development.

3.2.3 Step 3

The third step in the ontology development process involves the definition and creation of all classes based on the entities identified in step 2. This is typically done in a hierarchy structure with classes and sub-classes. The identified SoS needs of the analysed scenarios should also be defined as classes. Different approaches can be used when developing the class hierarchy as described in [15]. It is important that unequal classes are defined as disjoint for the later invocation of the description logic reasoner.

Step (B) is optional and can be used if a top-level ontology structure is to be utilized. Top-level ontologies include predefined classes that provide an overarching framework for organizing the knowledge of different domain ontologies [32]. The classes and eventual domain ontologies that represent the intended SoS should consequently be allocated to the existing framework of the chosen top-level ontology.

3.2.4 Step 4

This step is used to describe the internal structure, existing relationships and properties of the previously created classes of the ontology. The associations between the classes should come from the scenario analyses in step 2, where the relationships between the different entities are identified. Classes should also include properties that describe them. These properties can include values in the form of numbers and lists. Cardinality specifications are used to describe the number of relationships a class can have in a min., max. or exactly logic. Further explanations and details of available properties and cardinalities in OWL can be found in [15] and [25]. Finally, classes can be defined as either *primitive* or *defined*. The differences between these are so-called *necessary* and *sufficient conditions*. A *primitive class* only includes *necessary conditions* while a *defined class* includes at least one *necessary* and *sufficient condition*. The use of *defined classes* allows for the automated classification and computation of class relationships in the ontology by a description logic reasoner. This is particularly useful when building large ontologies [25]. The classes representing the SoS needs and boundaries should be specified as *defined classes* for the reasoner to associate available solutions to the *necessary* and *sufficient conditions* specified in them. This process is further explained in chapter 4.1.

3.2.5 Step 5

Instances should represent the lowest granularity of the modelled ontology domain. This is described in [15] as "*Indi-*
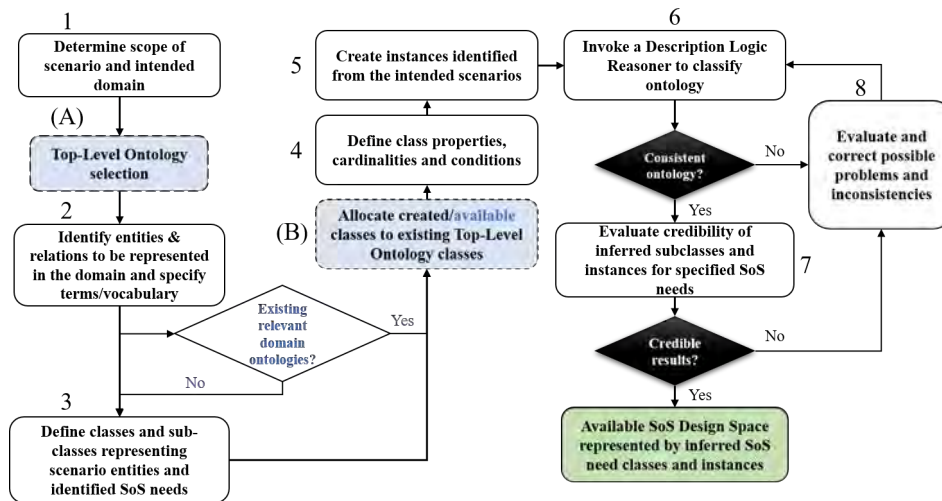
Figure 4: Ontology development process in the context of SoS design space generation.

vidual instances are the most specific concepts represented in a knowledge base". They represent the lowest level of detail for the intended domain and should be related to suitable corresponding classes of the ontology.

3.2.6 Step 6

The sixth step is used to invoke a description logic reasoner to check the consistency of the implemented ontology. The reasoner will classify the ontology and build an inferred ontology based on the previously defined relationships, properties and conditions of the different represented classes and instances. This process will associate the previously implemented classes representing the SoS needs of the ontology with all classes and individuals that have fulfilled the specified conditions. It will consequently show the available SoS design space as sub-classes and individuals to the inferred need classes. If the reasoner classifies the ontology as inconsistent, step 8 needs to be performed.

3.2.7 Step 7

It is recommended that the credibility of the inferred ontology and populated classes representing the needs should be evaluated. The ontology may prove to be consistent but include unreasonable or unwanted inferred relationships of the reasoner. If such unwanted relationships or results exist, step 8 needs to be carried out.

3.2.8 Step 8

This final step in the process presented in Fig. 4 is only to be performed if inconsistencies or unreasonable results or relationships exist in the inferred ontology. Possible corrections should be performed in the definition of classes and their relationships, properties and conditions. OWL supports the usage of various ontology debuggers that can prove to be useful for finding the source of any inconsistencies [33]. Unreasonable results can be further investigated by description logic querying, where the inferred ontology is used to answer "questions" about specific classes or instances [34].

The end result of the process shown in Fig. 4 is an inferred ontology where the defined need classes represent the available SoS design space. This process both creates and reduces the available SoS design space based on the description of needs identified from the possible scenarios for the intended SoS.

4 Implementation of case study

In order to test the proposed method of creating a SoS design space from specified needs with an ontology, a simple case study based on the operations and resources of the *Swedish Maritime Administration* (SMA) was performed. According to statistics, SAR at sea has accounted for approximately 60% of the operations performed by the SMA during the past three years and is a good example of a SoS where changing circumstances can affect the available solutions.

The scenario considered for the case study is fictitious and solely based on the available assets, regulations and capabilities of the SMA. It is also worth mentioning that the purpose of the case study is not to provide a complete ontology for Swedish SAR but instead to test the validity and usefulness of the proposed method in this context. Hence, the implemented scenario is kept simple in order to show the possibilities of the proposed method even at low levels of scenario detail. The use of a top-level ontology was also excluded for the purpose of this case study.

Following the process shown in Fig. 4, a SAR scenario was determined as the domain to be modelled in the ontology. This fictitious SAR scenario was created based on available information such as the assets, resources and regulations of the SMA [35] and is illustrated by a simple sketch in Fig. 5.

The scenario shown in Fig. 5 illustrates that there is a subject situated in the sea near Gothenburg, Sweden. The subject can make a distress call and indicate a position. Information
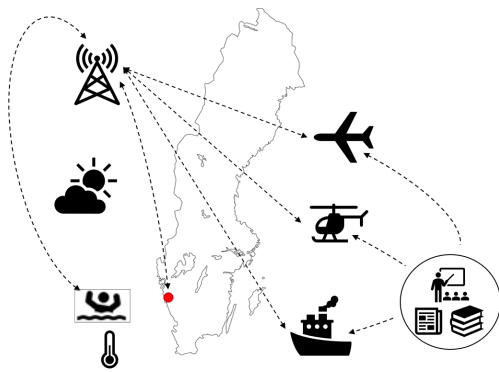
Figure 5: A fictitious scenario based on available assets and resources of the SMA.

Figure 6: The ontology class hierarchy for the intended scenario.

about the distress call is forwarded to the *Joint Rescue Co-ordination Centre* (JRCC) located in Gothenburg. The JRCC can dispatch suitable and available assets for the SAR of the subject. The available assets considered in this scenario are Bombardier Dash 8 Q300 aeroplanes, AW 139 helicopters and two fictitious kinds of sea vessels: fast and slow. The environment is associated with weather conditions, temperatures and time of year and day.

4.1   Ontology modelling

The process of identifying all entities and relationships to be modelled in the ontology involved detailed scenario analyses where available capabilities and assets were broken down into functions and means. Information flows between entities such as communication links were determined as well as the relationships between assets and their corresponding sub-components. After identifying entities and relationships to be represented in the ontology, the class and sub-class definition was initiated. The case study was implemented in the Protégé ontology editing software, which is based on OWL [36]. The class structure of the study was created using a top-down development process where the most general classes, such as System and Environment, were created first [15]. The class definition was done in a "is_a" manner, which meant that all sub-classes of an intended class inherited its properties. A depiction of the case study class hierarchy can be seen in Fig. 6. Classes not equal to each other are defined as disjoint.

The identified needs of the SoS are represented as the solution classes in Fig. 6 and are explained in more detail further down. The available assets of the case study are represented under the system classes which are defined by the various sub-system classes of the hierarchy structure. Figure 7 shows an example of how the *HelicopterSystem* class is defined.

The properties, cardinalities and conditions describing the various classes in the hierarchy were based on specifications of SMA assets, resources and available capabilities. The fictional sea vessel's properties were estimated based on high and low speed sea vessels for SAR. Mapping between capabilities, functions, systems and other classes was carried out to create the relational properties between all entities represented in the ontology hierarchy. Assets and resources
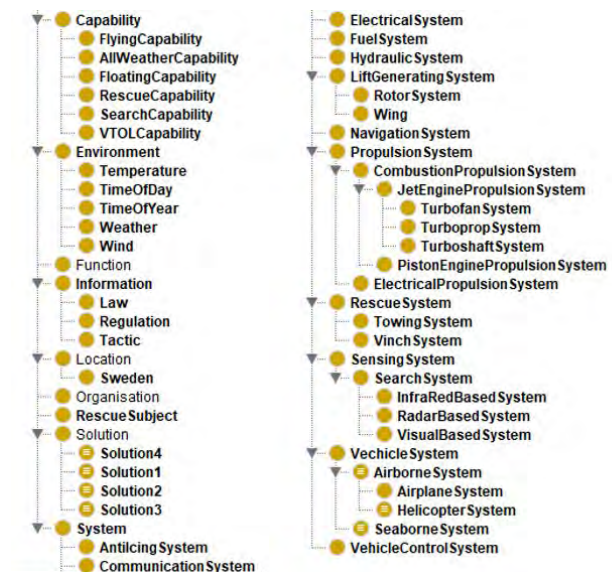
capable of performing the functions or capabilities from the former breakdowns were related to their respective function. Once a satisfactory mapping of the relationships between classes was achieved, the step of implementing the lowest levels, or instances, of the ontology was performed. The instances for this case study represented specific individuals such as an AgustaWestland 139 helicopter which is part of the SMA assets [35]. These individuals were assigned data properties describing their individual performances. Figure 7 shows the definition of a *HelicopterSystem* and the AW139 instance within the case study ontology.

It should be noted that the instance values described in Fig. 7 do not include any units. The data properties describing the relationships are instead assigned comments describing the intended units. The *hasOperationalRange* data property, for example, has the unit of kilometres in this case study. The instances of the sub-systems are defined for the sake of simplicity as various types indicating that resources can be composed of different alternatives such as different types of radar.

Finally, the solution classes describing the intended SoS needs were specified and given *necessary* and *sufficient conditions*. This was done so that the *Solution* classes were defined as equal to specified needs of the SAR scenario. The procedure was carried out by describing the required capabilities, functions and performances in the *necessary* and *sufficient conditions* of the *Solution* classes. An example of such a class can be seen in Fig. 8. The conditions described in the *Solution* classes of this case study are, as with the SAR scenario, fictitious and are only used to show the intended process of the method. It can be seen in Fig. 8 that the class describes a need for both a *search* and a *rescue capability*. It furthermore specifies that the rescue vessel needs to be able to carry at least 10 persons and have a length of 15 (meters) or above.

Figure 7: The definition of the HelicopterSystem class (top) and the AW139 instance (bottom), where both are defined by respective object and data properties.

4.2 Ontology usage

In order to classify the ontology, check consistency and obtain the available sub-classes and instances of the *Solution* classes, a description logic reasoner needed to be used. Protégé has a set of default reasoners as well as additional reasoners available through plug-ins. The default *Pellet* reasoner was considered suitable for this case study and was used throughout the implementation process. (For a detailed comparison of available reasoners in Protégé, see [26].) The reasoner was invoked to classify the ontology and consequently generate the logically inferred hierarchy. This process generated the available SoS design spaces as sub-classes and individuals of the defined *Solution* classes that could fulfil the *necessary* and
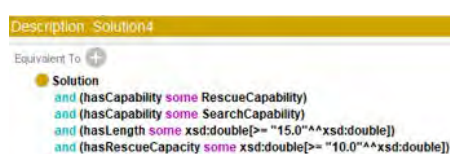
Figure 8: The necessary and sufficient conditions specified for the Solution4 ontology class.

sufficient conditions. These results are shown in Figs. 9 and 10.
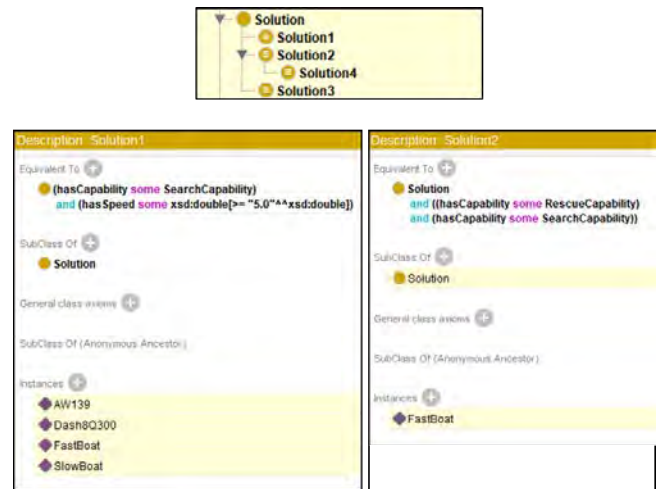
Figure 9: The inferred solution class structure (top) and the description of Solution classes 1 and 2 (bottom). Solution1 describes a need for a search capability and a minimum speed, while Solution2 requires both search and rescue capabilities.

Figure 10: The description of Solution classes 3 (left) and 4 (right). Solution3 specifies a need to keep the cost under a specified cost per hour.

It can be seen in Fig. 10 that the *Solution4* class from Fig. 8 has been populated with the *AW139* and *SlowBoat* individuals due to their fulfilment of the *necessary* and *sufficient conditions* specified in the class description. It should be noted that *Solution4* has been inferred as a sub-class of *Solution2* in Fig. 9. This is due to the solutions described in *Solution4* also fulfilling the conditions in *Solution2*. It is possible to carry out design space explorations by defining several solution classes based on different possible SAR scenarios and needs as illustrated in Figs. 9 and 10. Figure 11 shows the available design space for all *Solution* classes. The *Solution2* class describes needs for both search and rescue capabilities. All individuals in the *Solution4* class fulfil the conditions and are subsequently part of the solution for the needs described in *Solution2*.

By comparing all inferred design spaces, it can be seen that the individuals with the highest reoccurrence represent

DOI
10.3384/ecp19162004   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **41**

Figure 11: The inferred SoS design spaces where the Solution2 class is shown in the middle with corresponding individuals and Solution4 as a sub-class.

the most persistent solutions. The consequent SoS design space generated with the method is available in the OWL file of the inferred ontology and can be exported for further investigations and analyses. Such analyses can for example include determining the most suitable number of assets described by the inferred individuals based on *Measure of Effectiveness* (MoE), which is not provided by this method.

Finally, this case study has shown that the proposed method can be used to generate SoS design spaces for SAR and specified needs from a SoS scenario.

5 Discussion

The method proposed in this paper is intended to be used for any SoS design space creation and delimitation. The implementation of the case study was based on a fictitious existing SoS. This can be ref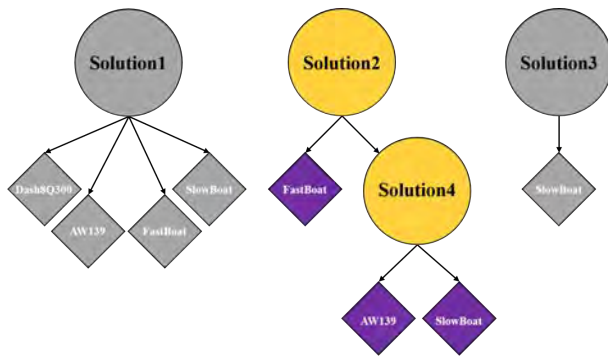erred to as a near-term SoS where available resources and how best to use them are investigated. The method and consequent use of ontology allows for the evaluation of new resources in combination with existing ones. New resources and capabilities can be introduced to the ontology structure and can be related to the existing entities. The use of a reasoner can once again prove useful for automatically expanding the ontology with new resources depending on their description. A new entity with a capability to fly can, for example, automatically be placed as a sub-class of *AirborneSystem* if an *AirborneSystem* has been defined with *necessary* and *sufficient conditions* specifying that an airborne system has the capability to fly. This ability makes ontologies highly scalable and flexible for introducing new entities to the modelled domain and building up taxonomies. However, this comes at the cost of increased computational time for automatic reasoning as described in section 2.2.

A combination of existing and new SoS solutions can be referred to as a mixed SoS. It is also possible to use the proposed method for analyses of long-term SoS where no existing resources are available. This typically corresponds to creating a SoS from scratch where only vague definitions of possible needs and capabilities are defined and related.

Assets and resources can be guessed and evaluated based on the scenario and epoch analysis of the intended SoS.

As mentioned in the frame of reference, there are supplementary methods for ontologies available for modelling and performing analyses on SoS. Enterprise architecture frameworks have been successfully used to describe SoS as well as intended scenarios in a SysML or UML language [13]. Relational databases have a similar structure to ontology, and it is possible to describe a SoS with all entities and the relationships that exists between them here as well. However, the advantage with ontologies is that they work under the previously mentioned open world assumption which, together with an ability to perform description logic reasoning, can be used to infer implicit knowledge of the intended domain. This implicit knowledge can reveal emergent behaviours of the SoS which are very important and desirable to understand at an early stage of development [17]. Nevertheless, ontology has some limits when it comes to inference by a reasoner. Such limits are shown in detail in [34], where inference is used to a high degree on an ontology. The presented method is, as previously mentioned, built upon existing methods for creating an ontology. There are many available methods and pieces of software for creating an ontology. Protégé was chosen as the implementation software due to its compatibility with OWL and RDF, as well as the support for description logics. The proposed method of this paper is however neutral in terms of the choice of software and could be used provided that the software supports a description logic reasoner.

An important delimitation of the presented case study is the exclusion of a top-level ontology structure. This was mainly excluded due to the fact that no existing ontology was to be used for the implemented example. The use of a top-level ontology would also imply a large increment in the class hierarchy structure and size, which would make the ontology less comprehensible to the human eye. Furthermore, the ontology for the case study was, as previously mentioned, implemented to show the possibilities for creating and exploring the SoS design space and not to provide an extensive ontology for SAR operations. However, the inclusion of a top-level ontology can be utilized at a later stage if desired due to the scalability and flexibility of ontology structures. This would require some restructuring of the class hierarchy to fit the top-level ontology properly, but would also enable re-usability and combinability with other ontologies that utilize the same top-level design. The scope of the modelled SAR case study was kept at a simple level to visualize the goal of the proposed process and method. It could however easily be expanded to include more details about the domain and available system components. The example components shown in Fig. 7 could be replaced with actual sub-system representations that have their own sub-classes and instances, leading to a question about fidelity levels in the ontology. The fidelity or detail level of the ontology is determined at the very beginning of the proposed method in Fig. 4. This is however also expandable at a later

stage due to the flexibility of ontologies. The results obtained in Figs. 9 and 10 show that the SoS needs can be represented as demands for the necessary performances and capabilities. The usefulness of this strategy becomes increasingly distinct in ontologies featuring a larger number of resources and possible solutions. The reasoner becomes a valuable asset for inferring knowledge and creating the available design space in larger ontologies.

The results in Figs. 9 and 10 show that the available solutions are instances of the *Solution* classes. It is possible to use the method in order to generate sub-classes representing the available design space as a complement to a set of individuals. The open world assumption that OWL ontologies work under can, however, complicate class inferences. A closed world assumption assumes that data which does not exist is false, while an open world assumption sees non-existent data as simply unknown. Ontologies consequently do not make any assumptions about incomplete data. This means that ontology classes can include more entities that are simply not yet known unless the class has been explicitly stated not to do so using closure axioms [25]. The open world assumption is also a reason why ontologies are so flexible and easy to expand. Instances are, however, easier to work with since they represent the lowest level of the intended ontology and are less affected by the open world assumption.

Finally, the implemented case study is still under development at the time of writing this paper. More SMA information and resources must be implemented in order to generate a larger pool of available solutions for the SoS design spaces. More detailed capability and functional breakdowns are subject to implementation to obtain a more comprehensive picture of the existing relationships in the domain. Additionally, more extensive design space explorations will be performed by defining more solution classes based on scenario and epoch analyses of the case study. As mentioned in section 3.1, varying the initial conditions to increase the understanding of the design space is an important future addition to the case study.

6   Conclusions and future work

The proposed method introduced in this paper has shown that a design space of possible SoS solutions can be generated using ontology. An implementation of a *search and rescue* (SAR) case study was carried out to test the method and show the process of building an ontology for SoS design space generation. Several ontology classes describing different needs were implemented and populated with suitable instances representing solutions by a description logic reasoner. This corresponded to a small design space exploration which was used to identify the most common solution elements of the SoS. The future work of this study involves the expansion of the case study and extraction of the SoS design space generated in the inferred ontology file. Similar approaches to those presented in [31] can be used to transfer the information from the ontology to matrix-based approaches. This transfer allows

for more advanced numerical calculations and optimizations needed for the remaining levels of interest proposed in [1]. Furthermore, the design space generation of this paper has only shown available types of solutions and no number of required assets. Future work includes the determination of the most suitable SoS architectures with regard to number and collaboration of assets. The work so far has, however, shown that ontologies provide a resilient way of exploring and generating SoS design spaces based on specified needs. It contribute to the ways of approaching the complex challenges of today's product development.

References

[1] Ingo Staack, Kristian Amadori, and Christopher Jouannet. "A Holistic Engineering Approach for Aeronautical Product Development". *Proceedings of the Congress of the International Council of the Aeronautical Sciences*, pages 1–15, Belo Horizonte, 2018.

[2] Office of the Deputy Under Secretary of Defence for Acquisition and Technology, Systems and Software Engineering. "Systems Engineering Guide for System of Systems". Version 1.0. Washington, DC: ODUSD(A&T)SSE, 2008.

[3] Mark W. Maier. "Architecting Principles for Systems-of-Systems". *Systems Engineering*, 1(4):267–284, 1998.

[4] Sjöfartsverket, "Sjö- och flygräddning", 2019. [online]. available: https://www.sjofartsverket.se/sv/sjofart/sjo–och-flygraddning/. [accessed 2019-07-02].

[5] INCOSE, "About Systems Engineering", 2019. [Online]. Available: https://www.incose.org/about-systems-engineering. [accessed 2019-07-02].

[6] David D Walden, Garry J Roedler, Kevin J Forsberg, Douglas R Hamelin, and Thomas M Shortell. "Systems Engineering Handbook". Wiley, San Diego, 4 edition, 2015.

[7] Jo Ann Lane. "What is a system of systems and why should i care?". Technical report, USC-CSSE-2013-001, University of Southern California, 2013.

[8] Charles Keating, Ralph Rogers, Resit Unal, David Dryer, Andres Sousa-Poza, Robert Safford, William Peterson, and Ghaith Rabadi. "System of systems engineering". *EMJ - Engineering Management Journal*, 15(3):36–45, 2003.

[9] CAE, "Capability Engineering", 2019. [Online]. Available: https://www.cae.com/media/documents/Defence_Security/Services_-_Documents/datasheet.capability.engineering.pdf. [accessed 2019-07-02].

[10] Jo Ann Lane. "System of systems capability to requirements engineering". *Proceedings of the 9th International Conference on System of Systems Engineering: The Socio-Technical Perspective, SoSE 2014*, pages 91–96, 2014.

[11] Jo Ann Lane and Tim Bohn. "Using SysML to Evolve Systems of Systems". Technical report, INCOSE, 2010.

[12] Marco Mori, Andrea Ceccarelli, Paolo Lollini, Bernhard Frömel, Francesco Brancati, and Andrea Bondavalli. "Systems-of-systems modeling using a comprehensive viewpoint-based SysML profile". *Journal of Software: Evolution and Process*, 30(3):1–20, 2018.

[13] Ron Williamson. "INCOSE (MBSE) Model Based System Engineering (SoS) System of Systems/Enterprise Activity Introduction", 2012.

[14] Thomas R Gruber. "A translation approach to portable ontology specifications". *Knowledge Acquisition*, 5(2):199–220, 1993.

[15] Natalya F Noy and Deborah L McGuinness. "Ontology Development 101: A Guide to Creating Your First Ontology". Technical report, Stanford, 2014.

[16] Leo Van Ruijven. "Ontology for systems engineering: Model-based systems engineering". In *Proceedings - UKSim-AMSS 6th European Modelling Symposium, EMS 2012*, pages 371–376. IEEE, 2012.

[17] Gary Langford and Teresa Langford. "The making of a system of systems: Ontology reveals the true nature of emergence". In *2017 12th System of Systems Engineering Conference, SoSE 2017*, pages 1–5. IEEE, 2017.

[18] Huseyin Dogan, Michael J de C. Henshaw, and Julian Johnson. "7.5.1 An incremental hybridisation of heterogeneous case studies to develop an ontology for capability engineering". In *INCOSE International Symposium*, volume 22, pages 956–971. INCOSE, 2012.

[19] John S. Osmundson, Thomas V. Huynh, and Paul Shaw. "Developing Ontologies for Interoperability of Systems of Systems". In *Conference on Systems Engineering Research*, 2006.

[20] Barry Smith, Mauricio Almeida, Jonathan Bona, and Mathias Brochhausen. et al. "Basic Formal Ontology 2.0 SPECIFICATION AND USER'S GUIDE". Technical report, 2015.

[21] Viviana Mascardi, Valentina Cordì, Paolo Rosso, and Paolo Rosso Viviana Mascardi, Valentina Cordì. "A Comparison of Upper Ontologies". Technical report, 2007.

[22] Markus Ast, Martin Glas, Tobias Roehm, and Bauhaus Luftfahrt eV. "Creating an Ontology for Aircraft Design". In *Deutscher Luft- und Raumfahrtkongress 2013*, pages 1–11, 2013.

[23] Peter Morosoff, Ron Rudnicki, Jason Bryant, Robert Farrell, and Barry Smith. "Joint doctrine ontology: A benchmark for military information systems interoperability". In *CEUR Workshop Proceedings*, volume 1523, pages 2–9, 2015.

[24] Diana Kalibatiene and Olegas Vasilecas. "Survey on Ontology Languages". Lecture Notes in Business Information Processing, 2011.

[25] Matthew Horridge, Holger Knublauch, Alan Rector, Robert Stevens, Chris Wroe, Simon Jupp, Georgina Moulton, Nick Drummond, and Sebastian Brandt. "A Practical Guide To Building OWL Ontologies Using Protégè 4 and CO-ODE Tools Edition 1.3". Technical report, Manchester, 2011.

[26] Sunitha Abburu. "A Survey on Ontology Reasoners and Comparison". *International Journal of Computer Applications*, 57(17):33–39, 2012.

[27] Thomas Lampoltshammer and Stefanie Wiegand. "Improving the computational performance of ontology-based classification using graph databases". *Remote Sensing*, 7:9473–9491, 07 2015.

[28] Volker Haarslev and Ralf Möller. "On the scalability of description logic instance retrieval". *Journal of Automated Reasoning*, 41:99–142, 08 2008.

[29] Kristin Stock, Didier G. Leibovici, Luciene Delazari, and Roberto Santos. "Discovering order in chaos: Using a heuristic ontology to derive spatio-temporal sequences for cadastral data". *Spatial Cognition & Computation*, 15:115–141, 2015.

[30] Kevin Lynch, Randall Ramsey, George Ball, Matt Schmit, and Kyle Collins. "Conceptual design acceleration for cyber-physical systems". In *11th Annual IEEE International Systems Conference, SysCon 2017 - Proceedings*, 2017.

[31] Matt Schmit, Simon Briceno, Kyle Collins, Dimitri Mavris, Kevin Lynch, and George Ball. "Semantic design space refinement for model-based systems engineering". In *10th Annual International Systems Conference, SysCon 2016 - Proceedings*, 2016.

[32] Robert Arp, Barry Smith, and Andrew D Spear. "Building Ontologies With Basic Formal Ontology". Massachusetts Institute of Technology, Cambridge, 2015.

[33] Konstantin Schekotihin, Patrick Rodler, and Wolfgang Schmid. "OntoDebug: Interactive Ontology Debugging Plug-in for Protégé". Technical report, Alpen-Adria-Universität, Klagenfurt, 2018.

[34] Robert Stevens, Margaret Stevens, Nicolas Matentzoglu, and Simon Jupp. "Manchester Family History Advanced OWL Tutorial". Technical report, The University of Manchester, Manchester, 2015.

[35] Lars Widell and Peter Hellberg. "Svenskt Program för Sjö- och Flygräddning". In *Sjöfartsverket Dr-nr 1199-13-01359 (in swedish)*, 2013.

[36] Mark A. Musen. "The protégé project. a look back and and a look forward". *AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence*, 1(4), 2015.

A study for a MDO Process Applied to Conceptual Design of a Remotely Piloted Aircraft

Maria Luiza Cassão Gatelli[1], Carlos Eduardo de Souza [2], and Marcos Daniel de Freitas Awruch[3]

[1]Aerospace Engineering Course
[2,3]Mechanical Engineering Department
[1,2,3]Federal University of Santa Maria (UFSM), Santa Maria, RS - Brazil.
E-mail: marialuizagatelli@gmail, carlos.souza@ufsm.br, marcos.awruch@ufsm.br

Abstract

The aircraft design process involves the interaction among disciplines with different nature. Finding an optimal concept is a trade off task that requires time and experience from designers and engineers. Multidisciplinary Design optimization (MDO) methods help to automatise and to simplify that complex task, aiming to obtain design solutions that comply with multiple requirements and constraints at once. The present work seeks to develop a conceptual design framework of a remotely piloted aircraft (RPA). A review on this subject is presented, identifying different MDO approaches. The analysed disciplines of aeronautical design are: geometry, aerodynamics, weight, performance, stability and flight dynamics. The relationship among them are presented in a block diagram. Afterwards, a technique of MDO is applied to the framework, assembling a mono objective problem using the Particle Swarm optimization (PSO) algorithm to minimise aircraft structural mass under stability constraints. This objective is obtained from dimensional parameters of feasible aircraft concepts as output solution of the process. Therefore, the MDO method speeds the design process up, without employing the traditional trial and error approach, which, as well as doesn't guarantee the achievement of an optimal configuration, turns the process slow and expensive.

Keywords: Conceptual design. MDO. optimization. PSO. RPA.

1 Introduction

The global market and international competition lead aeronautical industry to continuously improve its performance in the engineering design process [1]. As described by Raymer [2], the aircraft design process can be divided into three major phases: conceptual, preliminary and detail design. The design changes have a lower cost in the conceptual phase than in the subsequent ones, following a trend shown in Fig. 1. This figure presents the cost increasing as the project advances in time and also how the impact of design changes decreases as the level of detail increases.

Aiming to reduce time and improve design confidence level, the application of new design automated methods is welcome. "As the aircraft design is multidisciplinary by nature, the need to apply *Multidisciplinary Design optimization* (MDO) methods is well understood and accepted" [4]. "The main motivation for using MDO is that the performance of a multidisciplinary system is driven not only by the performance of the individual disciplines but also by their interactions" [5].

Furthermore, the operation of *remotely piloted aircraft* (RPA) is drawing an increasing attention in military and civil sectors because of the absence of human operators on board,
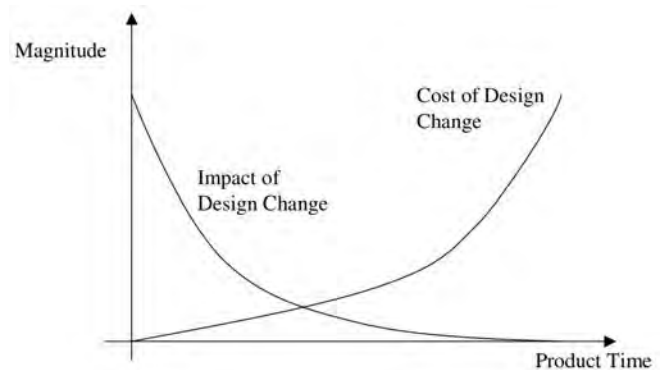
Figure 1: Cost and impact of design change over product time, adapted from [3].

what reduces operating costs and expands capabilities. These aircraft are also called *unmanned aerial vehicle* (UAV) and defined according to the US Department of Defence (DoD) as: "a powered, aerial vehicle that does not carry a human operator, uses aerodynamic forces to provide vehicle lift, can fly autonomously or be piloted remotely, can be expendable or recoverable, and can carry a lethal or nonlethal payload. Bal-

listic or semi ballistic vehicles, cruise missiles, and artillery projectiles are not considered unmanned aerial vehicles" [6].

The application of RPA, which had been largely military since the 1950s, has expanded into the civilian market. Tasks such as battlefield observation and surveillance have been replaced by monitoring of urban areas, plantations and power transmission lines or simply recreation, in the form of model aircraft with imaging systems. Table 1 lists examples of civilian uses of RPA.

Table 1: Examples of civilian uses of RPA, adapted from [7].

Area	Example
Aerial Photography	Films and advertising videos
Agriculture	Crop monitoring and spraying and herd monitoring and driving
Coastguard	Search and rescue, coastline and sea-lane monitoring
Conservation	Pollution and land monitoring
Energy Industry	Power line inspection, wind turbine inspection
Fire Services and Forestry	Fire detection and incident control
Gas and Oil Ind.	Land survey and pipeline security
Information Services	News information and pictures, feature pictures, e.g. wildlife
Lifeboat Institutions	Incident investigation, guidance and control
Local and Police Authorities	Survey, disaster control, search for missing persons, security and incident surveillance
Meteorology	Sampling and analysis of atmosphere for forecasting
Traffic Agencies	Monitoring and control of road traffic
Ordnance Survey	Aerial photography for mapping
Survey Organisations	Geographical, geological and archaeological survey
Water Boards	Reservoir and pipeline monitoring
etc.	

Another application of RPA is as research platform in academic and industrial spaces, aiming to develop aeronautical knowledge and employee qualification. Two factors are of key importance: the reduced dimensions allow the aircraft to be produced in small laboratories and by few people and the costs are very small when compared to the development of a manned aircraft. All electronic components, for example, are readily available from online stores, and a small specialization is needed to integrate a complete airframe.

The present work focuses on establishing a MDO process dedicated to the design of RPA and displays the recent advances in that purpose. The next section presents a brief literature review, followed by the proposed methodology. The numerical studies section details some simulations and optimization results applied to a small aircraft design.

2 Bibliography review

"A *remotely piloted aircraft system* (RPAS) comprises a number of subsystems which include the aircraft, its payloads, the control station(s) (and, often, other remote stations), aircraft launch and recovery subsystems where applicable, support subsystems, communication subsystems, transport subsystems, etc." [7].

Sobieszczanski-Sobieski and Haftka [8] presented a revision about multidisciplinary optimization in aerospace design, focusing on the multiple ways that engineers deal with the main challenges of a MDO process: computational cost and organisational complexity. According to the authors, it's possible to identify three categories of MDO problems. The first includes problems with two or three interactive disciplines, where a single analyst might acquire all necessary knowledge. At analysis level, this can lead to the creation of a new discipline responsible for the interaction between the involved disciplines, as aeroelasticity or thermoelasticity. The next category includes works in which the multidisciplinary optimization of the entire system is executed at a conceptual level by simple analysis tools. The third category includes problems that focus on organisation and computational challenges and develops techniques able of solving them.

Sobester and Keane [9] proposed a *Computer Aided Design* (CAD)-based UAV conceptual design framework and discussed some of the challenges arising at the conceptual level. The design workflow was described from the mission profile until the structural model and flow analysis. The basis of MDO study was the flight envelope resulted from the performance analysis of the aircraft. The results of six MDO iterations were achieved combining three different values of negative twist and two values of outboard sweep angle, showing the maximum wingtip deflections versus the maximum L/D ratios. The authors concluded that commercial, off-the-shelf, CAD tools could be integrated into the design process as early as the conceptual level, providing the models required by the various strands of multidisciplinary analysis.

Landolfo [10] outlined a structural and aerodynamic design of an UAV with multiple lift surfaces. The author investigated if a particular nonplanar wing concept could achieve the requirements of the mission of a small reconnaissance, surveillance and target acquisition UAV. The results, compared under varying assumptions specific to an equivalent monoplane and biplane, suggest that potential efficiency gains for the new configuration may be possible using the nonplanar wing configuration under explicit conditions. However, a conventional monoplane concept is suggested, given the areas of study covered in the paper.

Lundstrom et al. [11] proposed a distributed framework for *micro aerial vehicle* (MAV) design automation. A discrete propulsion system modelling is combined with a parametric CAD model and a panel code for aerodynamic performance prediction. The design optimization has been divided into two successive parts through a genetic algorithm. The first one runs without invoking CAD and aerodynamic analysis tools until convergence is achieved. The second part uses the pre-

vious optimum solution as the starting point to consider these time-costly software. The process runs until the results from the first and second part agree. The author suggests that the design framework is a helpful tool in MAV design. However, it is necessary to ensure the proper balance and size restrictions are met and the MAV must have large enough volume to accommodate its intended components.

Agte et al. [12] affirmed that the origins of MDO are found in the development of structural optimization. The inclusion of other disciplines into the process was a natural evolution, since the aerodynamics, propulsion and performance variables, for example, are intrinsically related to the structural efficiency. Initial processes were developed in sequential levels, where variables and objective functions were uncoupled, and later it was possible to study in a single optimization block with as much as possible coupling among disciplines.

Medeiros [13] developed an UAV for monitoring and sensing agricultural activities. The methodology applied to the prototype construction, the UAV development, the monitoring and processing of captured aerial images are presented in this master's thesis. The results are considered satisfactory due to the design requirements and to the image quality of tillage areas.

Martins and Lambe [5] performed extensive research about MDO architectures, classifying them in monolithic (a single optimization problem is solved) and distributed (the same problem is partitioned into multiple subproblems). A unified description with terminology, mathematical notation and diagram development was provided aiming to compare the benefits and drawbacks of each category. The authors suggest that it is necessary to test multiple architectures on a given MDO problem to determine which one is most efficient for each case. The work is a powerful tool to find the most appropriate architecture for the optimization problem.

Elmendorp et al. [14] developed a conceptual design method and analysis of conventional and unconventional aircraft. Initially, the top-level requirements (TLRs) in combination with a chosen design configuration and design objective are translated into a first estimate of the aircraft's geometry and its associated key performance indicators. A fully parametrized multi-model generator generates input data for the disciplinary analysis tools, including weight, aerodynamics, landing gear sizing, engine and cost. The verification of the software is performed by comparing the output data of a conventional configuration aircraft with another reference aircraft with similar TLRs, resulting in satisfactory error rates. The authors also analysed four different configurations demonstrating the variety of the configurations that may be designed and analysed as well as the validity of the comparison.

Bryson et al. [15] presented a multidisciplinary design analysis and optimization of quiet small unmanned aerial systems (SUASs). A multi-objective, non dominated sorting genetic algorithm (NSGA-II) was used to find the Pareto-optimal trade space for maximising range and minimising vehicle noise. The design variables were wing area, aspect ratio, taper ratio, maximum camber, maximum camber location,

maximum thickness and loiter speed. The constraints were angle of attack, propulsion weight, fuel weight and thrust required. The study provided suitable fidelity for early (conceptual to preliminary) design and ratified noise as a discipline integrated at the conceptual design level.

Hosseini et al. [16] applied a decoupling Uncertainty-based Multidisciplinary Design optimization (UMDO) method to design UAVs as a case study, using a genetic algorithm optimization method and sequential quadratic programming (SQP). The optimizer is located at system level within which UAV design parameters (a total of 37) are achieved in a way that, observing the problem's constraints (a total of 33), optimization criterion (overall mass of the UAV) becomes minimum. In order to provide an algorithm without uncertainties, it was necessary to identify the doubtful sources of the parameters, eg fuel consumption and aerodynamic coefficients. Then, the method proposed by the authors was applied to a study case, generating an increase in the objective function of the problem (mass). This result was justified as an algorithm way out to compensate for failure probability and elevate the chance of success to 100%.

Papageorgiou et al. [17] developed a MDO framework applied to UAV design focusing on radar signature and sensor performance considering also flight trajectory. The optimization problem of the case-study considered was about improving the performance of the aircraft over a specified mission while simultaneously considering the efficiency of the stealth and surveillance systems. One of the objectives was the minimisation of the Maximum Take-Off Weight (MTOW). The proposed MDO framework took into account aircraft's geometry, aerodynamics, trim, stability and simulation of the mission. The solution of the present optimization problem was achieved by implementing a single and a multi-level decomposition architecture. The proposed multidisciplinary optimization and analysis framework had the potential to significantly enhance the performance of the design when mission, stealth, and surveillance requirements must be considered.

Analysing the references cited previously, a work like Landolfo [10] elucidates the difficulties found during the unmanned aircraft design process, due to the lack of availability and reliability of historic regressions and design coefficients. "The potential consequences of early design decisions made with insufficient data include missed opportunities to increase vehicle performance and cost overruns to meet performance metrics or correct design defects" [15].

It is explicit the dependence between the mission requirements and the methodology adopted during the design process and the chosen optimization method. According to Martins and Lambe [5], there are many architectures available to solve the optimization problems, and they must be presented in diagrams, such as those found in Bryson et al. [15], Hosseini et al. [16] and Papageorgiou et al. [17]. It is notorious that the disciplines related to airworthiness, like propulsion, mass, structures, aerodynamics and stability, receive more attention in an optimization framework than disciplines that are not in the field of aeronautics. Hosseini et al. [16] and Papageorgiou et al. [17] integrated into their framework

non-conventional disciplines like noise and radar signature, respectively, demonstrating the potential of an embracing architecture integrated in an optimization process during the aircraft conceptual design phase.

2.1 Design phases

The division of the aircraft design into the three major phases, already cited above, is important to establish limits for decision making regarding to passing to subsequent phases and finally to manufacturing. In the first phase, the conceptual design, the objective is to develop alternative concepts that meet the problem solution. No excessive details should be taken into account in this phase since it would delay the process but also add many constraints. The second phase, the preliminary design, performs the definition of product layout configuration, technical and economically analysed. The detailed design, the third phase, is responsible for its final detailing, including meticulous documentation of the product and planning of the manufacturing process.

2.2 Design requirements

The customer's needs have to be translated into design requirements, mostly through customer plus engineering requirements. The first ones refer to objectives as articulated by the customer or client. The engineering ones refer to the design and performance parameters that can contribute to achieving the customer requirements [18]. The system engineering is fundamental in this process, because, as design requirements are a formalisation of all basic characteristics that involve an aircraft design, gross mistakes might occur if those are not well defined.

The aircraft mission defines the payload onboard in a determined velocity and range, utilising a certain type of engine. "The functional analysis is a process of translating system requirements into detailed design criteria and the subsequent identification of the resources required for system operation and support" [18]. Examples of functional criteria of RPA design includes maximum weight or span, operating altitude, range, fuselage volume, launching method, etc.

2.3 Multidisciplinary design optimization

The MDO technique was developed for complex engineering systems involving multiple disciplines or subsystems. The space design is explored in a form that the variables of interest are optimized under some constraints, simultaneously and automatically, until a satisfactory design point for all disciplines is found.

According to Agte et al. [12], the general formulation of an optimization problem can be written as

$$
\begin{aligned}
\min \quad & f(\mathbf{x}, \mathbf{p}) \\
& \mathbf{x} = [x_1, ... x_n]^T, \quad \mathbf{p} = [p_1, ... p_m]^T \\
& x_{i,L.B.} \le x_i \le x_{i,U.B.}, \ i = 1, 2, ... n \\
s.t. \quad & \\
& \mathbf{g}(\mathbf{x}, \mathbf{p}) < 0, \quad \mathbf{h}(\mathbf{x}, \mathbf{p}) = 0,
\end{aligned} \tag{1}
$$

where f is the objective function, $\mathbf{x}$ is a n-dimensional vector of design variables with lower and upper bounds, $\mathbf{p}$ is a

vector of fixed parameters that influence the behavior of the system but cannot be freely chosen (material properties, operating conditions,..), and $\mathbf{g}$ and $\mathbf{h}$ are vectors of inequality and equality constraints, respectively. However, when dealing with multiple disciplines, it might be a difficult task to solve the minimization problem at once, and different strategies must be sought.

One of the most important considerations when implementing MDO is to find a way of organising each discipline analysis models, simplified models and the optimization method agreeing with a correct problem formulation, as discussed by Martins and Lambe [5]. The MDO architecture is the combination of the problem formulation with the organisational strategy, allowing to identify how the different models are coupled and how the overall optimization problem is solved, depending if its structure is monolithic or distributed. This organisation itself is a challenge. The next section describes the strategy and framework proposed in the present work.

3 Proposed methodology

This works aims to develop a conceptual design tool of RPA integrated with an optimization algorithm. At this stage of the work, it is applied to a Mini UAV category, which is capable of being hand-launched and operating at ranges up to 30 km, according to Austin [7]. The tool considers aircraft with conventional configuration tractor or pusher with a payload bay with a single tail boom leading to a conventional, cruciform or "T" tail configuration without landing gear and electric propulsion system.

The first step to develop the design tool is to outline clearly the objectives of the conceptual design phase. Then, it's necessary to organise a framework aiming to reach a viable concept that meets the mission requirements. The framework structure depends upon the conceptual design process of unmanned aircraft. It can be expressed through a block diagram relating the various disciplines addressed in aircraft design such as aerodynamics, weight and stability. Afterwards, the optimization method can be applied, defining the project variables, objective function and constraints.

3.1 Remotely piloted aircraft conceptual design

The unmanned aircraft design process differ from the manned ones, once the information source, historic regressions and design coefficients are not available or are not reliable [19]. In this work, the following disciplines of aeronautical design are analysed: geometry, aerodynamics, weight, performance, stability and flight dynamics. The wing and tail geometry are calculated according to Gudmundsson [20]. Weight prediction is based on the volume calculation of each structural component and multiplied by a material density. Aerodynamics coefficients are calculated using Roskam Class II methodology [21]. Performance estimates are found from Traub [22] formulation. Stability derivatives are predicted from empirical data presented in Roskam [23]. Flight dynamics are analysed based on Cook [24] methodology.

3.2 Particle swarm optimization

"A PSO algorithm explores the design space of the objective function by adjusting the trajectories of the individual agents, also called particles, as these trajectories form piece wise paths in a quasi-stochastic manner" [25]. The movement of a swarm particle consists of two components: one stochastic and another deterministic. Each particle is attracted toward the position of the current global best $\mathbf{b}^*$ and its own best location $\mathbf{x_i}^*$ in history, while at the same time it has a tendency to move randomly. The aim is to find the global best among all the current best solutions until the objective no longer improves or after a certain number of iterations.

The position vector $\mathbf{x_i}$ and the velocity $\upsilon_\mathbf{i}$ for the particle i are utilised to determine the new velocity vector as

$$\upsilon_\mathbf{i}^{t+1} = \upsilon_\mathbf{i}^t + \alpha \varepsilon_\mathbf{1} \odot [\mathbf{b}^* - \mathbf{x_i}^t] + \beta \varepsilon_\mathbf{2} \odot [\mathbf{x_i}^* - \mathbf{x_i}^t], \quad (2)$$

where $\varepsilon_\mathbf{1}$ and $\varepsilon_\mathbf{2}$ are two random vectors that vary between 0 and 1 and the parameters α e β are acceleration constants.

The initial velocity for the particle may be taken as zero $\upsilon_\mathbf{i}^{t=0} = 0$. Then, the new position is updated as

$$\mathbf{x_i}^{t+1} = \mathbf{x_i}^t + \upsilon_\mathbf{i}^{t+1}. \quad (3)$$

According to Yang [25], the pseudo code of a PSO algorithm can be written as follows:

Initialise positions $\mathbf{x_i}$ of the n particles
Evaluate objective function $f(\mathbf{x})$, $\mathbf{x} = (x_1,...,x_p)^T$
Find $\mathbf{b}^*$ from $min\{f(\mathbf{x_1}),...,f(\mathbf{x_n})\}$ at $t = 0$
while (criterion)
  $t = t + 1$ (Iteration counter)
  for loop over all n particles and all p dimensions
    Generate new velocity $\upsilon_\mathbf{i}^{t+1}$ (Eq. (2))
    Calculate new locations $\mathbf{x_i}^{t+1}$ (Eq. (3))
    Evaluate objective function at new locations $\mathbf{x_i}^{t+1}$
    Find the current best for each particle $\mathbf{x_i}^*$
  end for
  Find the current global best $\mathbf{b}^*$
end while
Output the final results $\mathbf{x_i}^*$ and $\mathbf{b}^*$

In order to add constraints to the PSO problem, it's possible to transform it into an unconstrained problem by using the penalty method. When some constraint is not respected, a penalty parameter may be used, for example,

$$\lambda = e^{10|\mathbf{g}(\mathbf{x})|}. \quad (4)$$

Therefore, a corresponding value λ for each constraint will penalize the computed objective, turning the penalty function into the new objective function $\lambda f(\mathbf{x})$.

3.3 Definition of the optimization problem

This paper aims to investigate a mono objective constrained optimization utilising a PSO technique. The optimization problem must be assembled according to Eq. 1, resulting in

$$\begin{aligned}
\min \quad & f(\mathbf{x}, \mathbf{p}) = M_{tot} \\
& x_{i,L.B.} \le x_i \le x_{i,U.B.}, \ i = 1, 2, ...9 \\
s.t. \\
& C_{m_\alpha} < 0, \\
& C_{n_\beta} > 0, \\
& 12.5\% < SM < 17.5\%, \\
& \lambda_{Alp} < 0 \quad e \quad \lambda_{Alg} < 0.
\end{aligned} \quad (5)$$

In the above problem, $\mathbf{x}$ and $\mathbf{p}$ are the design variables and parameters vector, defined as

$$\mathbf{x} = \begin{bmatrix} AR \\ S \\ \lambda \\ \Lambda_{LE} \\ x_{r_w}/l_{fus} \\ \lambda_{fus} \\ AR_{HT} \\ AR_{VT} \\ l_{HT}/\bar{c} \end{bmatrix} \quad \text{and} \quad \mathbf{p} = \begin{bmatrix} H \\ V_\infty \\ V_{HT} \\ V_{VT} \\ Vol_{fus} \\ C_{Bat} \\ \rho_{Mat} \\ M_{PL} \\ \vdots \end{bmatrix}. \quad (6)$$

The components of the vector of design variables $\mathbf{x}$ are described in Table 2. The vector $\mathbf{p}$ of design parameters considers characteristics relating to tail and fuselage geometry, aerodynamics, propulsion, flight control and materials. Some of the parameters are listed in Table 3.

The constraints considered in this study are presented in Table 4 and correspond to flight quality parameters. These are just a small set of constraints, but play an important role in the final design configuration. This is a key condition for future improvements, since constraints such as aeroelastic conditions (divergence and flutter) or structural response (wing root stress) can be easily implemented.

Table 2: Design variables of the optimization problem in study.

Area/Nº	Design variable	Symbol
	Wing geometry	
1	Aspect ratio	AR
2	Area	S
3	Taper ratio	λ
4	Sweep of the leading edge	Λ_{LE}
5	Wing root location relative to fuselage length	x_{r_w}/l_{fus}
	Fuselage	
6	Fineness ratio	λ_{fus}
	Tail geometry	
7	Aspect ratio of horizontal tail	AR_{HT}
8	Aspect ratio of vertical tail	AR_{VT}
9	Distance of both horizontal tail $\bar{c}/4$ to wing $\bar{c}/4$ along the x-axis relative to $\bar{c}$	$l_{HT}/\bar{c}$

These disciplines are organised in a $N \times N$ design diagram in order to outline the inputs and outputs regarding to each dis-

Table 3: Design parameters of the optimization problem in study.

Area/Nº	Design parameter	Symbol
	Tail geometry	
1	Tail boom radius	r_{TB}
3	Horizontal tail volume coefficient	V_{HT}
4	Vertical tail volume coefficient	V_{VT}
	Fuselage	
10	Internal volume	Vol_{fus}
	Aerodynamics	
11	Operating altitude	H
12	Cruise velocity	V_∞
17	Wing airfoil	-
18	Tail airfoil	-
	Propulsion	
19	Electric motor mass	M_{Mot}
20	Electric motor power	P_{Mot}
	Flight control	
23	Battery capacity	C_{Bat}
25	Battery mass	M_{Bat}
26	Servomotor mass	M_{Srv}
	Materials	
31	Density	ρ_{Mat}
	Payload	
34	Payload mass	M_{PL}

Table 4: Design constraints of the optimization problem in study.

Area/Nº	Constraint	Equation
	Stability	
1	Pitching moment due to angle of attack derivative	$C_{m_\alpha} < 0$
2	Yawing moment due to sideslip derivative	$C_{n_\beta} > 0$
3	Limits of static margin	$12.5\% < SM < 17.5\%$
4	Real part of eigenvalues of matrix of longitudinal and lateral directional dynamics	$\lambda_{long} < 0$ and $\lambda_{lat} < 0$

cipline and also between them. This approach allows a multidisciplinary and integrated framework converging between subsystems.

The implemented code is organised according to the diagram shown in Fig. 2, that is an adaptation of a complete $N \times N$ diagram, with information running only in one direction. Here, the main design code calls each discipline from left to right, and thus the information to each block comes always in this direction. The final feedback comes from each block to the *Design* block, and then to the PSO algorithm. It means that there the result from each block on the right only comes back to a block on the left after updating data in the main design block.

The proposed framework is implemented completely in Python language and takes advantage of open source packages. All disciplines are implemented in specific functions, organised such that the input and output variables are available at the *Design* block. A 3D plotting function was also implemented to better visualise the final concept and identifies the general dimensions of the aircraft, as well as the centre of gravity position of the main parts.
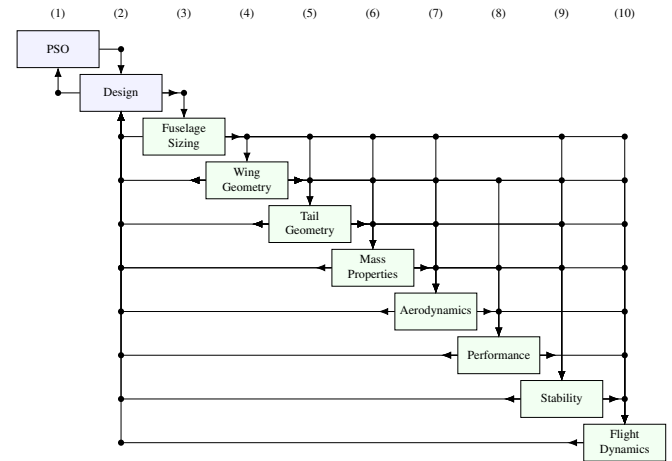

Figure 2: Proposed optimization diagram.

4 Numerical studies

The remotely piloted aircraft conceptual design tool is applied to the design of a small RPA. The aircraft described in 4.1 is considered as basis for the mission and design requirements for a case study. The application of the methodology is limited to conventional configurations, such as tube and wing, with electric propulsion.

4.1 Reference model

The chosen reference aircraft is the Spy Owl 200 Research Version presented in Figure 3. It is a commercial RPA designed exclusively for research. The main technical specifications of the Spy Owl 200 are presented in Table 5, from where stand out the Maximum Take-off Weight (MTOW) and span, including take-off and landing method, characteristics that have a great impact in the aircraft structural mass.

Figure 3: Spy Owl 200, a reference model for the present study, adapted from [26].

Table 5: Technical specifications of the reference aircraft, the Spy Owl 200, adapted from [26].

Characteristic	Value
Wing span	2.01 m
Length	1.53 m
MTOW (Hand-launched)	6.5 kg
Autonomy	2 hours
Typical mission altitude	75-1500 m AGL
Maximum payload mass including batteries	2.7 kg
Take-off method	Hand-launched
Landing method	Belly landing
Structure	Extremely robust full composite structure
Cruise speed	14 - 22 m/s
Maximum speed	40 m/s

4.2 Design requirements

The design requirements of the case study aircraft of a potential client proposed by the authors are presented below:

- The aircraft should be capable of being easily transported between operation sites;

- The aircraft should be light enough to be hand-launched;

- The aircraft should be recovered from a belly landing;

- The aircraft should be capable of being remotely piloted or to fly autonomously with interference of a human operator;

- The aircraft should transmit the flight data real-time for ground system.

The RPA mission is to carry a research payload with a determined mass utilising electric propulsion.

The functional requirements are:

- Maximum mass of 6.5 kg;

- Maximum span of 2 m;

- Cruise speed of 20 m/s;

- Hand-launched by the operator;

- Recovered by belly landing;

- Operating altitude of 100 m.

RPA used as research platform in academic laboratories, for example, may have a reduced number of subsystems. The launching requirements induce the absence of landing gear, reducing system complexity. The transporting requirement between operating sites induces to the facility of carrying RPAS by an operator between one place and another. The maximum mass requirement is associated with the physical resistance of an operator responsible for the launching.

4.3 Input data

Initially, it is necessary to define the set of input variables, and their lower and upper bounds. Those values are presented in Table 2. The values of some of the design parameters vector **p** are listed in Table 7. Those values are fixed for the entire process.

At this point, there is still a need for a careful interference of the designer, since the limit values should be chosen based on engineering knowledge, and especially considering experience or known references. The limits of the wing and tail aspect ratio are estimated according to Gudmundsson [20]. The limits of wing area are defined by the authors based on the reference model. Gudmundsson [20] also presents reference values for taper ratio and sweep angle of the wing leading edge. The fuselage fineness ratio is estimated based on Nicolai and Carichner [27]. The variable l_{HT} usually has a value between $3\bar{c}$ and $6\bar{c}$, where $\bar{c}$ is wing *mean aerodynamic chord* (MAC). Then, in tail geometry discipline, the variable l_{HT} is multiplied by $\bar{c}$. The same procedure is adopted for x_{r_w}. The maximum value of this variable is $(l_{fus} - c_r)/2$, where c_r is wing root chord.

Table 6: Lower and upper bounds of the design variables of the optimization problem in study.

Nº	Variable	$x_{i,L.B.}$	$x_{i,U.B.}$	Unit
1	AR	6	11	[-]
2	S	0.05	0.5	[m^2]
3	λ	0.3	1	[-]
4	Λ_{LE}	0	10	[deg]
5	x_{r_w}/l_{fus}	0	1	[-]
6	λ_{fus}	3	6	[-]
7	AR_{HT}	3	5	[-]
8	AR_{VT}	1	2.5	[-]
9	$l_{HT}/\bar{c}$	3	6	[-]

The values of V_{HT} and V_{VT} are the same utilised by Landolfo [10]. The tail taper ratio and sweep of the leading edge are chosen in order to simplify the design. The fuselage volume is based on the volume of propulsion and flight system components and in an estimate of payload mass.

In order to estimate the aircraft centre of gravity, a table with the estimated location in x, y e z and mass of each of the basic component is assembled and is updated through the process. This allows updating the aircraft mass and inertia properties for each design.

4.4 Minimisation of aircraft structural mass

Since propulsion and flight control systems are fixed, as well as payload, the objective is to minimise aircraft structural mass, finding the optimal geometry that respect all constraints. To start the optimization, the number of particles is 50. There are two stop criteria. The first one is related to a minimum number of iterations that the best value of the objective function is not updated. It means that if the best value of the global objective function remains the same for

Table 7: Fixed design parameters of the optimization problem
in study.

Area/Nº	Parameter	Value	Unit
	Tail geometry		
1	r_{TB}	0.015	[m]
3	V_{HT}	0.675	[-]
4	V_{VT}	0.0375	[-]
	Fuselage		
10	Vol_{fus}	0.0062	[m^3]
	Aerodynamics		
11	H	100	[m]
12	V_∞	20	[m/s]
17	Wing airfoil	NACA2412	[-]
18	Tail airfoil	NACA0012	[-]
	Propulsion		
19	M_{Mot}	0.082	[kg]
20	P_{Mot}	438	[W]
	Flight Control		
23	C_{Bat}	3	[Ah]
25	M_{Bat}	0.346	[kg]
26	M_{Srv}	0.1	[kg]
	Materials		
31	ρ_{Mat}	1420	[kg/m^3]
	Payload		
34	M_{PL}	0.5	[kg]

1000 iterations, the code considers it as a convergence. The
second stop criteria is related to the standard deviation of the
last 1000 best global values of the objective function. If this
standard deviation value is lower than 0.001, the code also
converges.

Objective function iteration histories for 10 runs are shown in
Fig. 4. The different initial values are due to random esti-
mates of the PSO. The evaluations of the objective function
(artificial value of the aircraft mass due to penalty factors)
varied within 1.64 and 1.8 kg, resulting in an error of approx-
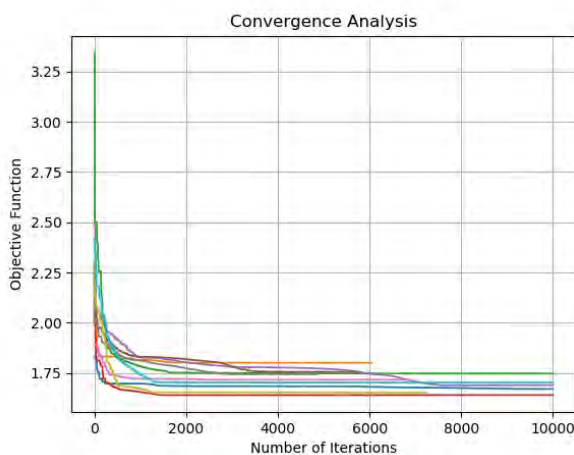imately 9.8% for this analysis.

Figure 4: Objective function iteration history.

4.5 Final concept

The final result obtained from the optimization defines an air-
craft concept, which should be the basis for the next design
phases. The history of iterations of the global objective func-
tion for the optimized concept is presented in Fig. 5. More
than 4000 iterations were necessary to achieve the optimal
result. That means one hour running in a personal notebook,
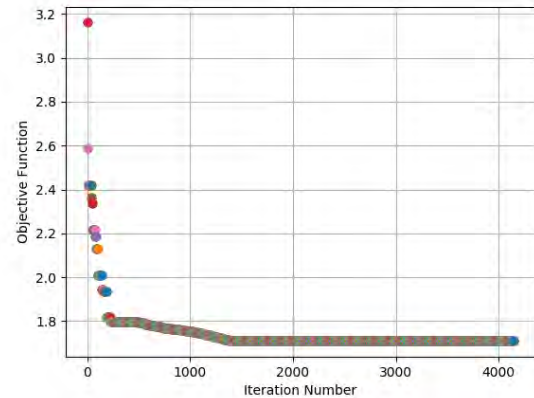with an Intel Core i5-4210U processor, at 1.7 GHz.

Figure 5: Objective function iteration history of the final
concept.

Figures 6 to 8 show the evolution of aircraft configuration at
different iteration steps of history shown in Fig. 5. The air-
craft configuration at iteration number 1 is presented in Fig. 6.
The length and span dimensions are larger than the following
cases, what impacts in the total weight. This configuration
doesn't respect all imposed constraints, especially the static
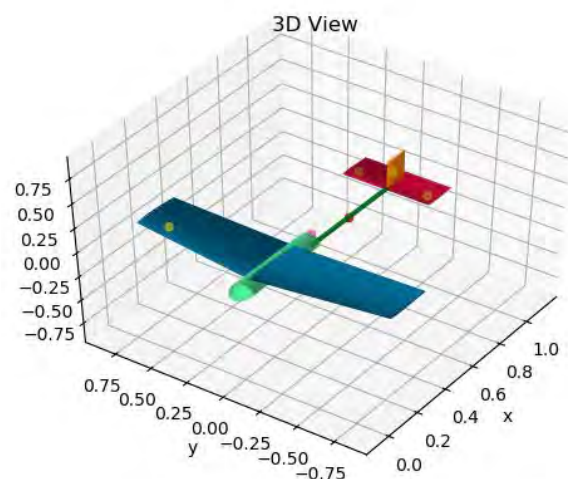margin, that is equal to 2.4 %.

Figure 6: Aircraft configuration at iteration number 1.

An intermediate condition can be seen in Fig. 7 obtained at
step 2000. It is a configuration already very similar to the final
concept, shown in Fig. 8. The sweep angle is already similar,
and the tail boom is shorter than the initial one. The final span
is 1.289 m, with a small taper, with chord varying from 0.150

m at the root to 0.116 m at the tip. The mass practically did not change after iteration 1400, reaching a value near 1.71 kg. The final convergence was achieved at iteration 4150 because no better value was obtained after 1000 steps.
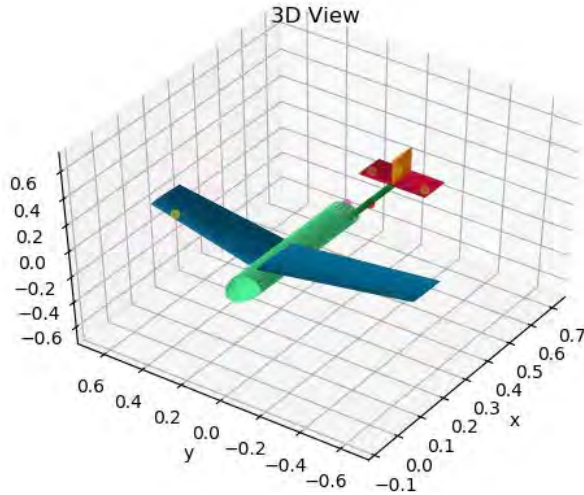

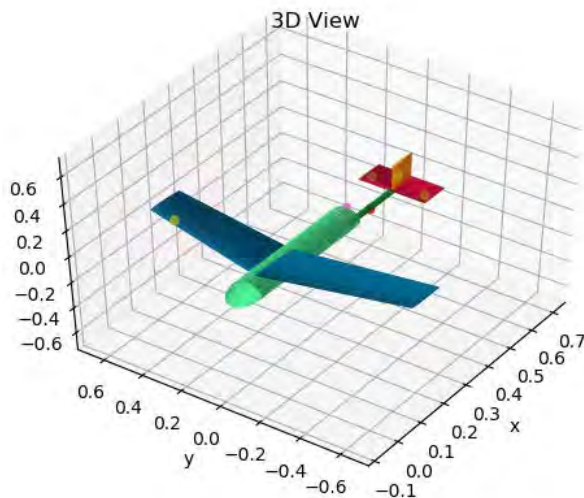Figure 7: Aircraft configuration at iteration number 2000.

Figure 8: Aircraft final configuration.

The optimized design variables vector is presented in Table 8. The star * indicates that the specific design variable stopped into the lower bound defined in Table 6. In this case, the fuselage fineness ratio gets lower because the volume is constant, and thus the MDO process tries to control the CG position and the inertia values, due to the static margin constraint.

Some of the output data from the optimization problem are presented in Table 9. In the aerodynamics block, i_w and i_{HT} are the incidence angles of the wing and horizontal tail, respectively, C_{D_0} is induced drag coefficient and k is defined as $1/(\pi e AR)$, where e is Oswald's efficiency. In the performance block, E is endurance and R is range estimate. In the stability block, C_{l_β} is the rolling moment due to sideslip derivative and C_{y_β} is the side force due to sideslip derivative.

Table 8: optimized design variables obtained in the case study.

Nº	Variable	x_{gbest}	Unit
1	AR	9.67	[-]
2	S	0.17	[m^2]
3	λ	0.77	[-]
4	Λ_{LE}	8.47	[deg]
5	x_{r_w}/l_{fus}	0.13	[-]
6	λ_{fus}	3*	[-]
7	AR_{HT}	3.69	[-]
8	AR_{VT}	2.17	[-]
9	$l_{HT}/\bar{c}$	3.87	[-]

Table 9: Output data from the optimization problem in study.

Area/Nº	Parameter	Value	Unit
	Mass properties		
1	M_{tot}	1.71	[kg]
	Aerodynamics		
11	i_w	3.23	[deg]
11	i_{HT}	1.31	[deg]
11	C_{D_0}	0.055	[-]
12	k	0.041	[-]
	Performance		
34	E	36	[min]
34	R	44	[km]
	Stability		
34	C_{m_α}	-0.584	[1/rad]
34	C_{n_β}	0.0	[1/rad]
34	C_{l_β}	-0.079	[1/rad]
34	C_{y_β}	-0.412	[1/rad]
34	SM	12.5	[%]

The solution respected all imposed constraints. The output data from aerodynamics are in the expected order of magnitude. The moment coefficient derivative in respect to the angle of attack C_{m_α} is negative, ensuring longitudinal static stability to the aircraft. The yaw moment coefficient derivative with respect to the sideslip angle C_{n_β} is zero. It respected the imposed constraint, but it was different from the expected ($C_{n_\beta} > 0$) to ensure directional static stability.

The eigenvalues of the longitudinal and lateral directional dynamics, describing short period characteristics for example, are presented below:

$$\lambda_{long} = \begin{Bmatrix} -6.8 \pm 7.32j \\ -9.21 \times 10^{-2} \pm 0.551j \\ -1.31 \times 10^{-4} + 0.j \end{Bmatrix} \quad (7)$$

$$\lambda_{lat} = \begin{Bmatrix} -116.43 + 0.j \\ -0.247 \pm 1.6j \\ -3.17 + 0.j \end{Bmatrix} \quad (8)$$

5 Concluding remarks

This paper describes the development of a tool for remotely piloted aircraft conceptual design considering a MDO framework integrated with a PSO optimization algorithm. The tool embraces diverse aeronautical engineering disciplines (aerodynamics, stability, flight dynamics, etc.), all coupled into a single driver code that handles all design information. This way, it's easier to perceive how an integrated process can facilitate decisions among a multidisciplinary team.

A numerical study is presented, aiming to obtain a concept of a small RPA. The case study considers the definition of multiple variable constraints, which requires certain expertise by the designer. Sometimes, the range of each variable differ considerably from the usual general aviation values, usually proposed by consecrated references, like Roskam [21] and Raymer [2]. The number of subsystems, and consequently, parts, decrease when the designer's team is dealing with unmanned aircraft. This happens, mainly, because it's not necessary to add life support and comfort systems. Also, dimensions are smaller, and the aerodynamics parameters are affected, for example. Therefore, to consider the constraints of RPA missions using known general aviation methodologies is still a challenging task.

This is an ongoing development work. The disciplines analyses are implemented in Python language in form of independent packages, called by the main code. So, new disciplines can be easily added or the ones already present can be improved. The same approach is applicable to the optimization method, once the RPA design code is also a module to be called in the routine.

List of Abbreviations and Acronyms

CAD *Computer Aided Design*.
MAV *micro aerial vehicle*.
MDO *Multidisciplinary Design optimization*.
RPA *remotely piloted aircraft*.
RPAS *remotely piloted aircraft system*.
UAV *unmanned aerial vehicle*.

References

[1] W. J. Vankan, W. F. Lammen, and R. Maas. Metamodeling and multi-objective optimization in aircraft design. Technical report, National Aerospace Laboratory, 2010.

[2] Daniel P. Raymer. *Aircraft Design: A Conceptual Approach*. AIAA education series. American Institute of Aeronautics and Astronautics (AIAA), Washington, D.C., 4th edition, 2006. ISBN 1563478293.

[3] Peter De Baets and Dimitri Mavris. Potential formulation for aeroelastic constraint analysis in a conceptual design environment. In *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Structures, Structural Dynamics, and Materials and Co-located Conferences. American Institute of Aeronautics and Astronautics, April 2002. doi: 10.2514/6.2002-1295.

[4] Plamen Roglev. MDO Framework for Conceptual Design of Closed Wing UAV. *International Journal of Scientific Research Engineering & Technology (IJSRET)*, 2(8):526–531, 2013.

[5] Joaquim R. R. A. Martins and Andrew B. Lambe. Multidisciplinary Design Optimization: A Survey of Architectures. *AIAA Journal*, 51(9):2049–2075, 2013. ISSN 0001-1452. doi: 10.2514/1.J051895.

[6] B. Leonard. *Department of Defense Dictionary of Military and Associated Terms: As Amended Through April 2010*. DIANE Publishing Company, 2011. ISBN 9781437938203.

[7] Reg Austin. *Unmanned Aircraft Systems: UAVS Design, Development and Deployment*. John Wiley & Sons, Inc., Chippenham, Wiltshire, 2010.

[8] Jaroslaw Sobieszczanski-Sobieski and Raphael T. Haftka. Multidisciplinary aerospace design optimization - survey of recent developments. In *AIAA 34th Aerospace Sciences Meeting and Exhibit*, Reno, NV, Jan 15-18 1996.

[9] Andras Sobester and Andy Keane. Multidisciplinary design optimization of uav airframes. In *47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Structures, Structural Dynamics, and Materials and Co-located Conferences. American Institute of Aeronautics and Astronautics, May 2006. doi: 10.2514/6.2006-1612.

[10] Giuseppe Landolfo. Aerodynamic and structural design of a small nonplanar wing uav. Master's thesis, University of Dayton, 2008.

[11] David Lundstrom, Kristian Amadori, and Petter Krus. Distributed framework for micro aerial vehicle design automation. In *46th AIAA Aerospace Sciences Meeting and Exhibit*, Aerospace Sciences Meetings. American Institute of Aeronautics and Astronautics, January 2008. doi: 10.2514/6.2008-140.

[12] Jeremy Agte, Olivier de Weck, Jarowlaw Sobieszczanski-Sobieski, Paul Arendsen, Alan Morris, and Martin Spieck. MDO: assesment and direction for advancement - an opinion of one international group. *Structural and Multidisciplinary Optimization*, 40: 17–33, 2010.

[13] Fabricio Ardais Medeiros. Desenvolvimento de um veículo aéreo não tripulado para aplicação em agricultura de precisão. Master's thesis, Universidade Federal de Santa Maria, 2007.

[14] R Elmendorp, R Vos, and Gianfranco La Rocca. A conceptual design and analysis method for conventional and unconventional airplanes. In *29th Congress of the International Council of the Aeronautical Sciences, ICAS*. International Council of Aeronautical Sciences, 2014.

[15] Dean E. Bryson, Christopher R. Marks, Ryan M. Miller, and Markus P. Rumpfkeil. Multidisciplinary design optimization of quiet, hybrid-electric small unmanned aerial systems. *Journal of Aircraft*, 53(6):1959–1963, June 2016. ISSN 0021-8669. doi: 10.2514/1.C033455.

[16] Majid Hosseini, Mehran Nosratollahi, and Hossein Sadati. Multidisciplinary design optimization of unmanned aerial vehicle under uncertainty. *Journal of Aerospace Technology and Management*, 2017. doi: http://dx.doi.org/10.5028/jatm.v9i2.725.

[17] Athanasios Papageorgiou, Johan Ölvander, and Kristian Amadori. Development of a multidisciplinary design optimization framework applied on uav design by considering models for mission, surveillance, and stealth performance. In *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA AVIATION Forum. American Institute of Aeronautics and Astronautics, June 2017. doi: 10.2514/6.2017-4151.

[18] M.H. Sadraey. *Aircraft Design: A Systems Engineering Approach*. Aerospace Series. Wiley, 2012. ISBN 9781119953401.

[19] Jay Gundlach. *Designing Unmanned Aircraft Systems: A Comprehensive Approach*. American Institute of Aeronautics and Astronautics, Inc., Reston, Virginia, 2012.

[20] Snorri Gudmundsson. *General Aviation Aircraft Design: Applied Methods and Procedures*. Elsevier, Waltham, Massachusetts, 1st edition, 2014.

[21] Jan Roskam. *Airplane Design*. Roskam Aviation and Engineering Corporation, 1985.

[22] Lance W. Traub. Range and endurance estimates for battery-powered aircraft. *Journal of Aircraft*, 48(2): 703–707, 2011. doi: 10.2514/1.C031027.

[23] Jan Roskam. *Methods for Estimating Stability and Control Derivatives of Conventional Subsonic Airplanes*. Roskam aviation and engineering corporation, 1983.

[24] M.V. Cook. *Flight Dynamics Principles: A Linear Systems Approach to Aircraft Stability and Control*. Elsevier aerospace engineering series. Elsevier Science, 2011. ISBN 9780080550367.

[25] Xin-She Yang. *Engineering Optimization: An Introduction with Metaheuristic Applications*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2010.

[26] UAS Europe. Spy owl 200, 2019. URL `http://www.uas-europe.se/index.php/products-mobile/spy-owl-200-mobile`. Online; Acessed: 10 set. 2018.

[27] Leland M. Nicolai and Grant E. Carichner. *Fundamentals of Aircraft and Airship Design*, volume Vol. I - Aircraft Design. American Institute of Aeronautics and Astronautics, Inc., 2010.

II Domain-specific System Analysis

Aerodynamic Performance of Natural Laminar Flow Aerofoils Applied to Low- and High-Speed Wings
Ramon Lopez Pereira, Jose Omar Martinez Lucci and Fermin Navarro Medina

Equations of State in Fighter Aircraft Oleo-pneumatic Shock Absorber Modelling
Arttu Heininen, Jussi Aaltonen, Kari T. Koskinen and Juha Huitula

Evaluation of Increase Weight in a Wing Fixed Leading Edge Design to Support UAS Impact
Tomaz Drumond, Marcelo Greco and Carlos Cimini

Aerospace Electric Generator Specification and Selection - Opportunities and Challenges
Avo Reinap

Understanding Loudness variations for Aircraft Landing Procedures at Arlanda
Anders Johansson and Karl Bolin

Aerodynamic Performance of Natural Laminar Flow Aerofoils Applied to Low- and High-Speed Wings

Ramón López Pereira*,[2], José Omar Martínez Lucci[2], and Fermín Navarro Medina[3]

* SAFRAN Engineering Services GmbH, Donauwörth, Germany
[2] Department of Mechanical and Aerospace Engineering, Universidad Europea de Madrid, Villaviciosa de Odón, Spain
[3] Department of Mechanical Engineering, Thermal Machines and Engines and Fluids, Universidad de Vigo, Vigo, Spain
E-mail: ramon.lopez-pereira@safrangroup.com (ramlopper@gmail.com),
joseomar.martinez@universidadeuropea.es, fermin.navarro.medina@uvigo.es

Abstract

The aim of this investigation is to assess if the use of Natural Laminar Flow (NLF) aerofoils alone improves the efficiency of a wing in comparison to their NACA equivalents.

To compare performance, the lift-to-drag ratio (Efficiency) of the wings has been considered. A cruise speed range from M0.1 to M0.3 was analysed for the low-speed wing, while a Mach sweep between M0.3 and M0.85 was assessed for the high-speed wing. A final analysis was performed to evaluate the effect of the presence of a sweepback angle in the wing, from straight leading edge to a 10-degree-sweepback wing.

For the low-speed wing, it was observed that the efficiency of the laminar wing is slightly decreasing with speed (up to 17% between M0.1 and M0.3) but is increasingly higher than in the NACA wing (from 5% at M0.1 to 16% at M0.3). This means that the aim of the laminar aerofoil is met, so that for a cruise speed between 100 and 300km/h lower drag is produced and therefore lower thrust (and fuel consumption) is required.

In the case of the high-speed application, it was found that the laminar flow wing had lower efficiencies when compared to its NACA equivalent. Analysing the results, it was noted that the ratio of the lateral forces to lift had a direct relation to the efficiency: when this ratio was increased, the efficiency was decreased, and vice-versa. It was observed that, in the case of laminar wings, not applying a sweepback to the leading edge (LE) is optimal and duplicates the efficiency with respect to adding any angle. Moreover, this is the only case observed where the efficiency of the laminar wing is higher than its NACA equivalent.

Keywords: Environmentally friendly technology; laminar flow; low-speed; high-speed; aerodynamic efficiency; performance

1 Introduction

Aircraft design has always tried to put the airlines at focus. This means that the wishes of the airlines need to be considered. Previously, this meant bringing more passengers further and faster, but the current path is to have greener aircraft, that is, the emissions of carbon and nitrogen oxides should be reduced as well as the noise made by the aircraft. This wish is not only driven by the airlines' corporate social responsibility guidelines, which is also the case, but predominantly due to more restrictions in terms of regulation, mainly in Europe, aiming for a cleaner sky [1].

Even though the main focus in this cleaner sky approach has been set for engine manufacturers (either directly by new regulations or by the aircraft builders), the main competitors in the aircraft-manufacturing field have been receiving pressure from the market to release improvements towards the clean sky objective.

This investigation tries to assess if the use of NLF aerofoils provides real benefit in terms of drag reduction or if it is still non-reliable for this use. A new design of high-speed NLF aerofoil is included to try to optimize its performance. Even though the usual wing design will include several aerofoils inside it, for the purpose of this research it was chosen to use a single aerofoil spanwise.

As a first approach, standard wing parameters have been used on both aerofoil types, and then a wing sweepback angle parametric analysis have been performed in order to assess its effect on high-speed wings. In the case of low-speed wings, the selected aerofoils are a NASA-designed NLF(1)-1015 aerofoil [2] and its 4-digit NACA equivalent (NACA4615), while for the high-speed case a new aerofoil has been designed (named Raw Laminar Profile for High Speed Flow, RLPHSF) and its NACA equivalent (NACA1709) has been compared against it.

To compare the performance of the wings, the efficiency has been obtained at different cruise conditions for each case. In the low-speed designs, this has been selected as 4900m (FL160) International Standard Atmosphere (ISA) and a speed range between M0.1 (30m/s, 60KTAS) to M0.3 (95m/s, 185KTAS). For the high-speed wings, two sweeps are performed as indicated before, though both at 11000m ISA (FL360). A speed variation between M0.3 (90m/s, 175KTAS) and M0.85 (250m/s, 485KTAS) was considered for the standard design, while a fixed M0.77 (230m/s, 445KTAS) condition with LE-sweepback angles ranging from zero to 10 degrees were also analysed.

2 Methodology

All cases have been modelled using the Lattice Boltzmann software XFlow. Dassault Systèmes XFlow [3] is a software package of new generation CFD, differing from classical CFD (e.g. ANSYS CFX) in that it uses the Lattice Boltzmann approach (transient and particle-based), instead of a Reynolds-Averaged Navier-Stokes (RANS) modelling. This software has been validated for benchmark problems in different industries, including the aeronautical field [4]. It can solve two- and three-dimensional problems indistinctively.

The time consumption in a full simulation (from setup to post-processing) is greatly reduced with the use of XFlow, as the pre-processing (including meshing) is much simpler. If wake effects want to be studied in detail, XFlow also introduces a dynamic meshing, which will follow the wake for further time steps. The last and main reason why this software has been used for this investigation is that turbulence need not be modelled, but solved, without a major increase in solver time. This eliminates the error-inducing situation of having to choose the appropriate turbulence model, which is of capital importance when considering the use of NLF aerofoils.

Several setup parameters can be considered when preparing the simulations in XFlow and will be discussed hereafter. These can be classified in solver and boundary conditions parameters.

2.1 Solver parameters

Among the several parameters that can be considered in terms of solver in XFlow, there is the main reason for selecting this software: turbulence modelling. Although turbulence can be solved as a Direct Numerical Simulation (DNS), for typical examples where other turbulence models have been validated, XFlow offers the possibility of using the following:

- Spalart–Allmaras: a one-equation model used for low-Reynolds and that is implemented in most CFD tools
- Smagorinsky / Dynamic Smagorinsky: these methods model the kinematic viscosity using a filter scale and the Smagorinsky constant. While the regular Smagorinsky method keeps the constant invariable throughout the simulation, the dynamic version adapts the values depending on the solutions that it obtains in the different steps.

- Wall-Adapting Local-Eddy (WALE): this is the standard setup option in XFlow, as it has been proven to be accurate both near and far from walls to simulate boundary layers.

Another setup condition for the solver is the definition of the fluid state equations, where the user should decide between:

- Incompressible: the option to be chosen for low speeds and fluids that are considered constant in density. This option was used for the low-speed wing models, as for up to M0.3 incompressibility can be considered to be acceptably accurate.
- Ideal gas: modelling the density using the ideal gas equation. This method was used in the case of the high-speed wings, as the analyses are mainly aimed at inlet velocities above M0.5.
- Boussinesq model: using the Boussinesq approximation for density, which requires the thermal expansion coefficient to be defined and should only be applicable in combination with a thermal model, as otherwise the results would be the same as using the incompressible approach.

Finally, the thermal modelling (that is, if the energy equation is used or not) needs to be defined. Out of the different options offered by XFlow, the isothermal approach was used for the low-speed wings, where the energy equation is not resolved. In the case of the high-speed wings, the coupled energy method was considered, where compressibility is considered for the energy equation. Two additional options are available, the segregated energy (similar to the coupled energy but not considering compressibility effects) and a supersonic model, which is still under validation and therefore was not considered although it would be the most accurate for transonic flight conditions.

Additionally, several viscosity models can be considered. In the case of the low-speed wing, a regular Newtonian viscosity model (temperature independent) has been considered as the energy equation was not used, while for the high-speed wings the Sutherland model using standard values for the constants required for modelling of standard air (C=120K; T=291.15K and μ=18.27·10^{-6}Pa·s) was taken.

2.2 Boundary conditions parameters

XFlow uses a "Non-equilibrium wall function" to model the boundary layer on the analysed surfaces. This means that the adverse pressure gradients are taken into consideration and the transitional effects (laminar to turbulent, detachment and reattachment) can be observed in the simulation, and as previously indicated turbulence effects are solved. The modelling options available allow a near-wall refinement and a wake-following refinement, in which the calculation points (the equivalent to the mesh in a RANS CFD) will be modified at each step depending on the results of the previous simulation step results. Moreover, a general resolved scale is defined which indicates the distance between calculation points in the free stream conditions.

The walls need thermal exchange boundary conditions if the energy equation is considered. The options available for XFlow include adiabatic, temperature law, heat flux, free convection and volume heat source. As in both wings no thermal exchange was taken into account, the adiabatic model (no heat flux through the solids) was used.

3 Low-speed model definition

Different aerofoils have been considered for this design. It was found that a moderately higher curvatures combined with thicknesses of over 10% were drawing the best efficiencies at levelled flight. This also produces setbacks in terms of range of pitch and critical Mach [5], but the analysis conditions for the performance comparison are levelled flight (no pitch) and low speed (thus far from the critical conditions). The selected aerofoil is the NASA-developed NLF(1)-1015, which has a maximum thickness of 15% at 40% chord and a maximum camber of 4% at 63% chord, distributed as shown in fig. 1. This means that the 4-digit NACA equivalent of this aerofoil would be the NACA4615, which is shown in direct comparison with the NLF(1)-1015 in fig. 2.
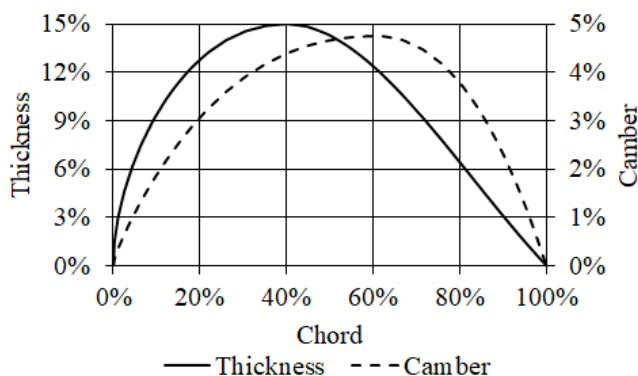
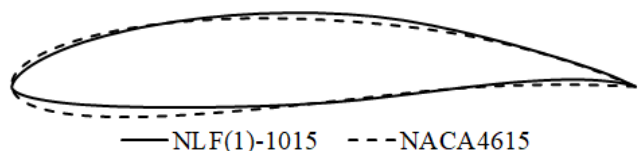Figure 1: Thickness/Camber distribution for NLF(1)-1015

Figure 2: Aerofoil Comparison: NLF(1)-1015 vs NACA4615

The wingspan has been defined as 13.7m, with a taper of 0.5 and a surface of 19m^2, which is representative of commuter aircraft (Part 23 certification). As usual in this type of aircraft, no sweepback angle nor twist or dihedral are introduced in the wing. Finally, no winglet has been considered. The final design of the wing can be seen in fig. 3.

Figure 3: Low-Speed Wing

For these analyses, the Isothermal (non-compressible flow) model has been used, with the air conditions set at ISA, 4900m above sea level. Turbulence has been resolved and not modelled, and a scale (grid size) of 20cm was used overall except near the wing, where it was refined to 10cm. The simulation time was set to 2s, with a time step of 0.02s. The flight speeds have been selected ranging between M0.1 (30m/s, 60KTAS) to M0.3 (95m/s, 185KTAS), in M0.05 steps.

4 High-speed model definition

In the case of the high-speed NLF aerofoils, not many have been designed yet and the geometrical details would probably not be undisclosed. Following the regular considerations for high-speed aerofoil design, though, it can be agreed that lower curvatures and thicknesses are the most efficient, and therefore the most used [6]. The main concerns are that at low speeds (e.g. take-off and landing) the use of non-laminar high-lift devices like flaps and/or slats is required, but for the analysed levelled, cruise conditions this is negligible. A new design has been developed considering the previous conditions, and was named RLPHSF. The maximum thickness used is 9% at 30% of the chord, with a maximum camber of 1.5% at 72% of the chord; the distribution of both parameters along the chord can be seen in fig. 4. Hence, the 4-digit NACA equivalent of this aerofoil is the NACA1709, and a comparison between both aerofoils is shown at fig. 5 below.
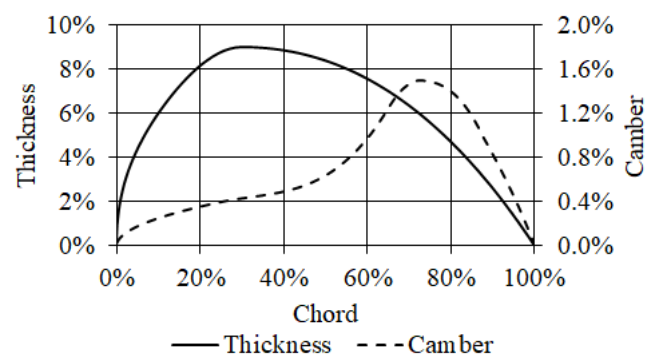

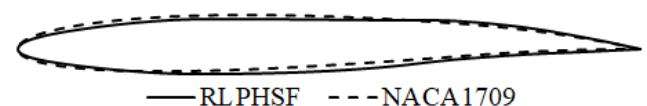Figure 4: Thickness/Camber distribution for RLPHSF

Figure 5: Aerofoil Comparison: RLPHSF vs. NACA1709

There are two analysis sweeps performed for this wing, a fixed-design flight-condition sweep and a fixed-flight-condition sweepback-angle modification keeping all other geometrical parameters fixed. For all cases, the wingspan has been defined as 26.3m, with a taper ratio of 0.6 and a wing surface of 69m^2, which is representative for business jets (Part 25 certification).

In the fixed-design case, a sweepback angle of 25 degrees without twist or dihedral has been introduced. The winglet was defined as 10% of the wingspan, with a dihedral of 10 degrees and an internal taper of 0.5. The final design of the wing can be seen in fig. 6.

Figure 6: High-Speed Wing

Finally, using the same basic geometrical parameters but removing the winglet, analyses with sweepback angles between zero and 10 degrees in 1.25 degrees steps have been considered.

For these analyses, the coupled energy (isentropic) model has been considered, with the air conditions set at ISA, 11000m above sea level. Turbulence has been resolved, and a scale (grid size) of 20cm was used overall except near the wing, where it was refined to 10cm, as it was in the previous cases. The simulation time was set to 2s, with a time step variating depending on the analysis Mach. This is changed due to the requirement for the time step to be able to accommodate the movement of the air between two analysis points inside it. This means that, if the time step is too high, a tracked particle would have gone through more than a single analysis point and therefore the results would not be accurate. A speed variation between M0.3 (90m/s, 175KTAS) and M0.85 (250m/s, 485KTAS) was considered for the fixed-geometry wing, while M0.77 (230m/s, 445KTAS) was taken for all sweepback variation analyses.

5 Low-speed results

The results for the low-speed wings show that, for any of the analysed Mach conditions, the NLF wing shows improved efficiency. Even though its value drops for the NLF wing as the speed is increased (as reflected in fig. 7), it can also be observed that the difference in efficiency between using the NACA and the NLF aerofoil is actually increased from roughly 5% at M0.1 to over 15% at M0.25 as shown in fig. 8. Moreover, it can be concluded that, on traditional designs of commuter aircraft, the use of NLF aerofoils can reduce the drag produced by the wing, and therefore the required thrust.
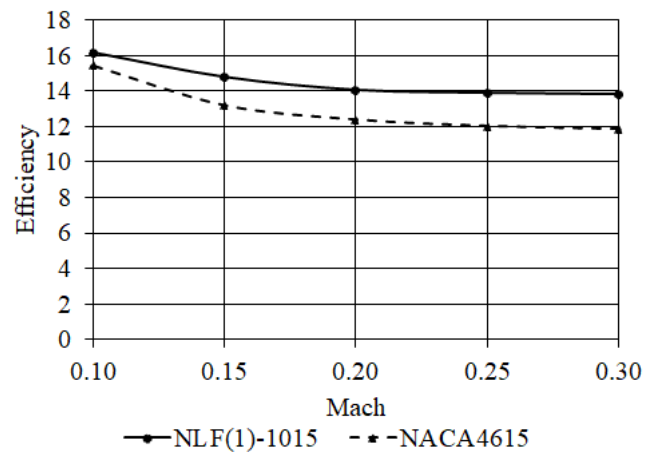

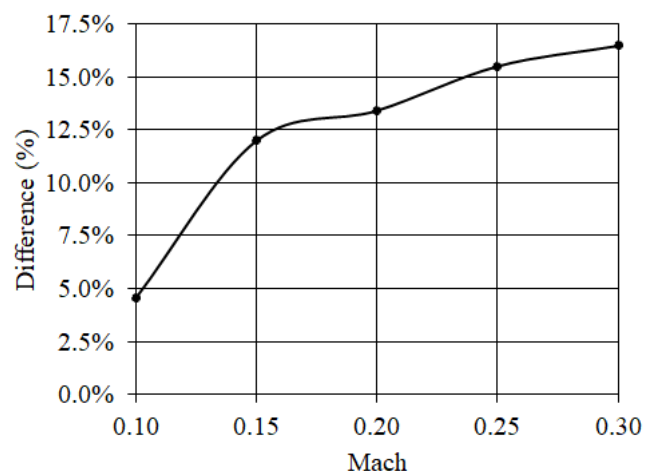*Figure 7: Efficiency for the low-speed wings*

Figure 8: Difference in efficiency between the low-speed wings

6 High-speed results – fixed wing design

In the case of the high-speed application, the NLF wing draws lower efficiencies when compared to its NACA equivalent. A deeper look into the results, not focusing only on lift and drag, shows that the ratio of the lateral forces to lift had a direct relation to the efficiency: when this ratio is increased, the efficiency is decreased, and vice-versa. This means that, in order to reach optimal efficiencies in high-speed wings, the lateral force needs to be restrained. Although this seems simple to comply with, the fact is that this lateral force is mainly produced by having a sweepback angle on the wing, which is usually required for flight at higher speeds [7]. It is noticeable and interesting to note that when the Mach increases from M0.6 onwards (up to near-transonic regions, as further analyses are not reliable) the efficiency is increased in the case of the NLF aerofoil.
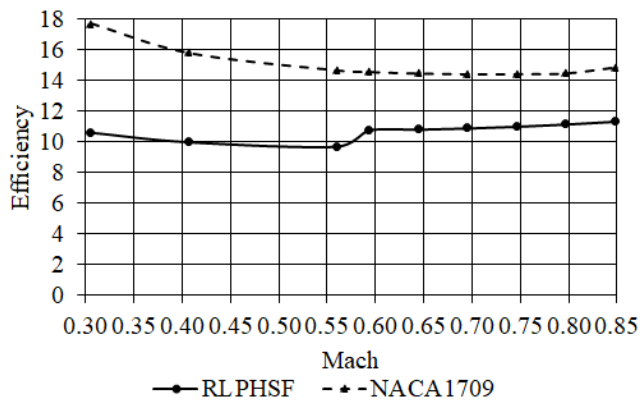
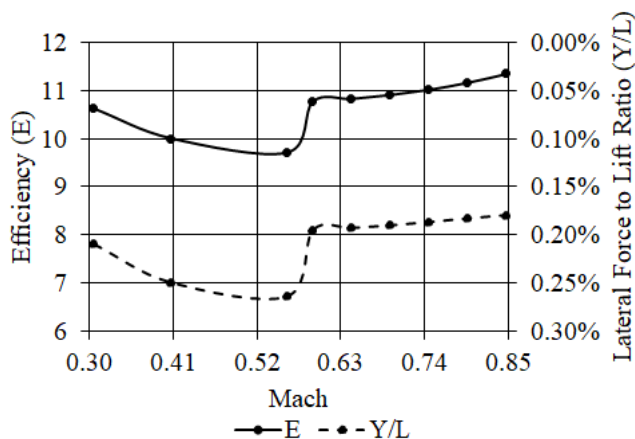Figure 9: Efficiency for the high-speed wings, Mach sweep

Figure 10: Relationship between lateral forces and efficiency for the high-speed NLF wing, Mach sweep

7 High-speed results – Sweepback angle effect

The final step on this investigation is to assess the effects of the sweepback on the efficiency of high-speed NLF wings. The results show that the efficiency at high speed is only higher for the NLF wing when no sweepback is present, dropping to less than half its initial value just by adding 1.25deg sweepback. As it can be assessed from fig. 11, there will be no significant differences between having low and high sweepback angles in terms of efficiency.
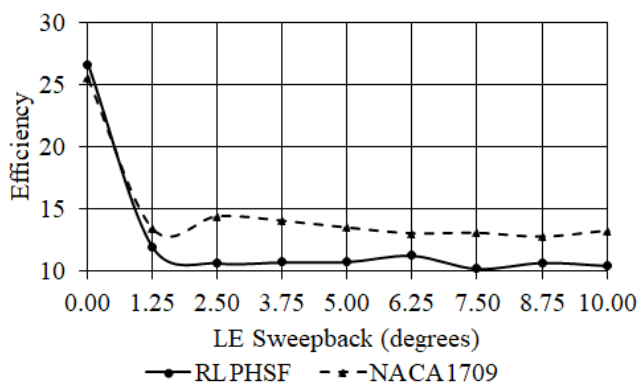

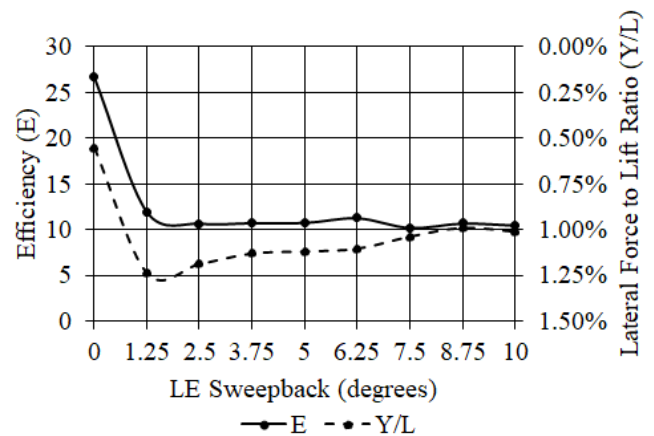Figure 11: Efficiency for the high-speed wings, LE sweepback variation

Figure 12: Relationship between lateral forces and efficiency for the high-speed NLF wing, LE sweepback variation

8 Conclusions

For Low-speed applications, it could be confirmed that the use of NLF aerofoils in a standard wing configuration is of advantage (up to 15% higher efficiency depending on the flight speed).

However, for a standard high-speed wing configuration (with sweepback), NLF aerofoils not only do not provide any improvement but show a 20% lower efficiency. This is seen to be due to the destabilizing effect of the lateral forces on the laminar boundary layer: for a concrete high-speed configuration, there is a direct relation between the lateral forces and the efficiency of the NLF wing when varying speed.

Finally, it was observed that, in the case of laminar wings, not applying a sweepback to the leading edge is optimal and duplicates the efficiency with respect to adding any angle at a high speed. Moreover, this is the only case observed where the efficiency of the laminar wing is higher than its NACA equivalent.

References

[1] Horizon 2020 European Union Funding for Research & Innovation, "Clean Sky," [Online]. Available: https://www.cleansky.eu/. [Accessed 08 07 2019].

[2] M. D. Maughmer and D. M. Somers, "Design and Experimental Results for a High-Altitude, Long-endurance Airfoil," *Journal of Aircraft,* vol. 26, no. 2, pp. 148-153, February 1989.

[3] Dassault Systèmes, "XFlow Brochure," 2019. [Online]. Available: https://www.3ds.com/. [Accessed 26 06 2019].

[4] D. M. Holman, R. M. Brionnaud and Z. Abiza, "Solution to industry Benchmark Problems with the Lattice-Boltzmann Code XFlow," in *European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2012)*, Vienna, 2012.

[5] J. D. Anderson, in *Fundamentals of Aerodynamics*, 5 ed., McGraw-Hill, 2011, pp. 728-737.

[6] C. D. Harris, "NASA supercritical airfoils: A matrix of family-related airfoils," NASA, 1990.

[7] A. Busemann, "Aerodynamischer Auftrieb bei Überschallgeschwindigkeit," in *5th Volta Conference*, Rome, 1935.

Equations of State in Fighter Aircraft Oleo-pneumatic Shock Absorber Modelling

Arttu Heininen , Jussi Aaltonen , Kari T. Koskinen , and Juha Huitula[*]

Automation Technology and Mechanical Engineering, Tampere University, Tampere, Finland
E-mail: arttu.heininen@tuni.fi, jussi.aaltonen@tuni.fi, kari.koskinen@tuni.fi, juha.huitula@mil.fi
[*]Finnish Defense Forces Logistics Command, Tampere, Finland

Abstract

Most of all modern commercial and military aircraft have oleo-pneumatic shock absorbers in their landing gear. An oleo-pneumatic shock absorber consists of a gas charge and an oil fill. During the stroke oil is forced through orifices which provides damping, while the gas charge is compressed and acts as a spring by increasing the stiffness of the shock absorber. Typically, when the gas behaviour is modelled, the ideal gas law is used as the equation of state as this provides in most cases adequate fidelity with relatively light computational load. However, in a fighter aircraft, especially in naval service, the gas pressure inside a shock absorber raises too high during landing for the ideal gas assumption to be valid. Therefore, other well-established equations of state have been considered. These are Van der Waals, Redlich-Kwong-Soave, and Peng-Robinson equation of state. This paper presents a multi-physics simulation model of a two-chamber oleo-pneumatic shock absorber based on fundamental analytical equations. Using this model, the behaviour of the aforementioned equations of state are studied in two cases: quasi-static and dynamical compression. The simulation results are compared to laboratory measurements. This comparison verifies that the ideal gas law should not be used when modelling naval fighter aircraft shock absorbers.
Keywords: Fighter aircraft, Shock absorber, Modelling, Simulation

1 Introduction

Essentially every modern aircraft, military or civilian, has some kind of landing gear. While there are many different landing gear configurations, most of them have a shock absorber to dissipate the kinetic energy related to landing.

Typically, shock absorber are either using a solid spring made of steel or rubber or a fluid spring. Solid spring shock absorbers are used in light aircraft and fluid springs in heavier aircraft. The most common shock absorber type is an oleo-pneumatic shock absorber. These have a high damping to weight ratio and are used in most modern fighter aircraft. An oleo-pneumatic shock absorber has both a gas charge, typically nitrogen or dry air is used, and a hydraulic oil fill. [1]

The gas charge acts as a spring when compressed. Becoming stiffer as the compression continues. Depending on the initial charge pressure, and the amount of gas and the stroke, the pressure increase can be very high. Especially in the case of naval fighter aircraft, because of their high sink speed during landing.

Usually, when modelling these shock absorbers using fundamental analytical equations, the gas and liquid volumes are modelled separately as control volumes. These can exchange mass and heat between control volumes that has the same substance, i.e., nitrogen or hydraulic oil. Between different substance control volumes only heat is transferred.

There are two ways to model the thermal behaviour of the gas control volumes. The first method is to model the gas compression as a polytropic process. This method depends on the chosen polytropic constant. For a slow rate of compression the polytropic constant approaches 1.0, which would indicate a isothermal process; whereas higher rates approach adiabatic process and a value of 1.4 for the constant could be used. Typically, if the gas and oil are separated, value of 1.35 is found to give accurate results, and 1.1 if the gas and oil is mixed. [1], [2]

Second method to model the gas is to use general internal energy model that is based on the first law of thermodynamics. This is a more accurate method, as in the first method the thermal exchange is only represented by one constant, and as such is a simple model.

The thermal behaviour of a gas volume based on the internal energy model uses a gas *equation of state* (EOS) to define the density and its derivatives. Therefore, the chosen *EOS* affects the predicted pressure and stiffness of the shock absorber. One of the simplest *EOS* is the ideal gas assumption. However, models based on this assumption behave poorly in high pressures [3]. For this reason, it is assumed that using the ideal gas assumption, when modelling a naval fighter aircraft shock absorber, produces poor results.

A schematic view of a naval fighter aircraft shock absorber is shown in fig. 1. It is a two-stage oleo-pneumatic shock

absorber, meaning that there are two chambers: primary and secondary chamber. The primary chamber is filled with gas and oil, just like a conventional oleo-pneumatic shock absorber, while the secondary chamber is filled with high pressure gas only.

As the shock absorber is compressed, the oil is forced from the bottom part of the primary chamber to the upper part through an orifice that is controlled with a metering pin, which has a variable cross-section. It is typically modelled using orifice flow model with a stroke dependent cross-section ([4], [5]). This approach is also used in this paper.

If the stroke is long enough, the primary chamber hits end-stops and the secondary chamber is engaged. The secondary chamber has a high stiffness due to the high pressure. Therefore, a rapid increase in the force is required for the shock absorber to further compress.
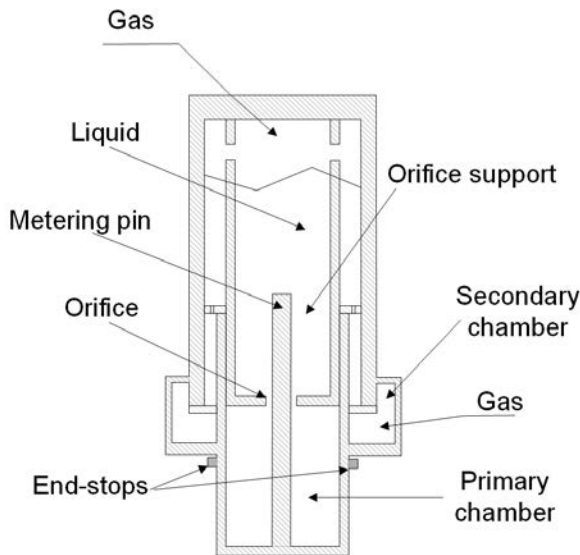

Figure 1: A schematic view of a two-stage oleo-pneumatic shock absorber.

The model presented in this paper is based on the aforementioned control volume approach. The gas volumes has been modelled using the general internal energy model using different *EOS*. These are the ideal gas assumption, Van der Waals, Redlich-Kwong-Soave, and Peng-Robinson *EOS*. Using the model two different cases are considered: quasi-static compression and dynamic compression. Both cases are simulated using the shock absorber model and the results are compared to those acquired by laboratory measurements. The aim of the measurements were to discover any abnormal behaviour and to provide more detailed information how the shock absorber behaves under different conditions. The aim of this paper is to provide a shock absorber model based on the internal energy model. The aim of this paper is to present a naval fighter aircraft shock absorber model based on fundamental analytical equations using the internal energy model. In the model, different *EOS* are considered, and it is shown that the pressure inside the shock absorber rises so high that the ideal gas law becomes inaccurate. Therefore, more accurate *EOS* should be used.

2   The shock absorber model

Balancing the forces that affect the moving parts of the shock absorber yields the following equation

$$
\begin{aligned}
F_{sa} = m_{sa}g - p_u A_{mp} - p_L(A_L - A_{mp}) \\
- A_{ph}(p_u - p_{ph}) - A_{HP}p_{HP} - F_{\mu 1} - F_{\mu 2} - F_{\mu 3}.
\end{aligned}
\tag{1}
$$

To solve this equation, the pressure must be solved inside the primary chamber, orifice support and secondary chamber. In addition, there are several sources for friction: friction of the primary chamber, viscous friction of the main cylinder and the friction of the secondary chamber. These must also be modelled. Figure 2 shows the pressure and friction sources inside the shock absorber. Due to the limited length, this paper only covers the most important parts of the model. A more detailed explanation of the model can be found in [6].
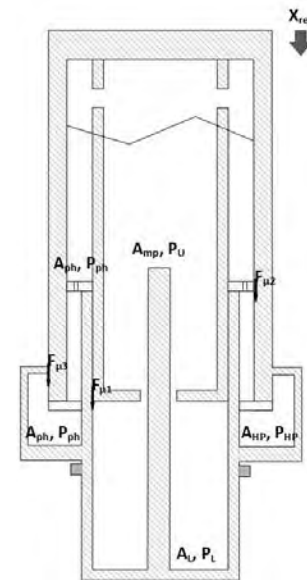

Figure 2: Pressure and friction sources inside the shock absorber.

2.1   Liquid volumes

The model is based on connected thermal-hydraulic and gas volumes together that exchange heat with each other. No mass is transferred between the two phases.

Considering a volume of liquid, its density is a function of pressure and temperature. Differentiating this we get

$$
d\rho = \left(\frac{\partial \rho}{\partial p}\right)_T dp + \left(\frac{\partial \rho}{\partial T}\right)_P dT
\tag{2}
$$

Solving for dp and taking its time derivative yields

$$
\frac{dp}{dt} = \frac{1}{\left(\frac{\partial \rho}{\partial p}\right)_T} \left[d\rho - \left(\frac{\partial \rho}{\partial T}\right)_P dT\right].
\tag{3}
$$

The isothermal bulk modulus of a liquid is

$$\beta_T(p,T) = \frac{\rho}{\left(\frac{\partial \rho}{\partial p}\right)_T} \tag{4}$$

and

$$\alpha_P(p,T) = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_P \tag{5}$$

is the volumetric expansion coefficient.

The continuity equation is

$$\frac{dm}{dt} = V\frac{d\rho}{dt} + \rho\frac{dV}{dt} \tag{6}$$

and solving for $d\rho/dt$ we get

$$\frac{d\rho}{dt} = \frac{\frac{dm}{dt} - \rho\frac{dV}{dt}}{V} \tag{7}$$

Combining equations 4, 5, and 7 with 3, the pressure change in a volume liquid is

$$\frac{dp}{dt} = \beta_T\left[\frac{1}{\rho V}\left(\frac{dm}{dt} - \rho\frac{dV}{dt}\right) + \alpha_p\frac{dT}{dt}\right]. \tag{8}$$

The change in the total energy of the volume, neglecting the kinetic and potential energies, is

$$dE = d(mu) = mdu + udm \tag{9}$$

where the change in the specific internal energy is

$$du = dh - \frac{dp}{\rho} \tag{10}$$

and in the specific enthalpy

$$dh = \left(\frac{\partial h}{\partial T}\right)_P dT + \left(\frac{\partial h}{\partial P}\right)_T dP = \\ c_P dT + (1 - \alpha T)\frac{1}{\rho}dP \tag{11}$$

so the change in the internal specific energy becomes

$$du = c_P dT - \frac{\alpha T}{\rho}dP \tag{12}$$

and the change in the total energy

$$dE = (h - \frac{P}{\rho})dm + mc_P dT - \frac{m\alpha T}{\rho}dP. \tag{13}$$

Now the change in the total energy is equal to the heat exchange with the volume's surroundings

$$dE = dQ. \tag{14}$$

Combining equations 9 and 14 and taking time derivative yields

$$\frac{dT}{dt} = \frac{\dot{Q} + (\frac{P}{\rho} - h)\dot{m}}{mc_P} + \frac{\alpha T}{c_P \rho}\frac{dP}{dt} \tag{15}$$

2.2 Gas volumes

Considering a gas volume, its rate of change of mass can be written

$$\frac{dm}{dt} = \rho\frac{dV}{dt} + V\left[\left(\frac{\partial \rho}{\partial P}\right)_T\frac{dP}{dt} + \left(\frac{\partial \rho}{\partial T}\right)_P\frac{dT}{dt}\right]. \tag{16}$$

Considering the same gas volume, its internal energy is

$$U = mu \tag{17}$$

and its total derivative is

$$dU = d(mu) = dmu + mdu. \tag{18}$$

Using the internal energy, the specific enthalpy of the gas volume is

$$h = u + Pv \tag{19}$$

and substituting this in equation 18 yields

$$\begin{aligned}
dU &= dm(h - Pv) + md(h - Pv) \\
&= dm(v - Pv) + mdh + m(-dPv - Pdv) \\
&= dmh - dmPv + mdh - mPdv - mvdP \\
&= dmh - dmPv + mdh - mPdv - Vdp \\
&= dmh + mdh - VdP - Pd(mv) \\
&= dmh + mdh - VdP - PdV.
\end{aligned} \tag{20}$$

On the other hand, using the first law of thermodynamics, the change in internal energy can be written as

$$dU = hdm + \delta Q + dW \tag{21}$$

where the work done by the system is

$$dW = -PdV \tag{22}$$

Combining equations 17 and 23 we get

$$\sum_i m_i h_i - \sum_i m_i h + \delta Q = mdh - VdP \tag{23}$$

Taking a time derivative of the above equation and considering the *state postulate* we have

$$\left(m\left(\frac{\partial h}{\partial P}\right)_T - V\right)\frac{dP}{dt} + m\left(\frac{\partial h}{\partial T}\right)_P\frac{dT}{dt}$$
$$= \sum_i \frac{dm_i}{dt}h_i - \sum_i \frac{dm_i}{dt}h + \delta Q \qquad (24)$$

Now we can form a system of equation using equations

$$\begin{bmatrix} V\left(\frac{\partial \rho}{\partial P}\right)_T & V\left(\frac{\partial \rho}{\partial T}\right)_P \\ m\left(\frac{\partial h}{\partial P}\right)_T - V & m\left(\frac{\partial h}{\partial T}\right)_P \end{bmatrix} \begin{bmatrix} \frac{\partial P}{dt} \\ \frac{dT}{dt} \end{bmatrix} =$$
$$\begin{bmatrix} \sum_i \frac{dm_i}{dt} - \rho\frac{dV}{dt} \\ \sum_i \frac{dm_i}{dt}h_i - h\sum_i \frac{dm_i}{dt} + \delta Q \end{bmatrix} \qquad (25)$$

Finally, the derivatives of ρ are solved from the chosen equation of state.

2.3 Equations of state

The most simple way to model the gas behaviour is to use the ideal gas law:

$$PV = mrT. \qquad (26)$$

Most of the real gases in mild temperatures or pressures behave like an ideal gas. However, it does not consider the volume that the gas molecules occupy or the intermolecular attraction forces. Therefore, its ability to accurately predict the gas state decreases as the gas pressure rises.

To take the interaction of the gas molecules into consideration van der Waals proposed an equation of state [7]

$$(P + \frac{a}{V^2})(V - b) - rT = 0. \qquad (27)$$

where

$$a = \frac{27r^2T_c^2}{64P_c} \qquad (28)$$

and

$$b = \frac{rT_c}{8P_c} \qquad (29)$$

Even though van der Waals *EOS* is an improvement over the ideal gas law, several more accurate *EOS* have been proposed after its publication. However, most of these, like Redlich-Kwong-Soave and Peng-Robinson *EOS* are based on the van der Waals *EOS*. So similar behaviour can be assumed.

Soave [8] proposed an *EOS* based on the works of Redlich and Kwong [9]. The main difference of these models is that the constant a has been replaced with a more general temperature dependent product of $a\alpha(T)$. It is similar to van der Waals equation but with some modifications:

$$(P + \frac{a\alpha(T)}{V(V + b)})(V - b) - rT = 0 \qquad (30)$$

where

$$a = 0.42748024\frac{r^2T_c^2}{P_c} \qquad (31)$$

and

$$b = 0.08664035\frac{rT_c}{P_c} \qquad (32)$$

and the temperature dependent adimensional factor α

$$\alpha(T) = \left[1 + m\left(1 - \sqrt{\frac{T}{T_c}}\right)\right]^2 \qquad (33)$$

where m is a substance dependent constant

$$m = 0.48 + 1.574\omega - 0.176\omega^2 \qquad (34)$$

defined by the acentric factor ω

$$\omega = -1 - 1log_{10}\left(\frac{P_{sat}}{P_c}\right)_{T=0.7T_c}. \qquad (35)$$

Peng and Robinson noticed that the Redlich-Kwong-Soave *EOS* predicted greater specific volumes than found in the literature [10]. To correct this behaviour, they proposed an *EOS* similar to Redlich-Kwong-Soave:

$$\left(P + \frac{a\alpha(T)}{V^2 + 2bV - b^2}\right)(V - b) - rT = 0. \qquad (36)$$

Here the adimensional factor $\alpha(T)$ is the same as in Redlich-Kwong-Soave EOS, but m has a different constants in it

$$m = 0.37464 + 1.54226\omega - 0.26992\omega^2. \qquad (37)$$

According to literature, Peng-Robinson *EOS* should be the most accurate of the two-constant *EOS* presented here. [10]

2.4 Friction model

Friction can be modelled using plethora of different friction models [11]. Here, a widely applied Karnopp friction model is used ([11], [12]) to model the friction between the orifice support and the secondary chamber, $F_{\mu3}$:

$$F_{\mu3} = \begin{cases} min(|F_E|, |F_S|)sign(F_E) & \text{and} \quad v = 0 \quad if|v| < dv \\ (F_C + (F_S - F_C)e^{-3|v|/V_S})sign(v) + F_VV & if|v| > dv \end{cases} \qquad (38)$$

2.5 End-stops

The viscous friction of the main cylinder, $F_{\mu2}$ is modelled as a moving cylinder within an envelope. It is based on the Karnopp friction model, equation 38, but also considers the viscous friction and the elastic end-stops.

The friction force between the cylinder and envelope is:

$$F_{\mu 2} = \frac{m_1 f_2 - m_2 f_1}{m_1 + m_2} \tag{39}$$

where

$$f_1 = F_{ext1} - F_{ext2} + F_{min} - F_{max} + F_{min3} - F_{max3}$$
$$+ m_1 g - R_{visc} V_{rel} \tag{40}$$

and

$$f_2 = F_{ext3} - F_{ext4} - F_{min} + F_{max} + F_{min2} - F_{max2}$$
$$+ m_2 g - R_{visc} V_{rel}. \tag{41}$$

The contact forces at the lower limit is F_{min}, F_{min2}, and F_{min3} are calculated from

$$F_{min} = Kb_{min}(X_{min} - X_{rel}) -$$
$$Db_{min}\left(1 - e^{\frac{-(X_{min} - Xrel)}{Pd_{min}}}\right) V_{rel} \tag{42}$$

if X_{rel} is smaller than the lower displacement limit and zero otherwise. The contact forces at the higher limit is:

$$F_{max} = Kb_{max}(X_{rel} - X_{max}) -$$
$$Db_{max}\left(1 - e^{\frac{-(X_{rel} - Xmax)}{Pd_{max}}}\right) V_{rel} \tag{43}$$

and zero if $X_{rel} \leq X_{max}$.

2.6 Orifice flow

A flow through an orifice can be described by

$$\dot{m} = \rho c_q A \sqrt{\frac{2|\Delta P|}{\rho}} \tag{44}$$

where A is the area of the orifice subtracted by the area of the metering pin.

3 Numerical solution

The aforementioned equations were modelled using a commercial multi-domain simulation software LMS Imagine.Lab Amesim 15, which uses bond graphs to represent systems. The software has many well-known algorithms to solve ordinary differential equations, based on linear multi-step methods.

4 Validation

The model was used to simulate two different cases: quasi-static and dynamic compression. In the quasi-static case, the shock absorber is compressed so slowly that the hydraulic damping is negligible, while the dynamic case considers damping also. In both cases simulation results are compared to the actual measured data.

4.1 Quasi-static case

The test setup of the quasi-static case has a hydraulic cylinder that is attached to the end of the shock absorber. The other end is rigidly supported preventing its movement. The motion of

the hydraulic cylinder is controlled using PID-controller, so that the compression is the same during each individual test.

The pressure is measured in the hydraulic cylinder during the test and the compression force is calculated from this pressure. The test system then produces a graph with the force as a function of the shock absorber stroke. The stroke, as an function of time, is given as an input to the model presented in this paper. The given input normalised with the maximum shock absorber design stroke is shown in 3.
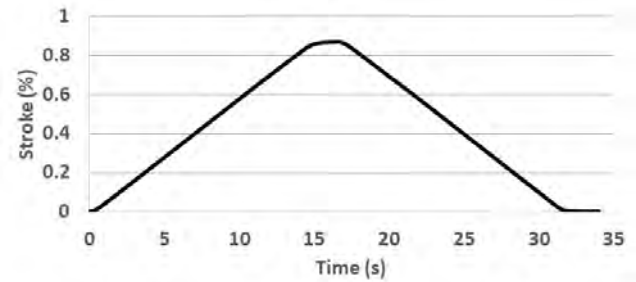
Figure 3: The stroke data given as in input to the model in the quasi-static case.

The force predicted by the model and the measured force is shown in fig. 4. The stroke-force plot is divided into two distinguishable parts. First, there is the initial part, where the primary chamber is compressed. This is followed by the secondary part, where the secondary chamber also activates. Between these two parts there is a vertical line. The secondary chamber has an initial stiffness and requires a certain force, before it activates.
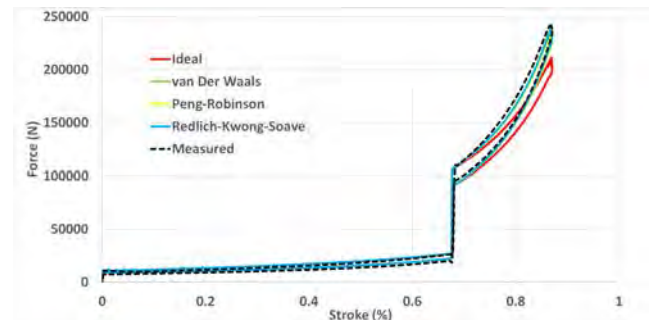
Figure 4: Force-stroke curve during quasi-static compression.

Figure 4 clearly shows that the force predicted using the ideal gas law is poor compared to the measurements. This is apparent, especially during the second part of the stroke due to the high pressure inside the secondary chamber. It shows that a shock absorber modelled using ideal gas law is less stiff due to the lower pressure. The other chosen *EOS* used in the simulation predict more accurate results. Their force-stroke curves are almost on top of each other. Even van Der Waals *EOS* could be used, although according to literature it's accuracy is limited. [3].

4.2   Dynamic case

The dynamic test setup is based on a nitrogen actuator. It has a tank of highly pressurised nitrogen, which operates a cylinder that is attached to the end of the shock absorber. The other end is rigidly supported. With this setup, the shock absorber can be compressed rapidly.

In the simulation, the measured force, as a function of time, was given as an input to the model and is shown in fig. 5. Then the stroke rate and the stroke of the shock absorber was calculated. These were then compared to the measured values.
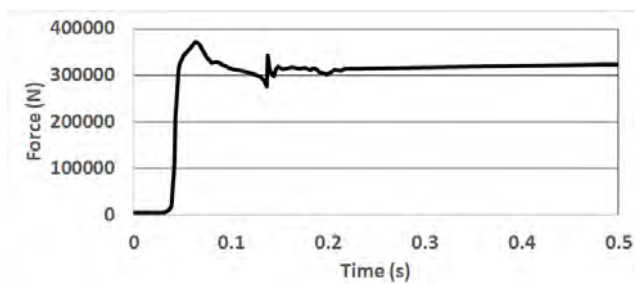

Figure 5: The given input force during the dynamical case.

Figure 6 shows the stroke during the dynamical compression. As expected, the maximum stroke that the ideal gas law produces is higher than the measured and maximum stroke produced by the other *EOS*. This is due to the lesser stiffness that was observed in the quasi-static compression. The other *EOS* behave similarly and produce a maximum stroke close to the measured.
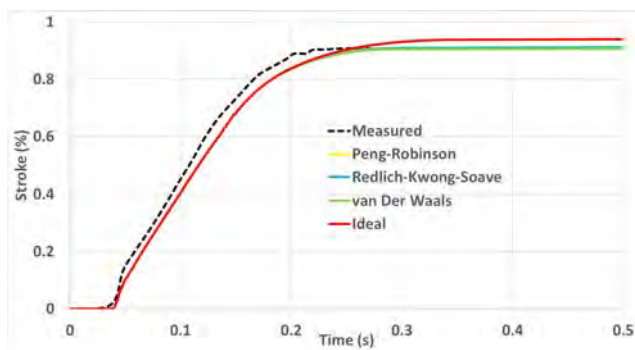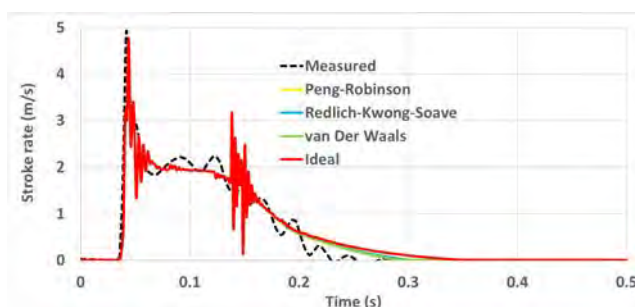

Figure 6: Stroke during the dynamical compression.

Figure 7: Stroke rate during the dynamical compression.

In addition, the stroke rate was modelled during dynamical compression. The results are shown in fig. 7. As the gas is responsible for the shock absorber stiffness, it has little effect on the shock absorber damping characteristics. Therefore, significant difference between the performance of different *EOS* cannot be observed. Ideal gas seems to reach zero stroke rate a bit slower than the other *EOS*. It is assumed that this is related to the larger maximum stroke observed.

In both cases the ideal gas law behaved poorly. Interestingly, the other *EOS* behaved similarly. However, the results were analyzed on a general level and more detailed analysis could show difference in the *EOS* behavior. In addition, according to literature, the Peng-Robinson is the most accurate *EOS* [6] and should be used.

5   Conclusion

In this paper, a model based on fundamental analytical mathematical equations of a two-stage naval fighter shock absorber was presented. In the model, gas volumes are modelled using general internal energy model. This requires derivatives of the gas density, which is solved from the chosen gas *equation of state* (EOS). Different *EOS* were used in the simulations. These were ideal gas, van der Waals, Redlich-Kwong-Soave, and Peng-Robinson. Two different cases were considered: quasi-static and dynamical compression. In the former case, the force of a given stroke was measured. In the latter the stroke rate and the stroke of a given force was measured. Using the model, both cases were simulated and the results from the simulation were compared to the measured data.

It was assumed that the pressure inside a naval fighter shock absorber rises so high that the accuracy of the ideal gas law is insufficient. This was seen in both cases. The ideal gas law predicted significantly lower reaction force during the quasi-static compression than the other three *EOS*. Especially, as the second chamber activates. Similarly, in the dynamic case, the shock absorber behaved less stiff, when the ideal gas law was used producing clearly higher maximum stroke.

Based on the aforementioned findings, it is suggested that the ideal gas law is not used when modelling fighter aircraft shock absorber. The other *EOS* were in good agreement with the measured values. However, the results were analyzed quite generally and a more detailed analysis is required. Also, there might be certain situations where the model differences can be clearly seen.

Designation	Denotation	Unit
α	Thermal diffusivity	m^2/s
α_P	Volumetric expansion coefficient	$K^{-}1$
β_T	Isothermal bulk modulus	Pa
ρ	Density	kg/m^3
ω	Acentric factor	
A	Flow area	m^2
A_{HP}	Secondary chamber area	m^2
A_L	Primary chamber area	m^2
A_{mp}	Metering pin cross-section	m^2
A_{ph}	Orifice support area	m^2
a	Constant	
b	Constant	
c_P	Specific heat	J/kgK
c_q	Flow coefficient	
Db	Damping coefficient	N/(ms)
dv	Relative velocity	m/s
E	Total energy	J
F_μ	Friction force	N
F_E	External force	N
F_{ext}	external force	N
F_{max}	higher limit contact force	N
F_{min}	lower limit contact force	N
F_S	Stiction force	N
F_{sa}	Force	N
F_V	Coulomb friction force	N
F_v	Coefficient of viscous friction	N/(m/s)
g	Gravitational acceleration	m/s^2
h	Specific enthalpy	J/kg
Kb	lower/higher limit stiffness	N/m
m	Mass	kg
m	Constant	
$\dot{m}$	Mass rate	kg/s
m_sa	Shock absorber mass	kg
Pd	penetration limit for full damping	m
P_c	Critical pressure	Pa
P_{sat}	Saturation pressure	Pa
p	Pressure	Pa
p_{HP}	Secondary chamber pressure	Pa
p_L	Primary chamber pressure	Pa
p_{ph}	Orifice support pressure	Pa
p_u	Orifice support pressure	Pa
Q	Exchanged heat	J
$\dot{Q}$	Heat exchange rate	W
R_{visc}	viscous friction coefficient	
r	Specific gas constant	J/kgK

T	Temperature	K
t	Time	s
T_c	Critical temperature	K
U	Internal energy	J
u	Specific internal energy	J/kg
V	Volume	m^3
V_{rel}	relative velocity	m/s
V_S	Stiction velocity	m/s
v	Specific volume	m^3/kg
W	Work	J
X_{max}	higher displacement limit	m
X_{min}	lower displacement limit	m
X_{ref}	Stroke length	m
X_{rel}	relative displacement	m

References

[1] Norman S Currey. *Landing gear design handbook*. Lockheed-Georgia Company, 1982.

[2] Benjamin Milwitzky and Francis E Cook. Analysis of landing-gear behavior. Technical report, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION WASHINGTON DC, 1953.

[3] Yunus A Cengel and Michael A Boles. *Thermodynamics: an engineering approach*, volume 1000. 2002.

[4] James N Daniels. A method for landing gear modeling and simulation with experimental validation. 1996.

[5] Lucas G Horta, Robert H Daugherty, and Veloria J Martinson. Modeling and validation of a navy a6-intruder actively controlled landing gear system. 1999.

[6] Arttu Aleksi Heininen. Modelling and simulation of an aircraft main landing gear shock absorber. Master's thesis, 2015.

[7] Johannes Diderik Van der Waals. *Over de Continuiteit van den Gas-en Vloeistoftoestand*, volume 1. Sijthoff, 1873.

[8] Giorgio Soave. Equilibrium constants from a modified redlich-kwong equation of state. *Chemical engineering science*, 27(6):1197–1203, 1972.

[9] Otto Redlich and Joseph NS Kwong. On the thermodynamics of solutions. v. an equation of state. fugacities of gaseous solutions. *Chemical reviews*, 44(1):233–244, 1949.

[10] Ding-Yu Peng and Donald B Robinson. A new two-constant equation of state. *Industrial & Engineering Chemistry Fundamentals*, 15(1):59–64, 1976.

[11] William S Levine. *The Control Handbook (three volume set)*. CRC press, 2018.

[12] Dean Karnopp. Computer simulation of stick-slip friction in mechanical dynamic systems. *Journal of dynamic systems, measurement, and control*, 107(1):100–103, 1985.

Evaluation of Increase Weight in a Wing Fixed Leading Edge Design to Support UAS Impact

Tomaz Drumond, Marcelo Greco and Carlos Cimini

Department of Structural Engineering, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais/Brazil
E-mail: tomdrud@ufmg.br, mgreco@dees.ufmg.br, cimini@ ufmg.br

Abstract

In recent years, the number of *Unmanned Aerial Systems* (UAS) operating in the sky of urban centres has quite increased due to the facilities to purchase this type of vehicle. Consequently, the risk of an airborne impact between these vehicles and manned aircrafts has been target of concern with incidents been reported recently. Although the current aircrafts had been designed to support bird strikes, a collision with an UAS could be more critical, as reported in a study published by the *Federal Aviation Administration* (FAA) in 2016. In the present paper, a commercial aircraft *Wing Fixed Leading Edge* (WFLE) was first designed to support a bird strike. Then, its skin and spar thicknesses were increased so the WFLE could withstand an impact with an UAS. The safety criterion adopted was that the lithium-ion battery from the UAS could not penetrate into the airframe when the impact occurs at the airliner cruising speed. Finally, the weight from both WFLE configurations were compared so the mass increase could be quantified. During simulation, the same mass was considered for the UAS and the bird, although they were modelled using different materials. Modelling procedures were verified with test data found in other papers and the results obtained presented acceptable correlation. The UAS considered was a quadcopter configuration intended for recreational and commercial aerial photography with a 1.8 kg mass. The results showed that spar thicknesses increase were more effective than increases in the skin thicknesses.

Keywords: UAS Impact, Bird Strike, Wing Fixed Leading Edge, Simulation

1 Introduction

The amount of Unmanned Aerial Systems (UAS) being operated all over the world on recent years has increased significantly. The *Association for Unmanned Vehicles Systems International* (AUCSI) estimates that in 2020 the number of UAS sold around the world could reach 4.7 million unities [1]. The *Federal Aviation Administration* (FAA) forecasts that in the United States, the sale of commercial UAS for private purposes will go from US$2.5 million in 2016 to US$7.0 million in 2020 [2]. Associated with these increase in the number of UAS being operated in airspaces, it raises the necessity to regulate how these vehicles are used. In Brazil for example, it was created the Brazilian Civil Aviation Special Regulation n° 94/2019 (RBAC-E n° 94/2017) by the *National Civil Aviation Agency* (ANAC); the ICA 100-400 – Remotely Piloted Aircraft Systems and Access to the Brazilian Airspace, by the *Department of Airspace Control* (DECEA), and the Drones Homologation Booklet, established by the *National Telecommunications Agency* (ANATEL).

Among the concerns raised by the growing number of UAS is the risk of mid-air impact occurring between this type of vehicle and an aircraft that is flying. In order to avoid such kind of occurrence, helicopters for rescuing and fighting fires have already been impeded to fly when UAS were seeing being operated near the take-off region. In addition, airports have been temporarily closed when UAS were identified operating in their proximities, causing discomfort for a large number of passengers and significant financial losses to airlines companies.

In addition to the troubles caused on airports, mid-air impacts between UAS and aircrafts has possibly occurred in the United Kingdom [3], Canada and, more recently, in Mexico.

Sharma [2] has performed an investigation into UAS incidents considering reports released by the FAA and the NASA. The results have shown an increase in the number of occurrences, as indicated in tab.1.

Year	Number of Events	Year	Number of Events
1993	5	2005	1
1994	4	2006	6
1995	7	2007	8
1996	3	2008	7
1997	5	2009	4
1998	10	2010	10
1999	3	2011	9
2000	10	2012	24
2001	5	2013	32
2002	1	2014	87
2003	3	2015	190
2004	2	2016	88

Although the incidents reported so far have not resulted in death of any of the crewmembers or passengers on board, the risk involved in such an event is extremely high. Aircrafts operating in the world today are not designed to withstand an UAS impact since there are no requirements for manufacturers to prove their aircraft's ability to land in secure after being hit by a stiffener component like an UAS.

Meng *et al.* [4] and Olivares *et al.* [5] has showed that an mid-air impact with UAS of equivalent mass causes more severe damage to an aircraft than a bird impact on empennages, wing, windshield and engine. This seems to be due to the fact that UAS components are more rigid than the organs of a bird. Studies such as the performed by the British Military Aviation Authority [6] have been carried out to better understand the phenomena and the criticality that UAS impact can lead to an aircraft.

In the future, new aircraft structures should be reinforced to withstand this kind of impact, resulting in an increase in aircraft and helicopters weights.

The objective of the present work is to quantify the mass increase in an aircraft wing fixed leading edge after being reinforced to withstand a 1.8 kg UAS impact, similar to the mass of a bird established on requirement FAR 25.571 (e).

Firstly, the *Wing Fixed Leading Edge* (WFLE) of a commercial aircraft is designed to withstand an impact with a 1.8 kg bird. The WFLE skin thickness is increased in a way that no rupture could be verified in the structure after being impact by the bird. Further, an impact simulation with an UAS quadcopter of similar mass is performed. Since the UAS impact induces larger damages in the aircraft than the bird, the WFLE skin and spar thickness will be increased one more time in order to increase the wing strength. Finally, the weight of the WFLE capable to withstand the UAS impact is compared with the configuration able to resist the bird strike.

2 Description of the FE model

2.1 Wing Fixed Leading Edge (WFLE)

The wing fixed leading edge of a commercial aircraft was modelled considering its FLE Ribs, Main Box Ribs, Skin and Spar as shown in fig. 1. Only shell elements were used and the WFLE components were connected using coincident nodes. Fixed constraints were set on the Skin, Main Box Ribs and Spar edges since only a segment of the aircraft was represented in order to reduce computational costs.

The materials considered on WFLE structures were the Aluminium 2024-T3 for the Skin and the Aluminium 7050-T7451 for both types of Ribs and Spar.

Gokhale *et al.* [7] and Olivares *et al.* [8] established a quality mesh criteria that were adopted for modelling the aircraft structures on the present work. Table 2 summarizes these quality mesh criteria.
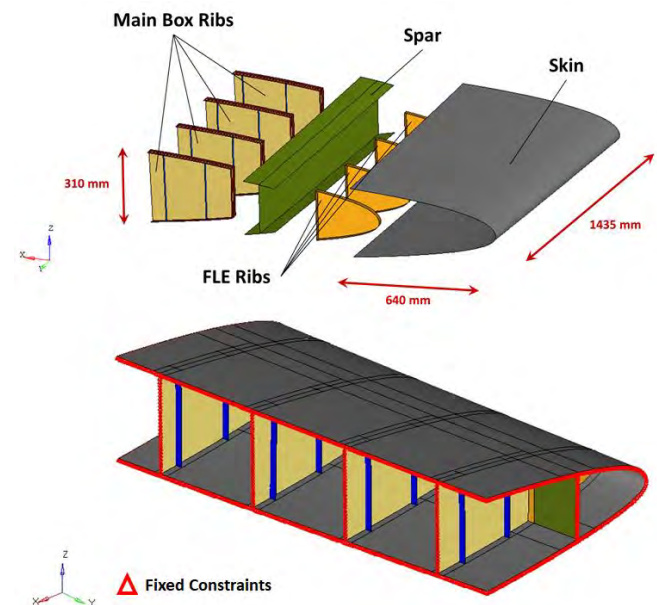

Figure 1: Wing Fixed Leading Edge FEM

Table 2: Criteria for Mesh Quality

	Minimum	Maximum
Warp Angle	-	10°
Aspect Ratio	-	5
Skew	-	45°
Jacobian	0,6	-
Distortion	0,6	-
Stretch	0,2	-
QUAD angles	45°	135°
TRIA angles	20°	120°
Taper	-	0,5
Element Size	5 mm	8 mm

2.2 The UAS Quadcopter

The UAS quadcopter model selected to be used on impact analysis on the WFLE was the DJI Phantom 3 Standard. Cairns *et. al* [9] reported that the DJI Phantom UAS family had been the most common model found in the market and for this reason, the DJI Phantom was selected. The referred UAS weights approximately 1.2 kg and its dimensions are about 290 mm x 290 mm x 185 mm.

To represent the UAS in a FE model only its most stiffener structures (battery, engines and camera) were detailed in the model. Solid elements were used to represent such components, which were connected by means of bar elements. The UAS model is shown on fig. 2 and its components materials were considered as follows:

- Battery cells constructed using lithium-ion polymer (LiPo) technology;
- Motors with an Aluminium Alloy A520.0-F rotor and a laminated steel AISI 4130 core stator;
- Camera made of Aluminium Alloy A520.0-F;
- UAS body and gimbal made of polycarbonate material, as reported in Olivares *et al.* [8]. The polycarbonate was attributed to bar elements that connect the other UAS components.

Although the UAS components have their own internal structures, they were considered as homogenous entities to simplify the simulation model. The bar elements had their material density adjusted in order to make the weight of UAS FE model be the same from a real UAS (~1.2 kg)
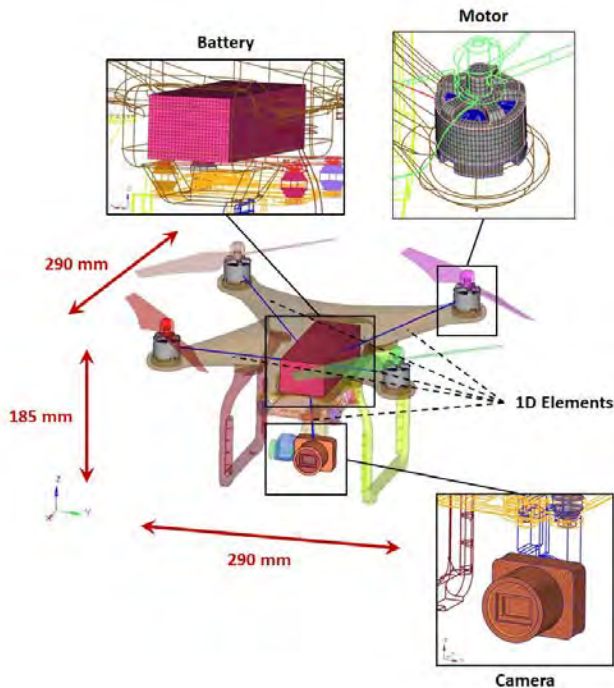

Figure 2: FE mesh for different UAS components

Further, with the UAS model completed, all its material densities were increased by a 1.5 factor (1.8kg/1.2kg = 1.5) in order to make the UAS model mass equivalent to the bird requirement from FAR 25.571(e).

2.3 Constitutive Model for Materials

The Aluminium alloy 2024-T3 on WFLE Skin was modelled as an elastic-plastic material with isotropic damage model, flow stress and strain rate dependent behaviour described by Johnson-Cook's law [10] defined as:

$$\sigma = [A + B(\bar{\varepsilon}^p)^n][1 + C ln(\dot{\varepsilon}^*)][1 - (T^*)^m] \qquad (1)$$

Where σ is the effective stress, ε^p the equivalent plastic strain, $\dot{\varepsilon}^*$ the dimensionless plastic strain rate and T^* the temperature rising during plastic deformation. Temperature influence was not consider during the analysis performed here, so only first and second terms of eq. (1) were taken in account. *A*, *B*, *C*, *m* and *n* are parameters obtained through static tests, which values for Al 2024-T3 were obtained on Kay [11] and are summarized in tab. 3.

Table 3: Johnson-Cook Parameters for Al 2024-T3 [11]

	A [MPa]	B [MPa]	C	m	n
Al 2024-T3	369.0	684.0	0.0083	1.7	0.73
	D_1	D_2	D_3	D_4	D_5
	0.112	0.123	1.500	0.007	0.0

During the collision process, most of the skin material failure was due to tensile stretching caused by the displacement of the impact region. The strain at fracture is defined by eq. (2) and the mesh elimination criterion is activated when damage parameter D defined on eq. (3) exceeds 1.0.

$$\varepsilon_{failure} = [D_1 + D_2 \exp(D_3 \sigma^*) |1 + D_4 \ln(\dot{\varepsilon}^*)|1 + D_5 T^*] \qquad (2)$$

Where σ^* is the ratio of the pressure to the effective stress and D_i are parameters obtained through static tests.

$$D = \sum \frac{\Delta \bar{\varepsilon}^p}{\varepsilon_{failure}} \qquad (3)$$

Where $\Delta \varepsilon^p$ is the accumulated incremental effective plastic strain.

The Aluminium 7050-T7451 applied on the WFLE Ribs and Spar, the Aluminium Casting Alloy A520.0-F and the Steel Alloy AISI4130 on UAS camera and motors were represented with a bi-linear elastic plastic material with criterion for element elimination being the maximum tensile strain value. Bi-linear elastic-plastic materials parameters are shown on tab. 4.

Table 4: Bi-linear Elastic Plastic Parameters

	ρ [kg/m³]	E [GPa]	ν	σ_y [MPa]	E_t [MPa]	ε_{max}
Al 7050-T7451 [12]	2770.	71.	0.33	462.	663.	0.10
A520.0-F [13]	2600.	66.	0.33	170.	1164.	0.14
AISI 4130 [12]	7850.	200.	0.32	483.	1174.	0.12

Based on the study published by Sahraei *et al.* [14], the UAS pounch lithium-ion batteries were considered as a crushable foam with a 60% porosity of the active particles and voids filled with electrolyte. The basic parameters obtained on Sahraei *et al.* [14] are listed in tab. 5.

Table 5: Mechanical Properties of the Battery [14]

Compressive Stress [MPa]	ρ [kg/m³]	E [GPa]	ν	ε_{max}
$\sigma = 276 \cdot \varepsilon^{1.8}$	1750	500	0.01	0.16

The polycarbonate bars that connect the UAS components were modelled with the elastic-plastic model and it was assumed to be ideal plasticity. The maximum tensile plastic strain was set to be the criterion for element elimination and the parameters listed on tab. 6 were set as found on Olivares *et al.* [8].

Table 6: Mechanical Properties of Polycarbonate [8]

ρ [kg/m³]	E [GPa]	ν	σy [MPa]	ε_{max}
1180.	2.35	0.3	62.	0.2

2.4 Computational Model

The impact analyses of the present study were made through explicit models of Finite Elements using Solution 700 from NASTRAN software. The generation of the mesh elements, the attribution of properties and the inclusion of the applicable parameters to the model were performed using the HYPERMESH software.

3 FEM Validation

3.1 Bird Strike

In this section, a series of simulations were performed in order to validate the bird model to be used on the design of the WFLE. The formulation adopted to represent the 1.8 kg bird was the meshless *Smooth Particle Hydrodynamics* (SPH), which can handle larger deformations compared to other formulations. In SPH technique, the fluid is represented as a set of particles which interaction between them is ruled by interpolation theory and smoothing kernel functions. Hedayati and Ziaei-Rad [15] reported that the hemispherical-ended cylinder geometry presented the closest results to experimental data and this is the reason why it was selected as the bird geometry used on the present work.

Bird dimensions and density were calculated using eqs. (4), (5) and (6) proposed by O'Connor [16] after analysing common characteristics of many bird species and different parameters of a bird impact. Bird dimensions are shown in fig. 3.

Bird Density

$$Density = 959 - 63 \cdot \log_{10}(Mass) \qquad (4)$$

Bird Diameter

$$Diameter = 0{,}0804 * Mass^{0,335} \qquad (5)$$

Bird Length

$$L = 4 * \left(\frac{Mass}{\pi \cdot Density \cdot Diameter^2} + \frac{Diameter}{12} \right) \qquad (6)$$
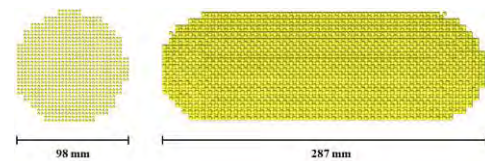

Figure 3: Bird Dimensions

The bird material behaviour is defined by the *Equation of State* (EOS) that describes the constitutive pressure-volume relationship of a fluid material. Welbeck [17] pointed out that bird material could be consider as a fluid for impacts in high velocities since pressure on bird tissues severely exceeds the tissues tolerable values. Hedayati and Sadighi [18] pointed out that most commonly used EOS for the water bird is a polynomial of degree 3 that corresponds to a hydrodynamic, isotropic, and non-viscous constitutive law, as eq. (7).

$$P = c_0 + c_1\mu + c_2\mu^2 + c_3\mu^3 + (c_4 + c_5\mu + c_6\mu^2)E_i \qquad (7)$$

$$c_1 = \rho_0 c_0^2 \qquad (8)$$

$$c_2 = (2k - 1)c_1 \qquad (9)$$

$$c_3 = (k - 1)(3k - 1)c_1 \qquad (10)$$

Where E_i is the internal energy, μ the change in density during impact, ρ_0 the density of the medium, c_0 the speed of sound in the medium and k an experimental constant (k=2 for water). For negligible initial equilibrium pressures, values of coefficients c_4, c_5 and c_6 are vanished.

To validate the SPH model used in the present work, it was performed simulations of the bird impacting rigid targets considering three different velocities (100 m/s, 200m/s and 300 m/s) and three different impact angles (90°, 45° and 25°). Values obtained for Stagnation Pressure and Impact Force were compared with parameter calculated with eqs. (11) and (12), while Peak Pressure was correlated with experimental results from Barber *et al.* [19]. Table 7 and fig. 4 present the obtained results.

$$P_s = \frac{1}{2} \cdot \rho \cdot v_b{}^2 \qquad (11)$$

$$F_{avg} = \frac{M \cdot v_b{}^2 \cdot \sin\theta}{L + D \cdot (\tan\theta)^{-1}} \qquad (12)$$

Where ρ is the bird density, v_b is the bird initial velocity, M the bird mass, L the bird length, D the bird diameter and θ the impact angle.

Since tab. 7 shows that values obtained during simulation were higher than the ones found on literature for experimental tests, the bird modelled with SPH methodology and

hemispherical-ended geometry was assumed to be conservative.

Table 7: Bird Impact Simulation Results

Impact Angle	Impact Speed [m/s]	Ps [MPa] Calculated	Ph [MPa] Literature	Ph [MPa] Simulation	Impact Force [N] Calculated	Impact Force [N] Simulation
25°	100	4.71	20.50	53.10	15411	38673
	200	18.85	64.28	102.84	61644	143091
	300	42.42	115.82	165.22	138700	302551
45°	100	4.71	42.90	121.14	33305	67282
	200	18.85	139.15	233.87	133220	238417
	300	42.42	257.06	397.69	299746	532227
90°	100	4.71	79.39	130.53	63207	113758
	200	18.85	240.48	331.44	252827	421963
	300	42.42	422.71	553.09	568860	867738
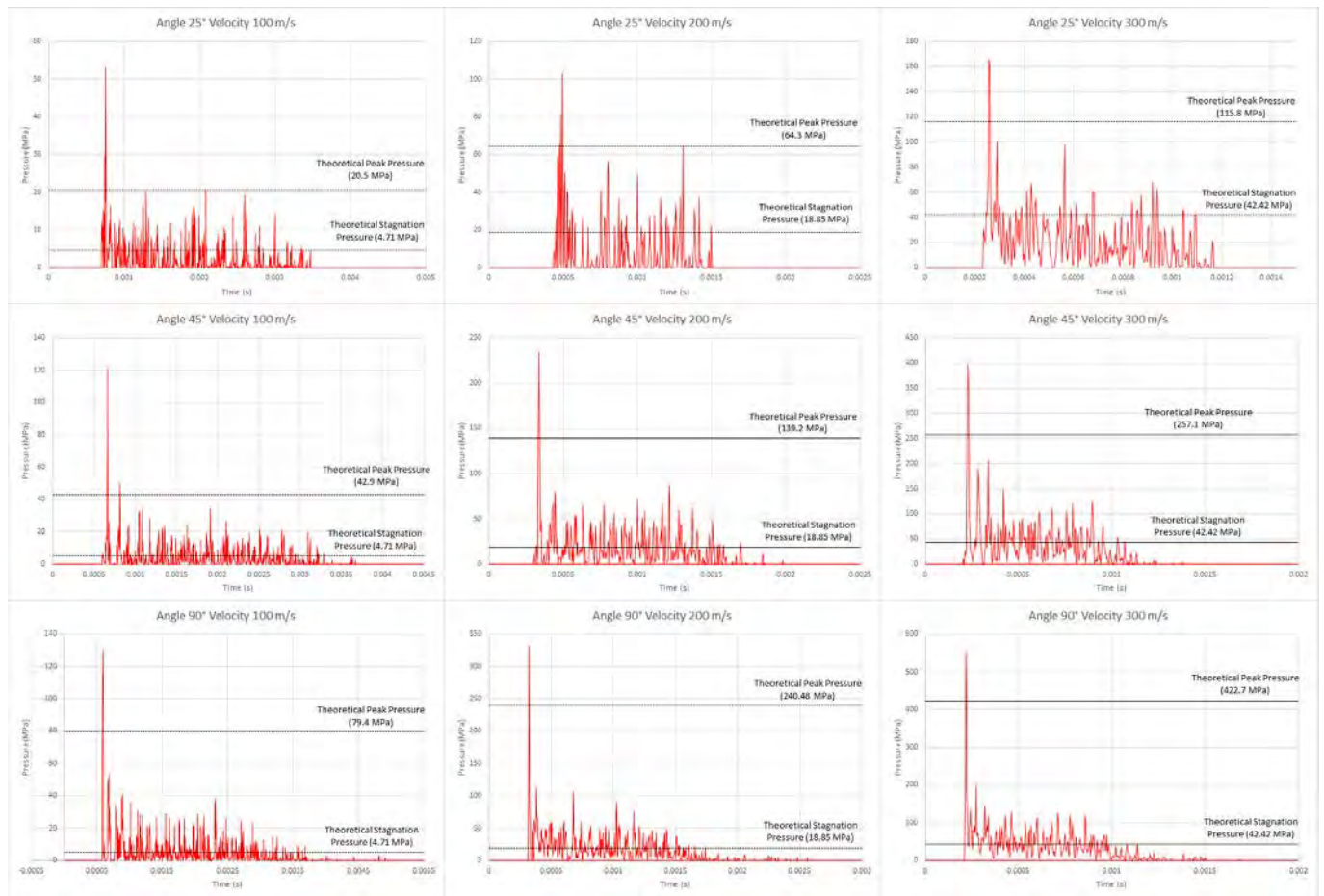
Figure 4: Bird Impact – SPH Model Pressures

3.2 UAS Impact

In order to validate the UAS model to be used, impact simulations of its rigid components (battery, motor and camera) were simulated separately. Each component was impacted against Aluminium 2024-T3 flat panels of different thicknesses at velocities of 250 knots and 100 knots. The final results were confronted with ballistic tests data performed on Olivares *et al.* [8].

Figure 5 shows the simulation performed for each UAS component and tab. 8 presents the correlation between simulation results and experimental data found in the literature.
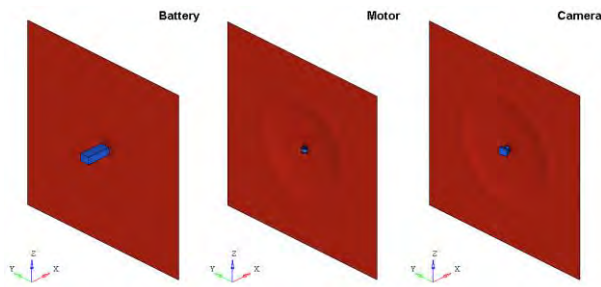
Figure 5: UAS Components Tests Simulation

On the ballistic test, the test setup consisted of two square steel frames bolted together sandwiching about a flat aluminium test panel, while in simulation translation and rotation constrains were attributed to nodes situated in the panel edge. Possibly, differences in the boundary conditions and simplifications adopted on the FE model, like not considering the plastic case that wrap the battery or the different materials on the camera parts, could have induced differences verified on Peak Loads and Maximum Deflection. Nevertheless, it was considered that simulations presented correlation with experimental data and the UAS model was validated.

Table 8: UAS Components Simulation Results

Component	Impact Velocity (knots)	Panel Type	Impact Energy (J)	Panel Penetration		Peak Load (N)		Max Deflection (mm)	
				Literature	Simulation	Literature	Simulation	Literature	Simulation
Battery	250	1.6 mm Al 2024-T3	2797.7	N	N	92483	106074	47.8	46.0
Battery	250	6.35 mm Al 2024-T3	3034.5	N	N	123892	167227	19.6	22.2
Battery	100	3.18 mm Al 2024-T3	424.6	N	N	48170	33603	14.5	12.7
Motor	250	1.6 mm Al 2024-T3	417.7	Y	Y	16160	21883	n/a	n/a
Motor	250	6.35 mm Al 2024-T3	473.2	N	N	58703	44637	7.9	9.1
Camera	250	1.6 mm Al 2024-T3	439.0	N	N	31142	28768	18.5	20.6

4 Aircraft Impacts

4.1 Bird Strike

The bird strike on the WFLE was performed with an impact velocity of 127.0 m/s (250 knots) and to be occur at a location between two FLE ribs, considered as a more critical region. Initially, it was set a WFLE skin thickness of 2.0 mm.

In the first attempt, skin presented rupture as shown in fig. 6. Therefore, another impact was performed increasing the skin thickness to 2.5 mm in order to evaluate if the aircraft would withstand the dynamic loads from the impact. In this second attempt, a small rupture was verified in the region close to the FLE Rib. A third simulation was performed with a 3.0 mm thickness on the skin and then all WFLE structure resist without suffer any failure, as shown in fig .7. The maximum effective plastic strain verified on the skin was 20.2%, while the Al 2024-T3 rupture strain is 26.0%, approximately, according to calculation performed with data from MMPDS [4].
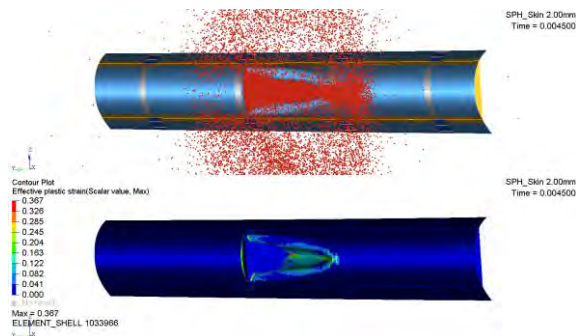

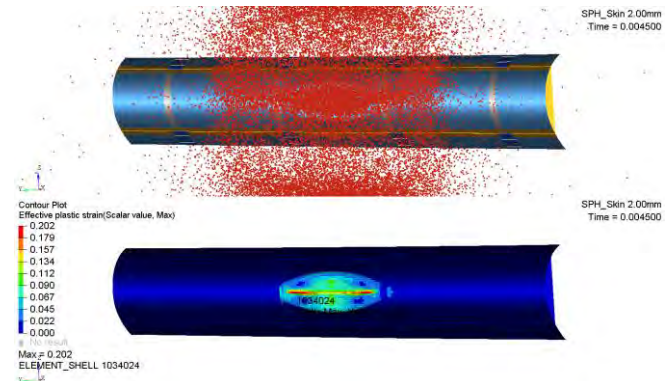*Figure 7: WFLE Bird Impact – thickness = 3.0 mm*

4.2 UAS Impact

In the same manner as in impact simulations with the bird, the UAS impact simulations considered the impact occurring in a velocity of 127.0 m/s (250 knots). This value was calculated from the difference between the maximum speed that the evaluated quadcopter UAS model can reach (20 m/s) and maximum speed that paragraph 5-3-8-j.2 (a) of Aeronautical Information Manual [20] establishes for a holding phase on a flight below 6000 feet (107.0 m/s).

4.2.1 UAS Position

Meng *et al*. [4] and Olivares *et al*. [8] performed studies to define the critical location where the UAS impacts a leading edge. Both works reported that when impacting close to the rib, the target structure mostly received damage to the rib and little penetration is verified into the airframe. In contrast, when the impact was centered between ribs, the skin absorbed most of the kinetic energy and the UAS perforated it

Figure 6: WFLE Bird Impact – thickness = 2.0 mm

penetrating into the airframe and possibly hitting the spar. Thus, in the simulations performed for the present work, the UAS was positioned to hit the WFLE centered between the two central FLE Ribs.

In order to identify the UAS orientation that could cause the larger damage to the aircraft, five different simulations were performed changing the UAS yaw angle from 0° to 90° as shown in fig. 8. The criterion adopted to evaluate which UAS orientation would be the worst was the energy absorbed by the spar and the extension of the damage verified in the WFLE skin. As can be seen in fig. 9, the simulation with the yaw angle of 0.0° presented the most critical results and for this reason, this position was selected to perform the following UAS impact simulations.
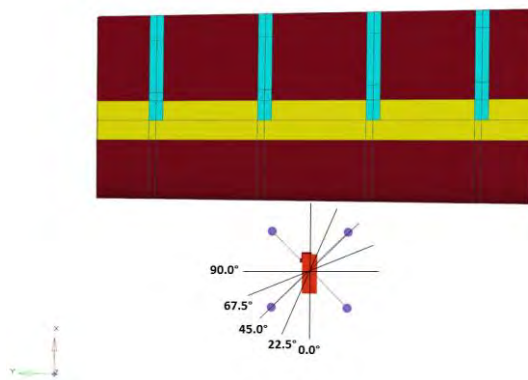
Figure 8: UAS Orientation
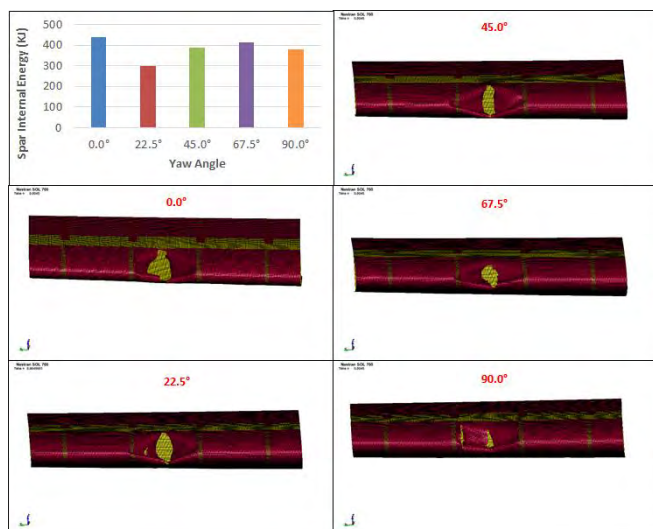
Figure 9: UAS Yaw Angle Simulation Results

4.2.2 Skin Reinforcement

As presented in section 4.1, the WFLE with skin thickness of 3.0 mm was able to withstand the impact with a 1.8 kg bird. In order to evaluated how critical would be an impact with an UAS, a simulation with the small aerial vehicle was performed considering the WFLE having the same skin thickness.

Simulation results presented in fig. 10 shows that the UAS easily perforated the WFLE skin. Such condition evidences a

very critical scenario, once that if the UAS penetrates the airframe, the front spar is exposed to impact with parts of the small aerial vehicle. In case of a front spar rupture, a catastrophic failure would occur. In addition to compromising a primary wing structure, UAS lithium batteries could ignite the fuel present in the wing main box causing a big explosion.
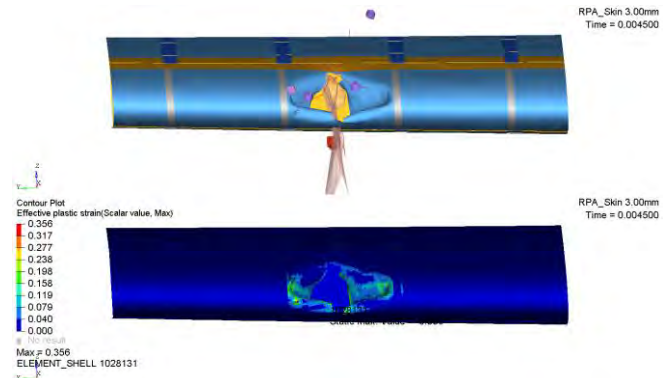

Figure 10: UAS impact in a WFLE with skin thickness of 3.0 mm

In order to evaluate if increasing the skin thickness would be an effective way to make the WFLE resistant to the UAS impact, a simulation was performed considering a skin thickness of 5.0 mm, even though a thick sheet of this order could represent serious difficulties for manufacturing the component. As can be seen in fig. 11, the UAS penetrates the airframe even in a WFLE with a thick Al 2024-T3 skin.
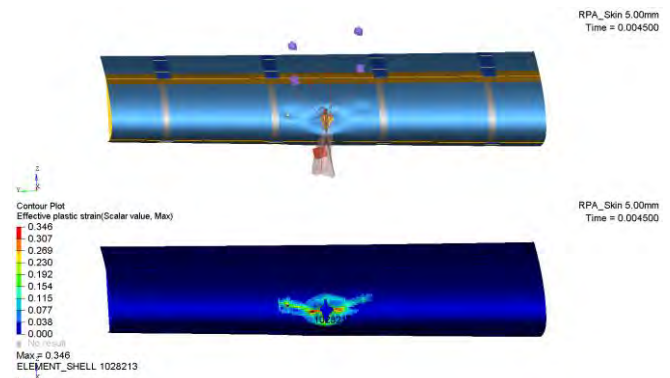

Figure 11: UAS impact in a WFLE with skin thickness of 5.0 mm

4.2.3 Spar Reinforcement

Since it was found out that increasing skin thickness would not be an effective way to make the WFLE resistant to UAS impact, an alternative solution would be design the front spar to withstand the impact. To this end, simulations were performed removing the skin from WFLE model and considering the UAS impact to occur directly in the front spar web. Different web thicknesses were investigated for the Al 7050-T7451 spar and results are shown in fig. 12 to fig. 14. From fig. 14, it can be noticed that no rupture occur on spar web with a thickness of 12.0 mm, although the effective plastic strain detected was very close to the material rupture strain.
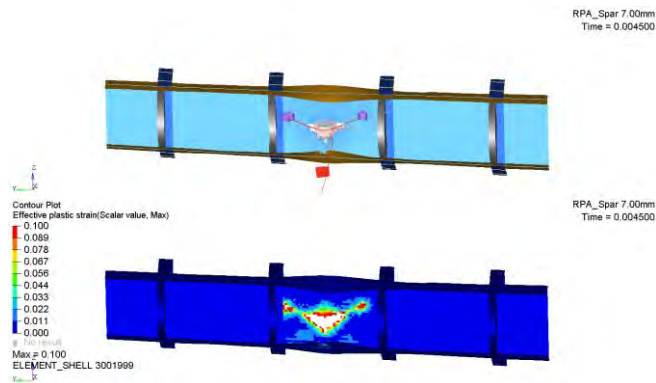
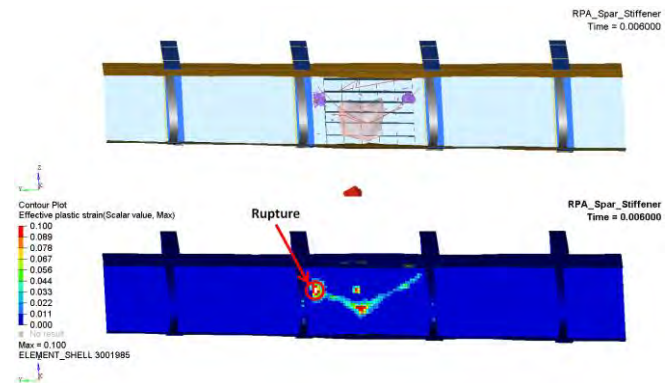*Figure 12: UAS Impact in Spar with Web Thickness of 7.0 mm*
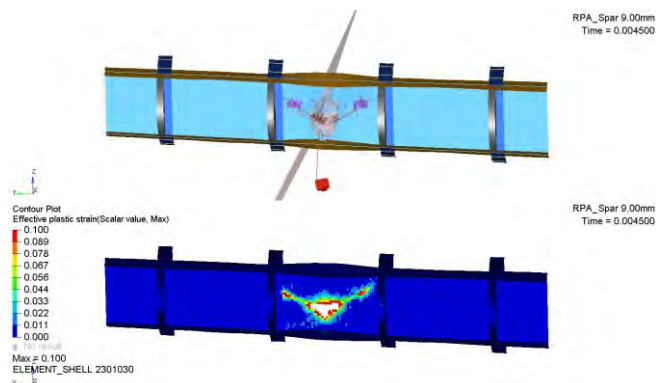

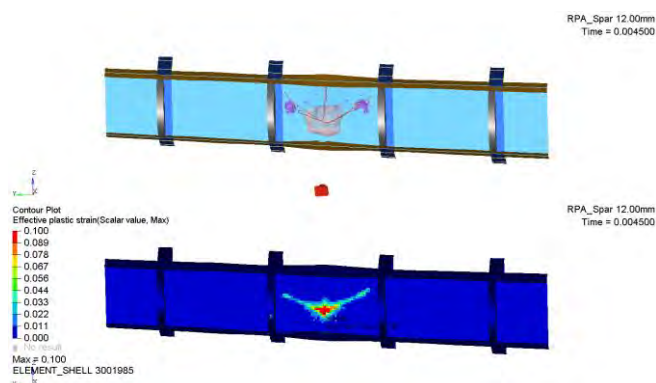*Figure 13: UAS Impact in Spar with Web Thickness of 9.0 mm*

Figure 14: UAS Impact in Spar with Web Thickness of 12.0 mm

Other reinforcement in the spar were also evaluated adding stiffeners on its back region in order to reduce the web deformation. The stiffeners dimensions were 26.5 mm x 4.0 mm and the web thickness considered was 7.0 mm. On fig. 15 it is shown that a small rupture occurred on region where stiffener was not added.
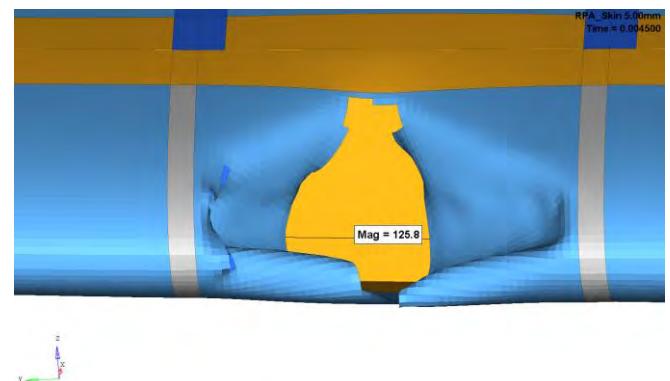

Figure 15: UAS Impact in Spar with Spines (26.5 x 4.0mm)

5 Discussions

5.1 Bird Strike X UAS Impact

During the present work it was verified that the mid-air collision with an UAS is much more severe for an aircraft than an impact with a bird. Even though the energy considered for both impacts (same mass and same velocities were considered on simulations), while the bird impact did not cause the rupture of a WFLE skin with a thickness of 3.0 mm, the UAS easily perforated the skin and caused a damage of 125.8 mm, as shown on fig. 16.

Figure 16: Damage cause by the UAS Collision

The first reason that could explain this difference refers to the material behaviour during impact. The bird was modelled with a fluid material that splashes when hit the aircraft, increasing the contact area and consequently reducing the pressure applied on the WFLE skin. Meanwhile, the UAS was represented with solid elements that apply concentrated loads on the WFLE, inducing higher local deformation and consequently causing the material failure.

The other reason that could explain the larger damage caused by the UAS impact is associated to its stiffness. While the bird was treated as a homogenous body during simulations validated with test results, the UAS was modelled with components made of different materials, all of them much harder and resistant than the one considered for the bird. The rigid structures battery, motors and camera induce higher loads in the WFLE and local stress regions during the impact. As can be seen on fig. 17, the contact force amplitude for the UAS impact is higher than the one for bird strike.
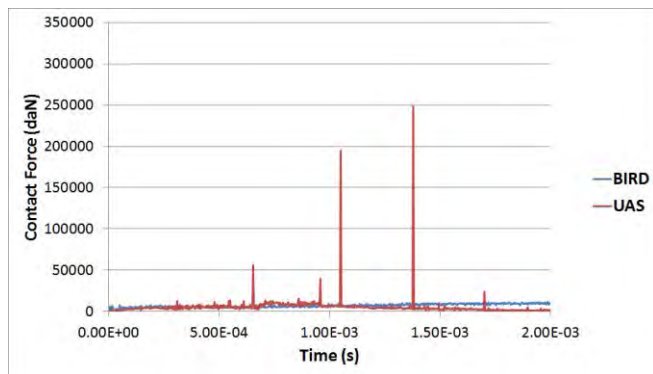
Figure 17: Contact Forces for Bird and UAS Impact

Table 9: WFLE Weight Comparison

	Weight (kg)	Added Mass (kg)
WFLE dimensioned with Bird Strike	31.6	-
WFLE dimensioned with UAS impact (spar web thickness)	35.8	+ 4.2
WFLE dimensioned with UAS impact (spar with stiffeners)	35.6	+ 3.6

5.2 Skin and Spar Reinforcements

The evaluation of reinforcements in the WFLE in order to make it resistant to UAS collision demonstrated that increase in spar thickness would be more effective than in skin thickness. In general, the spar web thickness is much thicker than WFLE skin due to necessity of spar withstand limit loads that could occur in the aircraft during a flight in case of a maneuver or a system failure. If the skin thickness were increased until it not rupture during UAS collision, the amount of weight added to the WFLE could be so large that would reduce drastically the performance of a commercial aircraft. In addition, other problems associated with thicker skins would also have to be evaluated, such as difficulty on forming thick aluminum sheets and excessive energy consumption for heating the leading edge in order to prevent ice formation.

For reinforcements in the spar, two configurations were evaluated: 1) Thicker web thickness; and 2) Additional stiffeners on its back region. The stiffeners would not be so effective since the UAS impact loads occur in small concentrated regions in a way that to the stiffeners work properly, they should be positioned exactly where the higher loads occur. Considering the impossibility to determining the exact point where the impact will occur in the wing, a large number of stiffeners would have to be added.

On tab. 9 it is presented a weight comparison between the WFLE dimensioned with skin thickness of 3.0 mm able to withstand the impact with the bird and the WFLE's that could resist to the UAS collision after adding reinforcements on its spar increasing the spar web thickness and adding stiffeners on its back region. Although the stiffeners are not very effective, the WFLE reinforced with them had less added mass than the one reinforced only with spar web thickness. It is important to notice that values on tab. 9 refers only to the part of WFLE considered on simulations, which have a length of 1435 mm. In large commercial aircrafts, the WFLE could reach almost 35.0 m in each half wing.

In future works, it could be evaluated the use of additional structures that split the UAS before it reaches the front spar, in the same way that it is already done with birds in some aircrafts. The applicability of alternative materials, such as honeycombs, could also be studied as means to make the WFLE withstand the impact with harder components. Other aircrafts regions as the windshield, engines and horizontal and vertical empennages also have to be analyzed with the UAS impact since all of them are exposed to this kind of occurrence. Finally, more experimental tests with UAS impact in a real aircraft structure have to be done in order to increase the validity of computational models and numerical simulations.

6 Conclusion

In the present work, it was analysed numerically a bird and a UAS impact against aircraft Wing Fixed Leading Edge. The bird strike simulations were developed and validated with experimental data obtained in literature, while the UAS collision was validated with results from ballistic tests found in the literature performed with different parts of the UAS. Both impact scenarios in the WFLE were compared and reinforcements were evaluated. The following conclusions f can be highlighted:

1) The impact behaviour simulated by the FE model developed in this work was in acceptable agreement with test data;
2) The commercial airliner must be reinforced so it could complete the flight safely when a UAS airborne collision happens at its cruising speed. The damage to the WFLE spar and fire risk of the lithium-ion battery must be take into account;
3) UAS impact would cause more damages in a WFLE than bird strike considering the same mass. The UAS penetrated the skin and could cause damages on primary structures of the aircraft, like the front spar. Relevant airworthiness standards should be developed to ensure the safety operation of commercial flights;
4) Reinforcements on the WFLE spar web proofed to be more efficient than reinforcements in the skin. For future works, additional structures could be evaluated in the WFLE to split the UAS before it reaches the spar.

Acknowledges

References

[1] JENKINS D.; VASIGH B. The economic impact of unmanned aircraft systems integration in the United States, Association for Unmanned Vehicle Systems International (AUVSI). P. 1-40, 2013

[2] SHARMA, R. S. Investigation into Unmanned Aircraft System Incidents in the National Airspace System. *International Journal of Aviation, Aeronautics, and Aerospace*, Article 2, Volume 3, Issue 4, 2016.

[3] WILD, G.; MURRAY, J.; BAXTER, G. Exploring Civil Drone Accidents and Incidents to Help Prevent Potential Air Disasters. *Aerospace* 2016, 3, 22.

[4] MENG, X.; SUN, Y.; YU, J.; TANG, Z.; LIU, J.; SUO, T.; LI, Y. Dynamic response of the horizontal stabilizer during UAS airborne collision. *International Journal of Impact Engineering*, 126, p.50-61, 2019.

[5] OLIVARES, G.; LACY, T.; GOMEZ, L.; ESPINOSA DE LOS MONTEROS, J.; BALDRIDGE, R. J.; ZINZUWADIA, C.; ALDAG, T.; KOTA, K. R.; RICKS, T.; JAYAKODY, N. UAS Airborne Collision Severity Evaluation - Executive Summary – Structural Evaluation. Federal Aviation Administration, *Report DOT/FAA/AR-XX/XX*, 2017a.

[6] BRITISH MILITARY AVIATION AUTHORITY, Small Remotely Piloted Aircraft Systems (drones) - Mid-air Colision Study. *Report-16*, 2016.

[7] GOKHALE, N. S.; DESHPANDE, S. S.; BEDEKAR, S. V.; THITE, A. N. Practical Finite Element Analysis. Finite to Infinite, 2008.

[8] OLIVARES, G.; GOMEZ, L.; ZINZUWADIA, C.; ESPINOSA DE LOS MONTEROS, J.; BALDRIDGE, R. J.; ALDAG, T. UAS Airborne Collision Severity Evaluation – Volume II – Quadcopter. Federal Aviation Administration, *Report DOT/FAA/AR-XX/XX*, 20 mm17b.

[9] CAIRNS, D. S.; JOHNSON, G.; EDENS, M.; ARNOLD, F. Volume I - UAS Airbone Collision Severity Evaluation – Projectile and Target Definition. Federal Aviation Administration, *Report DOT/FAA/AR-XXX/XX*, 2016.

[10] JOHNSON, G. R. A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures. *Proceedings of the 7th international symposium on Ballistics*, 1983.

[11] KAY, G. Failure Modeling of Titanium 6Al-4V and Aluminum 2024-T3 With the Johnson-Cook Material Model”, *Report DOT/FAA/AR-03/57*, Federal Aviation Administration, 2003.

[12] MMPDS-09, *Metallic Material Properties Development and Standardization* (MMPDS), Chapter 9, Ed. 7, Battelle Memorial Institute, 2014.

[13] ”520.0 (520.0-T4, formerly 220.0, LM10, G10A, A05200) Cast Aluminum”, URL:

https://www.makeitfrom.com/material-properties/520.0-520.0-T4-formerly-220.0-LM10-G10A-A05200-Cast-Aluminum

[14] SAHRAEI, E., MEIER J., WIERZBICKI, T. Characterizing and modeling mechanical properties and onset of short circuit for three types of lithium-ion pouch cells. *Journal of Power Sources*, Volume 247, p. 503–516, 2014.

[15] HEDAYATI, R.; ZIAEI-RAD, S. Effect of bird geometry and orientation on bird-target impact analysis using SPH method. *International Journal of Crashworthiness*, Volume 17, p. 445–459, 2013.

[16] O'CONNOR, J. The Hazard Posed to Aircraft by Birds – Attachment B. *Australian Transport Safety Bureau*, 2002.

[17] WILBECK, J. S. Impact behavior of low strength projectiles. *PhD dissertation*, Texas AM University, 1978.

[18] HEDAYATI, R.; SADIGHI, M. Bird Strike – An Experimental, Theoretical, and Numerical Inverstigation. *Woodhead Publishing in Mechanical Engineering*, 2016.

[19] BARBER, J. P.; TAYLOR, H.R.; WILBECK, J. S. Bird Impact Forces and Pressures on Rigid and Compliant Targets. Wright-Patterson Dynamic Laboratory – *AFFDL-TR-77-60*. May, 1978.

[20] AERONAUTICAL INFORMATION MANUAL. *Official Guideline to Basic Flight Information and ATC Procedures*. Federal Aviation Administration, 2017.

Aerospace Electric Generator Specification and Selection – Opportunities and Challenges

Avo Reinap

Div. Industrial Electrical Engineering and Automation, Lund University, Lund, Sweden
E-mail: avo.reinap@iea.lth.se

Abstract

The goal of this work is to specify and investigate suitability of a number of selected permanent magnet synchronous machine (PMSM) topologies to be mounted directly on a low pressure shaft (LPS) of a geared turbofan engine. The starting point of the work is to explore the status of electrical machines for more electric engine in order to study different design aspects when identifying and defining the designs. Even though the high flexibility and power density is anticipated from the system electrification the reliability is one of the most important aspects for the machine topology development and faultlessness can compromise desired performance. Large inner diameter of LPS and wide speed range are challenging the selection and specification of PMSM. The design initialisation is carried out by looking for more fault tolerant arrangements between machine windings and power converter topologies and appropriate rotor layouts that can provide excitation and constructional stability. Design space exploration is carried out on surface mounted Halbach array PMSM with distributed winding in the stator as one of the machine types that is best suited for high power density and high-speed application. Machine design aspects on modular design for faultlessness and cooling integration for enhanced thermal management and higher power density are the arguments behind the sensitivity study carried out on interior permanent magnet synchronous machine with non-overlapping concentrated winding. Six different magnetic core materials are compared to demonstrate that thin FeSi electromagnetic steel grades improve efficiency and FeCo materials mainly torque density but the machine may remain unsatisfactory in power and efficiency.

Keywords: More electric engine, electric generators for aircraft engine, electric drive and electric power system, permanent magnet synchronous machines, design of electrical machines, inset permanent magnet and Halbach array rotors, power losses in magnetic cores, cooling integration of electrical machines,

1 Introduction

Transport electrification gains uninterruptedly interest as this is apparently one of the promising ways towards more sustainable, energy efficient and environmentally friendly transportation in the future. All means of transports from road and rail to marine and airborne are under the investigations of the possible benefits that powertrain electrification can provide and accomplish compared to the existing solutions. The anticipated improvement in added functionality, increased efficiency, improved packing and power density is achieved through the advancements in control technology and power engineering which is applied on the electric drive system optimization and coincident development of system components: energy storage, power (micro) grids, power electronic converters and rotating electrical machines. No pain, no gain – all system components, in an electric drive system, are under

continuous research and examination in order to increase the functionality, power density, efficiency and reliability. The technology development is inspired by seeking for more suitable materials supported by rational production technologies and thereby looking for more affordable solutions and sustainable products. Since the transportation means motion, mechanical energy needs to be present and unquestionably the electrical machine or machines take place in this system. If the road and rail bound vehicles seek for traction motors, then aircrafts are after electric generators / starter motors that would fulfil the growing demand on electric power and added functionality. This demand is due to the electrification in the aircraft, where many existing hydraulic, pneumatic and mechanic actuators or related functions are replaced by electric counterparts.

The advancement of More Electric Aircrafts (MEA) hangs on the development on More Electric Engines (MEE) where the integration of electrical machines into the gas turbine is

foreseen and encouraged not because of supplying increased electric loads but as the possibility to reduce weight and increase efficiency of the turbomachinery. Due to harsh operating conditions and increased demand not only on power but also power density concern system developers and component designers accomplishing the challenging task of specifying an appropriate electric power generation and electric power system. As a matter of fact, the increase in power density or load is naturally not coaligned with the increase of durability, reliability or efficiency. In order to meet this challenge, the article explores the opportunities on 1) high efficiency magnetic core materials, and 2) power electronic converter to winding arrangement that are considered in the electrical machine design and development process. This paper addresses the design of electrical machine which is directly attached onto a low-pressure spool (LPS) of a turbofan engine. The design initialization is carried out in the first section where the anticipated perspectives and objectives for MEE [1]-[3] are faced to the recognized challenges and opportunities [4]-[7]. The aim is exploring several design aspects and specifying four generator topologies for further study. The following section explores the design space and compares the candidates in terms of size and material selection for the magnetic core since the large inner diameter of the LPS, the available geometric space and the speed range are some of the early recognized design concerns. The third section focuses on a single topology among four candidates and investigates some compromises between electromagnetic and mechanic and thermal design.

2 Electrical machine specification

First the jet-engine generator arrangement is studied, followed by electric drive system topologies with concern on power electronics, converters and microgrids [9]-[14] in order to specify the boundary conditions. Concerning to the performance and faultlessness capability, the converter to machine winding arrangement is explored [15]-[19] and the performance limits of the machine topologies are investigated [21]-[28]. Most commonly an electrical generator in a commercial transport aircraft, which is a three-stage wound-field synchronous machine, is externally coupled to high-pressure spool (HPS). Permanent magnet types of machines, which would provide engine starting function and higher power density, are continuously researched as a technological improvement for MEE.

2.1 Jet-engine generator arrangement

In the face of the recent growth in air transport and climate concern the powertrain electrification and weight reduction has been always one of the aircraft development targets. Already back in a half century the advantages on integrated engine-generator concepts were drawn and suitable electrical machine designs not only compared but also suggested [1]. The work concluded that 1) electrical machine should be mounted to High Pressure Spool or Shaft (*HPS*) so that the electrical machine (EM) is not only generating electricity but can be used at least *30 seconds* to start commercial turbofan engine *180 kN* thrust class, 2) the

use of electric energy reduces the need of engine shaft driven accessories and bleed air requirements, which improves packing and efficiency of the engine, and 3) the machine requires no extra space than already available and the reliability is higher compared to the machines used at that time. As a matter of fact, the power rating and the inner diameter of the machine, which is *Ø208 mm* and *200 kVA*, are nearly the same values targeted in this work. Even if full electrification is encouraged [2], which anticipates multi-spool embedded electrical machines with integrated active bearings in order to facilitate power management and dynamics in the turbomachinery, and technological development would support this growth [3] the published efforts on design and prototyping of embedded EMs and specifically EMs for LPS and for commercial transport aircraft are less common.

The evolution of the electric generation and power system perspective in context of MEA and MEE can be shortly summarized based on a few publications where the system development challenges [4] and some of the achievements [5] are reviewed. An extensive review of an electrical generator development is given in [6] and the generator development perspectives are collected in [7]. As the EM operates not only in generating operation but also motoring when starting up the engine, the engine architecture development gives both opportunities and challenges to electrical machine engineering to contribute to system efficiency, functionality and reliability. The anticipated features are that the electrical machines are attached to multiple engine shafts to provide not only distributed power generation but also enhance engine control at transient operation. According to [5] the distributed power generation endows with lower fuel conception at increased electrical loads meanwhile the wide speed range of the low pressure spool is one of the challenges and research subjects for an MEA paradigm.

The target engine configuration is a geared twin-spool turbofan (Figure 1) where Geared Turbofan Technology enables less fuel burn, acoustic noise and emission [8].
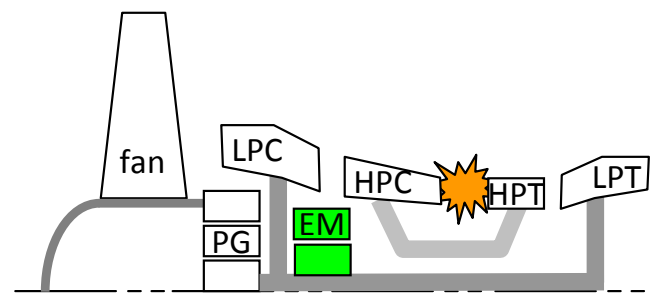
Figure 1 Schematic layout of a geared twin-spool turbofan.

The engine layout shows following components as a fan, planetary gear (PG), low pressure compressor (LPC), high pressure compressor (HPC), combustion chamber, high pressure turbine (HPT), low pressure turbine (LPT) and electrical machine embedded onto the low-pressure spool. Even if the EM for LPS is not able to start the engine it can contribute to the start up. However, the EM needs to operate over a wider speed range EM for HPS. The machine

operates mainly in generator operation including ram air turbine (RAT) function to generate emergency power at wind milling action of the gas turbine [4].

The design parameters of EM for LPS are selected inexact to the target system in order to pronounce the design challenges rather than integration constrains. The inner diameter of EM is Ø200 mm and the constant power speed range is defined from 1500 to 7500 rpm while the over speed is selected 15000 rpm for overload investigations.

The early design concerns of the large diameter and rotation speed is due to conflictive interest of reducing surface speed of the rotor of EM and the number of poles for a largest suitable supplied electric frequency. In case of inner rotor EM, the smaller outer diameter would result smaller structural stress due to centrifugal forces, but it would result unnecessarily high number of poles and magnetization frequency and related power losses or magnetically inefficient EM construction. On the contrary, large rotor would challenge not only the mechanical construction but also requires larger stator core for the best electromagnetic outcome. In either case, the EM may generate already unnecessary amount of no-load losses while operating at no-load and if the magnetization cannot be easily reduced, which would be the case of more torque efficient permanent magnet (PM) machines.

A high power-density and efficiency are anticipated from electrical machines. Consequently, PM machines may give advantage due to higher magnetization and electromagnetic torque at limited space. As the power is proportional to the rotation speed the machine becomes smaller and lighter for the given power. As the electrical machine becomes more attractive to provide rather the speed than the torque for given power the presence of mechanical transmission becomes rather practical in order to adapt the load torque and speed requirements. The jet-engine generator arrangement can be compared to the similar electrification development in powertrains for vehicular application, where large low speed direct coupled in-wheel machines can be compared to smaller high speed machines coupled via mechanical transmission (MT). As an outcome, the overall system becomes smaller when optimizing EM torque generation capability to MT torque conditioning capability. As a matter of fact, the most common solution, which is based to B787 and A380 example, the main engine generator is directly coupled to HPS via a gearbox. The advantage of the variable speed shaft (or shafts) driven by HPS as a mechanical power transmission is the effect of challenge segregation as the size of EM is not driven by jet-engine and EM is or EMs are by far accessible. Furthermore, the external EM may take advantage of position sensor for more common solution and superior drive performance. In case of embedded EMs sensorless control technique need to be considered. Even if this work focuses on a single large embedded EM for LPS, the presence of PG would give an option of distributed generator system integrated into a set of planet gears. This opportunity, if feasible, would reduce EM size in respect to the rotation speed.

2.2 Electric drive system

Concerning to EM, the electric power system (EPS) development moves from the constant input speed and constant output frequency in the EM terminals to variable input speed and variable output frequency in the EM terminals. In case, the magnetization of the EM cannot be changed like in PM machines then also the voltage becomes variable. This can be handled by an active front end rectifier / inverter together with field weakening control. Literally the challenges in mechanical engineering in a constant speed drive (CSD) unit is replaced by new challenges of electrical and control engineering in electric drive system (EDS) to perform safely and provide supply voltage, which is specified for the electric power system (EPS).

The EM coils or/and windings are terminated into a power electronic converter (PEC) (Figure 2). The purpose of the electric drive system (EDS), which consists of EM and PEC, is to connect variable frequency and voltage of the AC intermediate link to a constant voltage DC distribution system. The purpose of the PEC is to control power flow from EM to EPS. If the excitation of the electrical machine can be controlled like in a conventional three stage field-wound electrically magnetized synchronous machine (EMSM) then the voltage can be kept constant over the speed and load range as the magnetization current of the main exciter can be adjusted accordingly. This option provides also control over failure modes as the excitation can be switched off in the case of short circuit (SC) current or over voltage at open circuits (OC) of the machine windings. This would not be the case for permanent excitation or permanent magnet synchronous machine (PMSM). In generator and OC operation over voltage is induced at over speed and since the EM cannot be demagnetized it keep short circuit currents circulating at SC. If 3-stage EMPM includes unsophisticated PEC in rotor to control the excitation current, then PMSM must have more sophisticated PEC that can assure desirable faultlessness. A sketch of an EM winding to PEC arrangement is shown in Figure 2. The drawing demonstrates that the entire machine winding is divided into a number of separately driven coils from H-bridges or branches of PEC (upper half in Figure 2) or symmetric winding sections (lower half in Figure 2). The EM winding may have higher number of phases than 3 for the sake of higher redundancy or even efficiency [4].
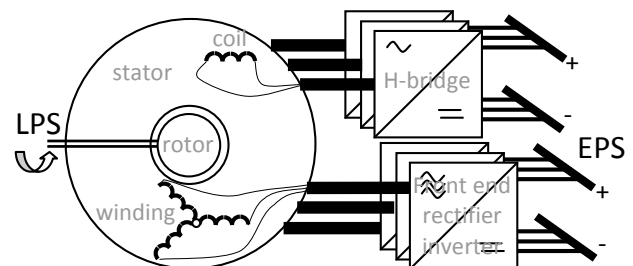

Figure 2 Sketch of electric drive system arrangement.

The electrical machine topology is selected in accordance to winding layout and PEC arrangement so that EDS is capable to provide desirable functionality and reliability. Different EPS architectures and MEA PEC are reviewed in [9]

whereas [10] compares efficiency and total harmonic distortion (THD) of current in 22 PEC PMSM arrangements, [11] explores deeper into aircraft on-board microgrids including control and distinguishes safety oriented performance optimized research topics between device level and system level as realization oriented and power/energy management related research. As an advantage, high power EM PEC topologies in wind power industry [12] and in DC microgrid arrangement [13] give good insight to aircraft EDS EPS development.

The selection of EM winding layout together with PEC 2-level or 3-level inverter topology while targeting higher efficiency, more reliable operation, reduced electromagnetic interference (EMI) results challenging compromises that influences machine design. Depending on reducing weight and power losses a 3-level topologies can be obvious choice while traditional 2-level inverter would offer higher reliability doe to simplicity and component count [14]. A dual fed machine where all phases are supplied independently allows highest grade of fault tolerance capability [10]. Reference [15] explores EMs for MEA load motors and looks into failure probabilities, modular single-phase and multi-channel three-phase approaches to more fault tolerant EDS as a compromise redundancy to added complexity. It is desirable that the system can sustain faults like SC without excessive heat generation or fault propagation and take this into consideration in EM PEC topology selection and arrangement [16]. The modularization of EM coils or winding-sections as well as multi-phase windings provide increased reliability [17]. The EM winding to PEC should be selected so that machine would be able to operate under fault, the occurrence of the fault is minimum, condition monitoring and post fault operation methods can be implemented and applied [18][19]. Furthermore, the practical transport electrification experience in automotive application concerning to EDS realization and reliability gives again good insight to the practical implementation and development [20].

2.3   Electrical machine types

From winding layout to converter arrangement and reliability point of view a switched reluctance machine (SRM) is attractive candidate as apart from aforementioned features it has also a high speed and a high temperature capability that because of simple mechanical construction. Likewise, an induction machine (IM) is classed to similar category as this type of machine is used in high speed high power applications. However, this machine, which is also excited via the stator like SRM has not found as common usage in aircraft similar to SRM [6]. This is partly due to the fact that SRM endurance depends on converter and the machine has taken advantage of concentrated windings (CW) while IM uses mostly distributed windings (DW) that have phases in vicinity due to overlapping and this is seen has higher risk for more reliable winding layout. Different machine types: IM, SRM, EMSM, PMSM has been reevaluated time and again shown that the latter, PMSM, has the highest specific torque and power capability per unit of machine while the other machine types are attractive as they

have lower cost and can operate over wider speed range [21][22]. Reference [23] organizes various prototypes and commercial products in order to explore suitability of different machine types for jet-engine generator and expresses the reliability concern of PMSM type of machines. The EM development and evaluation are based on appropriate design models [24]. The power density limits and design trends for PMSM are explored in [25] where a fast evaluation model is developed showing that cooling capability has higher influence over power density of EM than rotor surface speed, which is the mechanical design constrain and defines whether the EM is high speed machine or not. Also increasing power levels the power density reduces according to the model. If the previous study compared existing designed prototypes in order to validate the model the maximum surface speed of the rotor was no more than *200 m/s*. Reference [26] explores a number of high speed machines, where the circumferential surface speed *v* exceeds also *200 m/s* with the highest value of *280 m/s* for laminated core rotors and expresses a mechanical limit curve among collected power *P* (in *kW* in (1)) from few hundred Watts to few MW with respect to rotation speed *n* (in *1/s* in (1)),

$$\log_{10}(n) = 4.27 - 0.275 \log_{10}(P) \qquad (1)$$

If exploring the design space where the rotation speed *n* is in the range of *100* to *200 1/s* (*6* to *12 krpm*), and the rotor diameter *d* is in the range of *0.25–0.35 m* then the surface speed *v*,

$$v = \pi d n / 60, \qquad (2)$$

is in the range of *78.5–220 m/s*. If using the design "guide number" as a combination of speed and power in order to "account" for size and dynamic problems and when selecting $1.5 \cdot 10^5$ rpm√kW for interior permanent magnet synchronous machine (IPMSM) then the expected power capability for the machine would be ca 150kW [22][27].

All EM types can be developed so that they provide higher mechanical or electrical reliability and the focus is more on mechanical stress analysis of the rotor and selecting more suitable winding to converter configurations, respectively [27][28]. This compromises mechanical reliability to electromagnetic performance. In addition, all EM types have been subject to engine generator design initiation and development towards (embedded) EM and EDS prototypes. Concerning to high power *250 kW* machines there are few examples on prototype development and evaluation of: [29][30] RSM and [31][32] PMSM. Most often the prototypes are developed for lower power and higher speed with the intention to connect it externally to HPS via gearbox transmission [33][34][35][36][37][38] when considering the speed and size of the machines. Nevertheless, safety critical machine concepts for direct drive attached to engine main shaft are studied [39] and reduced scale prototype build for evaluation [40]. Likewise, this work proposes a few PMSM topologies for further suitability analysis for usage in LPS in an aircraft geared turbofan engine.

2.4 Permanent magnet machines

A few PMSM topologies are selected for further investigations, where PM rotor and stator with cooling jacket is shown in Figure 3. The initial designs, both outer rotor and inner rotor configurations, are shown so that it could fit into the engine (Figure 1).
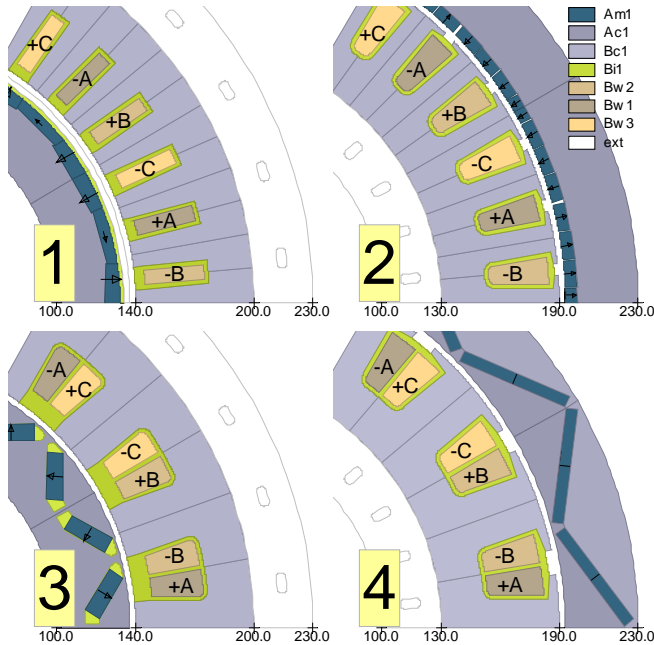

Figure 3 Overview of PMSMs: inner rotor (1,3) and outer rotor (2,4) machines, distributed (1,2) and concentrated winding (3,4) machines, surface mounted (1,2) and interior permanent magnet (3,4) machines.

The PMSM is often preferred due to high efficiency and torque capability. The EM design based evaluation is carried out on inner rotor machines:

- Interior Permanent Magnet (IPM) rotor (3, Figure 3) where the mechanical rotor integrity is provided by high strength rotor FeSi steel

- Surface mounted Permanent Magnet (SPM) rotor (1, Figure 3) where the mechanical rotor integrity is mainly provided by retaining sleeve

Both these rotors are capable to provide the large constant power speed range that depends on rotor design and the air-

gap between rotor and stator. Concerning to stator layout there are two types of winding layouts of interest:

- Distributed windings (1, Figure 3) that has a lower harmonic content but lengthy winding overhangs and end turns,

- Concentrated windings (3, Figure 3) as a more reliable windings due to larger space to neighbour phases

The specified machine configurations have 12-poles and the windings are arranged for 3-phase system (Figure 3) and the end turns are shown in Figure 4 with the exception that the distributed winding has two slots per pole and phase instead of one (Figure 3). As there is 2 stator and 2 rotor topologies of interest there is 4 PMSM machines in total under the design development. In addition to the performance estimation provided in this work, the detailed investigation is carried out in [42] while this article focuses on more general aspects on machine topology selection and specification.

Electric machine design consists of two main connected parts: 1) performance estimation and 2) fabrication preparation. Manufacturing aspects are vital in the EM design as the selection of materials and manufacturing aspects have a great impact to the power capability of the machine. As the machine operates at high frequency it is necessary to explore various power loss mechanisms as well as explore possibilities to improve heat dissipation and thermal management. Figure 3 demonstrates two types of windings that can either be premade and placed into the open slots (Figure 3 1,3) or loose coils/group of wires are inserted into semi-closed slots (Figure 3 2,4) that respectively would provide well defined insulation system and thermal contact from winding to core compared to windings with reduced speed dependent losses due to smaller slot leakage apart from strands and transposition. Consequently, the EM topology specification is defined from

- stator side – selecting winding type and assembling technique together with suitable magnetic core, and

- rotor side – specifying magnetization arrangement in respect to speed dependent structural loads and mechanical stresses.
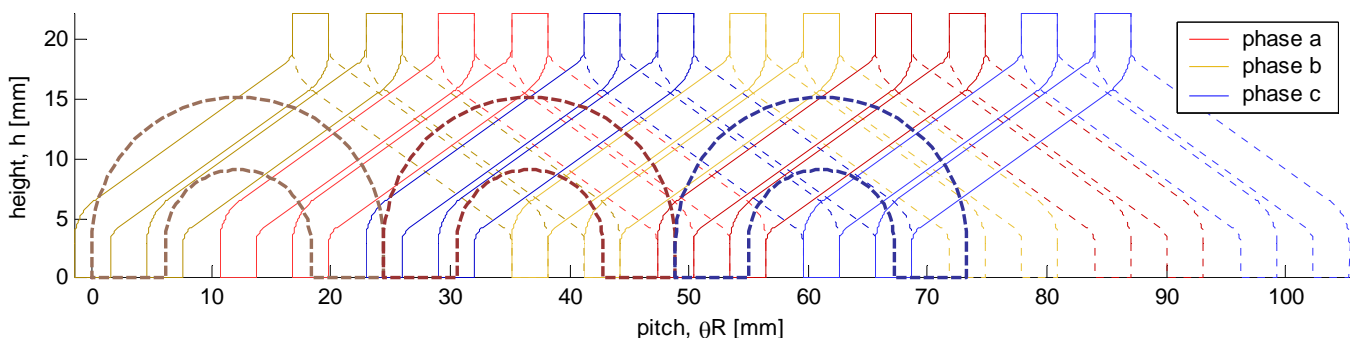

Figure 4 End turn of 12-pole 72-slot diamond type multi-layer winding and 18-slot concentrated type of winding. The coil distribution of the first pole-pair is drawn along inner stator periphery of Ø280 mm machine.

DOI
10.3384/ecp19162009
     Proceedings of the 10th Aerospace Technology Congress
October 8-9, 2019, Stockholm, Sweden
     85

3 Design space exploration

The machine design consists of number of compromises. The preference is given to electromagnetic design by choosing a narrow 1 mm air-gap and mechanically weaker rotor design: 2 mm thick retaining sleeve for SPMSM and 0.5 mm thick rotor bridges in IPMSM. The purpose of this selection is to identify the PMSM machine topology and selection of core material that suits best for large diameter high speed machine.

3.1 Parameter specification

The main EM design parameters exploring the design space is the number of poles, N_p, and stator radius difference, Δr, between the outer and inner rotor of the EM. The inner diameter is *280 mm* and the stator core is changed proportionally when increasing the outer diameter from *360* to *440 mm* and number of poles is increased from 8 to 16 (Figure 5).
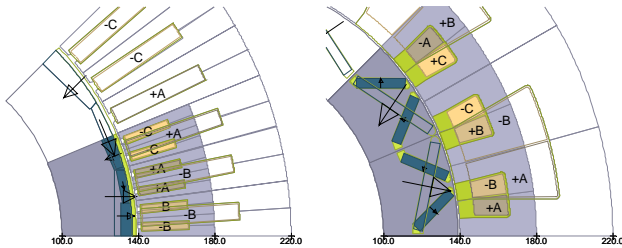
Figure 5 Model variations from small Δr and high Np (filled) to large Δr and low Np (unfilled contours).

Intentionally the rotor diameter is kept constant so that EM becomes magnetically limited by an available slot area to produce torque and a size of stator core to carry magnetic flux. Therefore, the machines that would be naturally disregarded due to magnetic saturation would be still choice for high saturation core materials.

3.2 Material selection

Concerning to the space and speed requirement the fundamental magnetization frequency becomes high. This is due to fact that the machine has large inner diameter and in order to keep the outer diameter and the weight reasonably low, also avoiding the stator core going into deep saturation, the magnetic pole pitch becomes short, number of poles high and the magnetization frequency high. If considering the nominal operation speed at *7500 rpm* and the range of number of poles from 8 to 16 then the electric frequency is changed from *0.5 kHz* to *1 kHz*, which means that high efficiency magnetic core material needs to be chosen for the stator core.

The list of material grades, which are used in design evaluation models, is shown in Table 1. The first three in the list are cobalt-iron *FeCo* materials with higher magnetic saturation capability. The rest of materials are silicon-iron *FeSi* materials, where two of them are more suitable for high frequency application due to thinner lamination thickness [43][44][45][46]. Since the coating thickness of the electromagnetic steel is unchanged the fill factor is changed.

Table 1 Properties of magnetic core materials

Grade	Density [g/cm^3]	Thickness [mm]	Core loss @1kHz 1.5T [W/kg]
Vacoflux 48	8.12	0.1	38.0
Vacodyr 49	8.12	0.15	100
Vacodur X1	7.9	0.2	146
Arnon-5	7.65	0.127	78.7
HI-LITE NO10	7.65	0.1	83.7
M250-35A	7.6	0.35	191

Six different core materials are compared using machine topology 1 in Figure 3. This specific topology has Halbach array magnet arrangement in the rotor and retaining sleeve. The rotor core is defined as laminated material in order to reduce rotor losses and not solid rotor that would be more suitable for high speed operation. The stator has two rectangular slots per pole and phase (Figure 4).

The electromagnetic torque capability as a function of number of poles and stator size is shown in Figure 6. The current density is *10 A/mm^2* and 60% fill factor is considered of the slot that includes a slot liner, which is the main electric insulation between the winding and the stator core. The torque contours reveal that *FeCo* materials have higher torque capability as these materials are less limited of the core saturation
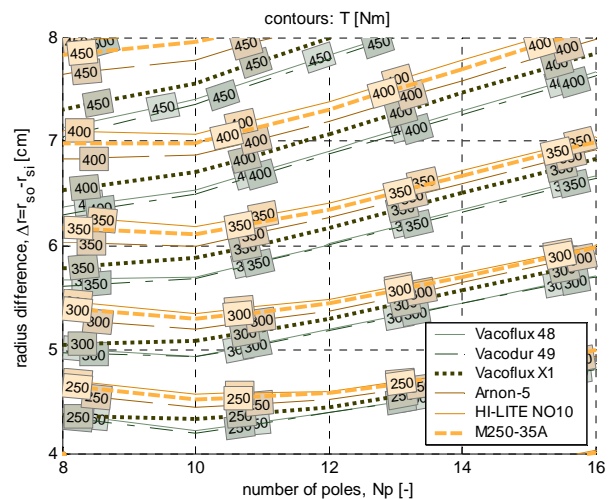

Figure 6 Torque as a function of stator size and number of poles

The high performance material that is capable for higher flux density may not outperform conventional materials as the weight increases [41]. In this comparison, EM designs at low number of poles also EM that have large radius benefit most of replacing the lighter saturated *FeSi* core to a weighty *FeCo* core. In this analysis the winding layout remains the same and also the stack length is not adjusted in respect to number of poles and stator size in order to simplify the multidimensional comparison.
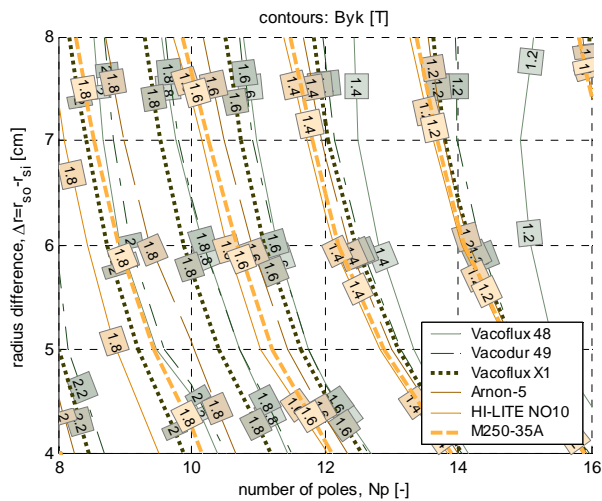
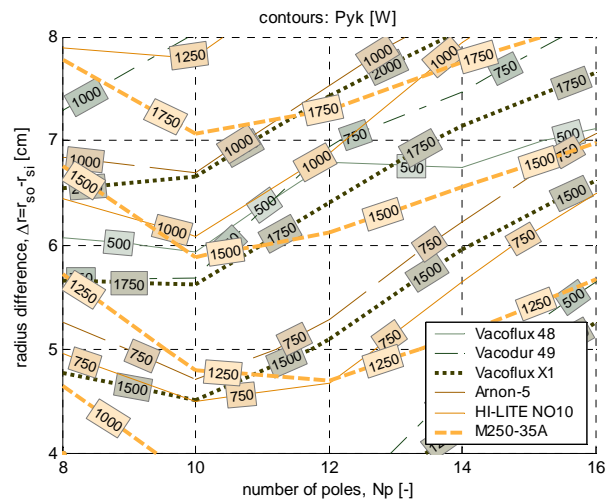*Figure 7 Flux density peak value in the stator yoke as a function of stator size and number of poles*

Figure 8 Flux density peak value in the stator teeth as a function of stator size and number of poles

Figure 9 Power losses in the windings excluding AC losses

Figure 10 Core losses in the stator yoke as a function of stator size and number of poles

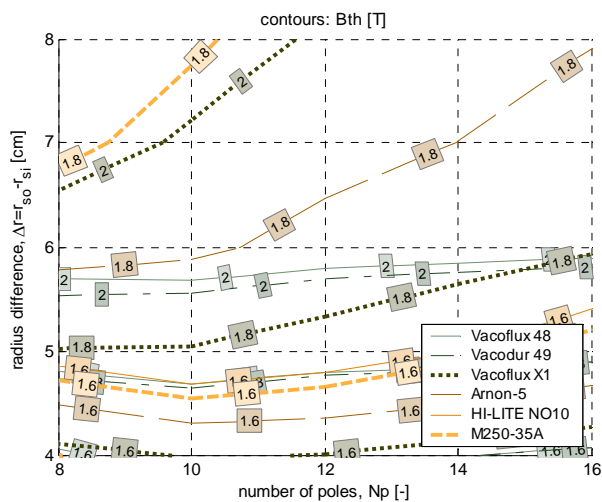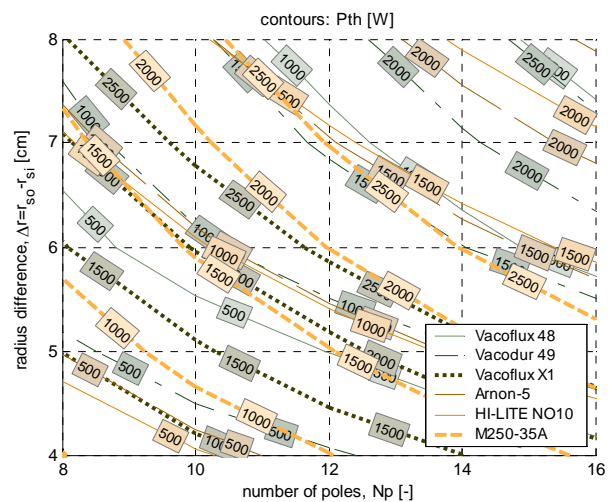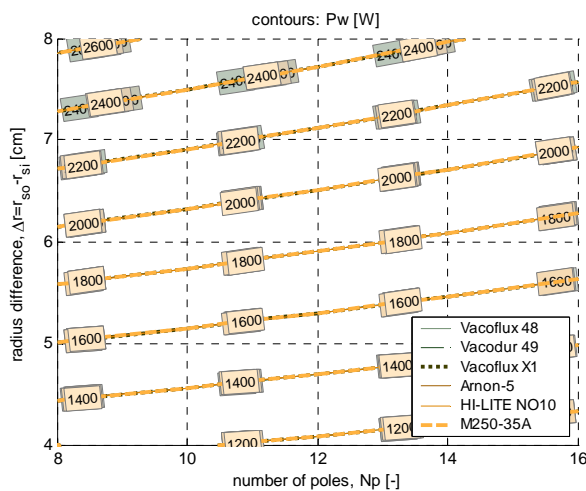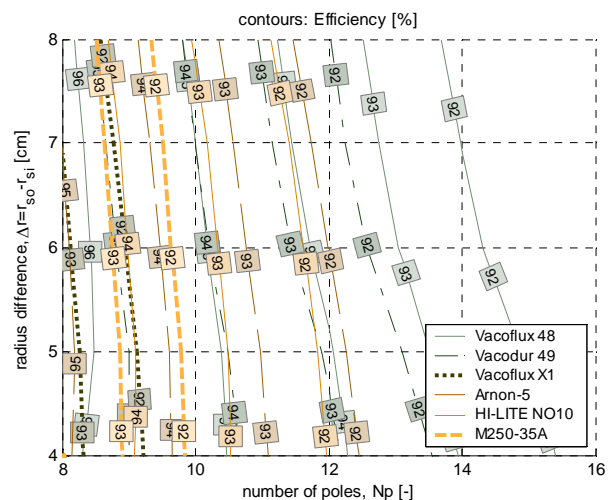*Figure 11 Core losses in the stator teeth as a function of stator size and number of poles*

Figure 12 Efficiency as a function of stator size and number of poles

The peak value of flux density of the loaded machine is shown Figure 7 and Figure 8, where a single representative node in the stator yoke and tooth is shown respectively. The estimated power losses in the stator winding (Figure 9) are in the same magnitude as in the yoke (Figure 10) and stator teeth (Figure 11). Stator teeth may have even higher losses due to high field derivative and harmonic content that is excluded from the power loss estimation as well as the AC losses in the winding. All these figures highlight conventional *0.35 mm* stator lamination *M350-35A* and recently developed *FeCo* material *VacoFlux X1* with compromising material properties and price. As this unsophisticated core loss estimation does not account for stray field power losses the estimated efficiency at 7500 rpm can be even lower than it is shown in Figure 12. However, based to optimistic estimation of power losses and efficiency comparison, the conventional FeSi core material may have even preference over high performance material (Figure 12).

4 Investigations on design specifications

The design space comparison is shown for the PMSM type that has a high amount of PM material and a high winding factor that expectedly is able to generate most power for given size compared to other topologies [42]. However, the machine has distributed windings that are not categorised as faultless due to overlaps, even if the machine can provide constant power over wide CPSR it has larger SC and generates more heat in the end turns. Therefore, an IPMSM is investigated that would take advantage of direct cooled windings [47] and have modular concentrated type of winding in order to contribute to fault tolerant operation.

4.1 Cooling integration

Conventionally, the EM windings have highest heat losses and highest heat resistance from the conductor to coolant. Therefore, it is appealing to bring dielectric coolant into direct contact to the winding or take advantage of hollow or profiled conductors, but this would cause additional stress and erosion factor on the electric insulation system (EIS).
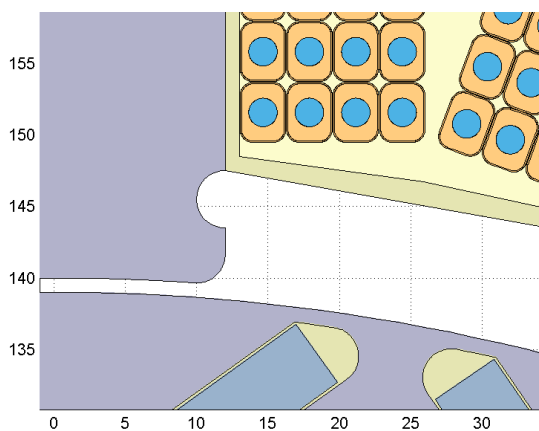

Figure 13 Detail view of IPMSM with direct cooled winding

Cooling integration, which is based on hollow conductors (Figure 13), is studied in [42] where distributed winding

with cooling jacket is compared to a concentrated winding machine where the individual coils are connected to corresponding PEC and cooling terminal. Apparently, the concentrated winding machine provides topology with more modular concept that it is easier to produce and assemble; however the serious drawback of this machine is the ripple. Since all the machine types suffer relatively high content of core losses the stator cooling jacket is needed to maintain the desirable core temperature. The direct winding cooling is desirable to provide superior thermal management and higher power capability.

4.2 Sensitivity study on selected design

Four different machine topologies are investigated in [42] where CPSR capability is achieved by larger air-gap, which is also needed for a retaining sleeve in the rotor (Figure 3 1) or rotor magnetic saliency (Figure 3 3). The concentrated winding machine (Figure 3 3) have narrower air-gap that contributes to higher efficiency but also considerably higher torque ripple. The electromagnetic performance is challenged by mechanical consideration where initial EM design (D00, Figure 14) is changed either thicker rotor bridges for increased mechanical strength (D01, Figure 14) or larger air-gap (D10, Figure 14), and in both (D11, Figure 14). These changes reduce flux density (Figure 14) and magnetic coupling, as it is seen from the change of .torque capability (Figure 15) and back emf (Figure 16). Even if these changes can reduce torque ripple, total harmonic distortion (in emf Figure 16) and SC current (Figure 17), the power capability of the machine is reduced (Figure 18).
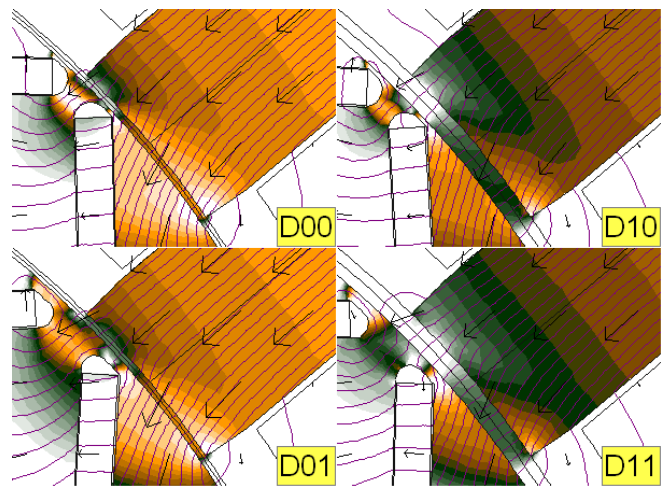

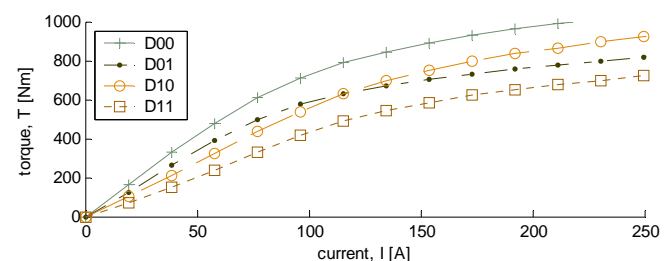*Figure 14 Detailed view of four different machine layouts*

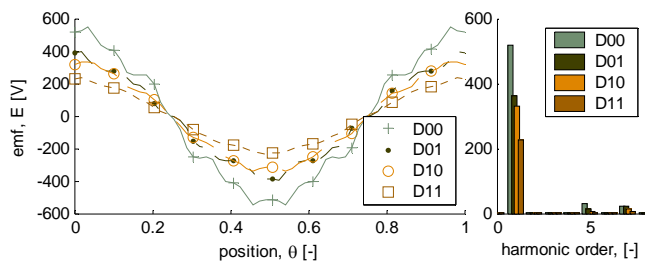*Figure 15 Electromagnetic torque as a function of current*

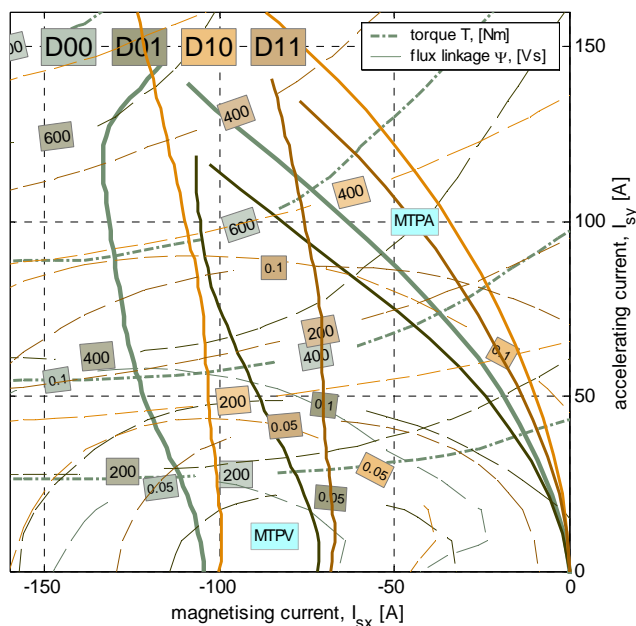*Figure 16 Back EMF waveform and spectrum at 7500 rpm*

Figure 17 Circle diagram – torque and flux capability as a function of vector current
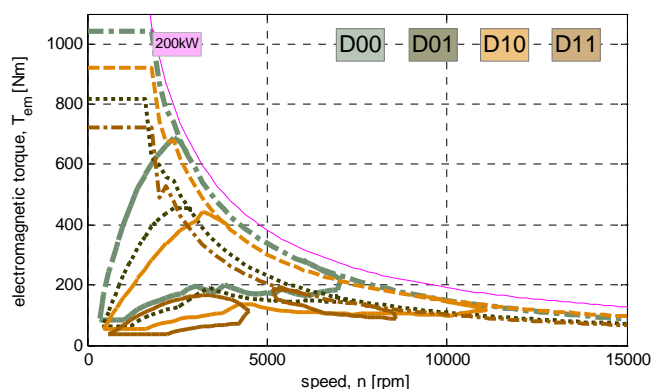

Figure 18 Torque speed characteristics including high-efficiency region above 92%

The machine performance (Figure 18) is estimated based to maximum torque per Ampere (MTPA) and maximum torque per Voltage (MTPV) lines (Figure 17) specified for *28 turn* coil at *Udc=300V*, which can be slightly improved by optimizing number of turns to practical dc-link voltage in respect to overvoltage and SC current. The increase in the air-gap reduces not only the magnetization, torque and power but also core losses and moves slightly the high

efficiency region towards higher rotation speed. However, the efficiency is relatively low at high speed operation of interest which brings to conclusion that the embedded machine design becomes unnecessarily large and inefficient for high speed operation compared to small geared machine. Nevertheless, this cannot be taken as general conclusion of this limited study provided in this work.

Conclusions

Electrifying propulsion targets higher system efficiency, weight reduction and all-in-all reduces fuel consumption. Increasing power density for weight reduction means that technology limits need to be pushed further. The ambition of pushing EM design to larger size, higher speed and increased extracted power, results higher structural stress in the rotor due to high rotational speed and in the stator EIS due to electric and thermal loads. More electrical engine and the radical change of replacing everything that mechanically wears out can include also electric wear due to multi-stressed loads and not just a list of improvements. The EM design based study on specifying EM for LPS 1) caused concern on machine size due to large inner diameter in respect to speed range and magnetizing frequency that causes unnecessary high core losses and low efficiency at desirable operation range which cannot be replaces simply by more efficient materials or technologies, 2) performance and power density need to be compromised to the improvement in faultless design: for high speed rotor and matched specification for EM winding layout, EIS and PEC topology, 3) cooling integration may provide solutions but that with new challenges.

Acknowledgement

References

[1]   R. R. Secunde, R. P. Macosko, D. S. Repas, "Integrated engine-generator concept for aircraft electric secondary power", Nasa technical memorandum x-2579, 1972.

[2]   M. J. Provost, "The More Electric Aero-engine: a general overview from an engine manufacturer", 2002 International Conference on Power Electronics, Machines and Drives (Conf. Publ. No. 487), June 2002, pp.246-251

[3]   J. A. Weimer, "The role of electric machines and drives in the more electric aircraft", IEEE International Electric Machines and Drives Conference, 2003. IEMDC'03., June 2003, pp.11-15 vol.1

[4]   C. R. Avery, S. G. Burrow, P. H. Mellor, "Electrical generation and distribution for the more electric

aircraft", 2007 42nd International Universities Power Engineering Conference, Sep. 2007, pp.1007-1012

[5]   B. Sarlioglu, C. T. Morris, "More Electric Aircraft: Review, Challenges, and Opportunities for Commercial Transport Aircraft", IEEE Transactions on Transportation Electrification, V 1 n 1, June 2015, pp.54-64

[6]   V. Madonna, P. Giangrande, M. Galea, "Electrical Power Generation in Aircraft: Review, Challenges, and Opportunities", IEEE Transactions on Transportation Electrification, V 4 n 3, Sep. 2018, pp.646-659

[7]   M. Henke, G.Narjes, J. Hoffmann, C. Wohlers, S. Urbanek, C. Heister, J. Steinbrink, W.-R. Canders, B. Ponick, "Challenges and Opportunities of Very Light High-Performance Electric Drives for Aviation", Energies 2018, 11, 344

[8]   Pratt&Whitney GTF-Engine https://www.pw.utc.com /products-and-services/products/commercial-engines/pratt-and-whitney-gtf

[9]   B. Rahrovi, M. Ehsani, "A Review of the More Electric Aircraft Power Electronics", 2019 IEEE Texas Power and Energy Conference (TPEC), Feb 2019, pp.1-6

[10]   G. L. Calzo, P. Zanchetta, C. Gerada, A. Gaeta, F. Crescimbini, "Converter topologies comparison for more electric aircrafts high speed Starter/Generator application", 2015 IEEE Energy Conversion Congress and Exposition (ECCE), Sep. 2015, pp.3659-3666

[11]   G. Buticchi, S. Bozhko, M. Liserre, P. Wheeler, K. Al-Haddad, "On-Board Microgrids for the More Electric Aircraft—Technology Review", IEEE Transactions on Industrial Electronics, V 66 n 7, July 2019, pp.5588-5599

[12]   V. Yaramasu, B. Wu, P. C. Sen, S. Kouro, M. Narimani, "High-power wind energy conversion systems: State-of-the-art and emerging technologies", Proceedings of the IEEE, V 103 n 5, May 2015, pp.740-788

[13]   D. Kumar, F. Zare, A. Ghosh, "DC Microgrid Technology: System Architectures, AC Grid Interfaces, Grounding Schemes, Power Quality, Communication Networks, Applications, and Standardizations Aspects", IEEE Access, V 5 n , 2017, pp.12230-12256

[14]   C. Gammeter. F. Krismer, J. W. Kolar, "Weight and efficiency analysis of switched circuit topologies for modular power electronics in MEA", IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Oct 2016, pp.3640-3647

[15]   W. Cao, B. C. Mecrow, G. J. Atkinson, J. W. Bennett, D. J. Atkinson, "Overview of Electric Motor Technologies Used for More Electric Aircraft (MEA)", IEEE Transactions on Industrial Electronics, V 59 n 9, Sep. 2012, pp.3523-3531

[16]   Z. Zhang, J. Huang, Y. Jiang, W. Geng, Y. Xu, "Overview and analysis of PM starter/generator for aircraft electrical power systems", CES Transactions on Electrical Machines and Systems, V 1 n 2, 2017, pp.117-131

[17]   R. Bojoi, A. Cavagnino, A. Tenconi, A. Tessarolo, S. Vaschetto, "Multiphase electrical machines and drives in the transportation electrification", 2015 IEEE 1st International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), Sep. 2015, pp.205-212

[18]   P. Arumugam, D. Barater, T. Hamiti, C. Gerada, "Winding concepts for ultra reliable electrical machines", IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society, Oct 2014, pp.959-964

[19]   D. Barater, F. Immovilli, A. Soldati, G. Buticchi and G. Franceschini, C. Gerada, M. Galea, "Multistress Characterization of Fault Mechanisms in Aerospace Electric Actuators", IEEE Transactions on Industry Applications, V 53 n 2, March 2017, pp.1106-1115

[20]   K. Rajashekara, "Parallel between More Electric Aircraft and Electric-Hybrid Vehicle Power Conversion Technologies", IEEE Electrification Magazine, V 2 n 2, June 2014, pp.50-60

[21]   M. van der Geest, H. Polinder, J. A. Ferreira, D. Zeilstra, "Machine selection and initial design of an aerospace starter/generator", 2013 International Electric Machines Drives Conference, May 2013, pp.196-203

[22]   D. Gerada, A. Mebarki, N. L. Brown, C. Gerada, A. Cavagnino, A. Boglietti, "High-Speed Electrical Machines: Technologies, Trends, and Developments", IEEE Transactions on Industrial Electronics, V 61 n 6, June 2014, pp.2946-2959

[23]   Ma Xiaohe, Su Rong, Tseng King Jet, Wang Shuai, Zhang Xiaolong, V. Vaiyapuri, G. Chandana, G. Amit, N. Sivakumar, "Review of high speed electrical machines in gas turbine electrical power generation", TENCON 2015 - 2015 IEEE Region 10 Conference, Nov 2015, pp.1-9

[24]   H. Lahne, D. Gerling, "Comparison of state-of-the-art high-speed high-power machines: Research study including a design example of a 50000 rpm induction machine", IECON 2015 - 41st Annual Conference of

the IEEE Industrial Electronics Society, Nov 2015, pp.003519-003524

[25] M. van der Geest, H. Polinder, J. A. Ferreira, M. Christmann, "Power Density Limits and Design Trends of High-Speed Permanent Magnet Synchronous Machines", IEEE Transactions on Transportation Electrification, V 1 n 3, Oct 2015, pp.266-276

[26] A. Binder, W.-R. Canders, "Tutorial High Speed Drives", Int. Conf. on Electrical Machines, 2014

[27] R. Benlamine, T. Hamiti, F. Vangraefschepe. D. Lhotellier, "Electromagnetic, Structural and Thermal Analyses of High-Speed PM Machines for Aircraft Application", 2018 XIII International Conference on Electrical Machines (ICEM), Sep. 2018, pp.212-217

[28] A. Yu, G. W. Jewell, "Systematic design study into the influence of rotational speed on the torque density of surface-mounted permanent magnet machines", The Journal of Engineering, V 2019 n 17, 2019, pp.4595-4600

[29] E. Richter, C. Ferreira, "Performance evaluation of a 250 kW switched reluctance starter generator", Proc. Conf. Rec. IEEE 13th IAS Annu. Meeting Ind. Appl. Conf. (IAS), vol. 1, pp. 434-440, Oct. 1995.

[30] A. V. Radun and C. A. Ferreira and E. Richter, "Two-channel switched reluctance starter/generator results", IEEE Transactions on Industry Applications, V 34 n 5, Sep. 1998, pp.1026-1034

[31] A. J. Mitcham, N. Grum, "An integrated LP shaft generator for the more electric aircraft", IEE Colloquium on All Electric Aircraft (Digest No. 1998/260), June 1998, pp.8/1-8/9

[32] Z. Sun, J. Ede, J. Wang, G. Jewell, J. Cullen, A. Mitcham, "Experimental testing of a 250-kW fault-tolerant permanent magnet power generation system for large civil aero engines", Proc. 5th Int. Conf. Energy Conversion Engineering (IECEC), pp. 1-8, 2007-June.

[33] P. Arumugam, Z. Xu, A. La Rocca, G. Vakil, M. Dickinson, E. Amankwah, T. Hamiti, S. Bozhko, C. Gerada, S. J. Pickering, "High-Speed Solid Rotor Permanent Magnet Machines: Concept and Design", IEEE Transactions on Transportation Electrification, V 2 n 3, Sep. 2016, pp.391-400

[34] A. L. Rocca, Z. Xu, P. Arumugam, S. J. Pickering, C. N. Eastwick, C. Gerada, S. Bozhko, "Thermal management of a high speed permanent magnet machine for an aeroengine", 2016 XXII International Conference on Electrical Machines (ICEM), Sep. 2016, pp.2732-2737

[35] A. Diab, M. Rashed, J. Li, C. Gerada, S. Bozhko, "Performance Analysis of PMSM for High-Speed Starter-Generator System", 2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles International Transportation Electrification Conference (ESARS-ITEC), Nov 2018, pp.1-7

[36] F. R. Ismagilov, V. E. Vavilov, D. V. Gusakov, "High-Speed Starter-Generator for Aerospace Applications: Design and Initial Testing", 2018 XIII International Conference on Electrical Machines (ICEM), Sep. 2018, pp.2593-2599

[37] F. R. Ismagilov, V. E. Vavilov, D. V. Gusakov, "Design Features of Liquid-Cooled Aviation Starter Generators", 2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles International Transportation Electrification Conference (ESARS-ITEC), Nov 2018, pp.1-5

[38] A. Balachandran, M. Boden, Z. Sun, S. J. Forrest, J. D. Ede, G. W. Jewell, "Design, construction, and testing of an aero-engine starter-generator for the more-electric aircraft", The Journal of Engineering, V 2019 n 17, 2019, pp.3474-3478

[39] A. Cavagnino, Z. Li, A. Tenconi, S. Vaschetto, "Integrated Generator for More Electric Engine: Design and Testing of a Scaled-Size Prototype", IEEE Transactions on Industry Applications, V 49 n 5, Sept 2013, pp.2034-2043

[40] R. Bojoi, A. Cavagnino, A. Tenconi, S. Vaschetto, "Multiphase PM machine for More Electric Aircraft applications: Prototype for design validation", IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, Oct 2012, pp.3628-3634

[41] N. Fernando, G. Vakil, P. Arumugam, E. Amankwah, C. Gerada, S. Bozhko, "Impact of Soft Magnetic Material on Design of High-Speed Permanent-Magnet Machines", IEEE Transactions on Industrial Electronics, V 64 n 3, March 2017, pp.2415-2423

[42] A. Reinap, "Aerospace electric generator design considerations" Electric drives production conference, EDPC 2019

[43] "Soft magnetic cobalt-iron alloys", Vacuumschmelze (available online)

[44] "Lamination Stacks made of VACOFLUX X1", Vacuumschmelze (available online)

[45] "Energy Savings with Thin Gauge for Motors, Generators, Transformers and Inductors", Arnold Magnetic Technologies, (available online)

[46] "Arnon thin and ultra-thin silicon iron datasheet", Arnold Magnetic Technologies, (available online)

[47] A. Reinap, M. Andersson, F.J. Márquez-Fernández, P. Abrahamsson, M.Alaküla, "Performance Estimation of a Traction Machine with Direct Cooled Hairpin Winding". IEEE Transportation Electrification Conference and Expo (ITEC), Novi, MI, USA, 19-21 June 2019. pp. 1-6.

Understanding Loudness variations for aircraft Landing procedures at Arlanda (ULLA)

Anders Johansson and Karl Bolin

KTH/MWL, Center for Sustainable Aviation Stockholm, Sweden
E-mail: aebjo@kth.se, kbolin@kth.se

Abstract

As of today air traffic is increasing at Arlanda airport, resulting in higher noise levels. This project has been launched to investigate the possibilities of optimizing flight procedures for a quieter approach. A network of acoustic measurement stations has been developed to measure single aircraft pass-by noise at ground level. The pass-by noise levels have been evaluated together with flight trajectory's and the current meteorological conditions in order to investigate how operational procedures impact the noise distributions on ground. This paper presents the result of from a snap-shot of measurements that show variations in generated noise levels for similar approach routes and weather conditions. Difference as large as 10 dB was found between, pass-by's performed by the same individual aircraft. The reduced velocity for the quieter flight, stands out as the major factor to reduce noise levels.

Keywords: Acoustical measurements, Aircraft noise

1 Introduction

Over the last few decades there has been a global increase in air traffic and Arlanda is no exception to the trend, rather the opposite, with far-reaching plans to expand and build additional runways to accommodate the increase in air traffic. This new situation urges the importance of investigating and evaluating the noise issues surrounding Arlanda in order to maintain a sound relationship with its nearby residents.

An epidemiological survey of 1240 residents around Arlanda carried out by KTH / MWL 2009-2010 in the EU project COSMA [1] provides that those who live near the approach routes are more disturbed than those who live near the departure routes. Of interest is therefor to perform long-term measurements around Arlanda to investigate how the sound sources vary during the approach and if this can be a basis for sound optimized approach procedures?

Aircraft noise are generated by both the engines and the aircraft body. The approximate relationship between engine and airframe noise at landing is about 50-50% with a slight overweight for sound generated by the aircraft body which will be even more accentuated with the ongoing development of more quiet engines. This means that noise from other components of the aircraft has increased in importance at landing [2]. It is therefore considered a priority to measure

sound for different landing procedures and meteorological conditions to analyze if there are differences in noise distributions related to flight operational variables.

Dose-response relationships between noise level and perceived annoyance for nearby residents are well established [3], but behind these connections it appears to be a significant variation in annoyance at the same noise level. Thus, annoyance from aircraft cannot only be explained by the noise levels the residents experience. The connection between sound and noise have to be explained by more variables and the sound quality is probably part of the explanation. Thus, there is a need to evaluate not only the sound level from different flight operations, but also how other sound aspects i.e. the sound quality is affected.

Considering the above aspects of aircraft noise and its components, sound measurements are of central importance for obtaining a detailed view of the noise situation around Arlanda approach routes. In order for this analysis to be possible, coordinated measurements are needed for several positions and preferably over a long period of time to ensure a solid statistical basis. Measurements of this kind generally involve high costs due to expensive measuring equipment. However, the development of *Single Board Computers* (SBC) together with cheap USB-interface microphones have enabled the composition of measurement equipment to a

considerable reduced cost. In this project, which is part of the *Centre for Sustainable Aviation at KTH* (CSA), five autonomous measurement stations have been built to form a network of measurement stations. The network covers an area equal to a circle of 2 km radius and is deployed to cover part of the approach to runway 26 to capture time and spatial variations in the sound field. Together with meteorological data and Secondary Surveillance Radar (SSR) the measurements are analyzed to investigate how sound field variations are related to the flight operational procedures.

2 Method

To enable the analysis, three sets of data is needed: the Aircraft trajectory, the current meteorological condition and the noise level on ground. How these are gathered is described below.

2.1 Aircraft trajectory

The aircraft trajectory is downloaded *live* from the online flight surveillance site at OpenSky Network [4]. The main technologies behind the OpenSky Network are the *Automatic Dependent Surveillance-Broadcast* (ADS-B) and Mode S which is a modern method for SSR. With this technology the aircraft determines its position via GPS and then periodically broadcasts it together with other aircraft information. The main difference to older systems of SSR is that no interrogation signal is needed and that the information is broadcast over the 1090 MHz radio frequency channel, making it publicly accessible. The OpenSky Network compiles ADS-B data into a set of variables which is updated in approximate real time every 5 seconds. The most important of these variables is data for position, altitude, speed and the ICAO24 address, which enables the identification of the aircraft. While the complete trajectory and movement of the aircraft is given by the ABS-B data there are no directly available variables that give information about the flight configuration i.e. thrust and flap settings, extension of landing gear, weight of aircraft etc. To directly collect such data access to the Flight Data Recorder (FDR) would be needed. There are however methods to estimate some of these parameters from the ADS-B data, for example the thrust setting, though this procedure is not covered in this paper.

2.2 Meteorological data

The meteorological data used in this project is provided from the numerical forecasting model AROME [5]. The model describes weather developments in time and space, in a grid of 2.5 x 2.5 km squares where the atmosphere has been divided into 65 levels in the vertical direction. The forecast is based on what the weather looks like at a certain time, the so-called initial state. Every third hour the forecast is updated with an analysis of the current weather situation. In this way, the forecast is modified to better match current observations. The model has been developed within an international collaboration involving a large number of European countries. The AROME analysis is performed over an area that covers Scandinavia, the northern part of the continent and the Baltic States.

2.3.1 Acoustic measurement stations

Acoustic measurements of aircraft pass-by's are warranted for an extended period of time, preferably covering the seasonal changes of the year and both day and night. To meet this demand, several autonomous measurement stations has been built, see fig.1.

Figure 1: Photo of the acoustic measurement station. Depicted is the solar cell and the microphone with fitted windscreen

The core of the station is a SBC that processes, stores and uploads measurement data. Connected to the computer, through a sound card, is a microphone of type electret-condenser (the performance of the microphone is described in section 2.3.2 below). The microphone output signal is sampled at 48kHz, with a bit-depth of 24 and the signal is filtered through 1/3-octave passbands, covering center frequencies from 25Hz to 20kHz. The slope and characteristics of the band-pass filters meets IEC 61672-1 standards for class-1 acoustical measurement equipment. From the output of each 1/3-octave band, A-weighted sound pressure levels are calculated and updated at 8Hz intervals equivalent to the time-weighting "Fast". In addition to the 1/3-octave band sound pressure levels, audio recordings are also performed to enable narrowband analysis and playback of single aircraft pass-by. The measurement station is powered by a solar panel and through GSM-connectivity measurement-data is uploaded to a central server for further compilation and analysis.

2.3.2 Microphone

The microphone used in the measurement station is a 6 mm electret-condenser microphone. In this particular microphone the preamplifier, soundcard and microphone capsule is built into one unit which connects directly to the computer via a USB-cable. Electret-condenser microphones are known for their robustness, linear response and cheap production costs. The noise-floor of the microphone and soundcard is estimated to be below 30 dB(A) and the dynamic range is about 80 dB.
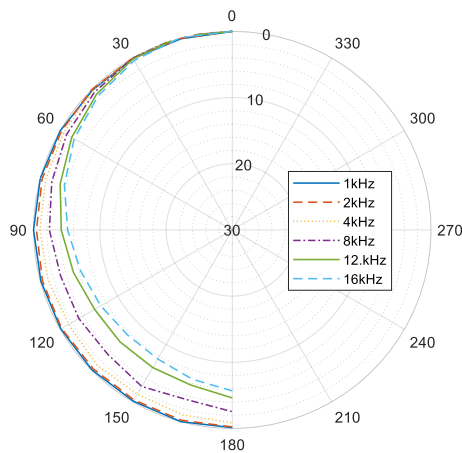
Figure 2: Polar plot of the microphones directivity for the 1/3-octave bands and center frequencies 1, 2, 4, 8, 12.5 and 16kHz

In fig. 2, the directivity of the microphone in the 1/3-octave bands and center frequencies 1, 2, 4, 8, 12.5 and 16kHz is shown. It can be seen that the response does not deviate more than 10dB for any angle compared to the on-axis response. The directivity pattern of the microphone therefore meets the requirements according to the IEC 61672-1 standard for electro acoustic sound level meters.   When acoustic measurements are made over long periods of time it is important that the performance of the microphone remains stable and that does not drift in sensitivity.
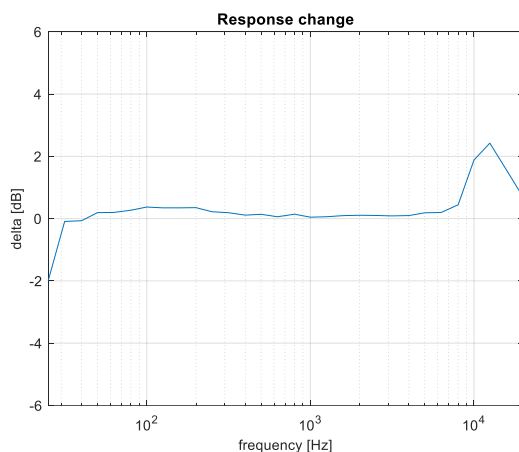

Figure 3: Change in frequency response of the microphone after 9 months of use in the field.

Figure 3. show the change in frequency response for a microphone used in the field for approximately 9 months. During this period, the weather changed from winter to summer, causing the microphone to be exposed to high temperatures and sub-zero degrees. However, as can be seen in figure 3 the response of the microphone remained essentially unchanged indicating that measurements taken over long periods produce reliable results. To mitigate wind-induced noise the microphone is equipped with a windscreen, see fig. 1. The windscreen, which is made of a foam core with a fur cover, also has a function as protection for rain and

moisture. When fitting the windscreen to the microphone higher frequencies become attenuated. However, the characteristics of aircraft noise is only relevant up to 5 kHz, due to the air absorption of the relatively large measurement distances. Above 5 kHz the sound absorption of the windscreen is insignificant. Rain and snow accumulating on the windscreen will likely change the acoustic properties and change the microphone's frequency response. As weather data is available, it is however possible to investigate statistically how the amount of precipitation affects the measurement data and then possibly omit measurements made under unfavorable weather conditions.

Several properties of the measurement equipment meet IEC 61672-1 class-1 standard, though the complete chain of components should not be considered a class-1 product. However, the authors consider the equipment to fulfill the purpose of survey measurements and to be adequate for the task of noise levels comparisons for different flights.

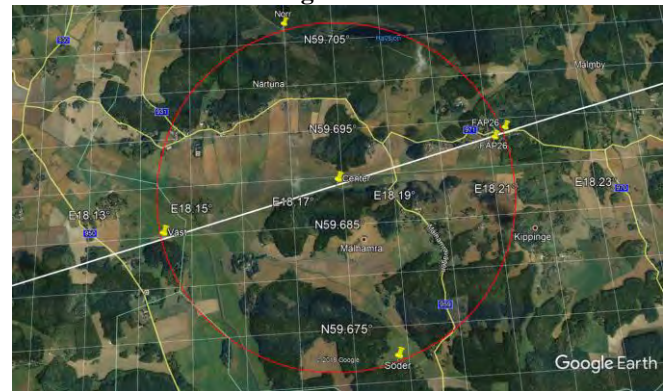### 2.3.3 Measurement arrangement

Figure 4: Positions of the measurement stations. The red circle has a radius of 2 km. The white line shows the approach path to runway 26R.

At an initial phase of the measurement campaign 5 measurement stations where placed beneath the final approach path to runway 26R to cover the sound propagation in lateral and longitudinal directions, see fig. 4. The most east bound station, is positioned beneath the *Final Approach Point* (FAP), where every aircraft approaching runway 26 R have to pass to enter the 3-degree glide-path down to the runway. The center and west station are positioned further down the glide path and spaced 2 km apart. The north and south station are positioned perpendicular to the glide path and 2 km out from the center station. The spacing and location of the measurement stations was chosen because it is at this distance and altitude the pilots usually deploy landing gear and flaps, which both are two major sources of air frame noise. The microphones on the measurement stations are positioned 1.2 meters above ground.

3 Results and Discussion

In the following section a snap-shot of noise measurement data will be presented that highlight how different pass-by noise levels can be generated by the same individual aircraft and for a seemingly similar approach route. The

measurements where made for approaching aircrafts to runway 26R at Arlanda airport during the period May to June-2019 and are presented in pairs where each pair show two pass-by events of the same aircraft. The events of the individual aircraft occurred during the same day and sometimes separated only by a few hours. In figures 5-10, in addition to noise levels recorded by the five measurement stations, ADB-S data for the geographical-altitude and velocity of the aircraft is also shown. The altitude is given in meters while the velocity is given as *meters per second* (m/s). (Perhaps an unorthodox quantity in the context of airplanes, but it facilitated the numbering properties of the figure axis. 1 m/s = 1.94384449 knots). All three aircrafts belong to the SAS fleet and are of the type: Boeing 737, Airbus A321 and A320.
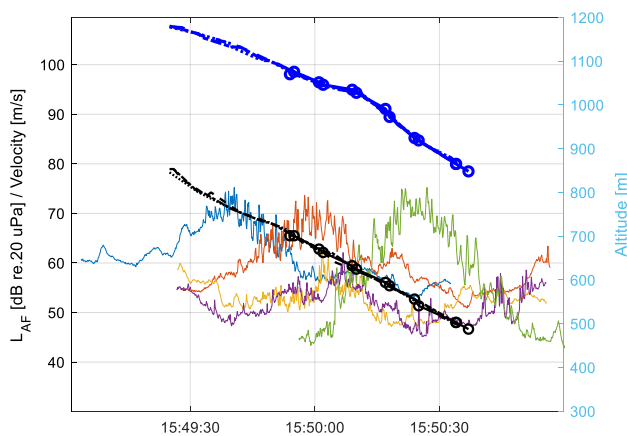
3.1 Airbus A320

Figure 1: First pass-by noise of the Airbus A320 at 15:49 flying over the measurement area. The blur thin right-most line shows the L_{AF} measured at the FAP-station. The red line is the Center-station and the green line is the West-station. Blue thick-dotted line shows the airplanes velocity in m/s while the black thick-dotted line shows the altitude in meters.

Figure 5 show noise level of the first pass-by of the Airbus A320 occurring at 15:49. The wind conditions at this time of the day was 4 m/s, heading 200 degrees. We can clearly see the noise level build up and the following level decrease as the aircraft pass over the measurement positions, creating characteristic noise contours resembling those of small hills. The most pronounced contours are of course seen for the stations directly under the flight path, were the distance: source-receiver is the smallest, i.e. at the FAP, Center and West station. For this particular pass-by, the noise levels reach about 70 dB(A), which would sort into the category: "high noise level pass-by". One could expect that the noise level on ground would increase as the aircraft drops in altitude and the source-receiver distance shrinks, in this case by over 200 meters from the FAP to the West station. However, no such effect can be observed as the contours for all three stations have a similar shape. The effect of noise-power distance is anyhow well observed for the lateral stations where the noise level only reaches a maximum of about 60 dB or lower. The blue thick dotted line shows the velocity and we can see that it is steady decreasing from 110 m/s in

the start of the pass-by to below 80 m/s in the final stretch, with an almost constant deacceleration. From the collection of pass-by measurements this kind of "high noise level pass-by" was rather common for the Airbus A320.
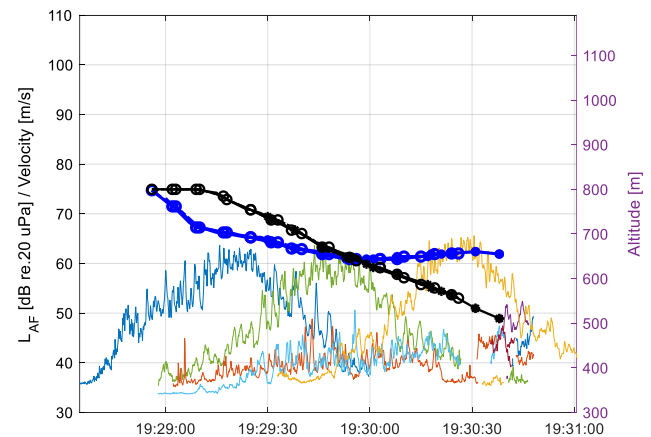

Figure 2: Second pass-by noise of the Airbus A320 at 19:29 flying over the measurement area. The blur thin right-most line shows the L_{AF} measured at the FAP-station. The green line is the Center-station and the yellow line is the West-station. Blue thick-dotted line shows the airplanes velocity in m/s while the black thick-dotted line shows the altitude in meters.

Figure 6 shows the second pass-by of the A320 at 19:29 later the same day. The wind was 4 m/s but the direction had slightly changed to 180 degrees heading. The noise levels generated by this pass-by is less common and would sort into the category: "low noise level pass-by". Compared to the previous result, there are some major differences. The noise levels are about 10 dB lower, which is a notable difference considering that it takes an eight times more powerful noise source to generate an increase of 9 dB. The other major difference is the reduced velocity for with which the aircraft travels. The velocity is also more constant, suggesting a smoother approach with less deaccelerating. The initial velocity is about 75 m/s which quickly drops to 65 m/s, before the aircraft reaches the FAP. The velocity then decreases slowly to 60 m/s and is almost constant for the reminder of the pass-by. Compared to the earlier pass-by the velocity is reduced by over one third for major parts of the glide path. The difference in noise levels and velocity profiles of these two flights, which were performed by the same aircraft, strongly suggest the obvious correlation between velocity and generated noise level.
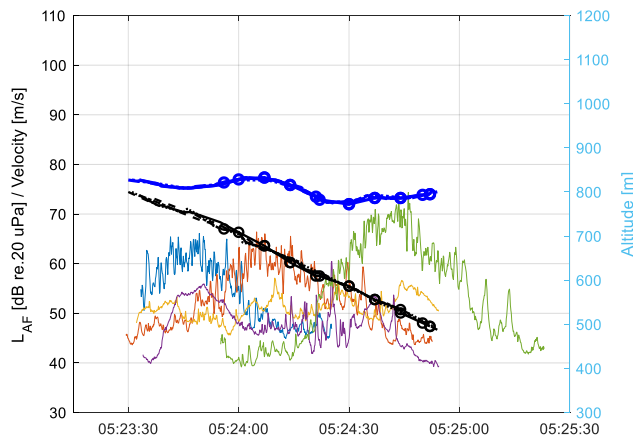
3.2 Boeing 737

Figure 3: First pass-by noise of the Boeing 737at 05:23 flying over the measurement area. The blur thin right-most line shows the L_{AF} measured at the FAP-station. The red line is the Center-station and the green line is the West-station. Blue thick-dotted line shows the airplanes velocity in m/s while the black thick-dotted line shows the altitude in meters.

Figure 5: First pass-by noise of the Airbus A321 at 16:07 flying over the measurement area. The blur thin right-most line shows the L_{AF} measured at the FAP-station. The green line is the West-station and blue thick-dotted line shows the airplanes velocity in m/s while the black thick-dotted line shows the altitude in meters.
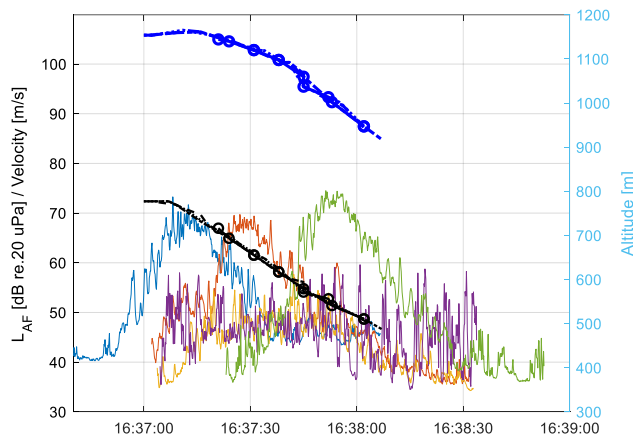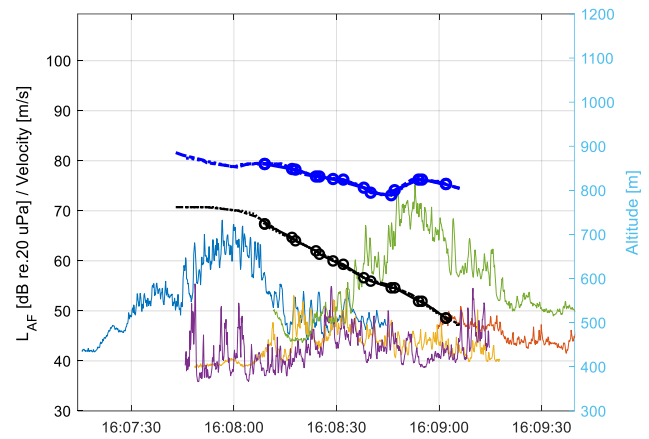

Figure 4: Second pass-by noise of the Boeing 737 at 16:37 flying over the measurement area. The blur thin right-most line shows the L_{AF} measured at the FAP-station. The red line is the Center-station and the green line is the West-station. Blue thick-dotted line shows the airplanes velocity in m/s while the black thick-dotted line shows the altitude in meters.

Figure 6: Second pass-by noise of the Airbus A321 at 19:55 flying over the measurement area. The blur thin right-most line shows the L_{AF} measured at the FAP-station. The yellow line is the West-station and blue thick-dotted line shows the airplanes velocity in m/s while the black thick-dotted line shows the altitude in meters.
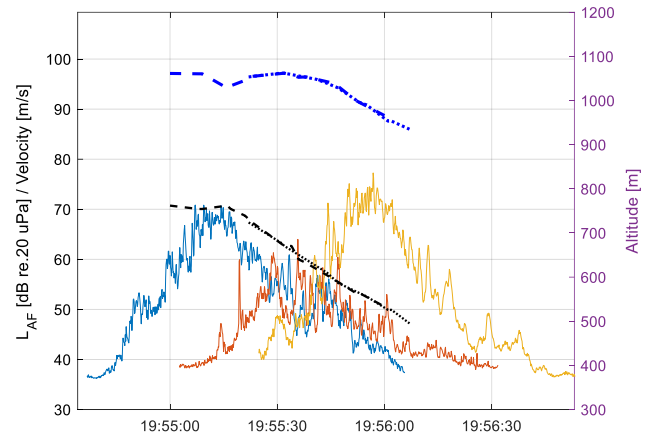
Figure 7 and 8 show a similar result for one Boeing 737, though not as drastic as for the A320. These two pass-by's occurred at 05:23 and 16:37. The wind this day was 5 m/s in the morning with direction 250 degrees and 3 m/s in the evening with direction 240 degrees. For these two pass-by measurements the noise level difference is about 5 dB. The velocity difference is also less but the acceleration for the low and high noise level pass-by's is similar to previous results. Note that West noise contour (the right-most curve) in fig.7-8, is higher than the other contours, showing the effect of noise-power distance.

3.3 Airbus A321

Figure 9 and 10 show a pair of additional measurements, this time for one Airbus A321 passing over the measurement area at 16:08 and 19:55. The wind condition for the earlier flight was 4 m/s with direction 300 degrees and 3 m/s for the later flight with direction 310 degrees. During this day the Center station was out of order leaving measurements only for the FAP and West station. Comparing the two measurements in fig. 9-10 we can see that it is about a 5 dB noise level difference between the flights and that it is the later flight that generates the highest noise level. The difference is mostly pronounced in the noise contour of the FAP station. A reason for the higher noise level at the West station could be that the landing gear or flaps are deployed. ADB-S data does not give any parameters for flaps or landing gear deployment. However, an indication of alternations in flight configuration could perhaps be seen in the flights velocity as sudden

DOI
10.3384/ecp19162010   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **97**

changes. Maybe this is what is seen in fig. 9 at 16:08:45? However, to verify such a statement ADB-S data would have to be run against FDR-data or other records of flight parameters, such as visual observations.

3.4 Further work

The above snap-shots in fig. 5-10 are taken from a batch of measurements containing about 100 passages of the A320, c. 40 of the Boeing 737 and 20 for the A321. No statistical calculations have been performed on the material and the estimates are to be considered as rough. The continued measurement analysis will contain a statistical analysis that also include a greater number of aircraft pass-by's. It is thus necessary to consider the above result only as preliminary observations. To minimize the number of error sources the comparison of pass-by noise levels have been restricted to flights of the same individual aircraft and to flights that occurred in close proximity. There are however, a number of parameters that can have changed during the time of the two passby's. One factor for instance, is the passenger load which greatly would influence the thrust-settings and directly alter the strength of the noise source. None the less, the large differences in generated noise levels does raise some questions. Is it the airframe or engine that dominates the noise? How does the duration of the pass-by influence the noise perception, is it better to have a slow and quiet pass-by than a fast and loud? These are questions that have to be answered in the continued work of this project in investigating the possibilities of optimizing flight procedures for a quieter approach

Conclusion

A system of acoustic measurement stations has been built and deployed in the field east of runway 26R at Arlanda airport to investigate how different aircraft operational procedures impact the noise distribution on ground. Together with measurements of the single aircraft pass-by noise at ground level ADB-S data of the flights trajectory have been recorded for a prolonged period of time. This paper presents measurements that compare the noise level on ground generated by the same individual aircraft at different occasions during the same day, for similar flight trajectories and meteorological conditions. The differences in generated noise levels was found to be as large as 10 dB. A reduced velocity for the less noisy flight, indicate that this is the main factor that govern noise generation. Velocity differences were as large as 90 knots. Although, the result of this paper is only preliminary and a first step in the analysis of the noise situation, the large discrepancies in noise emission shows possibilities to optimize for a quieter approach procedure.

References

[1] Community Oriented Solutions to Minimize Aircraft noise annoyance (COSMA), EU/FP7 Project reference: 234118, http://www.fp7-cosma.eu/

[2] Graham W. R., Hall C. A. och Vera Morales M., (2014) "The potential of future aircraft technology for noise and pollutant emissions reduction," Transport Policy, 34, 36 - 51

[3] Miedema, H. M. och Oudshoorn, C. G. (2001). Annoyance from transportation noise: relationships with exposure metrics DNL and DENL and their confidence intervals. Environmental Health Perspectives, 109(4), 409–416.

[4] Webpage: https://opensky-network.org, Last visited 2019-07-30

[5] Webpage: https://www.smhi.se/data/utforskaren-oppna-data/meteorologisk-modell-arome-prognosdata-realtid, Last visited 2019-07-30

III Simulation-based Analysis for HMI and Pilot Training

Multi-Agent Multi-Objective Deep Reinforcement Learning for Efficient and Effective Pilot Training
Johan Källström and Fredrik Heintz

Towards a Complete Co-Simulation Model Integration Including HMI Aspects
Ingo Staack, Jörg Schminder, Owais Shahid and Robert Braun

Initial Flight Simulation Testing of a Gesture-based Interface for Flight Control
Manuel Alejandro Rodriguez Diaz, Emilia Villani, Diego Arjoni and Petter Krus

HUMAER: A Test-Bed Environment for Human Factors Investigations in the Aeronautic Domain
Diego Hernandez Arjoni, Wesley Rodrigues de Oliveira, Emília Villani and Luís Gonzaga Trabasso

Multi-Agent Multi-Objective Deep Reinforcement Learning for Efficient and Effective Pilot Training

Johan Källström* and Fredrik Heintz**

*Saab AB and Department of Computer Science, Linköping University, Linköping, Sweden
**Department of Computer Science, Linköping University, Linköping, Sweden
E-mail: Johan.Kallstrom@liu.se, Fredrik.Heintz@liu.se

Abstract

The tactical systems and operational environment of modern fighter aircraft are becoming increasingly complex. Creating a realistic and relevant environment for pilot training using only live aircraft is difficult, impractical and highly expensive. The Live, Virtual and Constructive (LVC) simulation paradigm aims to address this challenge. LVC simulation means linking real aircraft, ground-based systems and soldiers (Live), manned simulators (Virtual) and computer controlled synthetic entities (Constructive). Constructive simulation enables realization of complex scenarios with a large number of autonomous friendly, hostile and neutral entities, which interact with each other as well as manned simulators and real systems. This reduces the need for personnel to act as role-players through operation of e.g. live or virtual aircraft, thus lowering the cost of training. Constructive simulation also makes it possible to improve the availability of training by embedding simulation capabilities in live aircraft, making it possible to train anywhere, anytime. In this paper we discuss how machine learning techniques can be used to automate the process of constructing advanced, adaptive behavior models for constructive simulations, to improve the autonomy of future training systems. We conduct a number of initial experiments, and show that reinforcement learning, in particular multi-agent and multi-objective deep reinforcement learning, allows synthetic pilots to learn to cooperate and prioritize among conflicting objectives in air combat scenarios. Though the results are promising, we conclude that further algorithm development is necessary to fully master the complex domain of air combat simulation.

Keywords: Pilot Training, Embedded Training, LVC Simulation, Artificial Intelligence, Autonomy, Sub-system and System Technology, Aircraft and Spacecraft System Analysis

1 Introduction

The tactical systems and operational environment of modern fighter aircraft are becoming increasingly complex. As a consequence, conducting training using only live, manned platforms is becoming increasingly difficult. Live training is related to high costs, and air space regulations as well safety restrictions place limitations on the type of training scenarios that can be realized. The logistics related to live training may also lead to poor availability of training. As the possibilities to do live training decrease, simulation-based training becomes more and more important.

In an ongoing project within the Swedish National Aeronautical Research Program 7 (NFFP7), we are studying how the next-generation pilot training systems should be designed to meet future training needs. In our research we are investigating how machine learning techniques can be used to construct advanced behavior models for synthetic, intelligent agents. The goal is to develop efficient methods to generate a wide range of intelligent, adaptive computer controlled allies and adversaries that can create realistic situations adapted for training of fighter pilots. We have identified two subfields of particular interest: Multi-Agent Reinforcement Learning (MARL) and Multi-Objective Reinforcement Learning (MORL) [1, 2]. MARL allows agents to learn how to achieve their goals in mixed cooperative and competitive multi-agent scenarios, such as an air combat scenario, while MORL allows agents to learn how to prioritize among multiple conflicting objectives, e.g. tactical mission goals, resource consumption and safety.

In this paper we discuss how these techniques can help address the challenges related to constructing high quality training simulations. We first give an overview of simulation-based training, and highlight aspects that motivate our work. We then present a proposed architecture for an intelligent, synthetic trainer, and machine learning techniques that could be used to implement it. Finally, we evaluate the approach in a number of experiments, with promising results, and give directions for future work.

2 Simulation-Based Training

Simulation-based training is an efficient way of training operators in complex, high-risk tasks. Scenarios with great variety can be realized at a low cost, without risk of injury. In our work, we model a typical training process for a simulation-based training system as illustrated in fig. 1. First, simulation contents are created to meet identified training needs. In the domain of air combat this could include vehicle models, behavior models for the synthetic operators of these vehicles, and definitions of the scenarios that they operate in. Then, in a training session, a briefing is conducted to present and discuss training objectives and scenario contents, followed by the actual execution of the scenario. Afterwards, trainee performance is evaluated in a debriefing. Over time, training needs are updated based on the learning progress of trainees, as well as input from the organization that they belong, e.g. due to changes in operational missions.
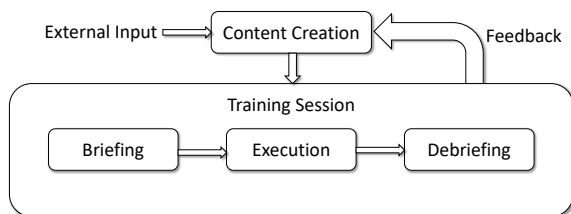

Figure 1: Training process.

User roles associated with the training process are illustrated in fig. 2. The Training Audience are those that we want to train, in the case of air combat training fighter pilots. The training environment is provided by a set of Training Providers. The Instructor is responsible for the pedagogical contents of a training session, and is supported by role-players and operators to provide it. Role-players participate in the training scenario, but they themselves do not receive training. Operators work behind the scenes of the training scenario, e.g. controlling simulation software, such as manual control of Computer Generated Forces (CGF), to make sure that the simulated scenario progresses in the right direction. In practice, one single person could act in several roles. For instance, due to limited resources, one person could act as instructor role-player and operator. This typically results in a high workload, and the desired training scenarios may not be achievable. It is desirable to reduce the need for training providers, to improve training efficiency as well as effectiveness.
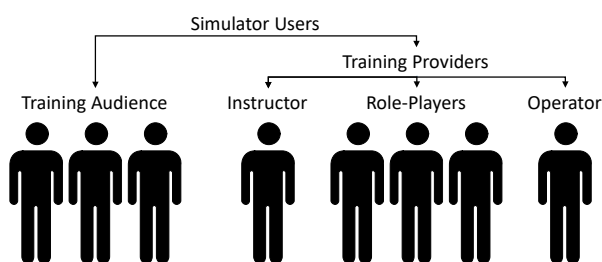

Figure 2: Users of simulation-based training systems.

2.1 Live, Virtual and Constructive Simulation

Computer simulations can be used to augment training in live systems. One approach is to replace some of the live training with training in simulators, e.g. using ground-based flight simulators instead of training in live aircraft. Another approach is to embed simulation capabilities in live systems, e.g. capabilities for generation of synthetic opponents in a fighter aircraft. It has been estimated that embedded training can improve training effectiveness of live training by 30% at the same cost [3]. In the Live, Virtual and Constructive simulation paradigm, the goal is to take things one step further, by seamlessly integrating live systems, manned simulators and computerized simulations in a distributed simulation. The three categories of simulations are defined as [4]:

- *Live*: Simulations involving real people operating real systems

- *Virtual*: Simulations involving real people operating simulated systems

- *Constructive*: Simulations involving simulated people operating simulated systems (possibly stimulated by real people)

In the domain of air combat simulation the goal is to integrate real aircraft, ground-based systems and soldiers (Live), manned simulators (Virtual) and computer controlled entities (Constructive). Such a simulation platform is valuable for training [5,6]. An example of an LVC simulation network for air combat training is illustrated in fig. 3.
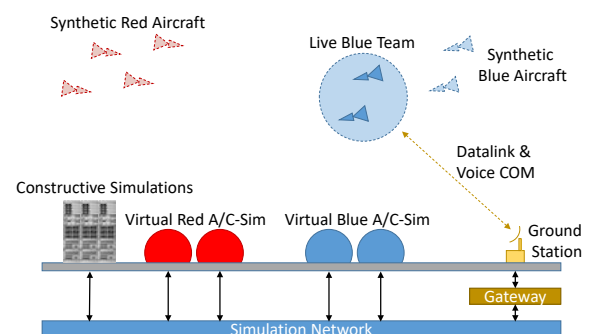

Figure 3: An LVC distributed simulation.

Constructive simulation enables realization of complex scenarios with a large number of autonomous friendly, hostile and neutral entities, which interact with each other as well as manned simulators and real systems. However, building realistic behavior models for CGF is a significant challenge [7–9], and consequently support from scenario operators, and possibly human role-players, is still required in many training scenarios. With improved behavior models, training systems with a higher level of autonomy could be built, and adaptive training (AT) with contents tailored to the current learning needs of individual trainees could be provided [10].

3 An Intelligent, Synthetic Trainer

To make pilot training more efficient and effective, we would like to increase the autonomy of air combat training systems, and minimize the dependence on human training providers. For this purpose, we propose to construct an intelligent, synthetic trainer, which can learn to understand the learning needs of trainees, and then act accordingly to provide the best possible training, as well as support for evaluation of trainees. The synthetic trainer should be able to represent allies as well as adversaries in a training scenario. We are investigating how machine learning could be used to create such an agent. A proposed architecture is shown in fig. 4.
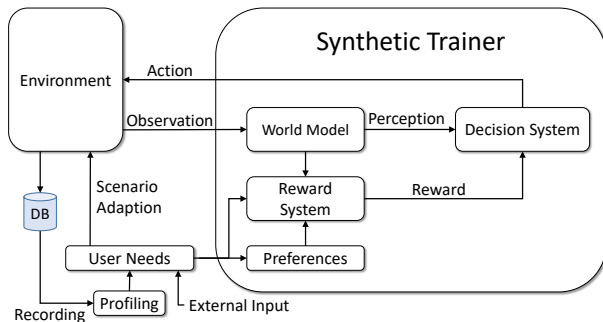

Figure 4: Architecture of a Synthetic Trainer.

The agent interacts with an environment, populated by human trainees as well as other, synthetic agents, who participate in an air combat training scenario. Users' training needs must be considered by a synthetic agent when acting in this scenario, and thus affect the goals of the agent, as well as preferences regarding how the agent should try to achieve those goals. In current training systems user needs are typically handled manually by an instructor, e.g. when constructing simulation contents or conducting training sessions, but we would like to automate this process, by using machine learning to model the training needs, progress and proficiency of the trainee. These models could be constructed between training sessions, based on recordings, or online during the execution of a training sessions, based on the observations of the agent. They can then be used to automatically adapt the contents and characteristics of the training scenario.

The agent's observations of the world are used as input to a world model, which will create the higher level perception of the agent. Such models can be used to predict the winner of a game given a certain state, as well as the skills, beliefs, long term goals and immediate actions of agents [11–17]. Parts of the model could be constructed by hand, based on domain knowledge, but we are primarily investigating learning approaches, such as Supervised Deep Learning or Unsupervised Learning [18, 19]. Since data from real or simulated air battles, with aircraft operated by human pilots, are not readily available, the intention is to use synthetic data for training of machine learning models.

The decision system has capabilities for learning policies suitable for training of trainees. These policies should consider the goals of the agent in the simulated air combat scenario, as

well as the learning objectives of the trainees. As mentioned, limited data is available from human pilots. Instead we have identified reinforcement learning as a promising technique for implementing this system, since it allows an agent to learn based on interaction with a simulation. Feedback regarding the agent's learning progress is then provided by a reward system. Reinforcement learning is discussed further in the following sections.

3.1 Reinforcement Learning

Reinforcement learning is a machine learning paradigm, which aims to allow synthetic agents to learn how to achieve their goals by interacting with their environment [20]. In recent years the technique has had great success in training agents to solve games, such as Go and StarCraft [11, 12, 21–23], as well as complex control tasks [24–26].

Reinforcement learning problems are modelled as Markov Decision Processes (MDP). An MDP is a tuple (S, A, T, R, γ), specifying:

- S: The finite set of states of the process
- A: The finite set of actions of the process
- T: The transition dynamics of the process
- R: The reward function of the process
- γ: The discount factor indicating the importance of immediate and future rewards respectively
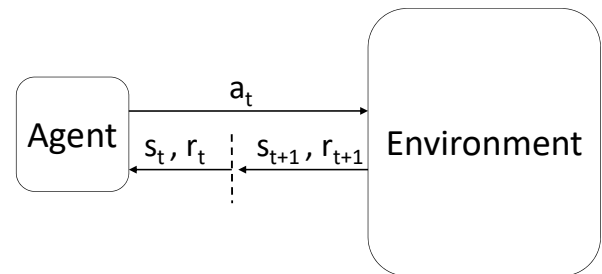

Figure 5: Markov Decision Process.

As illustrated in in fig. 5, in each time step the agent selects an action, observes the resulting new state of the environment, and receives a reward. The objective of the agent is to maximize its future expected return:

$$V_\pi(s) = E[R_t|s_0 = s] = E[\sum_{t=0}^{\infty} \gamma^t r_t|s_0 = s] \qquad (1)$$

The state value function $V_\pi(s)$ specifies the value of being in state s and then following policy π. Similarly, the value of being in state s and taking action a, and then following policy π is given by the state-action value function:

$$Q_\pi(s,a) = E[\sum_{t=0}^{\infty} \gamma^t r_t|s_0 = s, a_0 = a] \qquad (2)$$

One popular algorithm for reinforcement learning is Q-learning [27], which seeks to estimate the state-action value function $Q(s,a)$. This algorithm was extended to handle complex, continuous state spaces in the Deep Q-Networks algorithm [28], which uses deep neural networks to represent the agent's policy.

3.2 Multi-Agent Reinforcement Learning

In most air combat scenarios, pilots do not act on their own, but instead must cooperate with allies to achieve their goals, while competing with enemies. To train teams of agents, multi-agent reinforcement learning can be used. The single agent MDP can be extended to include multiple agents in so called Stochastic Games (SG), where multiple agents interact with the environment, and the environment state as well as the rewards of individual agents are determined by the joint actions of all agents [1]. Stochastic games can be characterized as fully cooperative when all agents have the same goal, and fully competitive when agents have opposite goals. Stochastic games that are neither fully cooperative nor fully competitive are called mixed games. A special case of cooperative stochastic games is the Decentralized Partially Observable Markov Decision Process (Dec-POMDP) [29,30]. A Dec-POMDP is a tuple $(S, \{A_i\}, T, R, \{\Omega_i\}, O, \gamma)$, specifying:

- S: The finite set of states of the process

- A_i: The finite set of actions of agent i

- T: The transition dynamics of the process

- R: The reward function of the process

- Ω_i: The finite set of observations of agent i

- O: The finite set of conditional observation probabilities

- γ: The discount factor indicating the importance of immediate and future rewards respectively
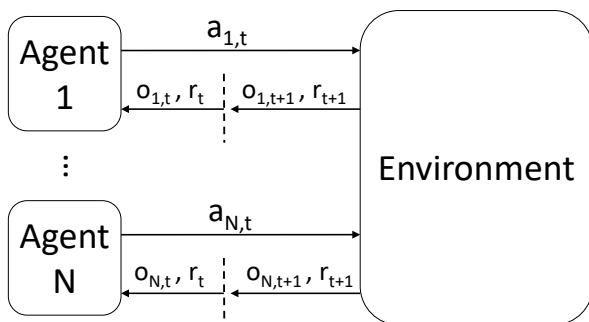

Figure 6: Decentralized Partially Observable Markov Decision Process.

Agents in the Dec-POMD, illustrated in fig. 6, can only observe parts of the environment state, since e.g. internal states of other agents may be hidden. In an air combat scenario observability could be affected by e.g. performance of sensors

and data links, which could also be affected by electronic warfare. In the illustrated Dec-POMDP, agents must try to coordinate their actions to maximize a shared reward. In other settings agents could have individual rewards, e.g. due to different priorities among members in a team, or due to a competitive scenario.

Multi-agent learning presents many challenges, such as coordination among agents and multi-agent credit assignment (i.e. determining a single agent's contribution to the success of a team of agents). As several agents learn concurrently, the environment may also become non-stationary from a single agent's point of view. The Multi-Agent Deep Deterministic Policy Gradient algorithm (MADDPG) [31] uses a centralized Q function to guide updates of decentralized policies. The algorithm supports continuous action spaces and mixed cooperative-competitive scenarios, and is thus suitable for applications in the air combat domain.

Multi-agent reinforcement learning can also be used as a framework for creating hierarchical policies for single entities in a simulation [32, 33], by placing agents in a hierarchy where higher-level agents try to reach abstract goals by issuing commands to lower level agents, with increasingly reactive behavior.

3.3 Multi-Objective Reinforcement Learning

Multi-objective reinforcement learning can be used to learn policies for problems where multiple, possible conflicting objectives must be considered [2]. Typical air combat scenarios fit this description, since they require that the participating pilots prioritize among objectives such as targets to attack, assets to protect, safety and resource consumption. In training scenarios, synthetic agents could also consider the learning objective of trainees, e.g. by adapting their behavior to fit the proficiency of the trainee. In multi-objective reinforcement learning the single-objective MDP is extended to a Multi-Objective Markov Decision Process (MOMDP). An MOMDP provides a vector-valued reward function, with each element representing the reward for one of the objectives. The user utililty of the vector-valued return of an MOMDP is given by using a scalarization function, which converts the vector to a scalar. One option is to use the weighted sum of all values:

$$V_{\mathbf{w}}^{\pi}(s) = f(\mathbf{V}^{\pi}(s), \mathbf{w}) = \sum_{n=1}^{n} v_i^{\pi}(s) w_i \qquad (3)$$

One approach for solving an MOMDP is to use scalarization directly on the vector-valued reward signal, to convert the MOMDP to an MDP for a set of preferences, and then use single-objective methods to find a set of policies [34, 35]. Then, at execution time, the user can select a suitable policy. By using a stochastic mixture of policies over time for episodic tasks, i.e. selecting one of several policies by random before the start of an episode, further parts of the solution space can be covered [36]. In our previous work, we proposed an approach for training a single, tunable neural network policy to prioritize among a set of objectives at execution time, by conditioning the network on user preferences [37].

4   Experiments

To evaluate the potential of machine learning, in particular multi-agent reinforcement learning and multi-objective reinforcement learning, as a tool for building CGF behavior models, we conduct a number of experiments. As simulation platform we use the tactical environment simulation software that is part of the Saab Gripen Flight Training Simulators. All simulation results are averaged over five runs with different random seeds.

4.1   Coordination of a Tactical Air Unit

We first study how multi-agent reinforcement learning can be used to coordinate the actions of agents that are members of the same Tactical Air Unit (TAU). For this purpose, we use the MADDPG algorithm. We also use environments that are similar to those used in the original paper [31], but implemented in our high-fidelity simulation engine. The increased complexity of the state space, as well as the increased number of time steps per simulated episode, add additional difficulty. We investigate how the algorithm performs in this setting, for different types of action spaces. The policy is represented by a multilayer perceptron (MLP), with 2 hidden layers, each with 64 neurons and the ReLU activation function. We use a learning rate of $\alpha = 10^{-2}$, a discount factor of $\gamma = 0.95$, and train using the Adam optimizer.

4.1.1   Coordinated Defense

In this scenario there are three high-value assets that should be protected by three learning agents. The assets are attacked by three enemy agents, which are controlled by handcrafted behavior models, implemented with Behavior Trees [38]. If a defending agent comes within 5 km of an attacking agent, the attacker will retreat to its home base, and then attack again. To protect all three high-value assets, the learning agents must learn to split up and escort one enemy each from the protected area. The defending agents are initialized with random positions and headings, while the attacking agents are initialized at random positions along their planned attack routes. The spawn area of blue aircraft and attack directions of red aircraft are shown in fig. 7.

The observation space of each agent is the relative position of all other agents, in a body-fixed coordinate system, for the last 4 time steps in the episode. We study three types of action spaces. The first two are continuous action spaces, which allow an agent to fly forward, or turn left or right with a load factor of 2-4 g. One of these action spaces also allows an agent to set its internal state as a three element, real-valued and normalized vector, which is then distributed to the other agents in the team in each time step. Previous work has shown that this type of mechanism can allow agents to develop a language for coordination of their actions [31]. The third action space is a hierarchical approach, with discrete actions that let the agent select a target and assign it as goal for a lower level controller. For the first two types of action spaces, the agent is executed at 1 s intervals, with episodes lasting for 600 time steps, while for the third type it selects actions at 10 s intervals, with episodes lasting for 60 time steps. To promote co-
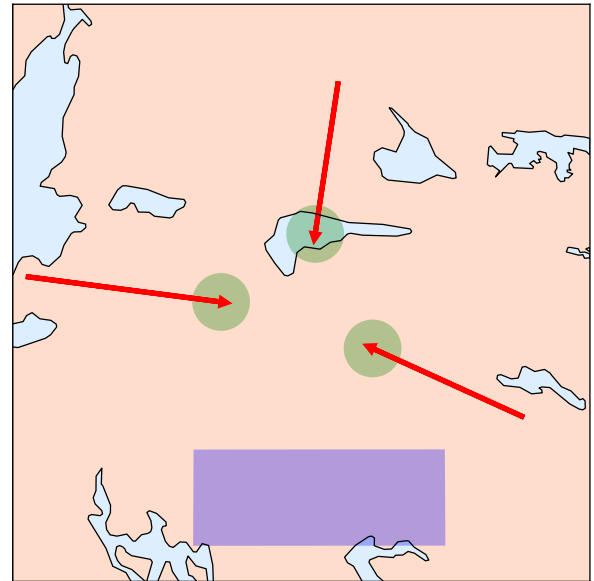

Figure 7: Rectangular spawn area of blue aircraft, and red arrows indicating attack directions of red aircraft towards high-value assets in green.

operation, the learning agents receive a shared reward defined as:

$$r_t = -\sum_{i=1}^{3} \min(\|p_{a_i} - p_{d_1}\|, \|p_{a_i} - p_{d_2}\|, \|p_{a_i} - p_{d_3}\|) \quad (4)$$

where p_{a_i} refers to the position of attacker i and p_{d_k} refers to the position of defender k.

The training progress for the low-level action spaces over 90k episodes is presented in fig. 8, and the training progress for the high-level action space over 30k episodes is presented in fig. 9. Agents that are hard-coded to always attack the same enemy are used as baselines, and their scores averaged over 1000 simulated episodes are also presented in the figures. The hard-coded baseline is strong, but not optimal, since it does not consider the initial positions of agents. To perform comparatively well, the learning agents must learn to coordinate their actions. Two of the agent types can coordinate based on only observations, while one of the agent types has the benefit of an explicit communication mechanism, provided that it can learn a protocol for coordination.

We can see that the low-level, silent controller makes fast initial improvement, but then reaches a plateau. This is because the agents must first learn to move as a team towards the protected area, before being able to learn the benefits and means of cooperation. The learning progress during the second stage of learning is quite slow, and varies among different runs, as can be seen by the increase in variance. The high-level controller, on the other hand, quickly converges to policies that perform close to the baseline. The high-level action space automatically moves the agents towards the protected area, so that agents can start learning cooperation strategies from the
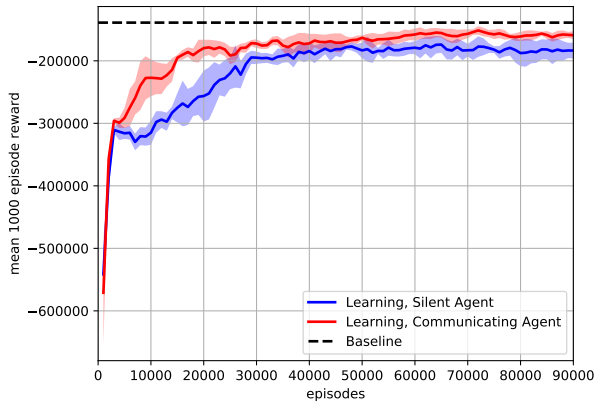
Figure 8: Mean and standard deviation for the training progress of coordinated defense with a low-level action space, for silent and communicating agents.
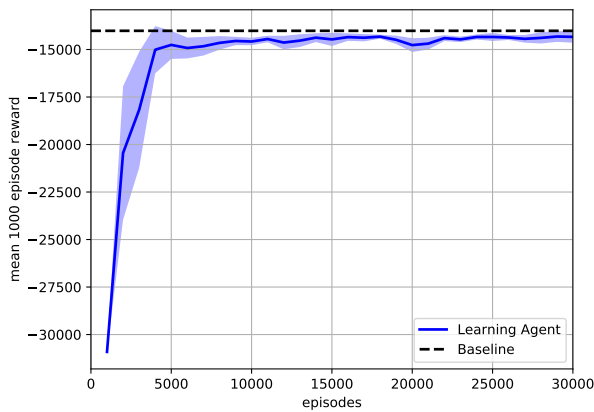

Figure 9: Mean and standard deviation for the training progress of coordinated defense with a high-level action space.

Figure 10: Rectangular spawn areas of aircraft and targets in blue and green respectively.

$$r_t = -\sum_{i=1}^{3} \min(\|p_{t_i} - p_{a_1}\|, \|p_{t_i} - p_{a_2}\|, \|p_{t_i} - p_{a_3}\|) \quad (5)$$

where p_{t_i} refers to the position of target i and p_{a_k} refers to the position of attacker k.

The agents are trained for 60k episodes. The training progress for the low-level action space is presented in fig. 11, and the training progress for the high-level action space is presented in fig. 12. Agents that are hard-coded to always attack the same target are used as baselines, and their scores averaged over 1000 simulated episodes are also presented in the figures. The hard-coded baseline is strong, but not optimal, since it does not consider the initial positions of agents. To perform comparatively well, the learning agents must learn to coordinate their actions based on only observations.

start. The low-level, communicating controller displays faster improvement, and also finds a policy that generates slightly more reward than the policy of silent agents. This indicates that explicit communication mechanisms can be valuable for efficient cooperation in air combat scenarios.

4.1.2 Coordinated Attack

In this scenario there are three targets that should be attacked by three learning agents. To carry out the task efficiently, the agents must learn to split up and attack one target each. The agents are initialized with random positions and headings in an area to the south, while the targets are initialized in random positions in a larger area to the north. The spawn areas of aircraft and targets are shown in fig. 10.

The observation space of each agent is the relative position of all other agents, as well as the targets, in a body-fixed coordinate system, for the last 4 time steps in the episode. We study the same two types of action spaces used by silent agents in 4.1.1. For the low-level action space, episodes last for 500 time steps, while for the high-level action space, episodes last for 50 time steps. To promote cooperation, the learning agents receive a shared reward defined as:
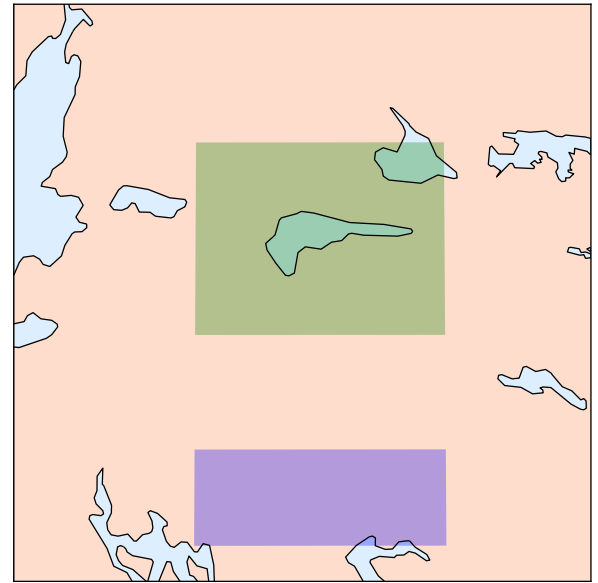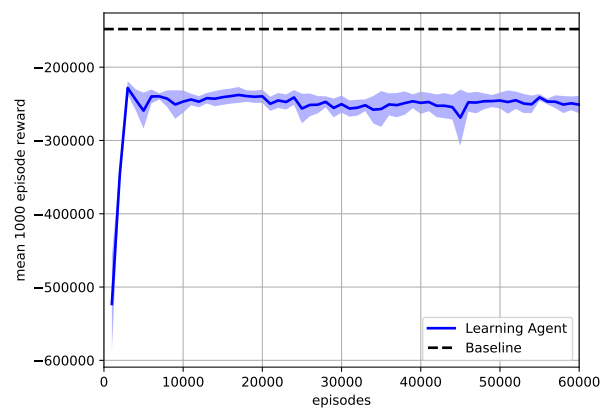

Figure 11: Mean and standard deviation for the training progress of coordinated attack with a low-level action space.
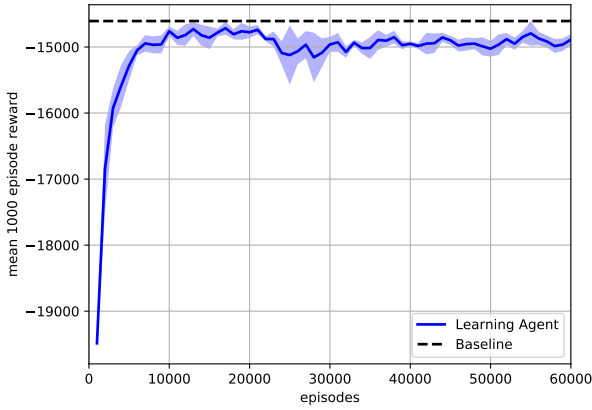
Figure 12: Mean and standard deviation for the training progress of coordinated attack with a high-level action space.

We can see that the low-level controller quickly converges to a sub-optimal policy, and then does not improve during the rest of training. Possibly further training episodes could eventually lead to an improvement of the policy. The high-level controller makes quite fast progress, but not as fast as in the experiment presented in 4.1.1. The learning also seems less stable. While some of the trained agents learn policies as good as the baseline, others struggle a bit in some episodes. This is possibly because in this scenario targets may spawn quite close to each other, which makes it difficult for the learning agents to cooperate based on observations alone.

To further study the performance of the low-level controller, and its dependence on the starting positions of agents, we conduct an additional experiment, where the aircraft are spawned in the green area in fig. 10. The training progress for this experiment over 60k episodes, with episodes lasting 300 time steps, is presented in fig. 13. We can see that the agent performs better for this scenario, since aircraft start closer to the targets, which simplifies the task of learning coordination among agents.
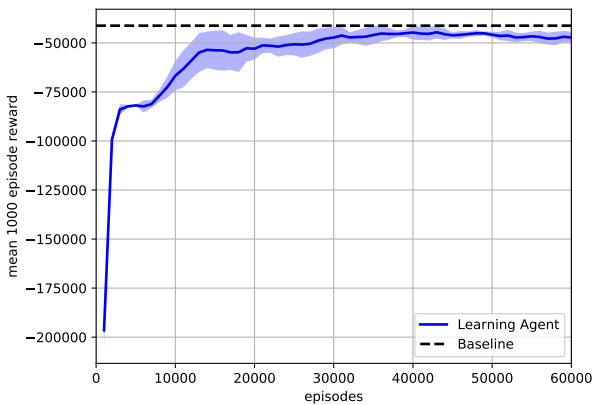

Figure 13: Mean and standard deviation for the training progress of coordinated attack with a low-level action space, when aircraft spawn close to the targets.

4.2 Risk Aware Attack

We now proceed to investigate how multi-objective reinforcement learning can be used to allow agents to learn how to prioritize among multiple conflicting objectives. We use two approaches: The outer-loop approach, where the MOMDP is converted to multiple single-objective MDPs, which are then solved with single-objective methods to produce a set of fixed policies [35], and our own approach using a single neural network conditioned on the objective priorities [37], which produces a tunable policy. We then compare the performance of the two approaches. In these experiments we use DQN as the core learning algorithm, as in the referenced papers. As previously mentioned, we use a high-fidelity simulation engine for the experiments, in contrast to the simple, low-dimensional gridworld environments studied in previous work. The policy is represented by an MLP, with 1 hidden layer with 64 neurons and the ReLU activation function. We use a learning rate of $\alpha = 10^{-4}$, a discount factor of $\gamma = 1.00$, a replay-buffer with 10^6 samples, and train using the Adam optimizer and prioritized experience replay [39].

In the studied scenario, the synthetic pilot must reach a target location in an attack mission, while avoiding enemy air defense systems. The agent must prioritize between time and safety when selecting a route. For simplicity, we place one air defense system between the agent's start position and the target. The agent always starts in the same position, with initial heading towards the target. The scenario is illustrated in fig. 16.

The observation space of the agent is the relative heading and distance to the center of the threat area, and the relative heading and distance to the target, for the last 8 time steps in the episode. Since DQN does not handle continuous actions, we must discretize the input to the controllers of the aircraft model. Thus, we define the agent's actions space as forward motion or right or left turns with a load factor of 2-4 g in discrete steps of 1 g. The agent selects actions at 1 s intervals. Each training episode is a maximum of 400 time steps long. The episode ends if the agent reaches the target. The reward vector of the MOMDP is defined as:

$$\mathbf{r}_t = [r_{goal}(t), r_{time}(t), r_{ad}(t)] \tag{6}$$

$$d_g(t) = \|p_g(t) - p_a(t)\| \tag{7}$$

$$d_{ad}(t) = \|p_{ad}(t) - p_a(t)\| \tag{8}$$

$$r_{goal}(t) = d_g(t-1) - d_g(t) \tag{9}$$

$$r_{time}(t) = -0.5 \tag{10}$$

$$r_{ad}(t) = \begin{cases} -\left(\frac{1}{10}(R_{ad} - d_{ad}(t))\right)^2 & \text{if } d_{ad}(t) \leq R_{ad}; \\ 0 & \text{if } d_{ad}(t) > R_{ad}; \end{cases} \tag{11}$$

where $r_{goal}(t)$ refers to the reward for the objective of moving towards the target, $r_{time}(t)$ refers to the reward for the objective of reaching the target fast, and $r_{ad}(t)$ refers to the reward for the objective of staying out of range of the air defense system. $p_a(t)$, $p_g(t)$ and $p_{ad}(t)$ are the positions of the agent, goal and air defense system, $d_g(t)$ and $d_{ad}(t)$ are the distances from the agent to the goal and air defense system, and $R_{ad} = 20$ km is the range of the air defense system. To scalarize the vector-valued reward of the MOMDP we define the parameterized vector of priorities among objectives $\mathbf{p}_\theta = [1, \theta, \theta - 1]$, with $\theta \in [0, 1]$. We then calculate a scalar reward as:

$$r_t = \mathbf{r}_t \cdot \mathbf{p} \qquad (12)$$

For the tunable agent we sample θ from a uniform distribution of $[0.75, 1.00]$ before each episode, and use it as input to the agent. We train fixed policy agents for 10M time steps, while tunable policy agents are trained for 30M time steps. The training progress for fixed policy agents is presented in fig. 14, for $\theta \in \{0.75, 0.85, 0.95\}$, and the training progress for tunable agents is presented in fig. 15.
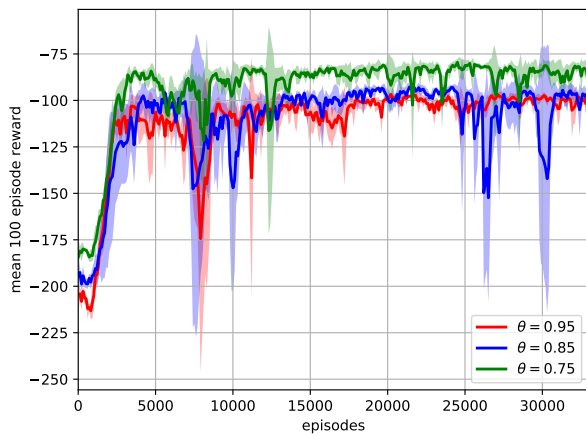

Figure 14: Training progress for a set of fixed policies with different priorities among objectives.

We can see that training is somewhat unstable, with spikes of high variance. The cause may be that small changes in policy have great effect on the accumulated reward, or that it is difficult for the agent to learn the characteristics of the reward function. It is also possible that the low frequency of the controller or the discretization of the action space has a negative effect on performance.

Three routes learned for different priorities, corresponding to high, medium and low risk exposure, are illustrated in fig. 16, for fixed policies and a tunable policy. The displayed routes are for single runs, not averaged over several runs or agents, since agents may choose to go on either side of the center of the threat area. The routes displayed for the tunable policy are from one trained agent. We can see that the tunable policy results in tighter routes around the center of the threat area, compared to those generated by the set of fixed policies. Finding an optimal route with the given reward system requires a
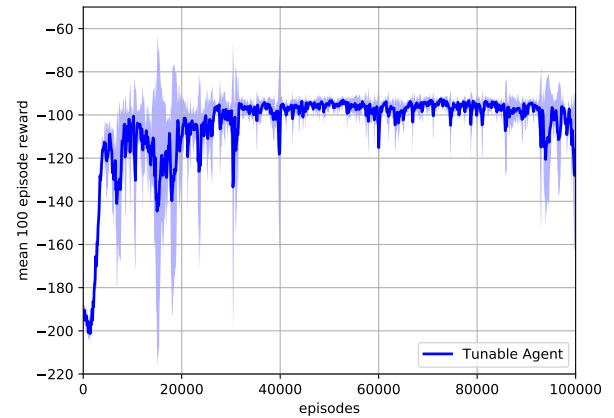

Figure 15: Training progress for single, tunable policy.

bit of tuning, and more exploration would have been required to achieve improved performance.
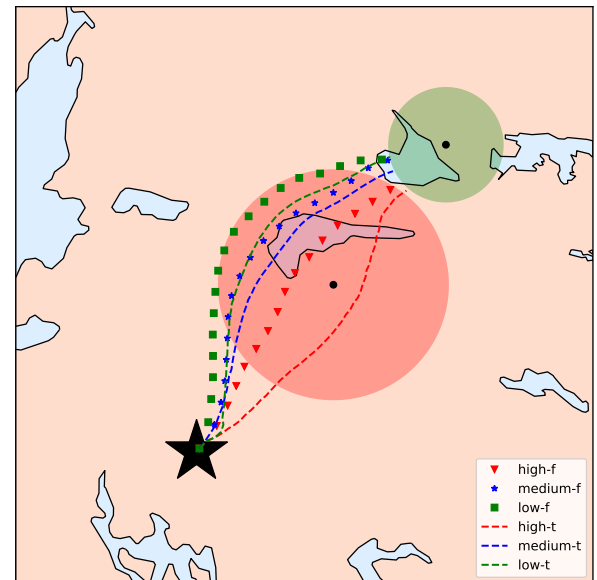
Figure 16: Learned routes to the target area with high, medium and low risk exposure, for fixed (f) and tunable (t) policies.

The mean and standard deviation for the rewards accumulated by fixed and tunable policies are presented in fig. 17. We can see that the tunable policies produce competitive results for $\theta = 0.95$ and $\theta = 0.85$, but perform worse for $\theta = 0.75$. The poor result is caused by one of the five trained agents, which fails to reach the goal for this configuration, which in turn heavily affects its accumulated reward. This also leads to high standard deviation for this case.

By extending multi-objective learning to more complex scenarios, with more objectives that must be prioritized, agents with diverse characteristics can be constructed. This can make training more interesting and stimulating, and by adjusting agents' objective preferences training contents can also be adapted to the training needs of specific trainees.

Figure 17: Mean and standard deviation for accumulated rewards for fixed and tunable policies.

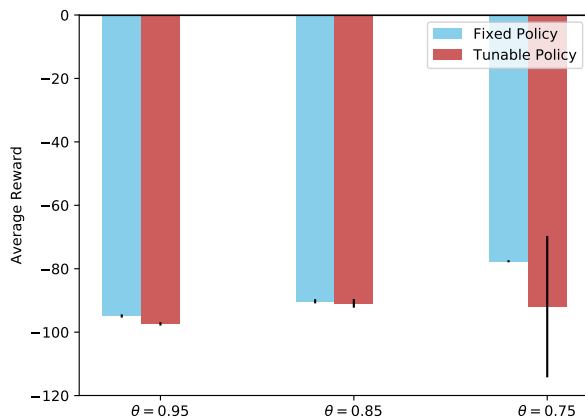## 5 Related Work

Over the years, there have been several attempts at using machine learning techniques for building behavior models for computer generated forces. Some approaches that have been studied are evolutionary algorithms [40–42], neural networks [43–46] and dynamic scripting [47–49], a technique originally developed for computer games. Still, the studied techniques have not been mature enough to include in commercial CGF software [9].

With the renewed interest in machine learning, sparked by e.g. AlphaGo [11, 21], there have been approaches using deep reinforcement learning [50–52]. However, we are not aware of any work that studies the recent advancements in multi-agent or multi-objective deep reinforcement learning in the context of air combat simulation.

6 Conclusions

In this paper we discussed the future of air combat training, and suggested an approach for building an intelligent, synthetic trainer for fighter pilots, using machine learning techniques. We also presented results of initial experiments, which indicate that state-of-the-art algorithms can allow agents to learn team coordination as well as prioritization among conflicting objectives in simple air combat scenarios. However, we also note some challenges posed by the complexity of the air combat domain. Learning high-level tactical behavior using a low-level action space may not be the best approach. As the complexity of scenarios grows, it will become more difficult for the agent to learn efficient policies. It may get stuck in a local optimum, or perhaps not learn any reasonable policy at all. We believe that a hierarchical approach to reinforcement learning, where the problem is decomposed into a number of sub-tasks handled by a hierarchy of agents, can help tackle this problem, as indicated by the results in 4.1.1 and 4.1.2.

In future work we would like to continue to study more complex scenarios, which more closely resemble those used in operational training systems, to facilitate experiments with manned simulators and studies of human-agent interaction.

We would then like to extend our study of multi-agent learning to include adversarial learning, where teams of agents compete against each-other. We would also like to combine multi-agent and multi-objective learning in an integrated architecture, using a hierarchical approach to reinforcement learning, in combination with learned models for predicting other agents' characteristics, goals and actions, to support decision making. Finally, we would like to study intelligent exploration schemes and other ways to achieve sample efficient learning in complex state and action spaces.

Acknowledgements

References

[1] Lucian Bu, Robert Babu, Bart De Schutter, et al. A comprehensive survey of multiagent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(2):156–172, 2008.

[2] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research*, 48:67–113, 2013.

[3] JJ Roessingh and GG Verhaaf. Training effectiveness of embedded training in a (multi-) fighter environment. Technical report, NATIONAL AEROSPACE LAB AMSTERDAM (NETHERLANDS), 2009.

[4] Ernest H Page and Roger Smith. Introduction to military training simulation: a guide for discrete event simulationists. In *1998 Winter Simulation Conference. Proceedings (Cat. No. 98CH36274)*, volume 1, pages 53–60. IEEE, 1998.

[5] Amy E Henninger, Dannie Cutts, Margaret Loper, Robert Lutz, Robert Richbourg, Randy Saunders, and Steve Swenson. Live virtual constructive architecture roadmap (lvcar) final report. *Institute for Defense Analysis*, 2008.

[6] Douglas D Hodson and Raymond R Hill. The art and science of live, virtual, and constructive simulation for test and analysis. *The Journal of Defense Modeling and Simulation*, 11(2):77–89, 2014.

[7] Jack Thorpe. Trends in modeling, simulation, & gaming: Personal observations about the past thirty years and speculation about the next ten. In *Interservice/Industry training, simulation, and education conference (I/ITSEC)*, 2010.

[8] TW van den Berg, NM de Reus, and JM Voogd. *LVC Architecture study*. Simulation Interoperability Standards Organization (SISO), 2011.

[9] Armon Toubman, Gerald Poppinga, Jan Joris Roessingh, Ming Hou, Linus Luotsinen, Rikke Amilde Løvlid, Christophe Meyer, Roel Rijken, and M Turcanık. Modeling cgf behavior with machine learning techniques: Requirements and future directions. In *Proceedings of the 2015 Interservice/Industry Training, Simulation, and Education Conference*, pages 2637–2647, 2015.

[10] Christopher Best and Benjamin Rice FLTLT. Science and technology enablers of live virtual constructive training in the air domain. *Air & Space Power Journal*, 32(4):59–73, 2018.

[11] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484, 2016.

[12] Oriol Vinyals, Timo Ewalds, Sergey Bartunov, Petko Georgiev, Alexander Sasha Vezhnevets, Michelle Yeo, Alireza Makhzani, Heinrich Küttler, John Agapiou, Julian Schrittwieser, et al. Starcraft ii: A new challenge for reinforcement learning. *arXiv preprint arXiv:1708.04782*, 2017.

[13] Tetske Avontuur, Pieter Spronck, and Menno Van Zaanen. Player skill modeling in starcraft ii. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2013.

[14] Niels Justesen and Sebastian Risi. Learning macromanagement in starcraft from replays using deep learning. In *2017 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 162–169. IEEE, 2017.

[15] Sid Reddy, Anca Dragan, and Sergey Levine. Where do you think you're going?: Inferring beliefs about dynamics from behavior. In *Advances in Neural Information Processing Systems*, pages 1454–1465, 2018.

[16] Neil C Rabinowitz, Frank Perbet, H Francis Song, Chiyuan Zhang, SM Eslami, and Matthew Botvinick. Machine theory of mind. *arXiv preprint arXiv:1802.07740*, 2018.

[17] Roberta Raileanu, Emily Denton, Arthur Szlam, and Rob Fergus. Modeling others using oneself in multi-agent reinforcement learning. *arXiv preprint arXiv:1802.09640*, 2018.

[18] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[19] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[20] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[21] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354, 2017.

[22] Tabish Rashid, Mikayel Samvelyan, Christian Schroeder Witt, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. Qmix: Monotonic value function factorisation for deep multi-agent reinforcement learning. In *International Conference on Machine Learning*, pages 4292–4301, 2018.

[23] Jakob N Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli, and Shimon Whiteson. Counterfactual multi-agent policy gradients. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[24] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[25] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[26] Martin Riedmiller, Roland Hafner, Thomas Lampe, Michael Neunert, Jonas Degrave, Tom Wiele, Vlad Mnih, Nicolas Heess, and Jost Tobias Springenberg. Learning by playing solving sparse reward tasks from scratch. In *International Conference on Machine Learning*, pages 4341–4350, 2018.

[27] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

[28] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.

[29] Daniel S Bernstein, Robert Givan, Neil Immerman, and Shlomo Zilberstein. The complexity of decentralized control of markov decision processes. *Mathematics of operations research*, 27(4):819–840, 2002.

[30] Christopher Amato, Girish Chowdhary, Alborz Geramifard, N Kemal Üre, and Mykel J Kochenderfer. Decentralized control of partially observable markov decision processes. In *52nd IEEE Conference on Decision and Control*, pages 2398–2405. IEEE, 2013.

[31] Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*, pages 6379–6390, 2017.

[32] Peter Dayan and Geoffrey E Hinton. Feudal reinforcement learning. In *Advances in neural information processing systems*, pages 271–278, 1993.

[33] Alexander Sasha Vezhnevets, Simon Osindero, Tom Schaul, Nicolas Heess, Max Jaderberg, David Silver, and Koray Kavukcuoglu. Feudal networks for hierarchical reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 3540–3549. JMLR. org, 2017.

[34] Kristof Van Moffaert, Madalina M Drugan, and Ann Nowé. Scalarized multi-objective reinforcement learning: Novel design techniques. In *2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pages 191–199. IEEE, 2013.

[35] Hossam Mossalam, Yannis M Assael, Diederik M Roijers, and Shimon Whiteson. Multi-objective deep reinforcement learning. *arXiv preprint arXiv:1610.02707*, 2016.

[36] Peter Vamplew, Richard Dazeley, Ewan Barker, and Andrei Kelarev. Constructing stochastic mixture policies for episodic multiobjective reinforcement learning tasks. In *Australasian Joint Conference on Artificial Intelligence*, pages 340–349. Springer, 2009.

[37] Johan Källström and Fredrik Heintz. Tunable dynamics in agent-based simulation using multi-objective reinforcement learning. In *Adaptive and Learning Agents (ALA) workshop at AAMAS*, 2019.

[38] Michele Colledanchise and Petter Ögren. *Behavior Trees in Robotics and AI: An Introduction*. CRC Press, 2018.

[39] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay. *arXiv preprint arXiv:1511.05952*, 2015.

[40] Sandeep Mulgund, Karen Harper, Kalmanje Krishnakumar, and Greg Zacharias. Air combat tactics optimization using stochastic genetic algorithms. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)*, volume 4, pages 3136–3141. IEEE, 1998.

[41] Magdalena D Bugajska, Alan C Schultz, J Gregory Trafton, Shaun Gittens, and Farilee Mintz. Building adaptive computer-generated forces: The effect of increasing task reactivity on human and machine control abilities. Technical report, NAVAL RESEARCH LAB WASHINGTON DC CENTER FOR APPLIED RESEARCH IN ARTIFICIAL INTELLIGENCE, 2001.

[42] Jian Yao, Qiwang Huang, and Weiping Wang. Adaptive human behavior modeling for air combat simulation. In *2015 IEEE/ACM 19th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*, pages 100–103. IEEE, 2015.

[43] Ervin Y Rodin and S Massoud Amin. Maneuver prediction in air combat via artificial neural networks. *Computers & mathematics with applications*, 24(3):95–112, 1992.

[44] Amy E Henninger, Avelino J Gonzalez, Michael Georgiopoulos, and Ronald F DeMara. Modeling semi-automated forces with neural networks: Performance improvement through a modular approach. In *The Ninth Conference on Computer Generated Forces and Behavioral Representation Proceedings*, 2000.

[45] Teck-Hou Teng, Ah-Hwee Tan, Yuan-Sin Tan, and Adrian Yeo. Self-organizing neural networks for learning air combat maneuvers. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2012.

[46] Teck-Hou Teng, Ah-Hwee Tan, and Loo-Nin Teow. Adaptive computer-generated forces for simulator-based training. *Expert Systems with Applications*, 40(18):7341–7353, 2013.

[47] Armon Toubman, Jan Joris Roessingh, Pieter Spronck, Aske Plaat, and Jaap Van Den Herik. Dynamic scripting with team coordination in air combat simulation. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pages 440–449. Springer, 2014.

[48] Armon Toubman, Jan Joris Roessingh, Pieter Spronck, Aske Plaat, and Jaap van den Herik. Transfer learning of air combat behavior. In *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, pages 226–231. IEEE, 2015.

[49] Armon Toubman, Jan Joris Roessingh, Pieter Spronck, Aske Plaat, and Jaap van den Herik. Rapid adaptation of air combat behaviour. In *Proceedings of the Twenty-second European Conference on Artificial Intelligence*, pages 1791–1796. IOS Press, 2016.

[50] Roel Rijken and Armon Toubman. The future of autonomous air combat behavior. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3089–3094. IEEE, 2016.

[51] Babak Toghiani-Rizi, Farzad Kamrani, Linus J Luotsinen, and Linus Gisslén. Evaluating deep reinforcement learning for computer generated forces in ground combat simulation. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3433–3438. IEEE, 2017.

[52] Bogdan Vlahov, Eric Squires, Laura Strickland, and Charles Pippin. On developing a uav pursuit-evasion policy using reinforcement learning. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 859–864. IEEE, 2018.

Towards a Complete Co-Simulation Model Integration Including HMI Aspects

Ingo Staack*, Jörg Schminder*, Owais Shahid**, and Robert Braun*

*Dept. for Management and Engineering, Linköping University, Sweden
**National Electric Vehicle Sweden AB (Nevs)
*E-mail: ingo.staack@liu.se

Abstract

Modern aircraft can be seen as heterogeneous systems, containing multiple embedded sub-systems which are in today's simulations split into different domain-specific models based on different modelling methods and tools.

This paper addresses typical workflow-driven model integration problems with respect to model fidelity, accuracy in combination with the selected abstraction methods and the target system characteristics. A short overview of integration strategies with the help of co-simulation frameworks including an analysis of the inherent problems that emerge because of different domain-specific modelling methods is being given. It is shown that huge benefits can be reached with the help of a smart system break-up.

In detail, the discrepancy between the cyber-physical system simulations and human-machine interaction (HMI) models are being analysed. Therefore, a close look on typical shortcomings of behavioural models are being discussed, too.

To enable an effort-less human-in-the-loop integration into a cyber-physical system simulation, the usage of flight simulation software, offering real-time capability and a graphical user interface is suggested. This approach is applied to overcome today's complexity and shortcomings in human psychological models. An example implementation based on a commercial flight simulator software (X-Plane) together with a high-performance system simulation tool (Hopsan) via UDP communication is presented and analysed.

Keywords: flight simulator, model fidelity, co-simulation, mission simulation, workflow-driven integration, human-machine interaction, behavioural model, psychological model

1 Introduction

The demand for more efficient airplane increases steadily and as a result, the complexity of these airplane escalates as well. System simulations are extensively used to design, handle and maintain the understanding of such a complex product. It supports the designer on various tasks, to early detect possible design errors, performing design optimizations and enables for complex design analysis such as operational and maintenance concepts. First with excessive use of simulations a holistic whole product life-cycle analysis becomes possible.

Aircraft include several on-board systems such as hydraulic, pneumatic, mechanic or electric subsystems which operate in unison to fulfil a mission. It is crucial that each subsystem is performing optimal, and any possible flaws or malfunctions can be recognized and resolved early in the design process. Simulating these subsystems in a virtual test environment already during the system architecture work allows for design modifications and improvements prior to any physical testing. This enhances the design process by making it more efficient by reducing early the design uncertainties, thus enhancing the level of credibility.

2 Multi-domain Co-Simulation Frameworks

2.1 Motivation

Nowadays whole life cycle-focused product development requires a vast number of simulations to be performed at various disciplines facilitating various modelling methods that differ between the modelling task, the analysis task and the type of analysis (see [1] for an overview of different level of interest during product development work).

Any holistic model-based product development work has to include multi-domain co-simulations addressing:

- detailed sub-system simulations to study certain domain-specific system characteristics with respect to the overall system architecture
- *human-machine interaction* (HMI): the human is part of the control system or influences directly the mode of operation or use of the product.

Furthermore, it is crucial to pay attention for easy model exchange or on-the-fly model replacements in order to enable:

- **rapidly model adaption** for different analysis and design studies
- **model fidelity alteration** (preferably step-less) that fits to the available design information to enable a task/workflow-driven design process
- **model reuse** with limited adaption effort to boost development efficiency, preferable also involving black box models from third party suppliers.

2.2 Co-simulation Strategies and the Problem of Complexity

Recent research projects on simulation and modelling frameworks do focus on the above mentioned topics flexibility, adaptability and model re-usability such as the AGILE [2] or the OM-simulator [3] projects. Also, commercial tool vendors support various inter-disciplinary integration environments (such as modeFRONTIER, LMS Amesim, RCE, TechnoSoft AML, ANSYS) that allow a effort-less tool integration by supporting communication protocols, workflow process control and optimization algorithms. These environments are often denoted as *multi-disciplinary design optimization* (MDO) tools, partly enabling distributed software execution via the internet [4].

On a multi-domain cyber-physical system – and especially on a *system of systems* (SoS) – a large number of different modelling methods are composed together depending on the analysis task, the required and reasonable model accuracy and uncertainty, the available level of input information (growing over the project time), the system domains and so forth. Thereby, cyber-physical simulation models make often use of continuous time methods to model the physics, like MODELICA or different *computational fluid dynamics* (CFD) and *finite element method* (FEM) solvers. For the the control and information flow (the cyber part), discrete time simulation methods might be preferred while the behavioural model may be realized by *agent-based method*s (ABMs). Consequently, in order to compose a complete cyber-physical co-simulation, this requires to interconnected models from different domains – typically with different time constants – such that they work seamlessly together with acceptable performance and numerical robustness.

A feasible solution, shown by Fritzson et al. [5], is to use the *transmission line method* (TLM), a modelling approach which decouples sub-models using physically motivated time delays. This ensures numerical stability, and is especially suitable for real-time simulations. TLM can also greatly reduce simulation time, as was shown by Braun et al. [6] and Sjölund et al. [7]. With such strategies, making advantage of the model method specific advantages and still at the same time allowing the user to use the most suitable modelling method for the problem at hand, significant improvements to the design process, time and user-friendliness can be achieved without the need to develop new modelling tools.

Connecting inherently different simulation tools is facilitated by the *functional mock-up interface* (FMI), a tool independent standard for co-simulation and exchange of dynamic models supported by more than 100 simulation tools [8]. FMI can be used in conjunction with the the upcoming companion standard *distributed co-simulation protocol* (DCP), which facilitates communication and integration of models and real-time systems [9]. The OMSimulator master simulation tool currently supports FMI and TLM using TCP/IP connection to external tools. Possibilities of combining FMI with TLM was investigated in [10].

2.3 Model Fidelity, Complexity and Characteristic

Both terms, model complexity and model/simulation fidelity[1] are vague and no cross-domain application-independent valid standard has been establish so far (see e.g. definitions by [12, 13]) which to a large extend depends on the vast number (and weightings) of *fidelity criteria* [14–16]. Both, model complexity (in terms of the size of the overall design space) and model refinement can be expressed by the design information entropy [17].

The concept of abstraction – thus the model method and modelling fidelity – has to fit both the analysis needs *and* the systems characteristic. At a glance, any complex system model can be split-up into an structural and an behavioural part [15]. More in detail, refined taxonomies of the system's (or SoS's) characteristics like Gideons et al. [18] can be used. More in detail, the systems characteristic can be defined by the target system's properties. A selection of relevant properties to describe the characteristic type of an system is given in fig. 1.

2.4 Model Type by System Properties

Most complex systems incorporate some kind of a behavioural model part which can result in a stochastic behaviour. This behaviour can be anything from an easy control system (e.g. fuel tank filling/emptying sequence), a complex control system (e.g. an autopilot or other driver assistance systems) as well as any user-interactions on or within the system (such as pilots, flight controller, etc.). On a higher level, these parts of the system are often responsible for the way of operation of a system including more complex tasks such as operational strategies, tactics and doctrine. As a consequence, there is a unique combination of the model fidelity for each system depending on the system type (stated by the system properties shown in fig. 1). Figure 2 shows this differences in the desired degree of detail of the physical- and the behavioural model of three different systems.

In reality however, the desired direction of refinement is not reached as straight, smooth and continuous as indicated by the arrows for the three systems. Model refinement occurs instead in discrete steps, often within a single domain only by either refining an existing model or replacing a modelling technique with another abstraction method (and thereby most probably also replacing a tool within the co-simulation framework by another). A typical path of model refinement (of system b) is indicated in fig. 2 by the black stars. Using a priori a high-fidelity model techniques already from the beginning (in order

---

[1] Fidelity definition by [11]: "The degree to which a model or simulation represents the state and behaviour of a real world object or the perception of a real world object, feature, condition, or chosen standard in a measurable or perceivable manner; a measure of the realism of a model or simulation."
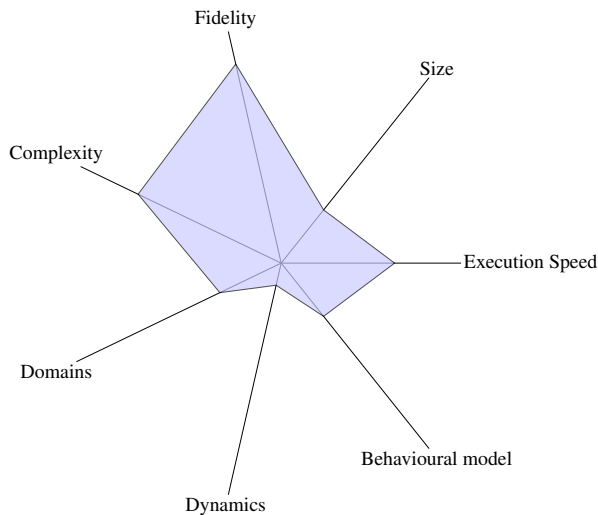
Figure 1: Radar plot of seven selected properties describing the type/characteristic of an system .

to skip the later tool transition) is often not applicable due to the absence of information to create the model and the usually lower execution speed of higher-fidelity models.
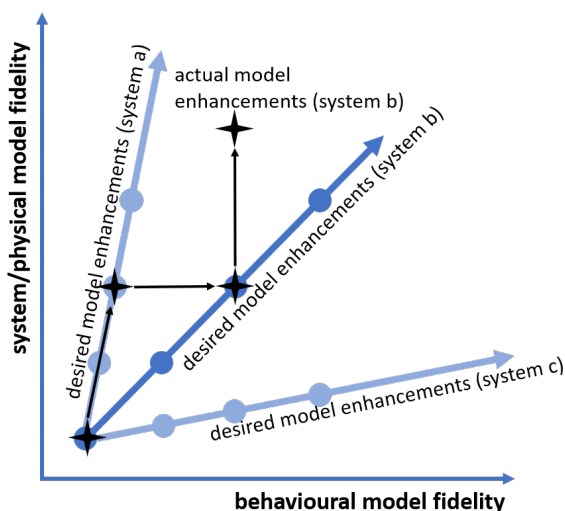

Figure 2: The ways of desired and actual model refinements within projects of different model type (characteristic).

2.5 HMI and the Need of a Real-time System

The above mentioned relationship of matching domain-specific models within a co-simulation environment means that each model domain has to be enhanced in fidelity thus reaching a certain level of certainty (or better said, reducing the level of uncertainty). The model fidelity of the physical part is rather simple to address: the model has to be able to capture the correct physics of the problem at hand. On an aircraft example, this might be already true for an object-oriented component-based (e.g. pumps, pipes, actuator, generator, etc.) approach. But what about the behavioural (control) model? The latter – often implemented in ABM – becomes already on moderate fidelities very complex, especially

if simulations also include the handling of abnormal situations such as component failure simulations. Additionally with one or several *human-in-the-loop* (HITL), the behavioural model has to represent different control action levels ranging from doctrine, strategy, tactic, operations down to the respective control schemes on system element level.

3   Human-machine Interaction Modelling

Until today, the aspects of the human's physiology as well as psychology are only insufficient considered in modern simulation models. In addition, the human-machine interface of modern system must not be the real interface of the target system as indicated in fig. 3. Today's technologies often do not return the actual behaviour of a system back to the user but a human adjusted *virtual* reality. While these systems, often denoted as assistant systems, improve in many cases safety and comfort, it requires at the same time a good understanding of how the operator perceive these inputs and react on them. These is gaining on importance if modern assistant systems are applied. For example, a decoupling of the input control stick forces, as it is common in any control-by-wire setup on modern aircraft could give the pilot a false understanding of the actual aircraft state. On the opposite side, there is a necessity that the system is interpreting the actual user inputs correctly under all circumstances.

3.1   Human Psychological Modelling

A common approach used to incorporate human physiological and mental factors in simulations is by treating the human as a machine-like stimulus-response (transfer function) system although this approach does not correspond to reality at all [19]. People's individual differences such as age, sex, physical fitness, ability to take decisions, reaction time, capacity for remembering, motivation, social interactions and creativity to solve a problem, require a much more appropriate representation of the human in simulation models.

To adequately forecast the effect of human performance in a complete human-machine simulation model, dynamic and integrated computational physiological and cognitive modelling, as depicted in fig. 3, is required. Even though this is the right approach to tackle the problem, it remains challenging to develop reliable models of physiological functions or cognitive architectures [20, 21]. Problems such as validation, scaling or model simplification must be solved first before these methods can meet their full potential [22, 23].

3.2   The Need of Humans in System Simulation

Although complete human modelling is in many cases still at a basic research level, this does not mean that the human-machine interactions cannot be simulated in a sufficient manner. An alternative solution is the traditional HITL approach where actual humans interact in real-time with simulation models. A major drawback of the HITL method is the need for access to probands, which can be in particular difficult when highly trained people are needed, such as pilots. Another drawback of HITL is the exact reproducibility of simulations. While a physiological/cognitive model could provide
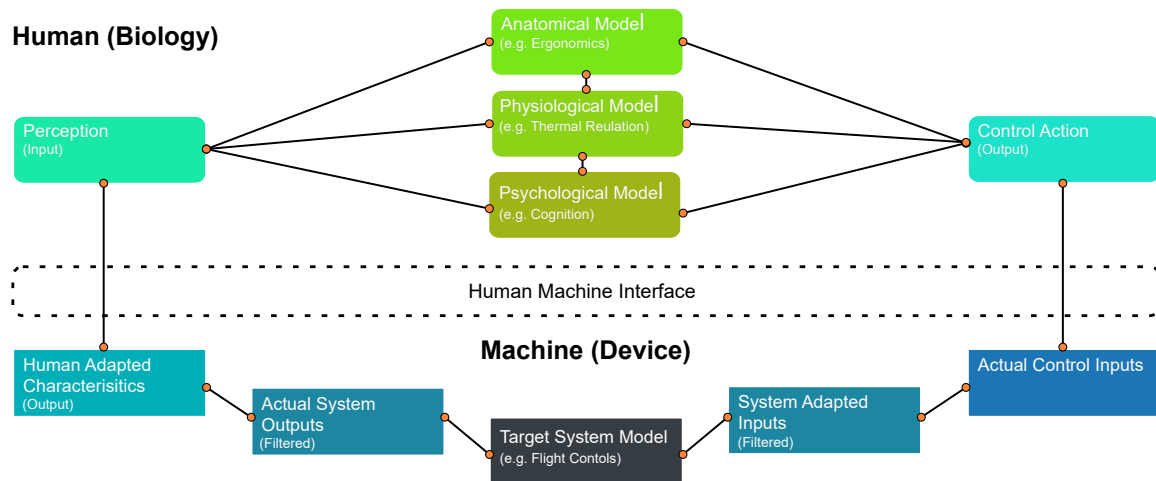
Figure 3: Human integration into the cyber-physical system context.

the same output under equal boundary conditions, is this for a HITL approach hardly to reach resulting in stochastic events. Furthermore, applying HITL, simulation speed is limited to real-time only, whereby this approach becomes inapplicable for high-fidelity co-simulations, which run often on much slower computational speed, low-fidelity simulations that run multiple times faster than real-time and any kind of optimization task that requires an enormous amount of simulation runs.

Nevertheless, no matter which of the two previous mentioned approaches is chosen, a realistic feedback from the human (model) back to the machine model can only be provided if the environment is realistic and rich in detail. A sufficient representation of the system's external and internal environmental conditions should apart from visual and aural information also include characteristic such as forces, thermal loads or odours. This high environment fidelity is necessary since the humans' reaction to this inputs are deeply context specific [14]. Integrating a flight simulator in a dynamic human-machine co-simulation can contribute to provide such a required environment.

4 Implementation Strategy & Realized Use Case Example

The following example shows a test case implementation aimed to assess a co-simulation setup of a flight simulator environment including aircraft on-board systems simulation to enable a HITL simulation and pave the way for a future human-model extension. Since the project aims to illustrate the possibilities of real-time co-simulation, the focus will be on the co-simulation integration of the example models into `X-Plane` rather than on the on-board system simulation models themself.

4.1 Related Work

A research project at Nanchang University developed a flight control model in `Simulink` for a helicopter implemented in `FlightGear` [24] including a detail description of the applied

communication protocol messages setup. The result was a simple data collection and a verification process illustrating the possibility of a real-time flight simulation including external tools. A somewhat similar project was published by the Instituto Tecnológico de Aeronáutica (ITA) which aimed to simulate a quadrocopter using `MATLAB/Simulink` and the flight simulator `X-Plane` [25]. Here the (hover) control system of the quadrocopter was implemented in `Simulink`, while the vehicle's physical properties were modelled within the flight simulator environment. At Vives University College the open source platform `FlightGear` had been used in a conceptual *unmanned aerial vehicle* (UAV) design project aiming to execute electro-thermal analysis in Simulink by applying HITL simulations [26]. Also other publications (such from A. Bittar et al. [27,28]) have utilized the `X-Plane` simulation environment for *hardware/power in the loop* (H/PITL) simulations incorporating `MATLAB/Simulink` with focus on new control/actuator concepts.
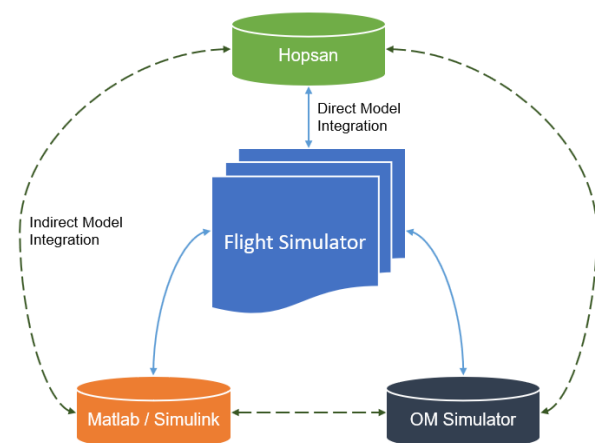

Figure 4: A simple flowchart illustrating the basic model exchange setup.

4.2 Framework Setup / Model Integration

Figure 4 shows the basic configuration of the realized co-simulation setup (exclusive the not yet realized `OMSimulator` part). The aircraft and the on-board systems are implemented in `Hopsan`, based on one example model slightly adapted to be co-simulated with `X-Plane`. It contains three main subsystems; propulsion-, control- and hydraulic actuation systems. A *six degrees of freedom* (6DoF) sub-/supersonic airplane simulation model from the standard component library is used to calculate the airplane dynamics (see component `6DOFSS` description in [29]).

On a Windows PC, `X-Plane 11` comes along with a *user datagram protocol* (UDP) communication to interact in pseudo real-time with other software. This enables co-simulations including aircraft subsystems and HITL together with `X-plane` (see fig. 5). UDP communication works with specific alloc-
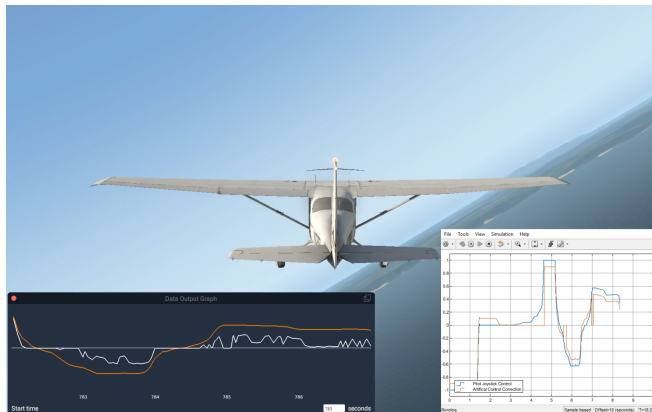

Figure 5: Screenshot of the user interface in X-Plane 11 with the X-plane datalogg to the left and the Simulink model parameter plot to the right.

ated ports from/to which messages, known as datagrams or packets, are sent or received [30]. `X-Plane 11` supports up to 138 output variables via the UDP connection. Each selection can consist of multiple variables. A complete list can be found in the `X-plane` UDP package description [31]. `Hopsan` models can be co-simulated in `Simulink` by con-
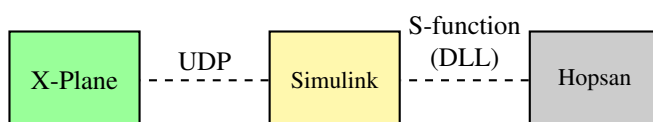

Figure 6: Communication setup of the simulation. Simulink communicates with X-plane using UDP sockets. Hopsan models are imported to Simulink as S-functions.

verting them into S-functions making use of executable, pre-compiled `Hopsan` *dynamic-link library*s (DLLs), see fig. 6. In order to establish the input and output ports of the `HOPSAN` model, S-function interface components must be used for the variables of interest. This allows `Hopsan` models to be integrated into `X-Plane` indirectly via `Simulink`, which supports UDP.

4.3 Model Analysis and Performance

The described UDP communication based framework has been tested in three different setups to identify the performance and bottlenecks of this integration method. To get valid results, the test have been performed on two different standard PCs running on Windows 10.

I. Pilot control test:
Time step delays and signal loss during transmission of data between the tools has been tested with a model setup of having once the pilot on the controls, recording all inputs. Afterwards, the model is fed with the recorded pilot commands again (open loop control).
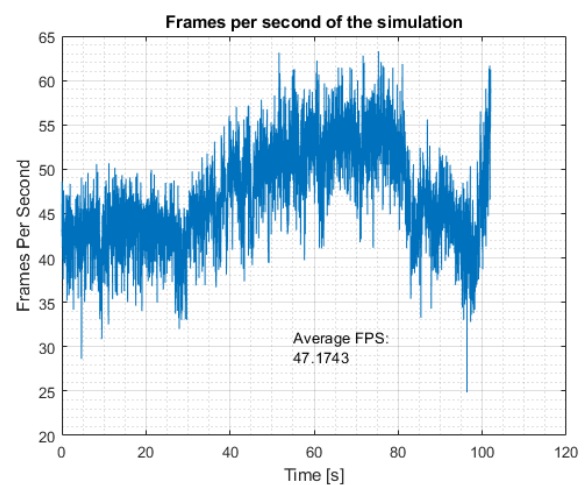

Figure 7: The frames per second model update frequency of the X-Plane model.

Figure 7 shows the performance of the `X-Plane` flight model in terms of *frames per second* (FPS). The average FPS rate during the simulation period is 47.2 with a significant fluctuation between 63 and 25 FPS. This is circa four times slower than the fixed calculation step size of 0.005 [s] or 200 [Hz] of the `Hopsan` model. The performance on the second PC was in the same range with an average FPS of 57 and even higher extrema (min: 20; max: 72).

II. Flight manoeuvrer model:
In this setup, the aircraft model including the subsystems, the flight controller and the mission controller were implemented in `Hopsan`. Consequentially, `X-plane` was only used in this setup to provide a graphical user interface, ensure (pseudo) real-time clocking. The model showed no evidence of package losses, however some flight controller parameter tuning problems arose. The reason of this model instability could not be found but may relate to the lesser update rate of the flight simulator (ca. 50[Hz]) than the simulation model time-steps (200 [Hz]). Also, round-off errors during the communication chain (due to truncations because of the different parameter normalizations in the applied tools) may have contributed to this model behaviour.

DOI
10.3384/ecp19162012
     Proceedings of the 10th Aerospace Technology Congress
October 8-9, 2019, Stockholm, Sweden
     116

III. Pilot-in-the-loop simulation:

In this test-setup, the pilot is in the loop giving command inputs (via joystick, pedals and power lever) to the flight controller which in this setup is a simple wing leveler only. The flight controller is implemented in `Simulink` while the aircraft is modelled within `X-Plane`.

Figure 8 shows the aileron signal of the pilot control and the correction from the flight controller. It can be noticed that the flight controller signal calculated in `Simulink` corrects the aircraft to level flight by compensating with additional aileron commands. The `Simulink` model is able to compute and transmit the control signal required to restore the aircraft to level flight even with sudden, high amplitude inputs (see fig. 8 time range between 25 to 35 seconds). Due to the
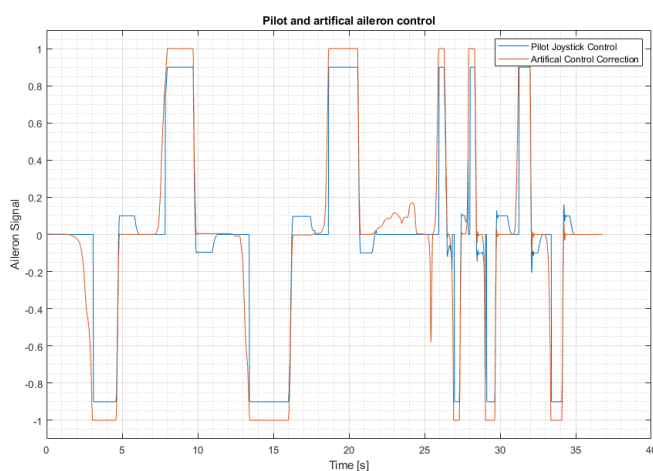

Figure 8: Aileron control signal from the pilot (red) and the Simulink roll control model (blue).

slightly higher calculation effort of the `Simulink` model in this setup, the update rate went down to 39.6 FPS in average (min: 32; max: 42.5).

4.4 Future Work

The above described implementation work is only the starting point towards a holistic scalable co-simulation model including HMI aspects. Future work (specific to this project) is going to address the following topics:

- (project specific) implement UDP communication in `Hopsan` and test the performance without `Simulink/MATLAB` in the simulation framework
- comparative studies on the existing UDP-based co-simulation framework with implementation strategies making use of a integration environment (including design process control, automation and optimization functionality) such as `RCE` [32], `OMSimulator` [33] or `modeFRONTIER` [34].
- investigate useful methods/tools to integrate H/PITL simulations into the framework, eventually making use of `dSPACE` or `LMS Amesim` (commercial tool vendors).

Figure 9 shows the planned future setup using the `OMSimulator` framework together with UDP, DLL and
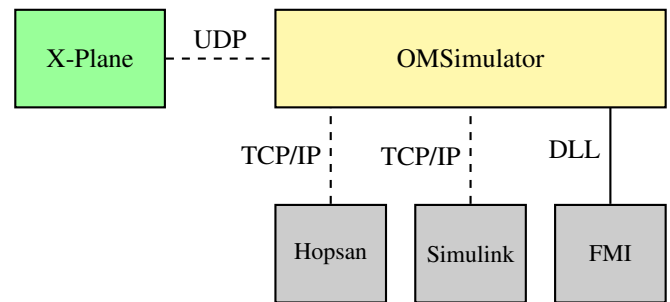
TCP/IP communication.

Figure 9: Possible future communication setup, using OM-Simulator as the master communication unit. Other tools can be connected to OMSimulator using TCP/IP sockets, and FMI units can be imported directly.

5 Conclusion

This paper shows that it is possible to execute a flight simulation with the human (pilot) in the loop including detailed on-board system simulations down to hardware component level (such as actuators, power electronics and fuel-pumps) by co-simulation on a standard PC.

In the presented test implementation, no dedicated co-simulation environment tool has been used. Instead, a work-around using `Simulink` with its multiple interface options (here UDP and S-function/DLL) were realized. In this framework, the flight simulator works as the real-time engine and the graphical (and haptic) user interface while most simulation model calculations were performed in the co-simulated cyber-physical simulation tool.

To enable easier model handling capabilities like motivated in section 2.1, a dedicated co-simulation environment as sketched in fig. 9 should be used. This enables also other benefits such as model stability and high execution speed through its model decoupling capability as shown in section 2.2.

The typical drawbacks of a human-in-the-loop simulations has been named in section 3.2. With the limitation to real-time execution (beside other reasons), it is not suitable for many tasks during the development work. Especially optimization task that require multiple simulations cannot be executed with human-in-the-loop setups. Remaining future work to reach complete full-system simulations including simulated human behaviour is mainly needed for the psychological model, but also other aspects such as the wished continuous model refinement are not in place yet. The authors of this paper therefore suggest to work further towards universal system characteristics and model fidelity and complexity descriptions as suggested in fig. 1 and section 2.3-2.4. Such a normalization could lead to simulation guidelines that would help the product developer to choose the right simulation and simulation integration approach(es) for the problem at hand without the need of being an expert in all included model domains.

DOI
10.3384/ecp19162012   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **117**

References

[1] I. Staack, K. Amadori, and C. Jouannet. A holistic engineering approach to aeronautical product development. *The Aeronautical Journal*, page 1–16, 2019.

[2] P. D. Ciampa and B. Nagel. Streamlining cross-organizational aircraft development processes: an overview of the AGILE project. In *Proceedings of the 31th ICAS Congress*, Belo Horizonte, Brazil, 2018. International Congress of the Aeronautical Sciences (ICAS).

[3] Lennart Ochel, Robert Braun, Bernhard Thiele, Adeel Asghar, Lena Rogovchenko-Buffoni, Magnus Eek, Peter Fritzson, Dag Fritzson, Sune Horkeby, Robert Hällquist, Åke Kinnander, Arunkumar Palanisamy, Adrian Pop, and Martin Sjölund. OMSimulator - integrated FMI and TLM-based co-simulation with composite model editing and SSP. In *Linköping Electronic Conference Proceedings*, pages 69–78. Linköping University, 02 2019.

[4] Petter Krus. Complete aircraft simulation for distributed system design. In *Proceedings of the International Conference on Recent Advantages in Aerospace Actuation Systems and Components*, 2001.

[5] Dag Fritzson, Robert Braun, and Jan Hartford. Composite modelling in 3-D mechanics utilizing transmission line modelling (TLM) and functional mock-up interface (FMI). *Modeling, Identification and Control*, 39(3):179–190, 2018.

[6] Robert Braun and Petter Krus. Multi-threaded distributed system simulations using the transmission line element method. *SIMULATION*, 92(10):921–930, October 2016.

[7] Martin Sjölund, Robert Braun, Peter Fritzson, and Petter Krus. Towards efficient distributed simulation in modelica using transmission line modeling. In *3rd International Workshop on Equation-Based Object-Oriented Modeling Languages and Tools; Oslo; Norway; October 3*, number 47, pages 71–80. Linköping University Electronic Press, 2010.

[8] T. Blochwitz, M. Otter, M. Arnold, C. Bausch, C. Clauß, H. Elmqvist, A. Junghanns, J. Mauss, M. Monteiro, T. Neidhold, D. Neumerkel, H. Olsson, J.-V. Peetz, and S. Wolf. The Functional Mockup Interface for tool independent exchange of simulation models. In *8th International Modelica Conference 2011*, Como, Italy, September 2009.

[9] Modelica Association Project DCP. *DCP Specification Document, Version 1.0*. Modelica Association, Linköping, Sweden, 2019.

[10] Robert Braun, Robert Hällqvist, and Dag Fritzson. TLM-based asynchronous co-simulation with the functional mockup interface. In *IUTAM Symposium on Solver-Coupling and Co-Simulation*, Darmstadt, Germany, September 2017.

[11] DMSCO. M&S glossary - terms and definitions. [Online; accessed 2019-08-11], 2019.

[12] Manfred Roza, Jeroen Voogd, and Paul va. Fidelity considerations for civil aviation distributed simulations. In *Modeling and Simulation Technologies Conference*, number AIAA-2000-4397, 2000.

[13] Edward Burnett. A proposed model fidelity scale. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit*, number AIAA-2008-6689, 2008.

[14] Dennis A. Vincenzi, John A. Wise, Mustapha Mouloua, and Peter A. Hancock. *Human Factors in Simulation and Training*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 2008.

[15] T. Caughlin, Donald and F. Sisti, Alex. Summary of model abstraction techniques. In *Proceedings of the SPIE 3083*, volume 3083. The International Society of Photo-Optical Instrumentation Engineers SPIE, 6 1997.

[16] T. Mowbray, Carol, C. Holter, Mark, B. Teague, Gregory, and Anddeborah Bybee. Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, 24:315–340, 09 2003.

[17] Petter Krus. Information entropy in the design process. In Amaresh Chakrabarti and Raghu V. Prakash, editors, *ICoRD'13*, pages 101–112, India, 2013. Springer India.

[18] James Gideon, Cihan H. Dagli, and Ann K. Miller. Taxonomy of systems-of-systems. In *Conference on Systems Engineering Research*, pages 356–363. Institute of Electrical and Electronics Engineers (IEEE), 2005.

[19] Reno Filla. *Quantifying Operability of Working Machines*. PhD thesis, Linköping UniversityLinköping University, Fluid and Mechatronic Systems, The Institute of Technology, 2011.

[20] Alessandro Vinciarelli, Anna Esposito, Elisabeth André, Francesca Bonin, Mohamed Chetouani, Jeffrey F. Cohn, Marco Cristani, Ferdinand Fuhrmann, Elmer Gilmartin, Zakia Hammal, Dirk Heylen, Rene Kaiser, Maria Koutsombogera, Alexandros Potamianos, Steve Renals, Giuseppe Riccardi, and Albert Ali Salah. Open challenges in modelling, analysis and synthesis of human behaviour in human–human and human–machine interactions. *Cognitive Computation*, 7(4):397–413, Aug 2015.

[21] M. Viceconti, D. Testi, F. Taddei, S. Martelli, G. J. Clapworthy, and S. V. S. Jan. Biomechanics modeling of the musculoskeletal apparatus: Status and key issues. *Proceedings of the IEEE*, 94(4):725–739, April 2006.

[22] Alex Kirlik. Conceptual and technical issues in extending computational cognitive modeling to aviation. In *International Conference on Human-Computer Interaction*, pages 872–881. Springer, 2007.

[23] Marc Halbrügge. Evaluating cognitive models and architectures. 2007.

[24] Jin Ying, Huazhu Luc, Jiyang Dai, and Haoman Pan. Visual flight simulation system based on Matlab/FlightGear. In *Proceedings of 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2017*, pages 2360–2363, 2017.

[25] Helosman V. Figueiredo and Osamu Saotome. Simulation platform for quadricopter: Using Matlab/Simulink and X-plane. *Proceedings - 2012 Brazilian Robotics Symposium and Latin American Robotics Symposium, SBR-LARS 2012*, pages 51–55, 2012.

[26] Yves C.J. Lemmens, Tuur Benoit, Rob De Roo, and Jon Verbeke. Real-time Simulation of an Integrated Electrical system of a UAV. *SAE Technical Paper Series*, 1, 2014.

[27] Adriano Bittar and Neusa Maria Franco De Oliveira. Hardware-in-the-loop simulation of an attitude control with switching actuators for SUAV. *2013 International Conference on Unmanned Aircraft Systems, ICUAS 2013 - Conference Proceedings*, pages 134–142, 2013.

[28] Adriano Bittar, Helosman V. Figuereido, Poliana Avelar Guimaraes, and Alessandro Correa Mendes. Guidance software-in-the-loop simulation using x-plane and simulink for UAVs. *2014 International Conference on Unmanned Aircraft Systems, ICUAS 2014 - Conference Proceedings*, pages 993–1002, 2014.

[29] HOPSAN. HOPSAN - multi-domain system simulation tool. [Online; accessed 2019-07-19], 2019. https://github.com/Hopsan/hopsan.

[30] UDP. User datagram protocol. [Online, accessed 2019-07-18], 2019. https://searchnetworking.techtarget.com/definition/UDP-User-Datagram-Protocol.

[31] X-Plane. X-Plane UDP data set output table. [Online, accessed 2019-05-22]. https://www.x-plane.com/kb/data-set-output-table/.

[32] RCE. Remote component environment (RCE): distributed, workflow-driven integration environment. [Online; accessed 2019-07-19], 2019. https://rcenvironment.de/.

[33] OMSimulator. OMSimulator - FMI-based co-simulation tool. [Online; accessed 2019-07-19], 2019. https://github.com/Hopsan/hopsan.

[34] modeFRONTIER. modeFRONTIER - multidisciplinary design optimization platform. [Online; accessed 2019-07-19], 2019. https://www.esteco.com/modefrontier.

Initial Flight Simulation Testing of a Gesture-based Interface for Flight Control

M. A. Rodríguez Díaz, E. Villani, D. Arjoni, P. Krus*

Aeronautics Institute of Technology, São José dos Campos, Brazil.
E-mail: manuelrd@ita.br, evillani@ita.br, diegoarjoni@gmail.com, petter.krus@liu.se
*Linköping University, Linköping, Sweden.

Abstract

As our knowledge in aeronautics has improved, so have aircraft. Modern flight control systems can use electronic systems to greatly enhance the controllability, efficiency and safety of even aerodynamically unstable aircraft, however the man-machine interface method has remained the same since the early days; sticks and pedals. This paper depicts the very first simulated test flights on a prototype flight control interface for manned aircraft under development, this interface uses the pilot's eye focus direction (known as "gaze") and hand movements to control the aircraft in flight by just looking into a desired direction, or by having the aircraft to "imitate" her or his hand like one would do to illustrate an aircraft flight maneuver: an open hand with thumb and pinky finger extended like "wings" to imitate an aircraft motion in flight, except that this time is the other way around, with the aircraft following the hand rolls and climbs as actual commands. Although Gaze-guided and Hand gesture-guided flight control has been proven feasible using this system, only the initial operative experience is covered, with a general overview on the systems functionality and of data collected, this exploratory research does not intends to result in a immediately-applicable system, rather as proof of concept of the interface method of the proposed system and as foundation for further development.

Keywords: Flight Controls, Eye Tracking, Body Tracking, Control Systems.

1 Introduction

Along more than one century of research and development in Aeronautics, many innovations have been introduced. Examples are the use of composite materials and the introduction of fly-by-wire control systems [1].

Fly-by-wire systems have made possible the use of new aerodynamic design configurations, with different shapes, purposes and capabilities [2]. Revolutionary as these aircrafts might be, a proficient pilot from an early biplane would have no issue controlling them. This is possible because he would still use a thrust lever for controlling power, a yoke or side stick for controlling pitch and roll, and a pair of pedals for controlling yaw. The main aircrafts control interfaces have not changed in concept since the first heavier-than-air airplanes.

Despite advances in technology and the introduction of fly-by-wire systems, aircraft control interfaces have remained largely unchanged. Current interfaces are limited to 3 degrees of freedom for each control as a maximum. Most commonly, they use only one or two degrees of freedom, be it a stick or a pedal.

A vehicle like a rotorcraft can move in 6 degrees of freedom in normal conditions. New control interfaces using current technologies could provide more controllability of the craft than existing interfaces.

They could also reduce the training required to operate a flying vehicle. This last point can be of particular interest for the personal flying vehicles current under development by major aerospace corporations [3] [4] [5]. While their conception calls for them to be self-piloted, the capability for a human to intervene in case of a navigation failure could be important for safety reasons.

Considering the above motivation, this paper presents the development of new control interfaces that do not require yokes, sticks or pedals for inputs on the aircraft roll, pitch, yaw and thrust controls. Instead, it explores the use of body motion and visual focus direction.

The paper is organized as follows. Section 2 presents the architecture of the proposed system. Section 3 presents the flight control modes, Section 4 describes the experimental simulated flights that were performed to collect data, while Section 5 Presents the results from these experiments. Finally, Section 6 presents the conclusions.

2 The System Architecture

The proposal presented in this paper uses two types of signal inputs for controlling the aircraft: **body motion** and **visual focus direction**.

In the first case, the movements of the pilot's hand are captured using components from the **HTC Vive** virtual reality kit. This equipment relies on a system of rotating laser beam base stations ("lighthouses", as these are known) to perform spatial tracking. It also uses accelerometers and gyroscopes to perform dead reckoning between laser beam sweeps [6].

The visual focus, referred to as "gaze", is captured using **Tobii Pro Glasses 2.** These eye-tracking glasses rely on retina geometry detection with infrared cameras to obtain a gaze vector for each eye. The resulting vectors from both eyes can be intercepted to obtain a spatial focus point or the individual vectors can be used separately [7].

The general architecture of the proposed system is presented in Figure 1. The 'model computer' (the computer system where the system's software is executed, it is referred to simply as "computer" in the next diagrams) receives information about the body gestures and gaze direction, as well as the aircraft states. Based on this information, it determines the aircraft control inputs.
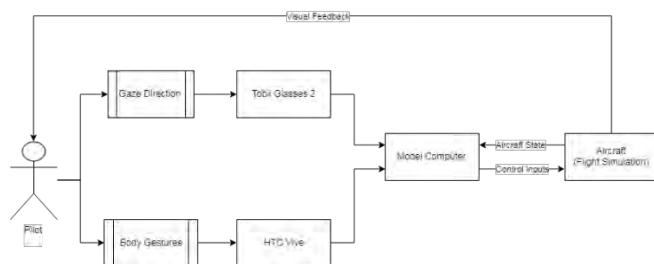

Figure 1: General diagram of the proposed system.

Next, we describe in detail each of the proposed aircraft interface.

2.1 Eye Tracking System

The Pro Glasses 2 measures the pilot gaze direction at a rate of 25 Hz. It also provides the glasses acceleration and angular rates, measured by *a Micro-Electro-Mechanical System* (MEMS) *Inertial Measurement Unit* (IMU).

These variables are sent, via wi-fi, to the Tobii *Software Development Kit* (SDK). It is then packaged in an *American Standard Code for information Interchange* (ASCII)-coded string and sent using the *Transfer Control Protocol* (TCP) to the model computer. There, a control system implemented in Simulink interprets and processes the signals and generates a command output for the aircraft. The diagram in Figure 2 illustrates this.

The gaze vector is a two-component vector comprising the horizontal gaze and the vertical gaze component. Both components vary from 0 to 1 depending on the pilot's focusing direction. The (0,0) coordinate is at the lower left

corner of the pilot's field of view through the glasses, while the (1,1) coordinate is at the top right corner.
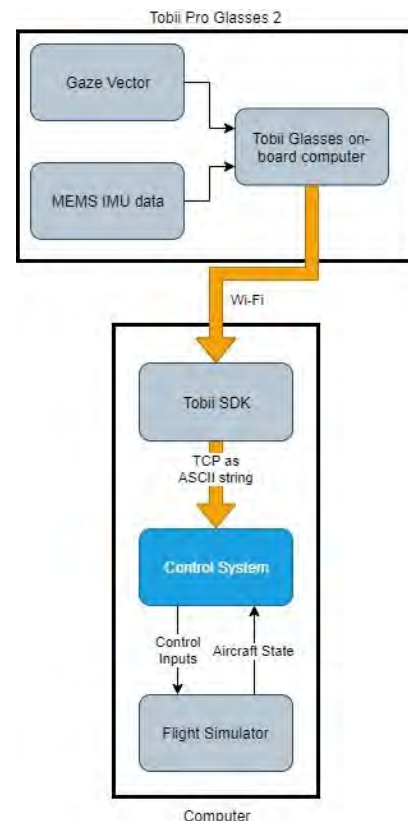

Figure 2: Gaze- based configuration architecture.

The gaze vector is fixed to the glasses reference system. Therefore, an additional reference is needed to convert this to a frame of reference aligned with the horizon. For this purpose, the IMU data is used to sense the orientation of the glasses with respect of the Earth's gravity. Combining the IMU data with the gaze vector provides us a gaze vector that is relative to the horizontal plane of the Earth instead of being relative to the glasses, irrespective of their inclination or orientation. Figure 3 illustrates the signal path in the control system, which is the blue block in Figure 2.

The current configuration allows the pilot to control the aircraft by looking into the desired trajectory. While focusing on the desire trajectory, he/she should press and release the 'command button'. This button assures that the gaze information is used to control the aircraft only when the pilot intends to do so.

When the pilot releases the command button, the system determines the gaze direction. It then processes it and defines the output for the aircraft' flight control surfaces, in order to guide the aircraft into the desired trajectory.

Before the first utilization of the system the pilot must define a "nose" vector for the aircraft. This is done by looking straight ahead to the display where the simulation will be shown and pressing a calibration button. The provided gaze vector is considered as a reference.

The above described operation mode is referred to as the "gaze-guided configuration" in the rest of the paper.
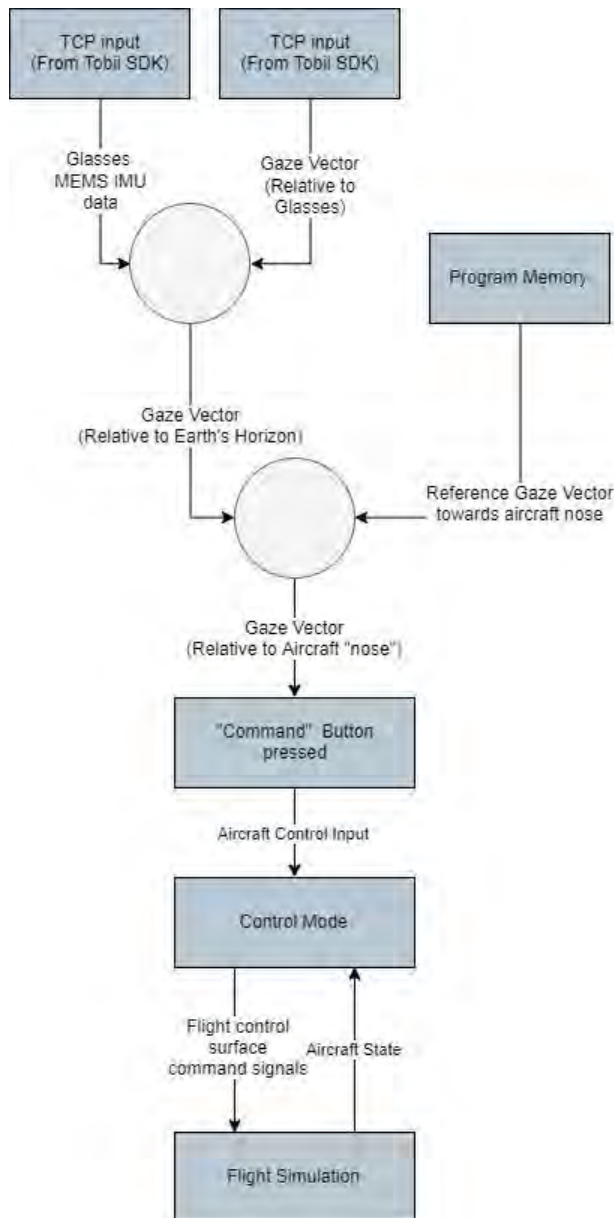
Figure 3: Gaze-guided configuration Control System.

2.2 Gesture Tracking System

The Vive is used at a 50 Hz tracking rate. Information from the lighthouse tracking stations is used to locate a Vive Controller in a reference frame. The horizontal axes of this frame are North-South and East-West. The vertical axis is in the Up-Down direction.

In order to obtain useful signals this reference is transformed to an arbitrary reference set by the pilot at the beginning of each usage of the system. This allows the roll input to be calculated using the longitudinal axis correspondent to the aircraft's own longitudinal axis instead of the North-South axis.

In the current setup, the controller is attached to the dorsal part of the pilot's hand eliminating the need to "hold" the controller. The Vive system then acquires tracking of this controller and feeds a 7-element vector to the SDK. This vector contains the three linear coordinates about the original

North-South East-West reference and the four quaternion rotation elements for the controller's current tracked state.

The SDK uses TCP to transmit this vector as a binary string to the Control System in Simulink, which performs the necessary transformations and calculates the input for the aircraft flight controls. This is illustrated in Figure 4.
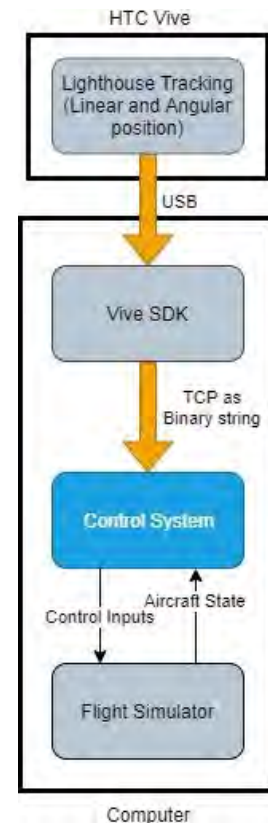

Figure 4: Hand-guided configuration architecture.

In the current configuration, the pilot "imitates" the aircraft using the hand, as a flight instructor would imitate aircraft rolls and climbs to illustrate maneuvers to students.

The system considers that the aircraft longitudinal axis extends along the open hand from the palm towards the middle finger. The nose of the aircraft is on the tip of this finger, while the left wing would extend towards the open thumb.

The system takes a roll of the hand as a roll input and a rise or lowering of the hand as a pitch input. This configuration will be referred to as the "hand-guided configuration" for the rest of the paper.

Like the gaze-guided configuration, the hand-guided configuration requires the pressing of a "command" button. The system reads the input only when the button is pressed and released. Figure 5 illustrates the signal path in the control system, which is the blue block in Figure 4.
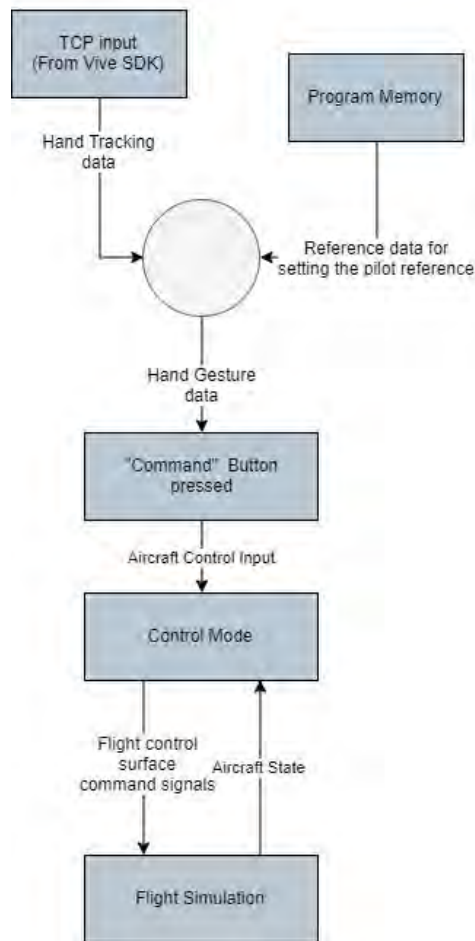
DOI
10.3384/ecp19162013   Proceedings of the 10th Aerospace Technology Congress
October 8-9, 2019, Stockholm, Sweden   **122**

Figure 5: Hand-guided configuration control system

3 Flight Control Modes

A series of input signal interpretation modes are implemented in the control system in order to produce meaningful flight control signals based on body gestures and gaze direction. This set of control modes are only for initial exploratory purposes and might not be the most efficient or the most comfortable control modes possible for the system. All the input signals that these modes receive vary from -1 to +1, where zero is a neutral input be it in roll or pitch.

Flight Path Angle Mode

This mode takes the processed input from either gaze or body tracking systems and makes this proportional to the aircraft's *Flight Path Angle* (FPA) or Flight Path Trajectory.

The input is passed through a gain that converts it to a target FPA for a controller to track it by commanding elevator deflection.

Flight Path Angle Change Mode

Similar to the previous mode, but it instead makes the new FPA to be a function of current FPA at the time of the command being issued. It adds an angle proportional to that of the received input signal.

Roll Angle Mode

In this case, the input signal is passed through a gain that converts it to a specific roll angle to either side. A controller will then track this reference generating aileron deflection signals.

Heading Change Mode

This mode allows for the pilot to look or point with the hand in a direction to either side of the aircraft's nose and set a new heading.

The system takes the current heading at the time of the command being issued and computes the angle between the nose of the aircraft and the direction that the pilot is gazing or pointing at. It then stablishes the new heading as the current heading plus this difference.

Finally, a controller tracks the set heading by generating outputs for the ailerons up to a specific maximum roll angle.

4 Experimentation

A series of simple simulated flights were performed to acquire an initial experience with the system and provide a better understanding of it. The information collected will be used to improve the system from its current state.

As a way of simplifying inputs, only either roll or pitch is controlled at a time during each flight. This allows the pilot to focus on a single type of input signal. It also results in a set of data that is more easily interpreted and analyzed for each configuration and control mode.

4.1 Experiments Set-up

Even though the system intends to explore the feasibility of a gesture-based system for piloting a manned aircraft, the initial experiment uses the *Flight Dynamics Model* (FDM) of a light cargo *Unmanned Aerial Vehicle* (UAV) for research time-efficiency purposes. Figure 6 provides a rendering of this aircraft.
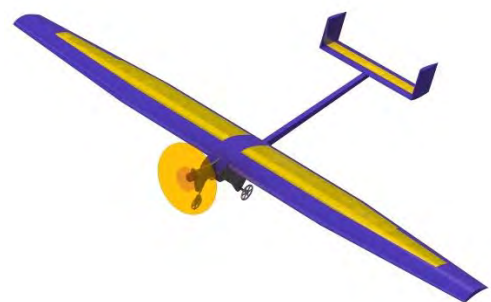

Figure 6: Rendering of the Aircraft model used during the experiment.

The Aircraft contained in this flight simulator had its FDM verified using wind tunnel and flight test data and is considered an appropriate representation of the real equivalent, it has Electric propulsion producing 40 Newtons of static trust, a mass of 15 Kilograms and a Wingspan of 3.2

meters. The fact that it is modeled in Simulink allows for a quick integration into the rest of the system.

Being a simulation, it is possible to place the pilot's perspective to be "inside" of the aircraft, therefore providing for the same type of perceptions as a manned aircraft. Further work on the next round of experiments is expected to be carried out using FDMs of manned transport or acrobatic aircraft.

The output spatial position and Euler angles representing the aircraft attitude in space is calculated in Simulink and sent to the FlightGear flight simulator. Flight Gear then provides a visual feedback of the aircraft's flight trajectory to the pilot.

The simulated flight tests presented in this paper take place near the Sao Jose dos Campos Airport (SBSJ) in Brazil.

4.2 Scenario 1: Aligned Approach for Landing

The aircraft begins the flight aligned with a runway at a distance of 2 km from the threshold. It has an altitude that places it near the typical 3-degrees glideslope for landing.

The pilot is tasked with controlling only the longitudinal axis of the aircraft in order to intercept and maintain the glideslope for landing as close as possible to the runway's aiming marks, as illustrated in Figure 7. The flight ends as soon as the aircraft contacts with the runway, recording the position where touchdown occurred for later analysis.

The roll and yaw in this flight are fixed to have no variation during the experiment while the airspeed is kept constant by an auto throttle system.
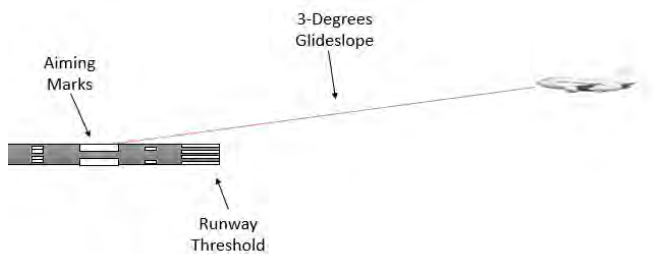

Figure 7: Aligned approach for landing.

The purpose of this flight is to evaluate the precision control of the longitudinal axis of the aircraft using both gaze-guided and hand-guided configurations.

The distance from the touchdown point and the aiming marks of the runway as well as the trajectory of the aircraft relative to the 3-degrees glideslope are used as indicators of performance. The closer that the touchdown occurs to the aiming point and the smaller the deviation from the glideslope are considered to indicate a higher performance.

4.3 Scenario 2: Non-Aligned Approach for Landing

The aircraft begins the flight 2.8 kilometers from the Runway threshold. It is at an altitude that allows it to align and then land on that runway on a 3-degrees glideslope. In this case the aircraft is heading towards the threshold at the beginning of the flight but is not aligned with the runway heading.

The pilot is tasked with controlling only the roll axis of the aircraft in order to intercept and maintain the runway heading while the vertical speed is adjusted to allow the aircraft to land on the runway.

If the pilot handles lateral guidance correctly, the flight ends when the aircraft touches down on the runway. At this point the position where touchdown occurred as well as the heading of the aircraft at that moment is recorded for later analysis.

Figure 8 illustrates the flight of this scenario, this flight was adapted from reference [8] where a similar scenario is presented for evaluation of a brain-controlled interface.
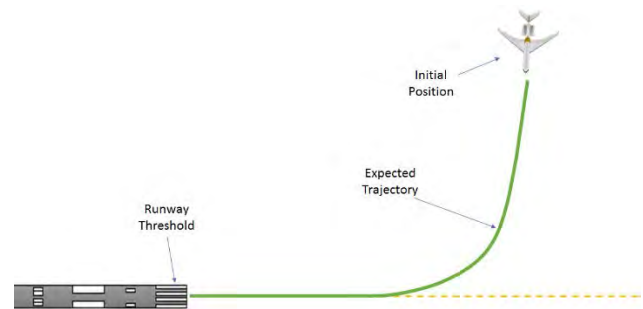

Figure 8: Reference flight for non-aligned approach.

The purpose of this flight is to evaluate the precision control of the aircraft in the roll axis using the proposed interfaces.

Touchdown position relative to aiming point and the aircraft's heading at the time of the touchdown are used as metrics for performance in this flight. Smaller deviations from expected touchdown point and runway heading at touchdown indicate better performance.

5 Results

The two scenarios described in section 4 where performed with the aircraft model of section 4.1.

The flights of Scenario 1 where performed with the pilot controlling the longitudinal axis only and is referred to as "Vertical Guidance". The flights of Scenario 2 were performed with the pilot controlling only the latero-directional flight and is referred to as "Horizontal Guidance" scenario. In each case the pilot used either the gaze or hand gesture guidance configurations.

Each flight scenario with each interface configuration was repeated 5 times by a single pilot, for a total of 20 flights performed with the intention of gaining insight of the system's operation.

5.1 Vertical Guidance

Firstly, the gaze-guided configuration was used with the FPA change mode. The results are presented in Figure 9, where the dashed red line represents the ideal glideslope.

It shows that the gaze-based configuration required constant corrections to be performed by the pilot. In some flights, the flight path is abruptly inverted. this problem could be related to the accuracy of the system calibration. Approaches to improve it are currently under investigation.
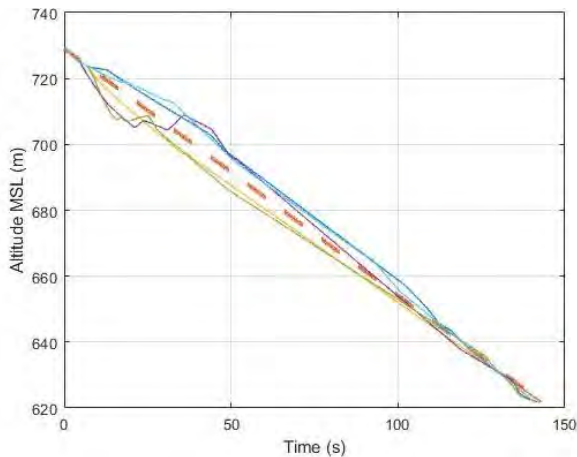
Figure 9: Altitude for five gaze-guided landings.

Next, the hand-guided configuration was used with the FPA proportional mode. Results are presented in Figure 10, where the dashed red line represents the ideal glideslope. The hand-based configuration allowed for a more precise control of the landing when compared to gaze-guided landing as can be seen in. It also presented a lower number of inversions of the flight path angle when compared to those on the gaze-guided configuration.
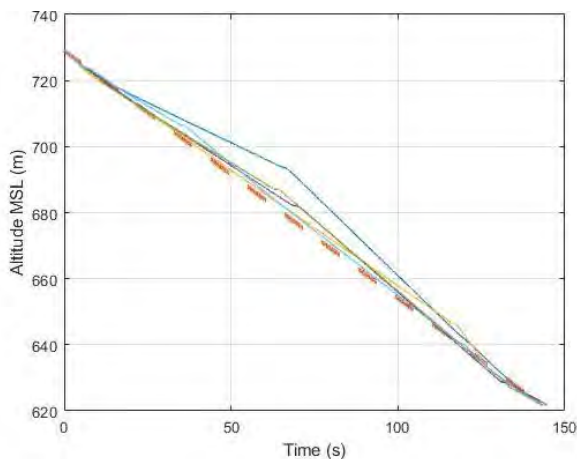

Figure 10: Altitude for five hand-guided landings.

Figure 11 shows an example of the FPA set by the pilot using gaze-guided input. Figure 12 shows the partially processed input of a hand-guided flight.
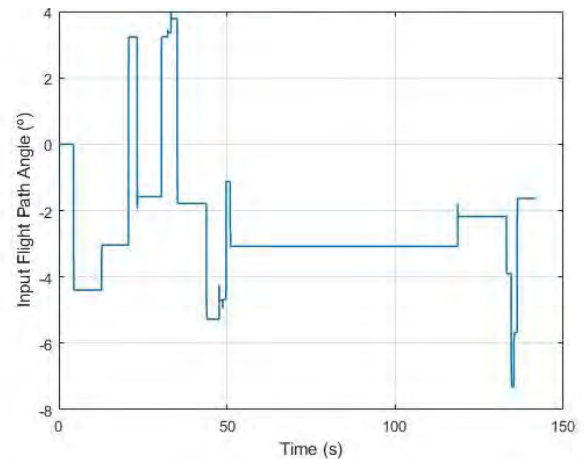
Figure 11: Gaze-guided FPA signal during one of the test flights.
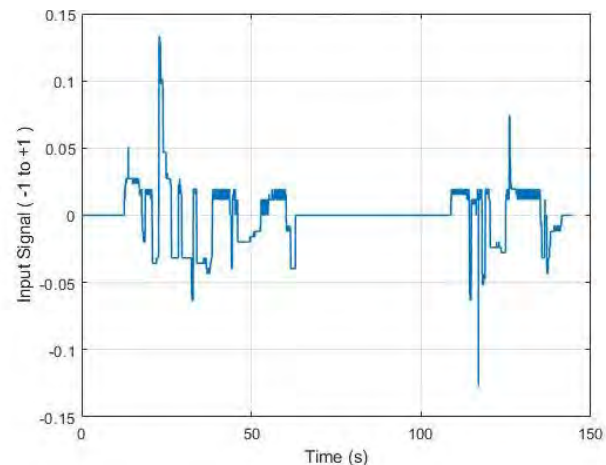
Figure 12: Hand-guided input signal before being converted to FPA during one of the test flights.

5.2 Horizontal guidance

The horizontal guidance flights had the limitation that the field of view on the computer's screen did not allow for visualization of the runway when the aircraft was perpendicular to it. Therefore, in both gaze-guided and hand-guided configurations, the pilot had to make turns into the runway to verify its position relative to it, and then return to its previous trajectory.

For this flight, the gaze-guided configuration used the heading change mode while the hand-guided configuration used the roll-proportional mode.

Both the gaze-guided and hand-guided configurations showed enough accuracy on the guidance of the aircraft into the runway heading to perform a landing on the aiming points of the runway. Figure 13 shows the trajectory for all five gaze-guided landings while Figure 14 presents the results for the hand-guided landings.
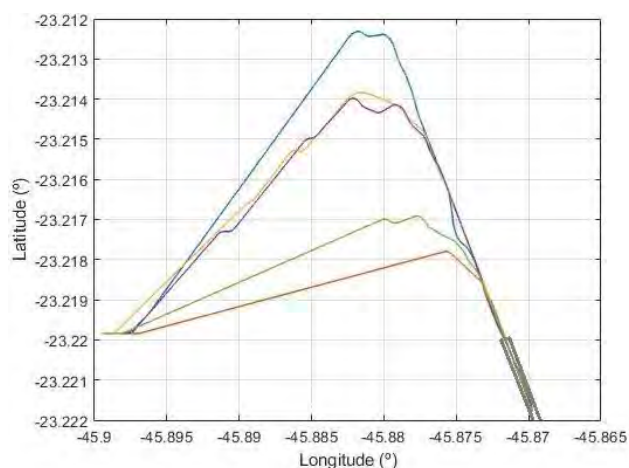
Figure 13: Horizontal trajectory of the aircraft during the gaze-guided landings.
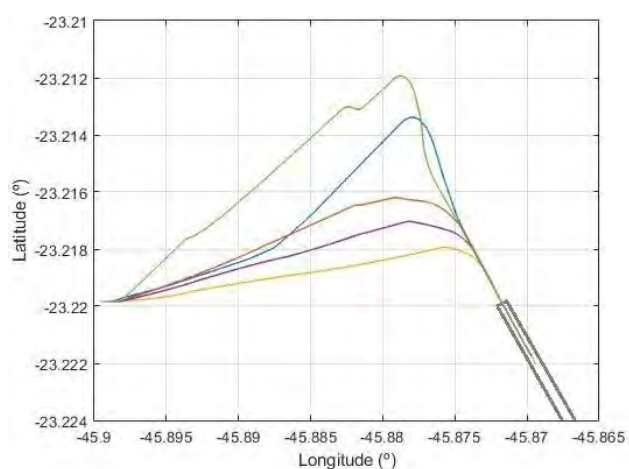
Figure 14: Horizontal trajectory of the aircraft during the hand-guided landings.

Signals recorded during these flights are presented in Figure 15 for the gaze-guided heading and in Figure 16 for the hand-guided configuration.
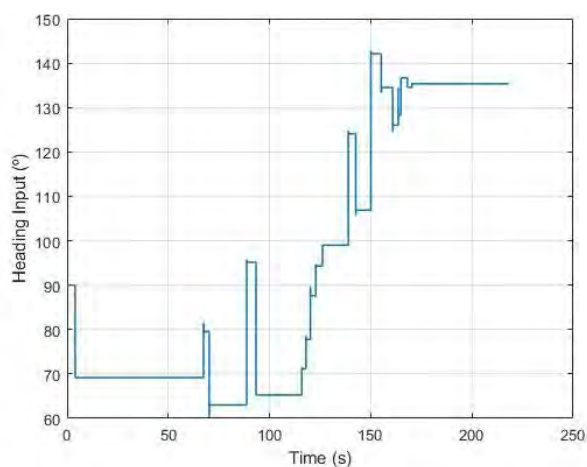
Figure 15: Gaze-guided heading input for landing during one of the test flights.
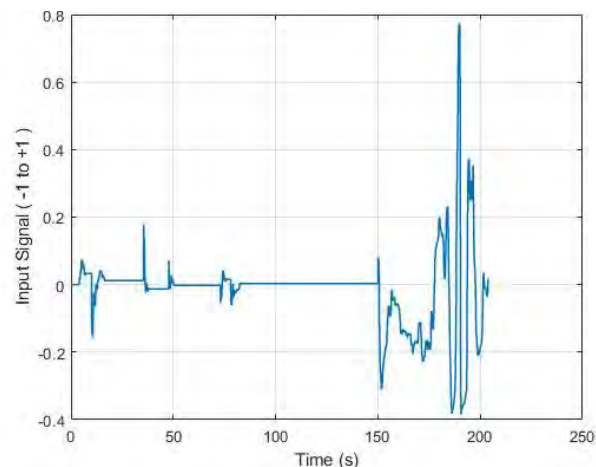
Figure 16: Hand-guided inputs before being converted to Roll inputs during one of the test flights.

5.3 Data Analysis

Figure 17 presents the error in the runway touchdown position when compared with the expected position, for the four cases analyzed in this paper. Although some variation can be observed in Figure 17, it is not possible affirm that one of them is significantly better than the other.
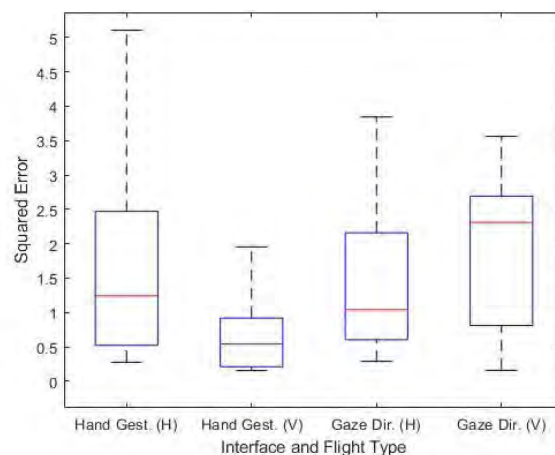
Figure 17: Touchdown position error for each case.

On the other hand, Figure 18 shows that in the current experiment the gaze-guided configuration presents a greater deviation from expected flight trajectories during approach (deviation from the expected glideslope) compared to those of the hand-guided configuration.
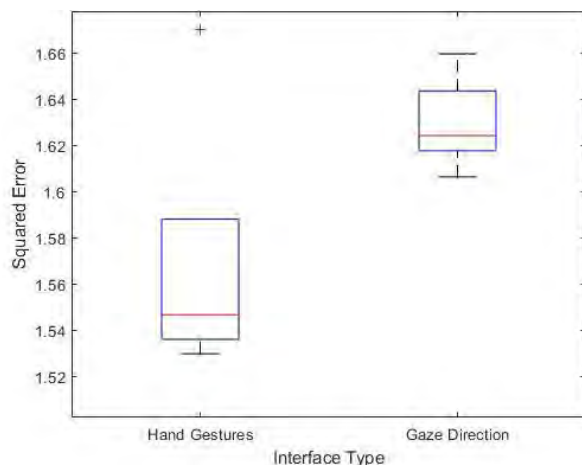
Figure 18: Cumulative glideslope deviation for the vertical guidance flight tests.

The final heading at touchdown also shows better performance with the hand-guided configuration as can be seen in Figure 19. This can also be attributed to the early development stage of the gaze-based configuration.
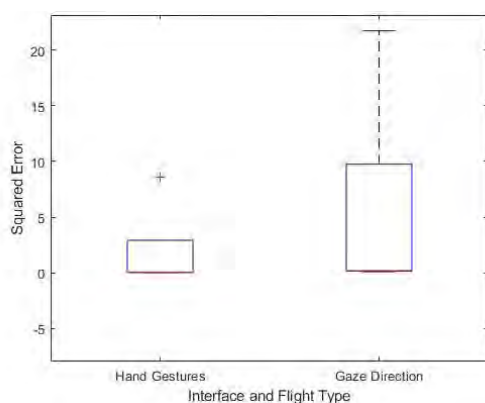
Figure 19: Aircraft heading error at touchdown for horizontal guidance test flights.

It is also clear from Figure 20 that the method of using both configurations is different. The gaze-guided configuration calls for a quick "stare" into the desired trajectory and then to issue a single pressing of the "command" button to allow the aircraft to follow that trajectory, making quick adjustments as required.

On the other hand, the hand-guided configuration allows the pilot to press the "command" button for longer periods, having direct and precise control of the trajectory for longer periods of time. In fact, it requires this method of input to achieve the landing.
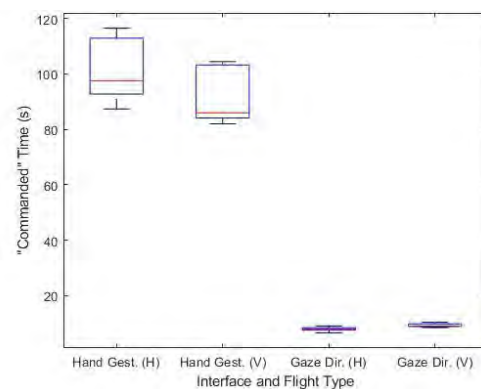
Figure 20: Cumulative "command" button time pressed for all the test scenarios and configurations.

It can also be seen that the necessity to use a "command" signal was confirmed. Figure 21 shows the uninhibited gaze signal as received by the system. Figure 11 shows the obtained command from that signal as the "command" button was pressed and released. Not having such an interruption makes the aircraft uncontrollable as the pilot would have to keep constant focus on the desired flight trajectory, even with the hand-guided configuration.
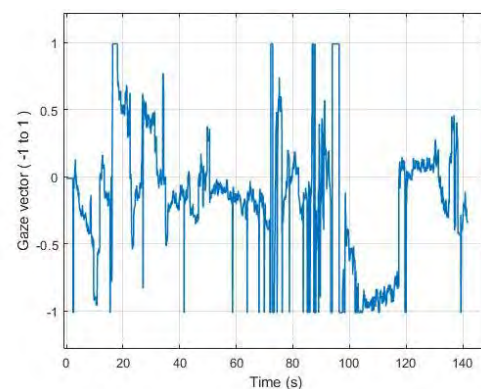
Figure 21: Example of a vertical gaze signal before being isolated by the "command" button during a whole flight.

6 Conclusions and Future Work

In this paper, the feasibility of gaze-guided and hand-guided control of an aircraft was demonstrated. Data collected in a first set of experiments will be used for further development of flight control modes, aiming at improving system performance.

Operational experience gained with this experiment allows for the planning of elaborate experiments using different pilots that are not yet accustomed to the system. New experiments would allow to properly measure the impact of the proposed interfaces on simulated flight.

Follow-up experiments will be conducted in a highly realistic flight simulator that includes wide field of view and a complete cockpit environment. It will also use a transport-class manned aircraft model and elaborate flight scenarios that would allow to compare the proposed interfaces with "classical" interfaces such as yokes and sidesticks in the same environment.

References

[1] J. Roskam, Airplane Design, Lawrence: DAR Corporation, 1985.

[2] NASA Armstrong, "X-29 Advanced Technology Demonstrator Aircraft," [Online]. Available: http://www.nasa.gov/centers/armstrong/news/FactSheets/FS-008-DFRC.html.

[3] Airbus, "Vahana, the Self-Piloted, eVTOL aircraft from A³ by Airbus, Successfully Completes First Full-Scale Test Flight," 02 February 2018. [Online]. Available: https://www.airbus.com/newsroom/press-releases/en/2018/02/vahana--the-self-piloted--evtol-aircraft-from-a--by-airbus--succ.html. [Använd 12 08 2019].

[4] Embraer, "EmbraerX unveils first eVTOL concept," 5 August 2018. [Online]. Available: https://embraer.com/global/en/news/?slug=906344-embraerx-unveils-first-evtol-concept. [Använd 12 08 2019].

[5] The Boeing Company, "Boeing Autonomous Passenger Air Vehicle Completes First Flight," 23 January 2019. [Online]. Available: boeing.com/features/2019/01/pav-first-flight-01-19.page. [Använd 12 August 2019].

[6] D. C. Niehorster, L. Li och L. Markus, "The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research," *i-Perception,* nr May-June, pp. 1-23, 2017.

[7] Tobii, "Pro Glasses 2 Product Description," Tobii AB, 2018.

[8] T. Fricke, "FIRST PILOT-IN-THE-LOOP SIMULATOR EXPERIMENTS ON BRAIN CONTROL OF HORIZONTAL AIRCRAFT MOTION," *Deutscher Luft- und Raumfahrtkongress,* 2014.

HUMAER: A Test-Bed Environment for Human Factors Investigations in the Aeronautic Domain

D. H. Arjoni[1], W. R. de Oliveira[2], L. G. Trabasso[3] and E. Villani[4] *

*Division of Mechanical Engineering, Aeronautics Institute of Technology, São José dos Campos, SP/Brazil
E-mail: arjoni@ita.br[1], wesleyro@ita.br[2], gonzaga@ita.br[3], evillani@ita.br[4]

Abstract

This work presents the *Human Factors in Aeronautics* laboratory (HUMAER), from the Aeronautics Institute of Technology (ITA). This laboratory is a test-bed environment built around the SIVOR – the flight simulator with a robotic motion platform from ITA – with the purpose of evaluating the different relations among the pilot cognitive, psychological and physiological states and the aircraft's human-computer interfaces and cybernetics. The lab is comprised of different virtual reality technology and non-intrusive physiologic sensor technology that can be installed on the pilot to monitor its psycho-cognitive state from a range of objective measures. Considering some characteristics of SIVOR's cockpit, like the motion-cueing ability and the haptics stimuli generated from the force-feedback inceptors and shakers, some of the main prospective research lines to be explored by the ITA's team are presented according to the key technologies made available at HUMAER. To exemplify some of the current research made within the lab, some findings on pilot perception and workload evaluation from previous referenced investigations are reported.

Keywords: Human factors, aeronautic training, flight simulation, physiologic sensors.

1 Introduction

Pilot training concerned with upset-recovery, pilot-induced oscillations (PIO), pre-stall conditions identification and new single-pilot cockpit paradigms are all some of the quite recent interests from the main aerospace community that have prompted academics worldwide towards human factors research [1-3]. Broadly speaking, human factors refer to a classical interdisciplinary research area concerned with the interaction of humans with other technologies, environments and systems [4-5]. In the aerospace context, this is translated into the evaluation of the diverse relations among the pilot cognitive, psychological and physiological states and the aircraft's human-computer interfaces and cybernetics [6].

With the aim of evaluating and developing solutions in this research field, there is HUMAER, the Human Factors in Aeronautics laboratory from the Aeronautics Institute of Technology (ITA), located in Brazil. This laboratory is a test-bed environment built around the SIVOR, a full-motion flight simulation system with a robotic motion platform from ITA (Figure 1). Besides the heavy-payload robotic arm and the embedded visual system, SIVOR presents a reasonable trade-off between the realism of the intern cockpit interface and its customization (Figure 2) [7]. The SIVOR cockpit is also enhanced by means of some haptics stimuli that can be generated from the force-feedback inceptors and unit shakers.

Figure 1: SIVOR – robotic flight simulator.

Figure 2: SIVOR's flexible cockpit with main flight deck, yoke and stick inceptors and virtual overhead panel.

HUMAER is based on the concept of monitoring the pilot's psycho-cognitive state from a range of objective measures gathered from non-intrusive physiologic sensors that can be installed on the subject. Some of the main sensor technologies that enable current research at HUMAER are presented in this work, providing a further overview on the research topics developed in the laboratory. Besides that, some selected results from previous investigations are discussed. To accomplish this goal, this work is organized as follows: *section 2* presents the laboratory's infrastructure; *section 3* details some of the main prospective research topics of interest of ITA's team; *section 4* provides some measurement results from previous investigations; and *section 5* brings up the main discussions and conclusions.

2 HUMAER's equipment and systems

HUMAER's main sensory equipment is built around sensors, modules and subsystems from Tobii ^{TM}, TEA^{TM} and Wearable Sensing ^{TM}. All of them are detailed in Table 1, Table 2 and Table 3, respectively.

Table 1- Tobii ^{TM} pro equipment.

Equipment	Image	Description
Tobiipro X3-120®		Eye-tracking bar
Tobiipro Glasses 2® Premium Package		Package: eye-tracking glasses with 50 Hz-record unit, calibration software, Prolab license
Tobii pro VR® Integration with HTC Vive		120 Hz eye-tracking glasses with HTC Vive integration

Table 2 – TEA^{TM} S.A.R.L. equipment, including CaptivePhysio and CaptivMotion.

Equipment	Image	Description
CAPTIV GSR		Galvanic Skin Response
CAPTIV Temperature		Body-temperature sensor
CAPTIV TCE Respiration		Thoracic contraction-expansion sensor for breathing monitoring
CAPTIV ECG		Electrocardiograph sensor

Equipment	Image	Description
CAPTIV EMG		Electromyography sensor

Table 3 - Wearable Sensing LLC^{TM} equipment.

Equipment	Image	Description
DSI-24 System EEG		Electroencephalography dry headband sensor with 24 channels.

Some general use for all of these sensors are [8-10]: ECG (electrocardiographic sensor) for heart-rate monitoring; ECG (electroencephalographic sensor) for brain activity monitoring; GSR (galvanic ski-response sensor) for sweating monitoring according to stress and awareness state; EMG (electromyography sensor) for muscular activity monitoring; TCE (thoracic contraction-expansion sensor) for respiration monitoring; eye-tracker glasses/bar for blinking, focused vision area and pupil dilation measurements.

They can be all synchronized together using the Captiv® application to help in the identification of the pilot workload and awareness in front of the flexible SIVOR's cockpit, which can also be adapted by means of some augmented reality technology detailed in Table 4 (Besides the Tobii HTC Vive in Table 1, which is also a virtual reality glasses).

Table 4 – Augmented and Mixed reality equipments.

Equipment	Image	Description
Moverio BT-350®		Epson® smart-glasses.
Meta 2®		Meta® Augmented Reality Glasses with development kit. Unity license.
HoloLens®		Microsoft ^{TM} HoloLens® for holographic mixed reality

The SIVOR's cabin can be further instrumented with an Inertial Measurement Unit (IMU) and unit shakers. In the following subsection, some research topics current under interest of the development team are detailed in regard to the technology made available at HUMAER.

3 HUMAER's research topics

Five research fronts are identified by the ITA's development team to use the aforementioned HUMAER's equipments within the SIVOR's investigations demands:

1) Evaluation of the effectiveness of the robotic motion platform with different strategies

Since some of the main approaches to evaluate the fidelity of motion-cueing in a flight simulator are based on qualitative pilot's feedback [11], the emphasis in this area is to use objective cognitive measures for this kind of investigation. Some topics are:

a) Comparative analysis of the washout filter algorithms used to control the robotic platform using additional excitation form the SIVOR's shaker units and measuring the platform's accelerations using the SIVOR's IMU (Inertial Measurement Unit) and the pilot's level of perception using the GSR, ECG and eye-tracking's pupil dilation measures.

b) Assessment of the correlation between the visual system projection and the platform's motion using the pilot eye-tracking and the IMU.

2) Evaluation of the cockpit's Human-Machine Interfaces (HMI) relations to the pilot workload and situation awareness

The idea of this research front is to evaluate pilot awareness upon different workload failure situations. Measurements of the pilot's response time, fatigue and flight performance are intended to be done using the HUMAER's sensory equipment. Some topics are:

a) Assessment of the current EMB505 (Legacy 500) symbology to failure, communication and reaction using parallel measurements from EEG, GSR, ECG and TCE. The eye-tracking technology can be used to evaluate the pilot's vision area, gaze blink, head-up and head-down times.

b) Evaluation of the more important contextual information in the Principal Function Displays (PFDs) and Multifunctional Displays (MFDs) during some failures using the eye-tracking bar, ECG, GSR and TCE.

c) Flight quality assessment in regard to the different profiles of inceptors' force feedback stimuli using EMG, GSR, TCE, ECG and eye-tracking system for pupil dilation measurements.

d) Pilot's fatigue and workload delimitation and quantification using the original cockpit's HMI's resources upon failure conditions. Equipments to be used: ECG, GSR, EEG, TCE, Temperature sensors and eye-tracking glasses.

e) Pilot's fatigue and workload delimitation and quantification using the additional mixed reality stimuli using further symbology from the HoloLens and Meta 2 glasses upon failure conditions. Equipments to be used: ECG, GSR, EEG, TCE, Temperature sensors, HoloLens and Meta 2.

f) Cockpit's preliminary design review using 3D virtual reality environments. Equipments: ECG, TCE and eye-tracking with Tobii Vive HTC glasses.

3) Identification of Pilot Induced Oscillations (PIO) events trough pilot cognitive feedback

Broadly speaking, this research area concerns with the identification of the pilot-induced oscillations mechanisms

and its respective detection by the pilot using the available cockpit's HMIs and contextual information feedback [12]. The main idea is to perform flight campaigns in SIVOR with pilots that present different experience inducing some characteristic inceptors' disturbances and trimming conditions that drive pilots to PIO. The following topics can be further investigated along the test:

a) Pilot behavior and physiologic state within the phases prior and after PIO events.

b) Pilot behavior and physiologic state during the attitude recovery and PIO mitigation.

c) Techniques for pilot training on PIO avoidance.

d) Development of flight control laws to detect and mitigate PIO events.

All of these research topics are supposed to use parallel measurements of: GSR; ECG; TCE; EMG; Temperature; EEG; and eye-tracker glasses (to monitor the focused area of the pilots vision, response times, gaze blink, head-up and head-down times).

4) Upset conditions identification and upset-recovery training.

An aircraft upset is defined when, during the flight, it is unintentionally driven to exceed the parameters (attitude or airspeed) normally experienced in line operations or training. It represents a dangerous condition since it may result in Loss of Controllability (LOC) of the aircraft according to some system failure, adverse environment or excessive inadvertent attitude command [13]. In the context of the SIVOR flight simulator, there exists some demands in evaluating possible training conditions where the pilot can identify an upset and recovery from that. The following topics can be further investigated within HUMAER test-bed:

a) Development of additional control laws and contextual HMI feedback to communicate upset conditions to the pilot.

b) Delimitation and quantification of pilot stress and fatigue in conditions of upset-recovery.

c) Analysis of correlation between qualitative psycho-behavioral feedback and physiologic measurements of pilots during trimmed upset-recovery maneuver.

d) Training on behavioral recovery techniques in emergency upset situations due to systems failures.

All of these research topics are supposed to use parallel measurements of: GSR; ECG; TCE; EMG; Temperature; EEG; and eye-tracker glasses (to monitor the focused area of the pilots vision, response times, gaze blink, head-up and head-down times).

5) Design of new cockpit concepts and HMI

a) Application of virtual and augmented reality tools for enhancing pilot's situation awareness through novel HMI and contextual symbology stimuli, interactive cockpit, single pilot cockpit. This topic is supposed to use: HoloLens Mixed Reality Glasses; Meta 2 Augmented Reality glasses; Moverio BT-350 Smart Glasses; VR Vive Tobii Pro; Unity

Engine and CAD tools for 3D modeling; eye-tracking glasses; EEG, ECG; GSR; TCE and EMG.

4 HUMAER's research samples

Let us present some samples of previous results [11, 14-17] from researches developed along the setup and development of the SIVOR flight simulator and HUMAER test-bed. In addition, to evince the potential of the sensory equipment, we provide some measurements from some sensors gathered along a regular flight training session in SIVOR.

4.1 Model-based evaluations

One of the major concerns of the SIVOR's development team during its implementation phase was to evaluate the effectiveness of the motion-cue strategy with a robotic arm and a linear rail unit for the fight training fidelity. Prior the setups of the HUMAER testbed, the analyses were eminently based on the qualitative pilots' assertions [11, 17] or model-based evaluations that further considered the human's physiologic perception system models [14-16].

In [14-16] full model-based approaches aimed at objectively evaluating the pilot perception on the robot motion are presented. All of them found on validated aircraft and robot dynamic models but lacked robustness on the final conclusions because of the use of a model of the pilot. Notwithstanding, some important results were presented, and one of them is depicted in Figure 3 [14].

Figure 3 depicts what is further understood as the pilot perception on the accelerations impelled by the simulator along a running take-off followed by its abortion.

4.2 Advantages of SIVOR for human factors research

According to the results presented in the previous subsection, since the SIVOR differs in its setup from "ordinary" full flight simulators (FFS) which usually use a Gough–Stewart platform [19], it has been of great importance to stablish an understanding of the influence of the robot arm and rail unit dynamics on the simulation quality. Basically, recent results developed by the SIVOR's team has evinced some advantages of this kind of approach for high-gain maneuvers.
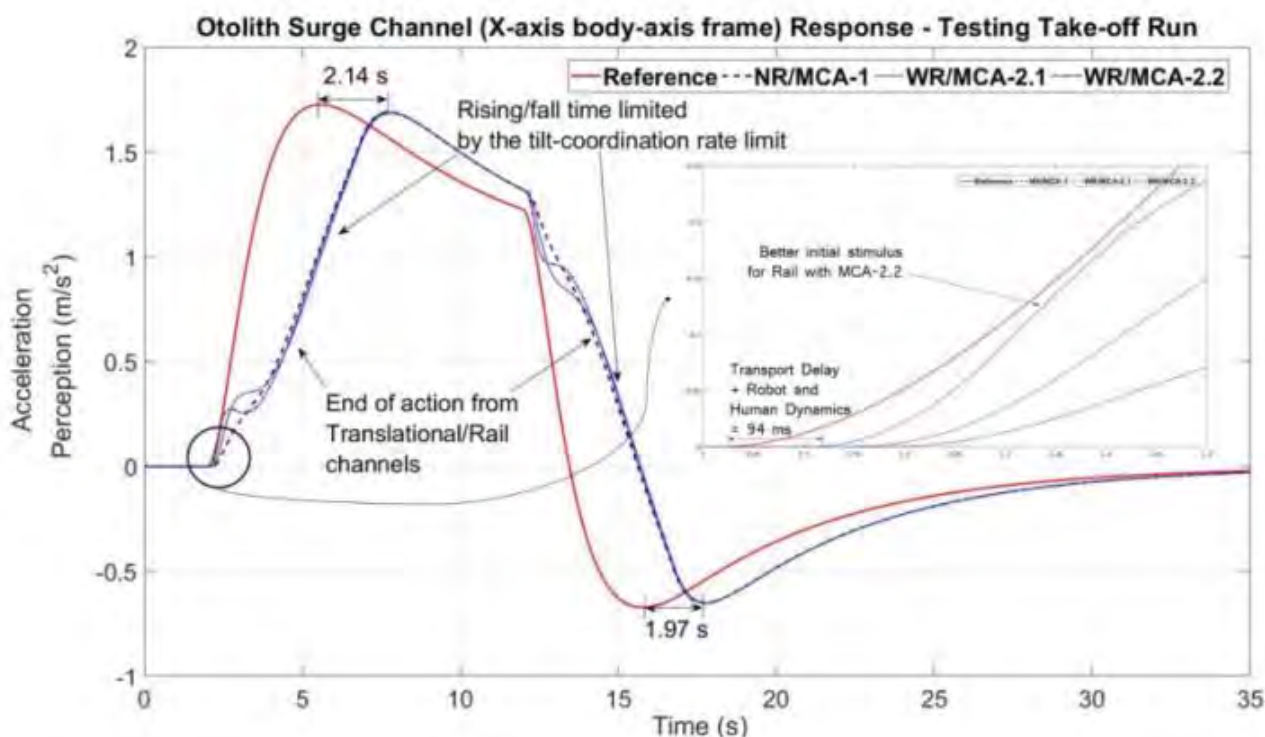
Figure 3: model-based comparison of pilot perceived surge acceleration under different simulator configurations [14].

These stand for those maneuvers where the aircraft is subject to more constant or continued accelerations, like in a reject take-off or a descent pitch-down, which may also precede a LOC event. This directly meets the objectives of the research lines on human factors presented in this paper. Quite newer results in the real SIVOR platform with qualitative pilot evaluations indicate that the rail unit being used as compared in [14], to excite low-frequency surge accelerations (WR/MCA-2.2 in Figure 3), has provided better stimulus. Besides of that, more objective measures

have been obtained in the real platform with an IMU (inertial measurement unit) from the HUMAER's infrastructure. They are depicted in Figure 4 for an aircraft Dutch-roll transient due to a rudder command performed within the simulator with the WR/MCA-2.2 active. This result evinces, at least from a first qualitative inspection, a good agreement between the surge acceleration generated by the aircraft and that gathered within the simulator from the IMU. The difference in scales are related to the scale factor used in the initial tests with the washout filter.
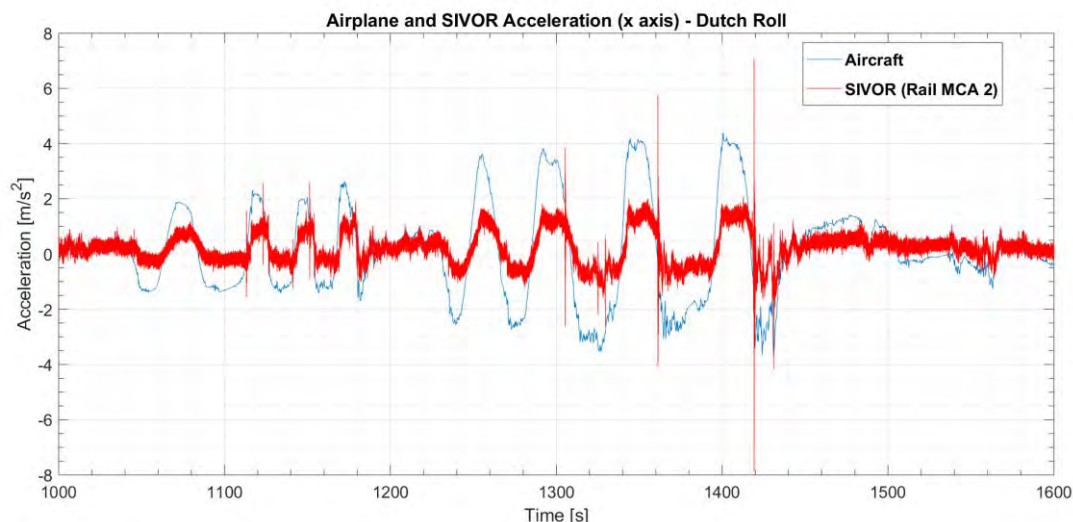
Figure 4: Comparison of the accelerations from the aircraft (blue) and the SIVOR (red). The latter is measured with an IMU.

4.3 Workload investigations

Some previous works from the SIVOR team aimed to use human factors related variables to access the effectiveness of the motion of the flight simulator [17]. In these works, performance and workload were estimated through a collection of aircraft dataset including joystick movements, altitude, attitude, airspeed, among others. The results pointed to a necessity of analysing these variables, considering a deeper psychological and biological approach.

To increase the practical understanding towards the workload and performance, the team designed an experiment to access how the pilot interact with the aircraft when exposed to different levels of workload [18]. The results detected changes of performance on variables regarded to the piloting command among workload levels, such as rate of climb, especially when the aircraft was experiencing an abnormal situation. With the setup of the HUMAER test-bed, some more objective measurements with real pilots have been collected, which are presented in the following subsection.

4.4 Samples of physiologic measurements

Some physiologic measurements using the ECG, GSR and eye-tracking glasses have been taken in a subject piloting the simulated EMB505 (Legacy® 500) with a simple regular session training comprised of (using the stick inceptor): take-off run, altitude stabilization and pitch and roll maneuvers. The pilot was asked to perform a coordinated bank at some point while keeping control of altitude and airspeed. Figure 5 presents the heat-map distribution of the focused vision area of the pilot during this occurrence. The redder is a given area the more the pilot focused in it.

Figure 5: heat map of the pilot's vision focused area during an attempt of compensating a system failure.

As it should be expected, the incipient result present in Figure 5 evinced the pilot focused its vision more in the feedback of the PFD and in position he was impelling to the stick grip, what shows his attempts for keep track of the established limits.

Additional data, such as event occurrence, and a simplified psychological state estimation, can be extracted from the ECG and GSR sensors curves (Figure 6), which were synchronized to the whole system and the eye-tracking glasses.

Although rather initial, these sample results evince the potential of the HUMAER's infrastructure to evaluate different abnormal conditions during regular flight training. All the investigations that are mentioned in this paper could be further developed considering more objective physiologic measures.
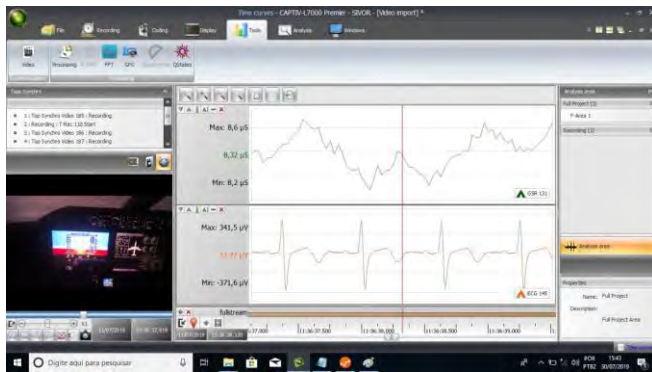
Figure 6: GSR (upper) and ECG (lower) signals acquired in the Captiv® software.

5 Conclusions

The main purpose of this work was to introduce the ITA's HUMAER lab, as a potential test-bed for Human Factors investigation in aeronautics. Besides of having been built around the flexible SIVOR flight simulation environment, the laboratory is equipped with relevant and modern physiologic sensor and virtual/mixed/augmented reality technology. Some research topics that has been emphasized by the ITA's team includes upset-recovery training, design of new cockpit concepts and HMI, identification of Pilot Induced Oscillations (PIO) events, evaluation of the motion-cueing strategies for flight simulation and of the cockpit's HMIs. Some results from previous works have been presented as well as some measurements gathered with some of the exampled sensors. This paper is expected to bring incitements to the aerospace community in Human Factors investigations.

References

[1] D. Acosta, Y. Yildiz, R. Craun, S. Beard, M. Leonard, G. Hardy and M. Weinstein, "Piloted Evaluation of a Control Allocation Technique to Recover from Pilot-Induced Oscillations," *JOURNAL OF AIRCRAFT,* vol. 52, pp. 130-140, 2015.

[2] J. Bürki-Cohen, "Technical Challenges of Upset Recovery Training: Simulating the Element of Surprise," em *AIAA Modeling and Simulation Technologies Conference*, Ontario, 2010.

[3] J. Liu, A. Gardi, S. Ramasamy, Y. Lim and R. Sabatini, "Cognitive pilot-aircraft interface for single-pilot operations," *Knowledge-Based Systems,* 10 9 2016.

[4] G. Salvendy, Handbook of Human Factors and Ergonomics, New Jersey: Wiley, 2012.

[5] F. H. Hawkins, Human Factors in Flight, New York: Routledge, 2016.

[6] J. Wise, D. Hopkin e D. Garland, Handbook of Aviation Human Factors, Taylor & Francis Group, 2009..

[7] Da silva, E. T. et. al. Flight simulator assisted by a robotic motion plataform. In: Proceedings of AIAA SciTech Forum, 7-11 January, 2019, San-Diego/USA. AIAA: 2019. DOI: 10.2514/6.2019-0435.

[8] TEAErgo, "TEA Ergo," [Online]. Available: http://teaergo.com/wp/?lang=en.

[9] "DSI 24," Wearable Sensing, [Online]. Available: https://wearablesensing.com/products/dsi-24/.

[10] Tobii, "Tobii," [Online]. Available: https://www.tobii.com/.

[11] Natal, G. S. et. al. Implementation Analysis of a Washout Filter on a Robotic Flight Simulator - a Case Study. J. Aerosp. Technol. Manag. [online]. 2019, vol.11, e0919. Epub 14-Jan-2019. ISSN 2175-9146. http://dx.doi.org/10.5028/jatm.v11.978.

[12] R. A. Hess, "Unified theory for aircraft handling qualities and adverse aircraft-pilot coupling," *Journal of Guidance, Control, and Dynamics,* vol. 20, nº 6, pp. 1141-1148, 1997.

[13] J. E. Wilborn and V. J. Foster, "Defining Commercial Transport Loss-of-Control: A," in *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Rhode Island, 2004. J. E. Wilborn and V. J. Foster, "Defining Commercial Transport Loss-of-Control: A," in *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Rhode Island, 2004.

[14] Oliveira et. al. Evaluation of the pilot perception in a robotic flight simulator with and without a linear unit. In: Proceedings of AIAA SciTech Forum, 7-11 January, 2019, San-Diego/USA. AIAA: 2019. DOI: 10.2514/6.2019-0713.

[15] Alves Jr. et. al. An Integrated Mode Based Development to Robotic Motion Flight Simulator. In: Proceedings of AIAA SciTech Forum, 7-11 January, 2019, San-Diego/USA. AIAA: 2019. DOI: 10.2514/6.2019-0435.

[16] Hidalgo, D.; Rodamillans, G. and Oliveira, W. R. Characteristic helicopter flight mechanics and motion cueing integration into the sivor robotic flight simulator: an objective approach. In: Proceedings of 31st Congress of ICAS, 09-14 Septembr, 2018, Beleo Horizonte/Brazil. ICAS: 2018.

[17] D. H. Arjoni, A. Kraemer e E. Villani, "Evolution of an Experiment to Assess Pilot Behaviour in Flight," em *International Congress of Mechanical Engineering*, Curitiba, 2017.

[18] D. H. Arjoni, E. Villani, M. Rodríguez, A. Matheus , A. Alemida , G. Rocha, L. Gonzaga and D. Hidalgo, "Experimental Analysis of Flight Performance Under Workload Variations," in *31st Congress of the International Council of the Aeronautical Sciences*, Belo Horizonte, 2018.

[19] Becerra-Vargas, M. and Belo, E. M., "Robust Control of Flight Simulator Motion Base," Journal of Guidance, Control, and Dynamics, Vol. 34, No. 5, 2011, pp. 1519–1528.doi: doi.org/10.2514/1.53301.
      .

IV Software and Avionics

Assurance Strategy for New Computing Platforms in Safety-Critical Avionics
Håkan Forsberg, Andreas Schwierz and Kristina Lundqvist

Triple Modular Redundancy based on Runtime Reconfiguration and Formal Models of Computation
Ricardo Bonna, Denis Loubach, Ingo Sander and Ingemar Söderquist

Lempel-Ziv-Markov Chain Algorithm Modeling using Models of Computation and ForSyDe
Augusto Horita, Ricardo Bonna, Denis Loubach, Ingo Sander and Ingemar Soderquist

Considerations on Domain-Specific Architectures Applicability in Future Avionics Systems
Denis Loubach, Johnny Marques and Adilson Adilson Marques da Cunha

Assurance Strategy for New Computing Platforms in Safety-Critical Avionics

Håkan Forsberg, Andreas Schwierz*, and Kristina Lundqvist

School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden[1]
E-mail: Hakan.Forsberg@mdh.se, Andreas.Schwierz@thi.de, Kristina.Lundqvist@mdh.se
*Research Center, Competence Field Aviation, Technische Hochschule Ingolstadt, Germany

Abstract

An assurance strategy for new computing platforms in safety-critical avionics has to be flexible and take into account different types of *commercial-of-the-shelf* (COTS) hardware technologies. Completely new COTS technologies are already being introduced and successfully used in other domains. Good examples are heterogeneous platforms, hardware-based machine learning and approximate computing. Current avionics certification guidance material cannot cope with next generation of devices. We suggest using the generic assurance approach of the *Overarching Properties* (OPs) together with assurance cases to argument that COTS assurance objectives are met and to achieve the flexibility required for future computing platforms. We introduce a novel assurance cased-based OP approach in [1] and refine the work into a framework in [2]. Within this framework we are able to integrate COTS technology specific assurance objectives using a five-step process. In this paper, we show through some representative examples of emerging computing platforms that our strategy is a way forward for new platforms in safety-critical avionics.

Keywords: safety-critical avionics, assurance strategy, assurance case, COTS assurance, Overarching Properties, computing platforms

1 Introduction

Over the last 15 years, RTCA/DO-254 [3] has been used as the guidance document to ensure design assurance for civilian airborne electronic hardware (AEH). Design/development assurance is "*All of those planned and systematic tasks used to substantiate, to an adequate level of confidence, that development errors have been identified and corrected such that the items satisfy a defined set of requirements*" [4]. AEH is often designed with several COTS components that are not developed according to RTCA/DO-254. The use of COTS components therefore requires other assurance guidance techniques to be used. The certification authorities have identified and produced several COTS assurance guidance documents, see Section IV in [1] for a literature review of these. When new technology has been introduced, new assurance activities have been suggested by the authorities. The latest guidance document from the certification authorities addressing COTS assurance (including COTS IP) is published in a Notice of Proposed Amendment [5], which is a joint EASA and FAA effort. This document is objective-based and is supposed to address all kinds of existing COTS components. It does not address new technologies such as e.g., hardware accelerated machine-learning. Objectives-based guidance benefits from being more flexible to adapt future technologies rather than activities-based documents.

Still, both objectives-based and activities-based guidance documents suffer from the assumption that, if it is followed it is sufficiently assured that the COTS component operates with integrity according to its specification.

EASA's Certification Memorandum SWCEH-001 [6] (non-binding guidance material) is a mixed level activities-based guidance document. Section 9, in [6], gives guidance for COTS integrated circuits and microcontrollers while Section 10, in [6], addresses COTS graphical processors (GPUs). For integrated circuits and microcontrollers, the assurance evidence depends on the amount of service experience, complexity of the component, and the design assurance level (DAL). For GPUs, the guidance material [6] assumes a discrete graphical processor (opposite to several of today's integrated GPUs) that has a very short lifespan with an increased possibility of design errors, is complex, contains configurable elements, and is only used for graphical applications. The guidance material also assumes that the component may exhibit performance variations over production time and may completely lack empirical data on the actual failure rates experienced in avionics applications. To cope with all these uncertainties, the guidance material explicitly assumes several low-level activities to be performed for all kinds of GPU devices [2]. Guidance on activity level is not suitable for new hardware technologies.

The suggested approach for new architectures and other new COTS technologies are instead an argumentative approach allowing for flexibility to use more appropriate methods and also directly show how these methods contribute to meeting the assurance objectives [2]. One such approach could be based on assurance cases. An assurance case is a structured argument, backed-up with evidence, that a system operates as intended for a defined application in a defined environment [7]. In [1] we demonstrated the use of an assurance case to structure COTS hardware components' assurance for safety-critical avionics and in [2] we refined our work and introduced a five-step process (see Section 2.3 in this paper) to provide a concept to connect the demonstration of assurance objectives. The use of assurance cases is in line with FAA's process to streamline the certification process by delivering an approach (Overarching Properties) usable for both software and hardware development to ease the use of alternative means of compliance [8]. The main contribution in this paper is the integration of existing assurance objectives and representative examples from new COTS-based computing platforms using Overarching Properties and assurance cases.

The remainder of this paper is structured as follows: Section 2 explains our COTS assurance case concept, the used graphical notation, the Overarching Properties, and shows how current COTS assurance objectives can be integrated. In Section 3 we integrate assurance objectives from emerging computing platforms and in Section 4 we discuss our assurance case concept. Finally, in Section 5 we conclude the paper.

2 COTS assurance case concept

In the framework of assurance cases, assurance refers to the proven confidence that a top-level claim of an argument is true [2]. Figure 1 shows a graphical presentation of an assurance case.
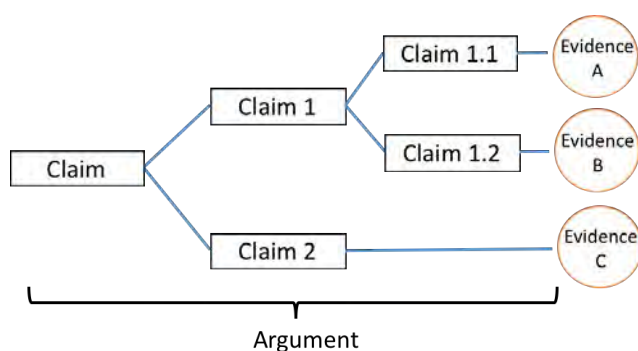
Figure 1. A graphical presentation of an assurance case. The top-level claim (leftmost) is decomposed until each sub-claim can be substantiated by evidence. The argument part, which is the focus of this paper, consists of strategies used to decompose claims and sub-claims.

Assurance cases in the specialized form as safety cases, have been successfully used for a very long time [9]. The strength with assurance cases is that it forces people to think deeper than usual [10] and motivates developers to formulate explicit arguments clearly targeting a top-level claim.

Structured assurances cases can be used to explain why a chosen assurance method is sufficient. They give the case writer the possibility to demonstrate (explicate) in a reviewable argument the assurance strategy in its entirety. This allows a third party (certification authority) to get the overview and full insight about how the item is assured and the justification why it is valid or acceptable to conclude that the item behaves with integrity in the system.

The question one can ask is - are assurance cases beneficial for emerging COTS-based computing platforms? Rinehart and Knight [9] have claimed several potential benefits for assurance cases in general. One of them is *assurance cases address modern certification challenges*. In [2] we interpreted all described benefits from Rinehart and Knight in the context of emerging computing platforms. Berthon [11] has used a structured assurance case for COTS AEH. Berthon suggests a design assurance level (DAL) based evidence approach for COTS hardware.

2.1 Assurance case notation

In this paper we use a graphical notation based on a subset of the Goal Structuring Notation (GSN). GSN is defined in [7]. Figure 2 shows the symbols we use in this article and Table 1 explains the used symbols.
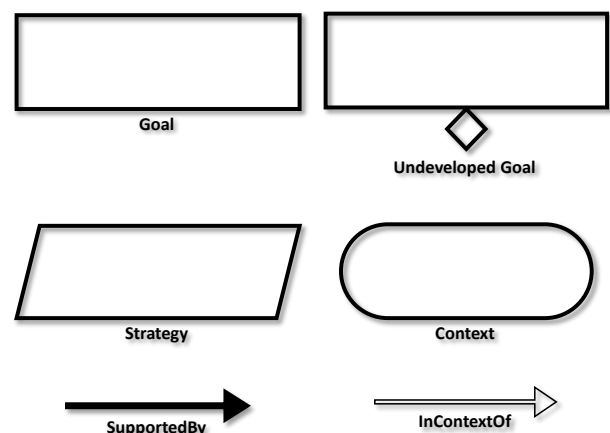
Figure 2. GSN symbol subset used in this paper

Table 1. Explanation of used GSN symbols

Goal	The *Goal* element illustrates claims and sub-claims supporting higher-level connected claims.
SupportedBy	The *SupportedBy* relationship creates a series of connected claims to establish an overall claim.
Undeveloped Goal	An *Undeveloped Goal* is left intentionally undeveloped for later investigations.
Strategy	The *Strategy* element helps explaining or argument the logic between a goal and its supporting goals.
Context	The *Context* element is used to clarify concepts mentioned in strategies.
InContextOf	A *Context* element has a corresponding *InContextOf* relationship.

2.2 Generic higher-level goals

In [1] we defined the top goal *COTS component operates demonstrably airworthy in its system context* for assuring a COTS integrated in safety-critical avionics. The top goal is based on applicable functional and safety certification specifications (CS) requirements, derived to the COTS component level by Berthon *et al.* [12]. Berthon *et al.* identified six key objectives applicable to all kinds of AEH

based on CS requirements. These objectives form the context to our top goal. We then used a strategy to decompose our top goal in "time", i.e. argument over initial airworthiness and argument over continuous airworthiness. From now on we only consider the former. For initial airworthiness, we then used the strategy *Argument over isolated COTS component and integrated COTS component* to decompose our case into two sub goals, see Figure 3.
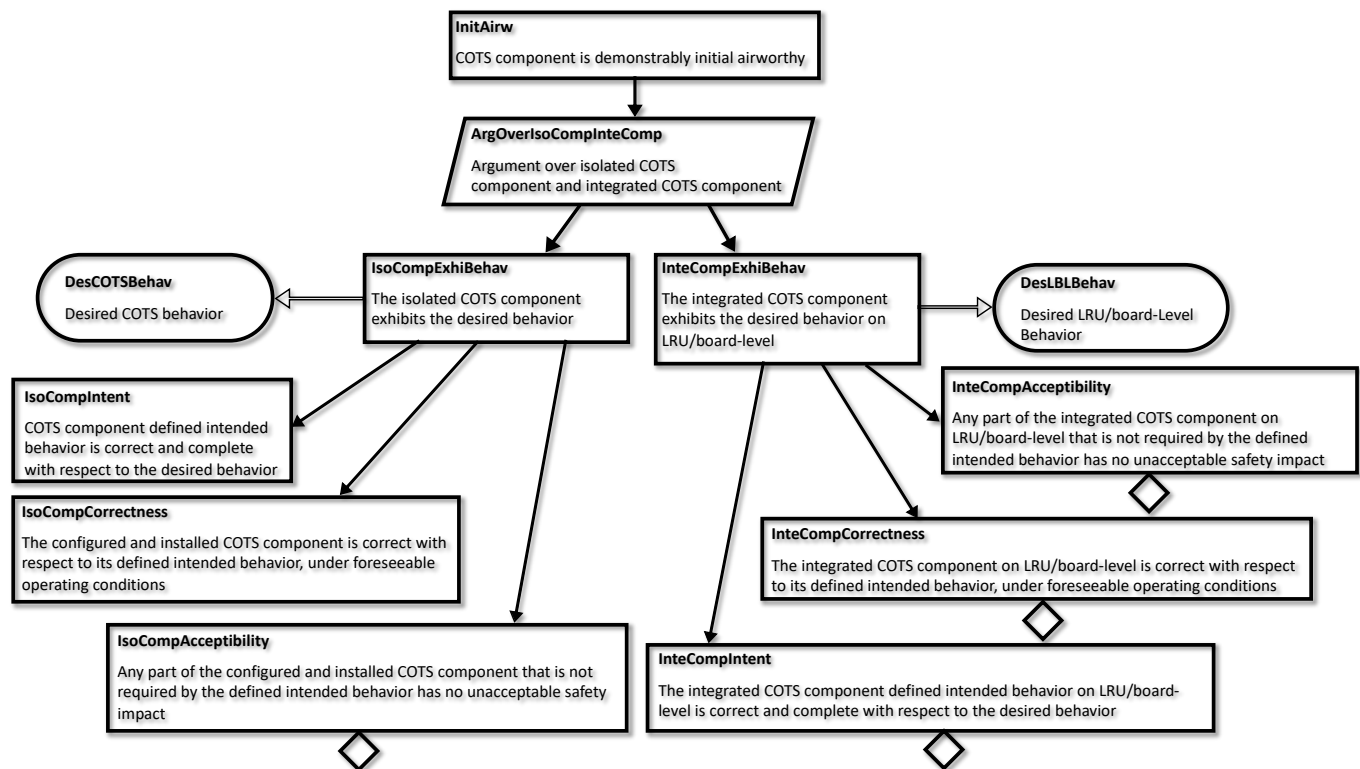
Figure 3. Initial airworthiness argument split into isolated COTS component and COTS component integrated on LRU/board level and further divided into the three Overarching Properties intent, correctness and acceptability.

Both sub-goals must be fulfilled to achieve the top-level goal for initial airworthiness, i.e.

1. the isolated COTS component exhibits the desired behavior and

2. the integrated COTS component exhibits the desired behavior on LRU/board level (note that certain behavior can only be verified on this level).

To build our next layer, we use the three Overarching Properties, informally written as [13]:

1. Intent – what the product is supposed to do is properly captured,
2. Correctness – the product does what it is supposed to do, and
3. Acceptability – the product does not cause harm since development decisions do not compromise the original safety assessment

See Figure 3 for proper implementation in the COTS element context. Note that both branches use the OP approach, i.e. the isolated and the integrated COTS component should demonstrate that the OPs are possessed on each level. With other words, it is not until all three properties on both levels have sufficient convincing arguments that the overall argument *COTS component is demonstrably initial airworthy* can be considered evidenced. The idea behind using OPs is to provide a unified method for the approval of different kinds of objects, i.e. such that this approach can be an alternative to already used assurance methods.

Up to this point the presented argument is aligned according to a generic layout which should be applicable for all kinds of COTS components. Even if the OPs are used to organize the assurance concept, it has to be shown that the COTS device meets the allocated specification adequately. The next step in the argument provides a strategy to enable the demonstration of the OPs together with considering the COTS technology dependent assurance.

To facilitate the demonstration of each OP, we use a separation into a primary and a confidence argument proposed by Hawkins *et al.* [14] and successfully demonstrated by Holloway and Graydon [15]. Compare it with safety cases where safety is the attribute of interest, where identification and mitigation of hazards to reduce risk should be in the primary argument.

Let us look at the goal *IsoCompCorrectness* from Figure 3. The primary argument should be "The configured and installed COTS component performs its intended behavior correctly, under foreseeable operating conditions," i.e., in this case, exactly the same as before the separation. The confidence argument should then produce the evidence that the primary conclusion is sufficiently creditable, i.e. the reviewer should believe the chain of transformation was correctly performed with sufficiently avoidance of errors. We thus define the confidence argument for *IsoCompCorrectness* as "Uncertainties in the correct transformation of the defined intended behavior to the configured and installed COTS component are sufficiently reduced."

2.3 Five step process to integrate COTS objectives

In [2], we created a framework that can connect assurance objectives directly with a new COTS assurance concept based on OPs and assurance cases. The framework should be performed in the following process steps:

1. Choose the level on which the assurance objective has to be demonstrated (isolated or integrated).
2. Assign the assurance objective to the relevant OP.
3. Reformulate it to a conclusion.
4. Demonstrate its satisfaction in the primary argument.
5. Explain in the confidence argument how you reduce the uncertainty in the primary argument.

Below we demonstrate the framework by integrating an already defined objective, *COTS-3*, from the Notice of Proposed Amendment [5]. In Section 3, we integrate two examples from emerging computing platforms.
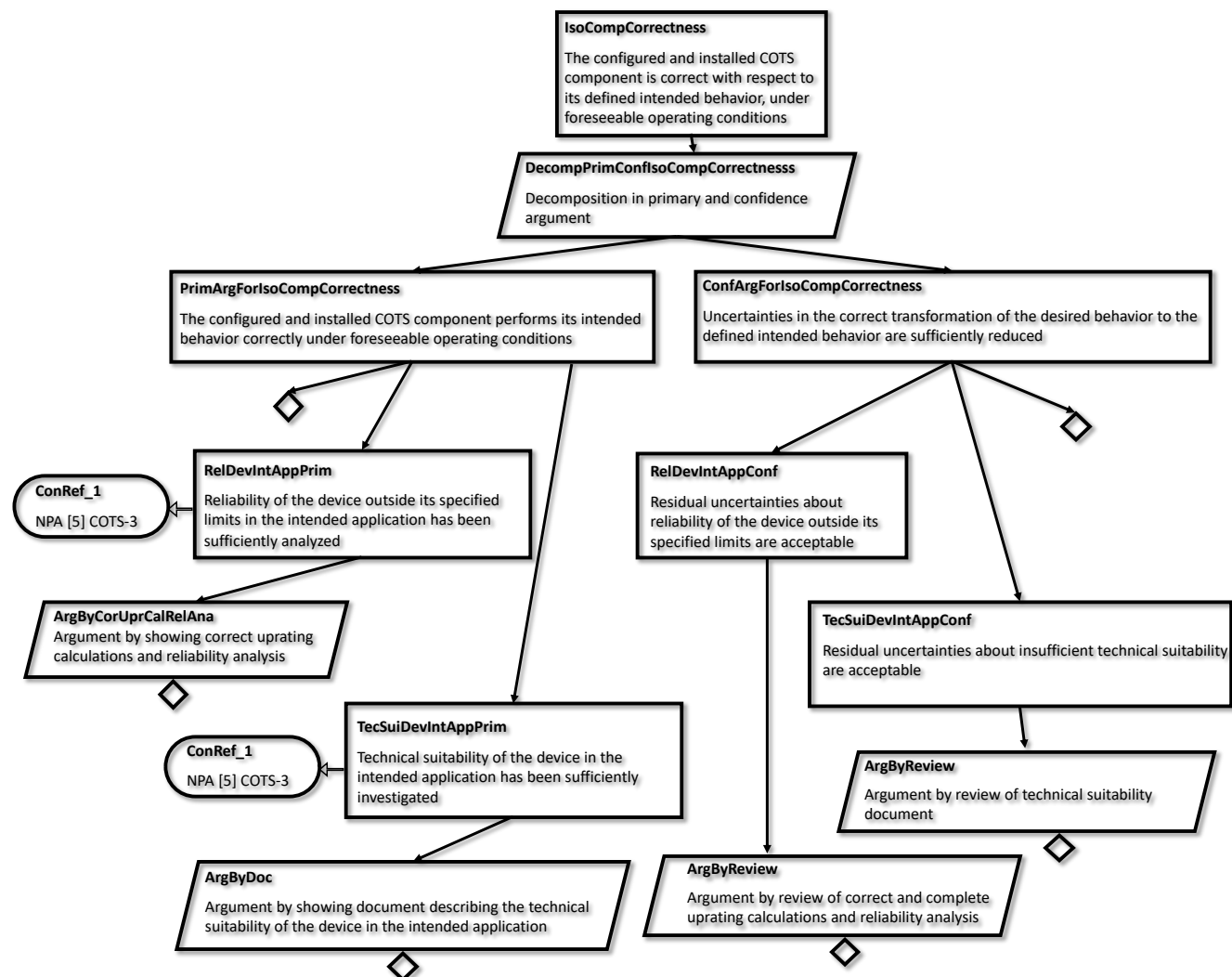

Figure 4. Integration example of the assurance objective COTS-3 from the Notice of Proposed Amendment [5]

The *COTS-3* objective in [5] states: "*When the complex COTS device is used outside the device manufacturer's specification (such as recommended operating limits), the applicant should establish the reliability and the technical suitability of the device in the intended application.*"

Step 1 - demonstration level – For *COTS-3*, the correct level of demonstration is the isolated COTS component level.

Step 2 – assignment to relevant OP – The appropriate OP is *IsoCompCorrectness*. See Figure 4. Even if the component is supposed to be used outside the device manufacturer's limits, it is under foreseeable operating conditions.

Step 3 – reformulate to conclusion – The objective should be reformulated to express a goal whose content can be substantiated with an assurance case. This concerns both the primary and the confidence argument. Here, we have chosen to split the *COTS-3* objective into two goals (the same two objectives as *COTS-3* actually addresses):

1. The reliability of the device outside its specified limits in the intended application has been sufficiently analyzed, and

2. The technical suitability of the device in the intended application has been sufficiently investigated.

Step 4 – demonstrate its satisfaction in the primary argument - In this case, we assume that the primary argument is to show correct uprating calculations and reliability calculations and by showing a document where the technical suitability has been documented and demonstrated.

Step 5 – explain in the confidence argument how you reduce the uncertainty in the primary argument - The uncertainty will be reduced by independent reviewing of uprating calculations and reliability analysis and by independently reviewing the document describing the technical suitability.

2.4   Integration of existing COTS assurance objectives

The latest proposed guidance document from the certification authorities (FAA and EASA) addressing COTS assurance is a Notice of Proposed Amendment (NPA) [5]. The COTS assurance part of the NPA consists of eight objectives (*COTS-1 to 8*) that must be fulfilled for the highest design assurance levels. In previous research we showed how to implement *COTS-8* in our framework. Above, we demonstrate the implementation of *COTS-3*. From a completeness standpoint, all eight *COTS-x* should cover the six sub-claims we defined, i.e. isolated COTS intent, correctness, and acceptability as well as integrated COTS intent, correctness, and acceptability. Our analysis show that all eight certification objectives indeed cover the three Overarching Properties on both isolated as well as integrated COTS level, with emphasize towards the former. Table 2 shows the mapping between the existing COTS assurance objectives in [5] and the relevant level and Overarching Property in our framework.

Table 2. Existing COTS objectives [5] and their mapping into isolated or integrated COTS level and Overarching Property

COTS objective – from [5]	Relevant level and Overarching Property
COTS-1 – assessment of complexity	Isolated - Intent
COTS-2 – electronic component management process	Isolated - Correctness
COTS-3 – usage outside manufacturer's specification	Isolated - Correctness
COTS-4 – non-qualified microcode	Integrated - Correctness
COTS-5 – assessment of errata	Isolated - Acceptability
COTS-6 – failure modes and common modes	Isolated - Acceptability & Integrated - Acceptability
COTS-7 -intended function of COTS device including interfaces	Integrated - Intent
COTS-8 – inadvertent alteration of critical configurations settings	Isolated - Intent

3   Emerging platforms implementation

In this section we introduce new emerging computing platforms that in the future might be introduced for the avionics industry and use two examples of potential assurance objectives from these platforms to show that our framework and five step process work for new objectives as well.

3.1   Heterogeneous platforms

Heterogeneous computing platforms use massive parallelism from non-traditional computing devices, e.g. GPUs or digital signal processors (DSPs) to achieve high performance computations at low energy. At the same time, they use traditional central-processing units (CPUs) for latency-sensitive serial parts of the code [16]. Medical imaging, computational photography and fluid dynamics are areas where heterogeneous platforms have been successful [16]. New programming models and compilers, hardware/software interface, run-time support, load balancing and scheduling policies are all challenges for heterogeneous architectures [17]. These heterogeneous COTS components do not have any specific certification guidance ready for the avionics market yet [2].

3.2   Hardware based machine learning

Machine learning in the form of deep neural networks (DNNs) has shown to be a promising alternative for object identification for several application domains [18]. In autonomous cars it is used as one of the primary sources for detection of pedestrians, cars, bicycles, animals, etc. Obstacle avoidance decisions are made from different types of objects and their movements. DNNs may also be successful in airborne systems. One such possible application is guided landing. To land autonomously without support from ground infrastructure requires advanced airborne systems including algorithms for detecting the runway. These systems are

safety-critical. The use of DNNs in safety-critical systems cannot rely on traditional design assurance techniques. Instead, other techniques have to be used. The main reason for this is that a DNN has to be trained with data sets of images (or other data) with objects it should be able to classify, but it cannot be trained with all possible inputs. Thus, misclassification of objects may appear. DNNs are also weak to adversarial inputs (the alteration of inputs which forces a trained DNN to misclassify) e.g. due to malicious attacks or external faults caused by the environment such as single event upsets. Several assurance techniques have been suggested for the use of DNNs in safety-critical applications [19, 20].

It is important for hardware-based machine learning to quantify the probability of an undetected misleading error and show that the error is appropriate to the function. It is not possible to perform this objective on isolated COTS component level since the component is trained for a certain purpose. We therefore define the following assurance objective ML-OBJ-1:

- *ML-OBJ-1* - the applicant should quantify the probability of an undetected, misleading error and show that the error is appropriate to the function.

We will now demonstrate this objective in our framework.

Step 1 - demonstration level – For *ML-OBJ-1*, the correct level of demonstration is isolated COTS component level.

Step 2 - assignment to relevant OP – The adequate OP is *InteCompAcceptability*, see Figure 5 below. Undetected misleading errors must be captured outside the device.

Step 3 - reformulate to conclusion – we reformulate the conclusion to "The probability for undetected misleading errors is quantified and the errors are appropriate to the function."

Step 4 - demonstrate its satisfaction in the primary argument – Our selected primary argument to fulfill the conclusion is to use statistical testing appropriate to the function.

Step 5 - explain in the confidence argument how you reduce the uncertainty in the primary argument – This one is really hard to cope with. How do you reduce the uncertainty of results from statistical testing? By testing more? We believe one solution can be the use of a diverse redundant architecture and to include timing aspects (to detect changes of objects in time). See Figure 5. This architecture is subject to future research.
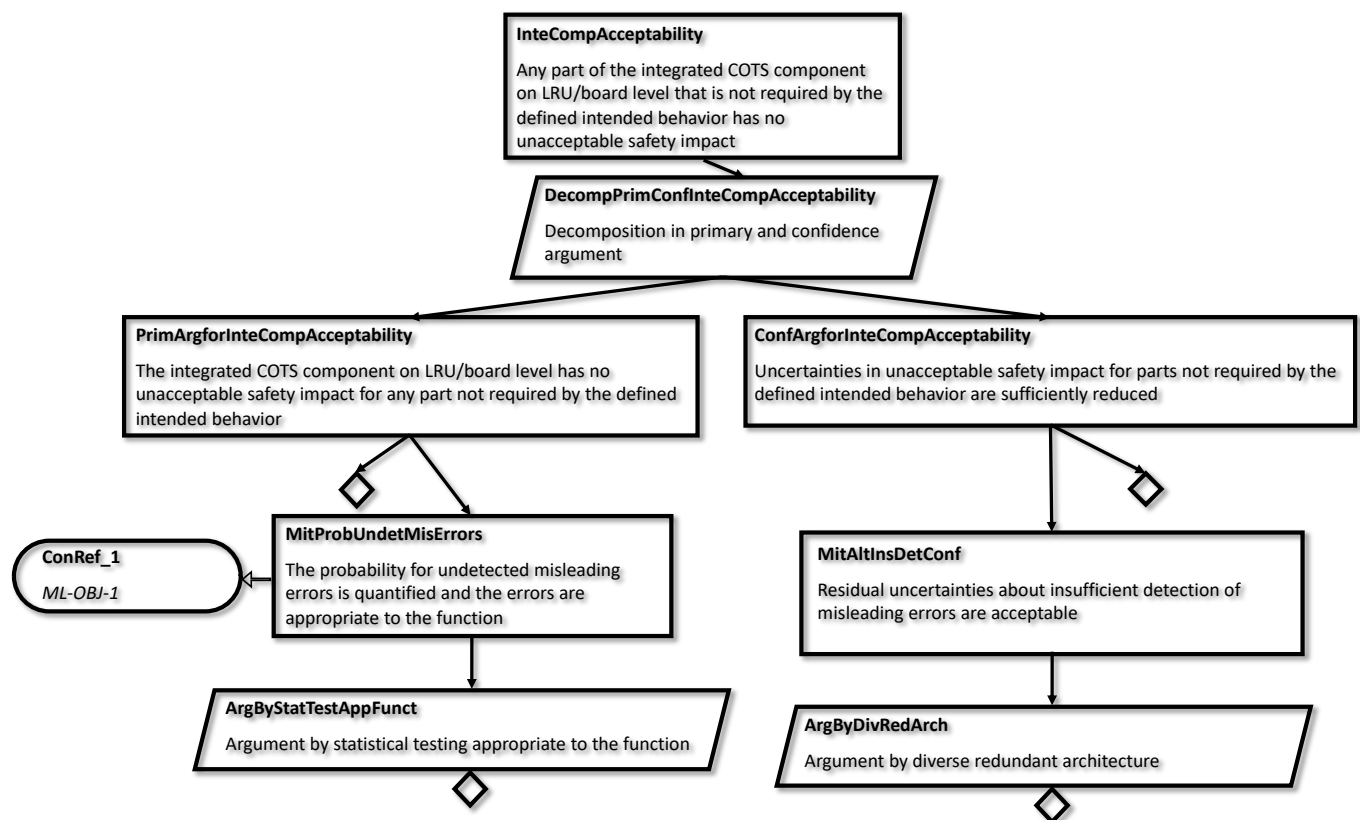
Figure 5. Integration example of an assurance objective most probable for machine learning architectures

3.3 Approximate computing architectures

Approximate computing is when computation accuracy is traded for better performance or energy consumption [21]. In approximate computing, different types of reduction of computation accuracy can be used, e.g. reduced number of bits in the arithmetic operations, approximate findings of results from expensive function calls (approximate memorization), reduced number of loops in loop constructions (loop perforation), or relaxed synchronization in parallel applications [22, 23]. These kinds of algorithms may be used for applications such as machine learning, computer vision, and speech.

In [2], we listed general guidelines that should be followed for architectures using approximate computing. We show one of them here:

- the computation should maintain integrity, i.e. using the same input data twice should show identical results (unless altered by physical phenomena, which must be detected)

We will now demonstrate this objective in our framework.

Step 1 - demonstration level – For *AC-OBJ-1*, the correct level of demonstration is integrated COTS component level.

Step 2 - assignment to relevant OP – The adequate OP is *IsoCompIntent*, see Figure 6. The device itself shall maintain integrity.

Step 3 - reformulate to conclusion – we reformulate the conclusion to "Using the same input data twice should show identical results."

Step 4 - demonstrate its satisfaction in the primary argument – Our primary argument is to use testing with equivalence classes.

Step 5 - explain in the confidence argument how you reduce the uncertainty in the primary argument – Here, you might be able to use advanced verification mitigation methods to ensure correct operation of the device, but you may also reduce the uncertainty by using an external monitor. In our case we used the latter. A simple monitor built upon a memory that stores previously received data and associated output results, and triggers only when a subsequent computation with identical input data results in different outputs.
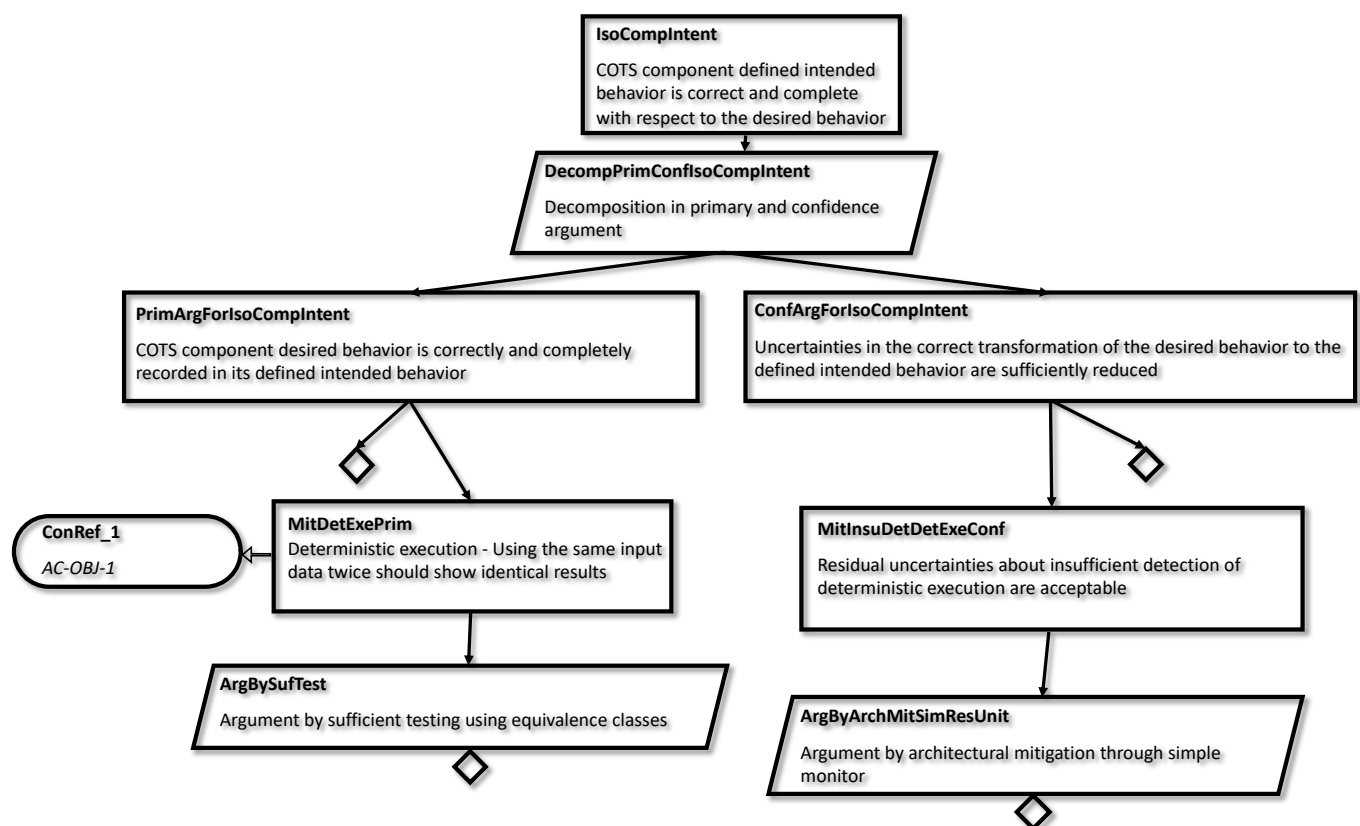

Figure 6. Integration example of an assurance objective most probable for approximate computing architectures

4 Discussions

In this section we particularly discuss two topics, the first regards the completeness of existing guidance and the second DAL modulation.

4.1 Existing certification guidance

One can argue that existing COTS assurance objectives in the NPA [5] alone is sufficient for implementing new COTS technology since all six sub-claims as described above are covered, but this is not true. New COTS technology may implement behavior which cannot be assured using existing guidance, e.g. statistical testing. The convincing arguments then need to address other assurance approaches such as architectural mitigation, see Section 3.2. The same is true for other new technologies. The convincing approach is thus first to implement the eight COTS objectives from [5] and then complement with sufficient arguments for the new (or emergent) technology to be used. This argumentation can be implemented using our framework. The hard part is to understand when the convincing arguments are sufficient.

4.2 Design assurance level modulation

In our approach we have not considered a DAL modulation so far. On a conceptual level it is reasonable to have such a mechanism so that the AEH manufacturer can indicate for which DAL the provided evidence is deemed to be sufficient. But the association of DALs to certain assurance activities is often debatable and a very subjective task that should be agreed with the certification authority. When the DAL modulation is integrated in our concept it will mainly affect only the confidence arguments.

5 Conclusions

We have in this paper, via different examples, shown that COTS specific assurance objectives can be dealt with through assurance cases using Overarching Properties. We showed how already existing objectives [5] were successfully implemented and then we showed how examples of assurance objectives for new emerging computer platforms can be implemented. Through our framework consisting of five process steps, the applicant will have flexibility to adapt the assurance task for the current project needs. We believe that our results are a way forward to address the assurance of future COTS-based computer platforms. In future work, we will continue working on representative examples to show the strength of our framework.

References

[1]   A. Schwierz and H. Forsberg, "Assurance case to structure COTS hardware component assurance for safety-critical avionics," in 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC), IEEE, 2018, Electronic ISBN: 978-1-5386-4112-5.

[2]   H. Forsberg and A. Schwierz, "Emerging COTS-Based Computing Platforms in Avionics Need a New Assurance Concept," to appear in 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC), IEEE, 2019.

[3]   RTCA, DO-254 Design Assurance Guidance for Airborne Electronic Hardware, 2000.

[4]   SAE Aerospace, ARP4754A: Guidelines for Development of Civil Aircraft and Systems , Rev. A, 2010.

[5]   EASA, Notice of Proposed Amendment 2018-09, "*Regular update of AMC-20:AMC 20-152 on Airborne Electronic Hardware and AMC 20-189 on Management of Open Problem Reports*," TE.RPRO.00034-006.

[6]   EASA, "*EASA CM – SWCEH – 001 Development Assurance of Airborne Electronic Hardware*," EASA, Issue 1, Rev. 2. 2018.

[7]   Origin Consulting, GSN Community Standard Version 1, 2011.

[8]   J. Wlad, Verocel, "Certification initiatives ongoing for unmanned aircraft systems," *Military Embedded Systems*, April 26th, 2018.

[9]   D. Rinehart and J. Knight, "*Understanding what it means for assurance cases to "work*"", NASA, Tech. Rep. NASA/CR-2017-219582, 2017.

[10] M. Holloway, *Understanding Assurance Cases: An Educational Series in Five Parts*, NASA, 2015. [Online]. Available: https://shemesh.larc.nasa.gov/arg/uac-all5.pdf [Accessed: 2019-07-08].

[11] G.-A. Berthon, "A Structured Assurance Case for Commercial Off-The-Shelf (COTS) Airborne Electronic Hardware (AEH)," SAE Technical Paper 2018-01-1939, 2018, doi:10.4271/2018-01-1939.

[12] G. A. Berthon, L. H. Mutuel, and C. Marchand, *DOT/FAA/TC-xx/xx: Final Report for System-Level Assurance of Airborne Electronic Hardware*, FAA, 2017.

[13] M. C. Holloway, DOT/FAA/TC-xx/xx: *Understanding the overarching properties: first steps*, Limited release document, September 2018.

[14] R. Hawkins, T. Kelly, J. Knight, and P. Graydon, "A New Approach to creating Clear Safety Arguments", in *Advances in Systems Safety*, C. Dale and T. Anderson, Eds., Springer London, pp. 3–23, 2011, ISBN: 978-0-85729-132-5.

[15] C. M. Holloway and P. J. Graydon, *DOT/FAA/TC-17/67: Explicate '78: Assurance Case Applicability to Digital Systems*, FAA, 2018.

[16] W-m. Whu, *Heterogeneous System Architecture: a New Compute Platform Infrastructure*. First edition, Amsterdam, Netherlands: Morgan Kaufmann, 2016. Print.

[17] H. Houcine, L. T. Yang, J. Xue, and E. Villar "Special Issue on: Heterogeneous Architectures for Cyber-Physical Systems (HACPS)," *Microprocessors and Microsystems,* vol. 52, pp. 333–334, 2017.

[18] LeCun, Y., Bengio, Y., & Hinton, G., "Deep learning". *Nature*, No. 521, 2015, pp. 436-444.

[19] C. Schorn, A. Guntoro, & G. Ascheid, "Efficient on-line error detection and mitigation for deep neural network accelerators," in International Conference on Computer Safety, Reliability, and Security, Springer, Cham, 2018, pp. 205-219.

[20] T. Gonschorek, M. Filax, & F. Ortmeier, "A very first glance on the safety analysis of selflearning algorithms for autonomous cars," in International Conference on Computer Safety, Reliability, & Security, 2018.

[21] M. Ammar Ben Khadra, "An Introduction to Approximate Computing," arXiv:1711.06115v2, 2017.

[22] A. Agrawal *et al.*, "Approximate computing: Challenges and opportunities," 2016 IEEE International Conference on Rebooting Computing (ICRC), San Diego, CA, 2016, pp. 1-8.

[23] M. Samadi, D. A. Jamshidi, J. Lee, and S. Mahlke, "Paraprox: Pattern-based approximation for data parallel applications," in *ACM SIGPLAN Notices*, vol. 49, no. 4, Feb., pp. 35-50, 2014.

Triple Modular Redundancy based on Runtime Reconfiguration and Formal Models of Computation

Ricardo Bonna[1], Denis S. Loubach[2], Ingo Sander[3], and Ingemar Söderquist[4]

E-mail: rbonna@fem.unicamp.br; dloubach@ita.br; ingo@kth.se; ingemar.soderquist@saabgroup.com
[1]Advanced Computing, Control & Embedded Systems Lab, University of Campinas – UNICAMP,
Campinas, SP, Brazil - 13083-860
[2]Department of Computer Systems, Computer Science Division, Aeronautics Institute of Technology – ITA,
São José dos Campos, SP, Brazil - 12228-900
[3]Division of Electronics/School of EECS, KTH Royal Institute of Technology, SE-164 40, Kista, Sweden
[4]Business Area Aeronautics, Saab AB, Linköping, Sweden

Abstract

Runtime reconfiguration is one promising way to mitigate for increased failure rate and thereby it fulfills safety requirements needed for future safety-critical avionics systems. In case of a hardware fault, the system is able, during runtime, to automatically detect such fault and redirect the functionality from the defective module to a new safe reconfigured module, thus minimizing the effects of hardware faults. This paper introduces a high level abstraction architecture for safety-critical systems with runtime reconfiguration using the triple modular redundancy and the synchronous model of computation. A modeling strategy to be used in the design phase supported by formal models of computation is also addressed in the paper. The triple modular redundancy technique is used for detecting faults where, in case of inconsistency in one of the three processors caused by a fault, a new processor is reconfigured based on a software or hardware reconfiguration, and it assumes the tasks of the faulty processor. The introduced strategy considers that no other fault occurs during the reconfiguration of a new processor.

Keywords: safety-critical systems, triple modular redundancy, runtime reconfiguration, formal models of computation.

1   Introduction

The safety and reliability of modern avionics may be threatened by trends that are largely driven by high-volume commercial applications, *e.g.* environmental concerns as restriction of hazardous substances (RoHS) directive that forced the removal of lead from commercial electronics and solders.

Another trend arises from technological innovation in commercial electronics. The effort to place more functionality and performance in smaller packages and lower power has led to ever-shrinking device geometries down to deep submicron dimensions with new physical failure mechanism that affect the wear out of semiconductor devices. Additionally, small geometries negatively affect the susceptibility of the semiconductor device to atmospheric radiation.

One of the next big challenges for the avionics industry is to address these trends that increase failure rate and thereby affect safety and reliability. Bieber *et al.* [1] points out runtime reconfiguration as one of the big challenges for the future generation of integrated modular avionic (IMA) systems. In the event of a hardware failure, the system is able to reallocate the functionalities from the faulted module into a safe module, thus limiting the effects of a hardware failure on aircrafts.

Perhaps the most important component of a runtime reconfigurable safety-critical system is the *fault detection mechanism*. One of such mechanisms is the *triple modular redundancy* (TMR), capable of detecting and mask possible faults in a system, improving reliability [2]. In such architecture, depicted in Figure 1, three processes execute the same functionality, and a majority voting mechanism selects the output that most occurs. If one of the processors fails to produce the correct output, possibly due to a single event upset (SEU), the voting mechanism masks such fault with the output of the other two processes. Any number of processors can be used in modular redundancies, however the minimum number of redundant processors necessary to detect and mask a fault is three.

Systems with triple modular redundancy are tolerant to both *transient faults*, *i.e.* faults that appear for a very short period of time and then disappear, and *single permanent fault*, *i.e.* faults that remains active for a long or possibly indefinite
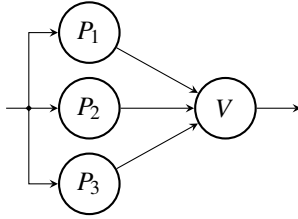
Figure 1: Triple modular redundancy architecture. V represents the voting mechanism and P_1, P_2 and P_3 represent three processes with the same functionality.

amount of time. However, faults in the voting mechanism lead to errors, making the voter a *single point of failure*. To improve reliability, three voters can be used instead of one. In that case if a fault occurs on the voters, the system can mask such fault, thus eliminating the single point of failure.

In view of this, this paper proposes a *high level abstraction architecture for safety-critical systems* with *runtime reconfiguration* (RTR) using the triple modular redundancy and the synchronous (SY) model of computation (MoC). Such architecture is composed of one fault detection mechanism, several runtime reconfigurable processes, and a control device to manage the reconfiguration process. Differently from the traditional triple modular redundancy architectures, the proposed architecture can mask multiple permanent faults, provided that no two faults occur in a small time interval defined by the reconfiguration time of a new module.

2   Models of Computation

Models of computation are a collection of rules dictating the semantics of execution and concurrency in computational systems. A common framework to classify and compare different MoCs is the *tagged signal model* [3]. In such framework, MoCs are a set of processes acting on signals, according to the following definitions.

Definition 1 (Signal). *In the tagged signal model, a signal $s \in S$ is a set of events $e_i = (t_i, v_i)$ composed by a tag $t_i \in T$ and a value $v_i \in V$. The set of signals S is a subset of $T \times V$.*

Definition 2 (Process). *In the tagged signal model, a process P is a set of possible behaviors that defines relations between input signals $s_i \in S^I$ and output signals $s_o \in S^O$. The set of output signals is given by the intersection between the set of input signals and the process $S^O = S^I \cap P$. A functional process is a process described by a single value mapping $f : S^I \to S^O$ and describes either one behavior or no behavior at all.*

The tagged signal model classifies MoCs as being *timed* or *untimed*. In a timed MoC, all events in all signals can be ordered based on its tags, *i.e.* the set of tags T is totally ordered. In an untimed MoC, the set of tags T is partially ordered, *i.e.* events can only be locally ordered.

2.1   Synchronous (SY) MoC

The synchronous MoC belongs to the class of timed MoCs and it is based on the *perfect synchrony hypothesis*, which

states that neither computation nor communication consumes time. As a consequence, every signal is synchronized, meaning that for any event in any signal, there is an event with the same tag in every other signal. This allows the representation of signals as a list of values in which the position of each value in the list represents its tag, *i.e.*, $s[k] = v$ with $k \in T$ and $v \in V$. Another important property of the synchronous MoC is that the absence of an event is well defined. Such phenomenon is defined as an event, with some tag $t \in T$, whose value is the absent value $\perp \in V$, *i.e.* $e = (t, \perp)$.

Although the perfect synchrony hypothesis is not physically feasible, the synchronous MoC works well when modeling clocked-based systems, provided that both computation and communication are fast enough to fit within one evaluation cycle.

3   Modeling TMR with RTR

A triple redundancy architecture proposal using runtime reconfiguration is illustrated in Figure 2.
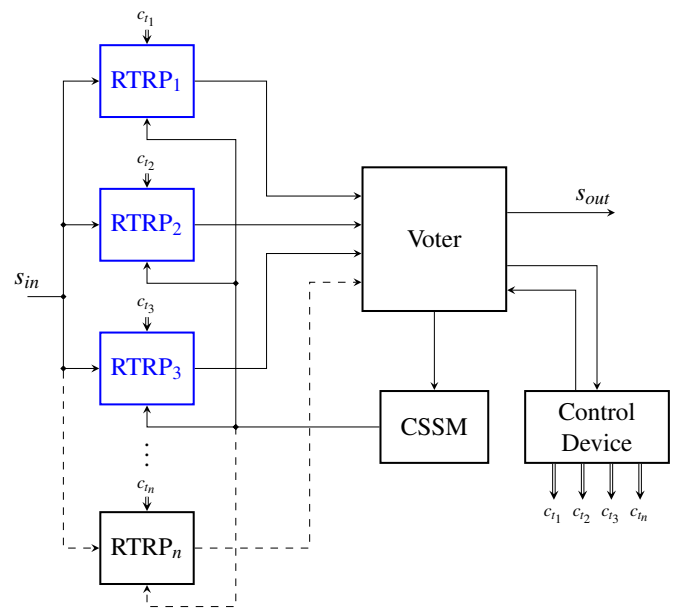
Figure 2: Triple redundancy architecture with runtime reconfiguration applicable to safety-critical systems.

It works as follows: three runtime reconfigurable processors, RTRP$_{1..3}$, are configured with the same functionality and, given the same input signal, should provide the same output signal. Knowing this property, the Voter compares the results and possibly the states outputted by each processor. If, by any chance, one of the processor's output differs from the other two, the Voter assumes that there must be a fault in such processor and, therefore, a new processor must take its place. In view of this, the Voter sends a signal to the Control Device informing which processor is malfunctioning, so that the Control Device can allocate a new RTRP$_x$ to assume the failed processor's task.

The newly allocated processor must then synchronize its states with the two remaining RTRP that are still executing

in order to mask the fault. To do that, the first time a processor executes, it loads the current states from a *current state shared memory* (CSSM), which is represented as a *delay* using the SY MoC and can be physically implemented as a set of processor registers.

Similarly to N-modular redundancy (NMR) with spares, when one of the processors becomes unreliable, *i.e.* starts to produce inconsistent results, it is replaced by a spare processor. However, here the *spare processors*, represented by RTRP$_n$, with $n > 3$, can be initially loaded with less critical applications that can be overloaded when needed, providing better usability of resources.

For such architecture to work properly it is assumed that neither the Voter nor the Control Device fails, thus both of these devices are single point of failure for this system. Although it is possible to eliminate the Voter's single point of failure by using three voters, there can only be one reconfiguration manager, represented by the Control Device.

When two RTRPs fail either at the same time, or in a time window smaller than the necessary time to reconfigure a new processor, we say the triple redundancy system fails. To show how likely such failure occurs, consider that all RTRP have the same failure rate and, for every cycle, the probability of failure of an RTRP$_j$ is $p(F_j) = \rho$. Consider also that it takes m clock cycles to reconfigure a new RTRP in case of a failure. Then, the probability $p(F_a|F_b)$ of some RTRP$_a$ to fail in a time window of $m+1$ cycles (including the cycle in which the fault was detected), provided that some RTRP$_b$ has already failed, is given by

$$p(F_a|F_b) = 1 - (1-\rho)^{2(m+1)} \tag{1}$$

Therefore, the probability of failure of our triple redundancy architecture with RTR is given by

$$\begin{aligned} p(F_a \cap F_b) &= p(F_b)p(F_a|F_b) \\ &= \rho(1 - (1-\rho)^{2(m+1)}) \end{aligned} \tag{2}$$

We define the *ratio of improvement R_I* as being the probability of failure of a single RTRP divided by the probability of failure of our triple redundancy architecture, given by (2). The larger the ratio of improvement, the more fail-safe the triple redundancy with RTR is when compared to an architecture with a single processor. Such ratio of improvement is given by

$$R_I = \frac{1}{1 - (1-\rho)^{2(m+1)}} \tag{3}$$

The ratio of improvement R_I shows the importance of the reconfiguration time m in the robustness of the triple redundancy architecture with RTR. In case of a fault in one of the processors, the system can still continue to perform correctly with two processors while another RTRP is being reconfigured, however it becomes vulnerable to a second fault in this time window. Therefore, the fastest the reconfiguration, the less vulnerable the system is. Traditional triple redundancy architectures (without RTR) are immune to a single processor permanent fault, however they are vulnerable to multiple faults.

3.1   Runtime Reconfigurable Process (RTRP)

We start to model the triple redundancy architecture by modeling what we are calling a runtime reconfigurable process, similar to the architecture presented in [4]. We consider that for such process to be in its most general form, it must have some internal memory to store its states, and it must take into account reconfiguration time. When a new RTRP is being reconfigured, it takes a number $m \in \mathbb{N}$ of clock cycles, proportional to the size of the functionality bitstreams, to perform reconfiguration before it is able to execute for the first time. When the RTRP executes for the first time, it must synchronize its internal memory with the internal memory of the other two RTRP executing the same functionality. In order to achieve that, we added an extra input, a *synchronization input $\bar{x}$*, so that when the RTRP executes for the first time, it gets its initial state from the synchronization input.

We model an RTRP as a finite state machine with $x[k] \in \mathbb{S}$ being the state vector, $s_{in}[k] \in V^I$ the inputs, and $y[k] \in V^O$ the outputs at an instant k, with V^I and V^O the set of values from the input and output signals respectively. The functionality of an RTRP is represented by a state transition function $f : \mathbb{S} \times V^I \to \mathbb{S}$ and an output function $g : \mathbb{S} \times V^I \to V^O$. Such functionality is stored in a configuration memory and can be changed by a control input signal c_t that is responsible for the reconfiguration. The control signal c_t is responsible for changing both f and g when needed, and this change takes m clock cycles to finish. Finally, the processor can execute for the first time with the new configuration.

A representation of an RTRP is shown in Figure 3. Feedback loops, along with delay blocks (represented by z^{-1}), are used to represent memories following the pattern: the blue rounded delay represents configuration memory, the squared black delay represents RTRP's internal memory, and the dashed delay represents a virtual count down to simulate reconfiguration time.
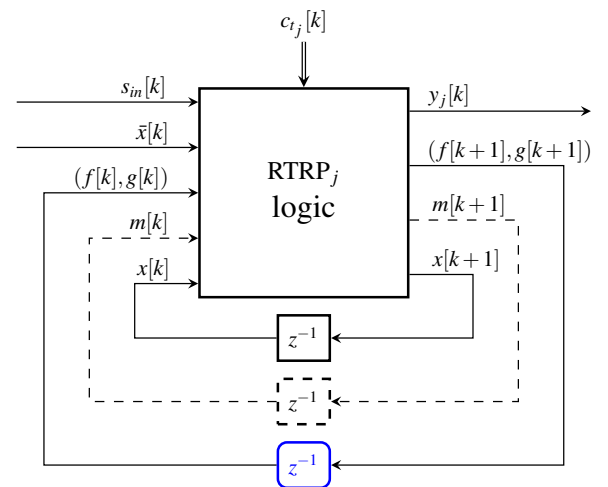

Figure 3: RTRP internal schematics, i.e. function application logic. The delays represented by z^{-1} are used to store state vector x, the pair of functions (f,g) and the reconfiguration countdown m.

Let $c_t[k]$ be the control signal in the instant k, which can be either the absent value $\perp$ or a 3-tuple $(\mathbf{f}, \mathbf{g}, \mathbf{m})$, with $\mathbf{f} : \mathbb{S} \times V^I \to \mathbb{S}$, $\mathbf{g} : \mathbb{S} \times V^I \to V^O$ and $\mathbf{m} \in \mathbb{N}$. $c_t[k] = \perp$ indicates no reconfiguration is needed and, therefore, the process can execute the current configuration normally, if able to. When $c_t[k] = (\mathbf{f}, \mathbf{g}, \mathbf{m})$, a reconfiguration must be performed and the functions $\mathbf{f}$ and $\mathbf{g}$ will replace the current configuration. Such reconfiguration process takes $\mathbf{m}$ cycles to finish.

At any instant k, the functionality of a runtime reconfigurable process is given by the pair $(f[k], g[k])$. Such pair is stored in the configuration memory until a reconfiguration request is received via c_t. To represent such behavior, the functionality transition function is given by

$$(f[k+1], g[k+1]) = \begin{cases} (\mathbf{f}, \mathbf{g}) & \text{if } c_t[k] = (\mathbf{f}, \mathbf{g}, \mathbf{m}) \\ (f[k], g[k]) & \text{if } c_t[k] = \perp \end{cases} \quad (4)$$

To represent the time spent to perform the reconfiguration of a process, the countdown variable $m[k] \in \mathbb{N}$ stores how many cycles are left to finish the reconfiguration. $m[k] > 0$ indicates that the process is reconfiguring at the instant k and, therefore, cannot execute. (5) represents the behavior of the countdown signal m.

$$m[k+1] = \begin{cases} \mathbf{m} - 1 & \text{if } c_t[k] = (\mathbf{f}, \mathbf{g}, \mathbf{m}) \\ m[k] - 1 & \text{if } c_t[k] = \perp \text{ and } m[k] > 0 \\ 0 & \text{if } c_t[k] = \perp \text{ and } m[k] = 0 \end{cases} \quad (5)$$

When a reconfiguration is being performed, *i.e.* when $m > 0$, the RTRP outputs the absent value $\perp$ for both the next state $x[k+1]$ and the output $y_j[k]$. The first time the RTRP executes after reconfiguration, it uses the state input $\bar{x}$ as initial states. Afterwards, it keeps executing with its internal state x. The state transition function at any instant k is given by

$$x[k+1] = \begin{cases} \perp & \text{if } c_t[k] \neq \perp \text{ or } m[k] > 0 \\ f[k](\bar{x}[k], s_{in}[k]) & \text{else if } x[k] = \perp \\ f[k](x[k], s_{in}[k]) & \text{otherwise} \end{cases} \quad (6)$$

with $\bar{x}[k]$ being the value of the states stored in CSSM in the instant k. The output of a runtime reconfigurable process at any instant k is given by

$$y_j[k] = \begin{cases} \perp & \text{if } c_t[k] \neq \perp \text{ or } m[k] > 0 \\ g[k](\bar{x}[k], s_{in}[k]) & \text{else if } x[k] = \perp \\ g[k](x[k], s_{in}[k]) & \text{otherwise} \end{cases} \quad (7)$$

Initial values $f[0]$, $g[0]$, $m[0]$ and $x[0]$, indicating the initial configuration and states of each RTRP, must be provided.

3.2 Voter

The `Voter`'s task is to compare the outputs of the three RTRPs that are currently active, and alert the `Control Device` when one of the outputs differs from the other two. Figure 4 shows the voter inputs and outputs.

The input c_v is responsible to select the three currently active RTRP, so that the voter can compare their results and, in case
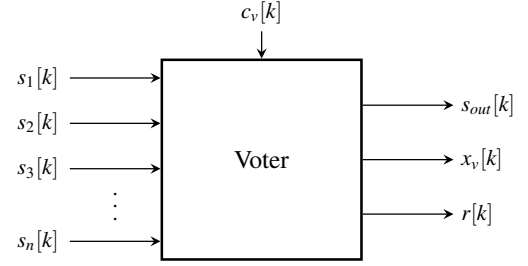
Figure 4: `Voter` *with inputs and outputs.*

of any inconsistency, it informs the `Control Device` about the failed RTRP through the signal r. The `Voter` outputs the most occurring of the RTRP results through s_{out} and send the current RTRP state to CSSM via x_v. Let s_j be the signal that carries the output and the states of RTRP$_j$. Events from c_v and s_j are defined as follows.

$$c_v[k] = (a, b, c), \quad a, b, c \in \{1, 2, \ldots, n\} \quad (8)$$
$$s_j[k] = (y_j[k], x_j[k]), \quad j \in \{1, 2, \ldots, n\} \quad (9)$$

The outputs s_{out} and x_v are modeled as follows.

$$(s_{out}[k], x_v[k]) = \begin{cases} (y_a[k], x_a[k]) & \text{if } y_a[k] = y_b[k] \\ & \text{or } y_a[k] = y_c[k] \\ (y_b[k], x_b[k]) & \text{if } y_b[k] = y_c[k] \end{cases} \quad (10)$$

The signal r is used to inform the `Control Device`, in case of a failure, which RTRP failed. If the results from the three active RTRPs are consistent in instant k, $r[k]$ assumes the absent value, otherwise it assumes the number of the faulted RTRP. Thus, the output r is modeled as follows.

$$r[k] = \begin{cases} \perp & \text{if } y_a[k] = y_b[k] = y_c[k] \\ a & \text{if } y_a[k] \neq y_b[k] = y_c[k] \\ b & \text{if } y_b[k] \neq y_a[k] = y_c[k] \\ c & \text{if } y_c[k] \neq y_a[k] = y_b[k] \end{cases} \quad (11)$$

3.3 Control Device

Finally, the `Control Device` is responsible for reconfiguring new RTRP based on the signal r received from the `Voter`, indicating which RTRP is not producing a correct answer. The `Control Device` keeps track of which RTRPs are active and, depending on the value received through the signal r, it performs a state transition to a new state indicating the active RTRP. Figure 5 shows the `Control Device` internal schematics.

Every time the `Control Device` performs a state transition, meaning an inconsistency was detected by the `Voter`, it sends the new configuration to the RTRP through signals c_{t_n}, and it waits m clocks, representing the time it takes to finish the reconfiguration of an RTRP, before being able to reconfigure a new RTRP in case of another inconsistency. The output/input signal m keeps track of how many clock cycles are left to finish a reconfiguration of a new RTRP, and it is modeled as

DOI
10.3384/ecp19162016
        Proceedings of the 10th Aerospace Technology Congress
October 8-9, 2019, Stockholm, Sweden
        148

Figure 5: `Control Device` internal schematics.

Figure 6: State chart ruling the behavior of the `Control Device`.

4 RTR Modeling with SY

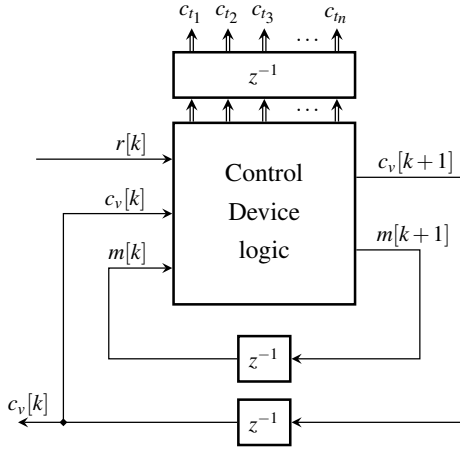An strategy and comparison of frameworks supporting formal-based development and models of computation is presented by Horita *et al.* [5]. Based on their result, we opt here for the use of ForSyDe [6] to model our TMR system. As ForSyDe is implemented in Haskell, a functional language, implementing (4) to (15) is considered an easy task, and we one does not need to worry about side effects either. Another advantage of functional languages is that functions can be used as normal data, allowing the exchange of control events such as $(\mathbf{f}, \mathbf{g}, \mathbf{m})$.

The ForSyDe SY library possesses a collection of process constructors, as well as delays, to implement all the processes presented so far. We use the process constructors comb$n$SY, with n indicating the number of inputs, unzip$m$SY, with m indicating the number of outputs, and delaySY.

Listing 1 shows the ForSyDe implementation of an RTRP process, where `rtrpFunc` is a Haskell function that implements (4) to (7). For this implementation, we consider an architecture with 5 RTRPs (three initially operating RTRPs and two spare ones). As mentioned in Section 3.1, the initial values $f[0]$, $g[0]$, $m[0]$ and $x[0]$ must be provided and are represented as f0, g0, m0 and x0.

follows.

$$
m[k+1] = \begin{cases} \mathbf{m} & \text{if } r[k] \neq \perp \text{ and } m[k] = 0 \\ m[k] - 1 & \text{if } m[k] > 0 \\ 0 & \text{otherwise} \end{cases} \quad (12)
$$

The output c_v carries the current three active RTRPs as a tuple, such as in (8). The behavior of the `Control Device` regarding the output c_v is modeled as follows.

$$
c_v[k+1] = \begin{cases} c_v[k] & \text{if } r[k] = \perp \text{ or } m[k] > 0 \\ h(c_v[k], r[k]) & \text{if } r[k] \neq \perp \text{ and } m[k] = 0 \end{cases} \quad (13)
$$

with $h((a,b,c),r)$ given by

$$
h((a,b,c),r) = \begin{cases} (\max(a,b,c)+1, b, c) & \text{if } r = a \\ (a, \max(a,b,c)+1, c) & \text{if } r = b \\ (a, b, \max(a,b,c)+1) & \text{if } r = c \end{cases} \quad (14)
$$

(12) and (13) define a behavior that is represented graphically in Figure 6. While $m[k] > 0$, the `Control Device` is in the "Reconf" state, meaning a new RTRP is being reconfigured, and any reconfiguration request that is sent through r is ignored while in this state. After the reconfiguration is finished ($m[k] = 0$), the `Control Device` returns to the "Ready" state, awaiting for a new reconfiguration request.

Finally, the control outputs c_{t_j} behave as follows: when the transition from Ready to Reconf is *taken*, i.e. $r[k] \neq \perp$ and $m[k] = 0$, the `Control Device` outputs a reconfiguration signal given by the 3-tuple $(\mathbf{f}, \mathbf{g}, \mathbf{m})$ to the output $j = \max(c_v[k]) + 1$ (the next available spare RTRP); in any other case, it outputs $\perp$. Such behavior is given by

$$
c_{t_j}[k+1] = \begin{cases} (\mathbf{f}, \mathbf{g}, \mathbf{m}) & \text{if } r[k] \neq \perp \text{ and } m[k] = 0 \\ & \text{and } \max(c_v[k]) + 1 = j \\ \perp & \text{otherwise} \end{cases} \quad (15)
$$

Initial conditions to m, c_v and c_{t_j} must be provided. As a general rule, we use the following initial conditions: $m[0] = 0$, $c_v[0] = (1, 2, 3)$ and $c_{t_j}[0] = \perp$. These initial conditions indicate the system starts with the three first RTRPs already configured and the `Control Device` in the Ready state.

Listing 1: `RTRP` process implemented in ForSyDe.

```
1 rtrp (f0,g0,m0,x0) ct s_in x' = out
2     where (out, fb) = unzipSY $ comb4SY
         rtrpFunc ct s_in x' fb'
3         fb' = delaySY (f0,g0,m0,x0) fb
```

In a similar way, Listing 2 shows the ForSyDe implementation of the `Voter` process, where `voterFunc` is a Haskell function that implements (10) to (11).

Listing 2: `Voter` process implemented in ForSyDe.

```
1 voter cv s1 s2 s3 s4 s5 = unzip3SY $ comb2SY
     voterFunc cv (zip5SY s1 s2 s3 s4 s5)
```

Listing 3 shows the ForSyDe implementation of the `Control Device` process, where `ctrlDevLogic` is a Haskell function that implements (12) to (15), and `prosopon1` is a Haskell implementation of a 3-tuple $(\mathbf{f}, \mathbf{g}, \mathbf{m})$.

Listing 3: `Control Device` process implemented in For-SyDe.

```
1 ctrlDev r = (cv, cts)
2     where (cv, m, ct) = unzip3SY $ comb3SY (
          ctrlDevLogic prosopon1) r m' cv'
3         cts = unzip5SY ct
4         m' = delaySY 0 m
5         cv' = delaySY (1,2,3) cv
```

Finally, Listing 4 shows the TMR process network from Figure 2 implemented in ForSyDe.

Listing 4: TMR process network implemented in ForSyDe.

```
1 tmrPN s_in = (r, s_out, out2, out4)
2     where out1 = rtrp1 ct1' s_in x'
3           out2 = rtrp2 ct2' s_in x'
4           out3 = rtrp3 ct3' s_in x'
5           out4 = rtrp4 ct4' s_in x'
6           out5 = rtrp5 ct5' s_in x'
7           (s_out, x, r) = voter cv' out1 out2
                out3 out4 out5
8           x' = delaySY (Prst 0) x
9           (cv, (ct1,ct2,ct3,ct4,ct5)) =
                ctrlDev r
10          cv' = delaySY (1,2,3) cv
11          ct1' = delaySY Abst ct1
12          ct2' = delaySY Abst ct2
13          ct3' = delaySY Abst ct3
14          ct4' = delaySY Abst ct4
15          ct5' = delaySY Abst ct5
```

4.1 Simulation Results

To simulate the TMR architecture, we first need to define the functionalities of each RTRP. The first three RTRPs are implemented to behave as accumulators, *i.e.* each input is added to the result of the previous execution. Functions **f** and **g**, from (16) and (17), are used to implement such accumulator. To simulate a failure in one of these three RTRPs (in this case, we chose to be $RTRP_2$) we implemented a faulted accumulator, replacing **f** for **f̃** given by (18), in which when the result of the previous execution is 3, instead of adding the input to it, it will subtract. RTRPs 4 and 5 are implemented using **f̄** given by (19). We assume that it takes 2 clock cycles to reconfigure a new RTRP, *i.e.* **m** = 2.

$$\mathbf{f}(x,u) = x + u \tag{16}$$

$$\mathbf{g}(x,u) = x \tag{17}$$

$$\tilde{\mathbf{f}}(x,u) = \begin{cases} x - u & \text{if } x = 3 \\ x + u & \text{otherwise} \end{cases} \tag{18}$$

$$\bar{\mathbf{f}}(x,u) = x - u \tag{19}$$

Table 1 shows the simulation results considering a constant input stream of ones. When $k = 4$, $RTRP_2$ outputs the wrong result, as seen in y_2. At the same instant, the `Voter` detects such fault and signals the `Control Device` that an error occurred in $RTRP_2$ and, therefore, it needs to be replaced. Then, the `Control Device` starts the reconfiguration procedure for $RTRP_4$, which takes 2 cycles to complete. When $k = 7$, $RTRP_4$ is fully reconfigured and matches the $RTRP_1$ and $RTRP_3$ outputs. As we can see, the output given by s_{out} is not affected

by the fault in $RTRP_2$, nor the reconfiguration of $RTRP_4$. We can also notice that $r = 4$ when $k = 5$ and $k = 6$, indicating that the `Voter` is signaling the `Control Device` about an error in the output of $RTRP_4$. As $RTRP_4$ is being reconfigured in this interval, the `Control Device` is in the Reconf state and, therefore, is ignoring the values arriving through r.

Table 1: Simulation results in ForSyDe

k	0	1	2	3	4	5	6	7	8	9
s_{in}	1	1	1	1	1	1	1	1	1	1
s_{out}	0	1	2	3	4	5	6	7	8	9
y_1	0	1	2	3	4	5	6	7	8	9
y_2	0	1	2	3	2	3	2	3	2	3
y_3	0	1	2	3	4	5	6	7	8	9
y_4	0	-1	-2	-3	-4	⊥	⊥	7	8	9
y_5	0	-1	-2	-3	-4	-5	-6	-7	-8	-9
r	⊥	⊥	⊥	⊥	2	4	4	⊥	⊥	⊥

5 Related Work

The idea of using triple modular redundancy with runtime reconfiguration is not new. SRAM field programmable gate arrays (FPGAs) must protect its configuration memory from SEUs, and TMR techniques are applied to such devices. However, when a majority voter is fed with two wrong answers, possibly caused by multiple independent SEUs, it produces the wrong result. One way to solve this issue is to periodically write back the whole bistream of each module, which is time consuming and leaves the modules inactive during this period. [7] proposes an optimization of the reconfiguration time in order to cope with this problem.

Another application of TMR using RTR is presented by [8], where an adaptive reconfigurable voting mechanism whose main goal is to extend the dynamic and partial reconfiguration SEU mitigation to the voter, which is usually the single point of failure in TMR architectures.

A novel technique for synchronizing the states of a newly reconfigured module is presented in [9]. Such technique consists on predicting the future state to which the system will soon converge (check point state) and presetting the reconfigured module to it. Therefore, only the reconfigured module will be set on hold until the check-point state is reached.

The research introduced in [10] claims an improvement of fault resilience, on up to 80%, by composing and applying space and time redundancy, *i.e.* multiprocessors and scheduling, with task migration among processors in hard real-time systems design. That architecture follows the multiple instruction, multiple data (MIMD) taxonomy, as proposed by [11].

6 Conclusion

This paper introduced a high level abstraction architecture for safety-critical systems with runtime reconfiguration (RTR) using the triple modular redundancy and the synchronous (SY) model of computation (MoC).

The triple modular redundancy was chosen to be the mechanism for detecting and masking faults. While the triple modular redundancy is a classic way to implement fail mitigation in safety-critical systems, in the event of a permanent fault, the system can mask such fault. However it gets vulnerable to a second fault.

A triple modular redundancy using RTR provides a way for the system to circumvent failures in the presence of multiple permanent faults, provided that no "two faults" happen in a time interval defined by the reconfiguration time of a new module.

We implemented the proposed high level architecture model in the framework ForSyDe and verified that a new `RTRP` can be correctly reconfigured in **m** cycles and can have its states synchronized with the other two `RTRP`s.

Acknowledgments

References

[1] Pierre Bieber, Frédéric Boniol, Marc Boyer, Eric Noulard, and Claire Pagetti. New Challenges for Future Avionic Architectures. *AerospaceLab*, (4):p. 1–10, May 2012.

[2] R. E. Lyons and W. Vanderkulk. The use of triple-modular redundancy to improve computer reliability. *IBM Journal of Research and Development*, 6(2):200–209, April 1962.

[3] E.A. Lee and A. Sangiovanni-Vincentelli. A framework for comparing models of computation. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 17(12):1217–1229, Dec 1998.

[4] Denis S. Loubach. A runtime reconfiguration design targeting avionics systems. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, pages 1–8, Sacramento, USA, September 2016. IEEE.

[5] Augusto Y. Horita, Ricardo Bonna, and Denis S. Loubach. Analysis and comparison of frameworks supporting formal system development based on models of computation. In Shahram Latifi, editor, *16th International Conference on Information Technology-New Generations (ITNG 2019). Advances in Intelligent Systems and Computing, vol 800*, pages 161–167, Cham, 2019. Springer International Publishing.

[6] Ingo Sander, Axel Jantsch, and Seyed-Hosein Attarzadeh-Niaki. ForSyDe: System design using a functional language and models of computation. In Soonhoi Ha and Jürgen Teich, editors, *Handbook of Hardware/Software Codesign*, pages 99–140. Springer Netherlands, 2017.

[7] L. Sterpone and A. Ullah. On the optimal reconfiguration times for tmr circuits on sram based fpgas. In *2013 NASA/ESA Conference on Adaptive Hardware and Systems (AHS-2013)*, pages 9–14, June 2013.

[8] F. Veljković, T. Riesgo, and E. de la Torre. Adaptive reconfigurable voting for enhanced reliability in medium-grained fault tolerant architectures. In *2015 NASA/ESA Conference on Adaptive Hardware and Systems (AHS-2015)*, pages 1–8, June 2015.

[9] Conrado Pilotto, José Rodrigo Azambuja, and Fernanda Lima Kastensmidt. Synchronizing triple modular redundant designs in dynamic partial reconfiguration applications. In *Proceedings of the 21st Annual Symposium on Integrated Circuits and System Design*, SBCCI '08, pages 199–204, New York, NY, USA, 2008. ACM.

[10] D. S. Loubach and A. M. da Cunha. Avionics hard real-time systems' concerning fault tolerance. In *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC)*, pages 6A2–1–6A2–18, Oct 2012.

[11] M. J. Flynn. Some computer organizations and their effectiveness. *IEEE Transactions on Computers*, C-21(9):948–960, Sep. 1972.

Lempel-Ziv-Markov Chain Algorithm Modeling using Models of Computation and ForSyDe

Augusto Y. Horita[1], Ricardo Bonna[1], Denis S. Loubach[2], Ingo Sander[3], and Ingemar Söderquist[4]

E-mail: ahorita@fem.unicamp.br; rbonna@fem.unicamp.br; dloubach@ita.br; ingo@kth.se;
ingemar.soderquist@saabgroup.com

[1]Advanced Computing, Control & Embedded Systems Lab, University of Campinas – UNICAMP,
Campinas, SP, Brazil - 13083-860

[2]Department of Computer Systems, Computer Science Division, Aeronautics Institute of Technology – ITA,
São José dos Campos, SP, Brazil - 12228-900

[3]Division of Electronics/School of EECS, KTH Royal Institute of Technology, SE-164 40, Kista, Sweden

[4]Business Area Aeronautics, Saab AB, Linköping, Sweden

Abstract

The data link is considered a critical function of modern aircraft, responsible for exchanging information to the ground and communicating to other aircraft. Nowadays, the increasing amount of exchanged data and information brings the need for network usage optimization. In this sense, data compression is considered a key approach to make data packages size smaller. Regarding the fact that avionics systems are safety-critical, it is fundamental not losing data nor performance during the compression procedures. In this context, manufacturers and regulatory agencies usually follow DO-178C guidance. Targeting model-based embedded design guidelines, DO-178C includes a supplement document, named DO-331. In this paper, we describe a widely used data compression algorithm, the Lempel-Ziv-Markov Chain algorithm (LZMA). Regarding formal model-based design, we argue that the synchronous dataflow model of computation captures the algorithm behavior more directly. The Formal System Design (ForSyDe) methodology is used to model the LZMA.

Keywords: data compression algorithm, avionics data link, DO-178C, DO-331, formal models of computation, synchronous dataflow.

1 Introduction

A wide range of aircraft functions with different safety requirements is present in the avionic system of today's modern aircraft. One of these aircraft functions is the *data link*. It is responsible for exchanging information to the ground, besides communicating to other aircraft during flight.

As the number of aircraft functions increases, the complexity of avionics systems exponentially grows. The number of processors, transducers and exchanged data and information also increase. This brings the need for network usage optimization. In this sense, *data compression* is considered a key approach to making data packages size smaller.

Considering that the avionics systems are safety-critical, it is fundamental not losing data nor performance during the compression procedures [1]. In this scenario, manufacturers and regulatory agencies follow the DO-178C [2] guidance. That document is developed and maintained by the Radio Technical Commission for Aeronautics. DO-178C aims to ensure that software development for avionics systems is dependable, safe, and meet the specified requirements.

Because of this, formal models of computation (MoC) have been used to *model*, *simulate* and *verify* algorithms in the system design and implementation phases, considering the embedded systems area. DO-178C includes a supplement document describing model-based design guidelines, named DO-331 [3].

Given this context, we address in this paper the formal modeling and simulation of a widely used data compression algorithm, named Lempel-Ziv-Markov Chain algorithm (LZMA). It is included as a CPU benchmark by the Standard Performance Evaluation Corporation (SPEC) [4].

Regarding formal model-based design, we argue that the synchronous dataflow (SDF) model of computation captures the algorithm behavior more directly. The LZMA intends to generate a compressed file based on the processing of a general data stream input. The Formal System Design (ForSyDe) methodology [5] is used to model the LZMA in our paper.

2 Background

This section presents the concepts used along with this paper including *models of computation* (MoC), *synchronous dataflow* (SDF) MoC, and the *Lempel-Ziv-Markov Chain* algorithm (LZMA).

2.1 Model of Computation (MoC)

An MoC represents an abstraction of a real computing device [6], which may have different behaviors and applications. It includes the relevant characteristics and properties for that particular model. Thus, different MoCs are used for modeling different systems depending on their behavior.

MoCs can be defined as an abstract rules collection stating the semantics of execution and concurrency in computational systems.

In this context, the *tagged signal model* (TSM) is also present. It is introduced as a meta-model, or even a framework, defining systems as compositions of processes acting on signals [7].

A *signal* is a set of events $e_i = (t_i, v_i)$, which are elementary units of information composed by a tag $t_i \in T$, and a value $v_i \in V$. A signal can be viewed as a subset of $T \times V$. A *process* P is a set of possible behaviors acting on a signal. The set of output signals S^O is given by the intersection between the input signals S^I and the process: $S^O = S^I \cap P$. A process is *functional* when there is a single value mapping $f : S^I \to S^O$ which describes it. Therefore, a functional process has either one behavior or no behavior at all.

TSM divides MoCs basically into *timed* and *untimed*. In a *timed* MoC, the set of tags T is *totally ordered*, *i.e.* one can order every event included in the model based on its tag. On the other hand, in an *untimed* MoC, the set of tags T is *partially ordered*, *i.e.* only local groups of events can be ordered based on its tag, *e.g.* the ones belonging to the same signal.

2.2 Synchronous Dataflow (SDF) MoC

A subclass included in the untimed MoCs comprehends the *dataflows*, which are directed graphs where each *node* represents a process and each *arc* represents a signal path. When a process is activated, *i.e.* fired, it consumes a certain amount of data, denominated *tokens*, from its input ports and generates another amount of tokens for its output ports. The amount of tokens consumed and produced by the ports in a single activation cycle is denominated *token rate* [8].

In SDF, the token rates are fixed and predefined, represented by a natural number associated with each input and output ports. An actor can fire only if the input signal paths have enough tokens to supply the amount needed by all input ports of the actor. As a consequence, no signal path can have a negative amount of tokens.

The predefined and fixed token rates of SDF actors allows efficient modeling of systems like signal processing or finding a static schedule for single and multi-processor implementation and also buffer size definitions [8].

2.3 Data Compression Algorithms

Data compression has been used to optimize storage and communication buses. One of the compression techniques is based on dictionaries, which consists of saving strings of a previously read input stream interval, composing a dictionary. When new groups of symbols are being read in, the algorithm searches through the dictionary, then the matches are encoded as pointers and sent as the output.

Abraham Lempel and Jacob Ziv have presented their first dictionary-based compression method in 1977 [9], which is referred to as LZ77. That method limits the previously read interval to a determined size window, which follows the data processing, creating the concept of *sliding window encoding* (SWE), which makes the algorithm simple and faster [10]. In 1978, LZ78 was presented [11]. The difference is that its dictionary is composed of all the symbol strings from previously read stream and it builds a single character at a time, making this algorithm longer but with a higher compression rate.

The algorithm analyzed in this paper, LZMA, was first used in 7z file format [12] and is presented as a benchmark by SPEC [4], comparing its implementation using different architectures.

2.3.1 Lempel-Ziv-Markov Chain Algorithm (LZMA)

LZMA was created as an LZ77 optimization, providing higher compression rate and fast decompression, with lower memory requirements [13]. It is open source and implemented in several different programming languages, including C, and Java, as an SDK [12].

LZMA can be divided into two steps, although an additional filtering step can be added before LZMA, aiming the optimization of the compression algorithm, as illustrated in Fig. 1.
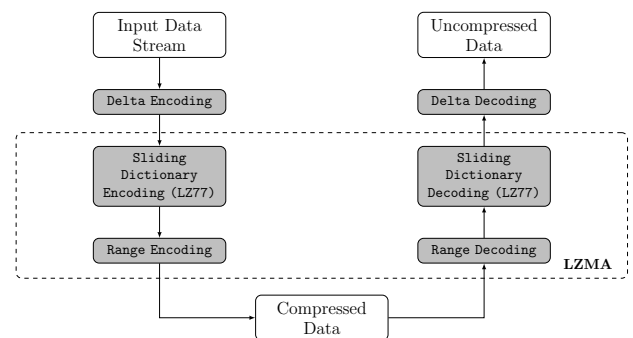
Figure 1: LZMA compression and decompression block diagram, based on [10].

$Step_1$ *Delta encoding* – this step consists of an optional data filtering that encodes the input stream outputting, for each byte, a byte representing the difference between the current and the previous bytes. The first byte output is itself. This process makes the sliding window more efficient;

$Step_2$ *Sliding dictionary encoding* – this step is based on the LZ77 algorithm, although it supports larger dictionaries.

Optimized search algorithms were implemented to perform faster searches in these dictionaries, named *hashed chain* and *binary trees* [13]. The output is, as in LZ77, a sequence of triplets, composed by the *distance* from the string in the look-ahead buffer to its match in the search buffer, the *length* of the string, and the next input *symbol*; and

$Step_3$ *Range encoding* – this step is context-based. The compressed range in each iteration is estimated based on probabilistic algorithms and can form a set of predefined types of packages depending on the input range size. This was first presented by [14]. In LZMA, the compression adopted context is the output of $Step_2$ in each firing cycle.

3 Lempel-Ziv-Markov Chain Modeling

This section presents the LZMA modeling based on the SDF MoC. Next, a brief model description is presented, followed by the model design using ForSyDe.

3.1 Model Description

The LZMA compression has as input a data stream, which is processed by a sequence of defined steps, outputting a compressed and encoded byte stream. Such behavior can be modeled using a dataflow MoC. Towards simple modeling, our paper presents the LZMA compression model based on SDF MoC, considering the decompression to be a similar process.

LZMA has a set of parameters that are user-defined at the compression beginning. These parameters configure the algorithm behavior, data structures, and performance. For an initial formal modeling purpose, this paper considers some specific configurations. The following introduces a list of assumptions taken into consideration to model the simplified LZMA version.

A_1 the absence of the delta encoding filter;

A_2 the sliding window encoder step is changed to use a similar algorithm, *i.e.* LZ78 sliding window encoder [11] was adopted instead; and

A_3 a fixed probability range encoder.

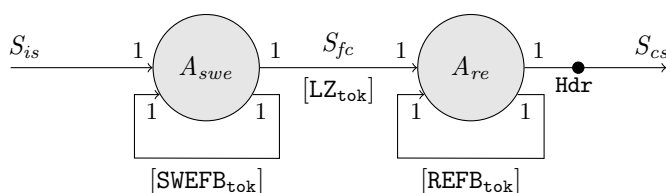Figure 2 illustrates the LZMA modeled as a SDF functional graph.

Figure 2: LZMA SDF MoC graph

The model blocks and communication paths are defined as follows:

- A_{swe} – Sliding window encoding block. It encodes the input stream based on a dictionary structure, outputting a literal, a match or an absent value in each process firing;

- A_{re} – Range encoding block. It performs a bit-wise compression of the literals and matches, generating a byte for each one, resulting in the final compressed stream;

- S_{is} – Input stream. It comprehends the input to be compressed;

- S_{fc} – Compressed stream first step. It contains the tokens already compressed using the sliding window encoding method;

- S_{cs} – Compressed stream. It represents the final LZMA output;

- SWEFB$_{tok}$ – Sliding window encoding feedback token. It is composed of the dictionary data structure, updated at each A_{swe} firing;

- REFB$_{tok}$ – Range encoding feedback token. It is composed of the considered range, lower limit and encoded character cache memory, updated at each A_{re} firing;

- LZ$_{tok}$ – Sliding window encoder output token. It can have an absent value or a tuple, depending on the A_{swe} read input token. In the latter case, the tuple is composed by a read character and another entry. If the read character is not in the dictionary, the first tuple entry is an absent value, and the token represents a *Literal*. Otherwise, the first entry represents the number of characters that was repeated inside the sliding window, and the token represents a *Match*; and

- Hdr – Initial bytes contained in the output compressed stream. It comprehends the configuration setup, besides the dictionary and uncompressed input stream sizes.

3.2 Modeling supported by ForSyDe

The ForSyDe framework was chosen to be used in the present paper based on the frameworks comparison presented in [15]. In view of that, we use ForSyDe to model LZMA as described in the previous section.

In ForSyDe, the SDF MoC actors, *i.e.* blocks presented in last section, are modeled as *processes*, and the communication paths as *signals*.

Listing 1 presents the model processes definitions, using the process constructors from the ForSyDe SDF library. For brevity, this paper only presents the main processes and signals models and definitions. However, the full functional implementation of the LZMA model can be found in [16].

Regarding the sliding window encoding, the A_{swe} actor was modeled as lzA, using the ForSyDe SDF process constructor actor22SDF, which indicates that the process has 2 inputs and 2 outputs. The SWEFB$_{tok}$ dictionary was modeled as fb. The LZMA input S_{is} was represented by the sIs signal, and the S_{fc} by the sFc signal.

When modeling the range encoding step, the actor A_{re} was represented as rgA based on the process constructor

actor22SDF. The REFB_tok range encoding variables tuple was modeled as sFb. The output compressed stream S_{cs} was modeled as sCs, and the initial header Hdr as initHdr.

Listing 1: LZMA code snippet in ForSyDe SDF/Haskell

```
1  -- LZ (Sliding Window) Encoding actor definition
2  -- Input is the lzma input stream Sis
3  -- Output is the first step compressed stream Sfc
4  lzA :: Signal Char -> Signal (Maybe (Maybe Int, Char))
5  lzA Sis = Sfc
6    where (Sfc, fb) = actor22SDF (1,1) (1,1) lzF Sis fb'
7          fb' = delaySDF [([],"",0)] fb
8
9  -- Range Encoding actor definition.
10 -- Input is the first step compressed stream Sfc
11 -- Output is the compressed LZMA stream output Scs
12 rgA :: Signal (Maybe (Maybe Int, Char))
13        -> Signal [Char]
14 rgA sFc = sCs'
15   where (sCs,sFb) = actor22SDF (1,1) (1,1) rangeFunc
      sFc sFb'
16         sCs' = delaySDF initHdr sCs
17         sFb' = delaySDF [(rangeInit,0, chr 0)] sFb
18         initHdr = [([dictsize] ++ [inpLen])]
19         inpLen = (intToDigit (lengthS sFc))
```

Listing 2 presents how the simplified LZMA *process network* lzmaSdf is assembled to compose the algorithm completely written using ForSyDe MoC. The range encoding process rgA is executed having the output of sliding window process lzA as input.

Listing 2: LZMA process netword in ForSyDe SDF/Haskell

```
1  module Lzmasdf (
2      lzmaSdf
3    ) where
4
5  import ForSyDe.Shallow
6  import Rangesdf
7  import LZsdf
8  import Data.Char
9
10 -- Simplified LZMA process network
11 lzmaSdf :: Signal Char -> Signal [Char]
12 lzmaSdf sLzmaIs = sLzmaOut
13     where sLzmaOut = rgA (lzA sLzmaIs)
```

4   Summary

This paper presented a simplified Lempel-Ziv-Markov Chain algorithm (LZMA) modeling based on the synchronous dataflow (SDF) model of computation (MoC) using the ForSyDe framework. Some configurations and behaviors assumptions were adopted towards the definition of actors ports fixed token rates. Those assumptions allowed for a first simple LZMA formal modeling and simulation.

As future work, we plan to implement LZMA based on the scenario-aware dataflow (SADF) modeling and simulation framework introduced in [17], which supports variable token rates, resulting in a more advanced LZMA model.

References

[1] Paul Berthier, Corentin Bresteau, and José Fernandez. On the security of aircraft communication networks. In Gregorio D'Agostino and Antonio Scala, editors, *Critical Information Infrastructures Security*, pages 266–269, Cham, 2018. Springer International Publishing.

[2] Radio Technical Commission for Aeronautics - RTCA. *DO-178C - Software Considerations in Airborne Systems and Equipment Certification*, 2012.

[3] Radio Technical Commission for Aeronautics - RTCA. *DO-331 - Model-Based Development and Verification Supplement to DO-178C and DO-278A*, 2011.

[4] Standard Performance Evaluation Corporation (SPEC). 657.xz_s spec cpu 2017 benchmark description. http:/www.spec.org/cpu2017/Docs/benchmarks/657.xz_s.html, 2019.

[5] Ingo Sander. The forsyde methodology. In *Swedish System-on-Chip Conference*, 2002.

[6] Axel Jantsch. Models of embedded computation., 2005. Embedded Systems Handbook, chapter Models of Embedded Computation. CRC Press.

[7] E.A. Lee and A. Sangiovanni-Vincentelli. A framework for comparing models of computation. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 17(12):1217–1229, Dec 1998.

[8] E.A. Lee and D.G. Messerschmitt. Synchronous data flow. *Proceedings of the IEEE*, 75(9):1235–1245, 1987.

[9] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory*, 23(3):337–343, May 1977.

[10] Epiphany Jebamalar Leavline. Hardware implementation of lzma data compression algorithm. In *International Journal of Applied Information Systems (IJAIS)*, 2013.

[11] J. Ziv and A. Lempel. Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory*, 24(5):530–536, Sep. 1978.

[12] Igor Pavlov. 7z format. http://www.7-zip.org/7z.html, 2019.

[13] David Salomon. *Data Compression: The Complete Reference*. Springer, 2007. With contributions by Giovanni Motta and David Bryant.

[14] G. Nigel Martin. Range encoding: an algorithm for removing redundancy from a digitised message. In *Video and Data Recording Conference*, 1979.

[15] Augusto Y. Horita, Ricardo Bonna, and Denis S. Loubach. Analysis and comparison of frameworks supporting formal system development based on models of computation. *Springer - Advances in Intelligent Systems and Computing*, 2019.

[16] A. Y. Horita. Forsyde simplified lzma sdf model source code. https://github.com/AugustoHorita/LzmaForSyDeSdf, 2019.

[17] Ricardo Bonna, Denis S. Loubach, George Ungureanu, and Ingo Sander. Modeling and simulation of dynamic applications using scenario-aware dataflow. *ACM Transactions on Design Automation of Electronic Systems*, 2019.

Considerations on Domain-Specific Architectures Applicability in Future Avionics Systems

Denis S. Loubach[1], Johnny Cardoso Marques[2], and Adilson Marques da Cunha[2]

[1]Department of Computer Systems
[2]Department of Software and Information Systems
Aeronautics Institute of Technology – ITA, São José dos Campos, São Paulo/Brazil - 12228-900
E-mail: dloubach@ita.br; johnny@ita.br; cunha@ita.br

Abstract

The community related to the avionics systems domain tends to be conservative mainly due to the safety-critical requirements this kind of system has to comply with. With respect to the presented context, this paper addresses and discusses the next challenges of future avionics systems always considering the safety-critical aspects and the possibility of formal models of computation application together with domain-specific architectures. Runtime partial reconfiguration, for example, is one promising way to mitigate the increasing failure rate and thereby fulfill safety requirements. N-modular redundancy implementation with different hardware architectures is also analyzed considering reconfigurable computing.

Keywords: domain-specific architectures, avionics systems, safety-critical systems, runtime reconfiguration, reconfigurable computing, formal models of computation.

1 Introduction

The industry community related to the avionics systems domain tends to be conservative mainly due to the safety-critical requirements this kind of system has to comply with. In that case, the majority of development follows the DO-178 [1] for the software, and the DO-254 [2] for the hardware guidance.

On the other hand, high-volume commercial electronics usually dictates some trends such as environmental concerns like the restriction of hazardous substances and packing more and more functionality with enhanced performance into smaller packages also taking into account lower power-consumption. Along the time, this led to chip geometries in the order of nanometers.

Nowadays, computer architecture improvement to deliver performance enhancements is getting harder to achieve, compared to the renascence period from about 50 years ago. The more relevant facts are the end of Dennard scaling and the slowdown of Moore's Law. Chien [3] points out the end of Dennard's scaling, shift to multicore, and slowing in the clock rate growth as *disruption from bellow*. That author also mentions the considerable architectures customization in mobile and embedded devices. Those facts have impact on the avionics systems domain as the system complexity increases exponentially from one generation to the next.

In view of this, one current trend in high-performance computer architecture is the use of *domain-specific architectures* (DSA) instead of the general-purpose ones [4]. DSA is able to provide performance and power benefits to face general-purpose architectures that are not having significant performance improvements in the last years. Potential research areas include high-quality implementations of *open source architectures* and *reconfigurable computing* which are nowadays supported by system-on-chip (SoC) composed by a hard processor and FPGA with runtime full and partial reconfiguration enabled [5].

With respect to the presented context, this paper addresses and discusses the next challenges of future avionics systems always considering the safety-critical aspects and the possibility of formal models of computation application together with DSA. Runtime partial reconfiguration, for example, is one promising way to mitigate the increasing failure rate and thereby fulfill safety requirements. N-modular redundancy implementation with different hardware architectures is also analyzed considering reconfigurable computing.

2 Software & Hardware Working Together

Notice that software and hardware always needed to work together, especially with respect to embedded systems.

One of the basic concepts of embedded systems is hardware and software tightly coupled to address an specific function. However, the modern concept may be slightly different in the sense embedded systems are becoming more "general-purpose" concerning its hardware architecture and computing power.

In this scenario, the following sections introduce the four standards used to develop a safety-critical class of real-time

embedded systems, *i.e.* avionics. Next, it is presented a general overview of domain-specific architectures, reconfigurable computing, and models of computation.

2.1 System, Software & Hardware Avionics Standards

Typically, safety-critical systems are developed in regulated environments by norms and standards. Examples are found in domains such as: aviation, automotive, medical, railway, space, and nuclear [6].

In avionics systems development, there are four standards typically required to be used as part of the aircraft certification process: SAE ARP-4754A [7], RTCA DO-178C [1], RTCA DO-254 [2], and RTCA DO-297 [8]. The relationship among these standards is presented in Fig. 1.
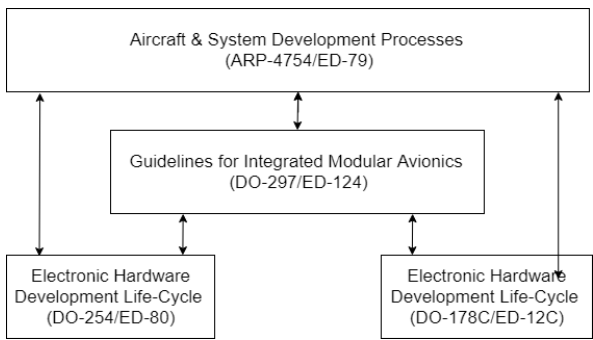

Figure 1: Relationship among certification standards

All those standards use *safety* classification. The system failures are classified within five categories: catastrophic, hazardous, major, minor, and no safety impact. The products (systems, software, or hardware) which malfunction, cause or contribute to a system failure occurrence have an attributed design assurance level (DAL) according to the most critical system failure condition.

2.1.1 SAE ARP-4754A

According to Xiaoxuna *et al.* [9], the SAE ARP-4754A [7] has been prepared primarily for electronic systems which, by their nature, may be complex and are readily adaptable to high levels of integration. However, the guidance is also applicable to engine systems and related equipments. It provides updated and expanded guidelines for processes used to develop civil aircrafts and systems that implement aircraft functions.

The SAE ARP-4754A has 47 objectives presented in Table A-1, published in Appendix A of the standard. Those objectives are organized in 8 processes: Planning; Aircraft and system development process and requirements capture; Safety assessment; Requirements validation; Implementation verification; Configuration management; Assurance; and Certification authority coordination. These processes are organized in process model as presented in Fig. 2.

2.1.2 RTCA DO-178C

The first version of RTCA DO-178 was released in 1982. The European Organization for Civil Aviation Equipment
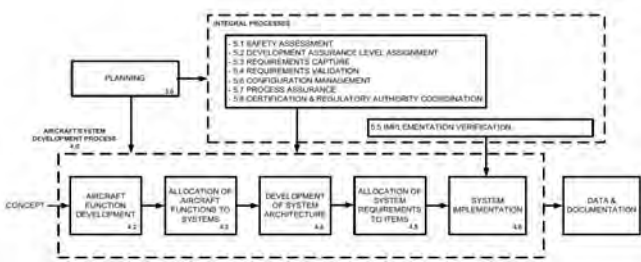

Figure 2: General processes available at SAE ARP-4754A

(EUROCAE) has an Equivalent Document ED-12. In 2011, the current RTCA DO-178C [1] was released.

The RTCA DO-178C has five software levels broken down into objectives to be satisfied. The satisfaction of applicable objectives enables the software approval as part of the aircraft certification process. Among the five existing software levels (A, B, C, D, and E), level A is the most rigorous and requires compliance with all objectives of the standard. The level E refers to software products which malfunction does not result in loss of safety margins.

As presented in Table 1, the classification of the failure condition is associated with the defined software levels. For each software level, a set of objectives are required for compliance demonstration.

Table 1: RTCA DO-178C software levels

Failure Condition	Software Level	Objectives
Catastrophic	A	71
Hazardous	B	69
Major	C	62
Minor	D	24
No Safety Impact	E	None

The 71 RTCA DO-178C objectives are presented in 10 tables, published in Annex A of the standard. The tables identify software process objectives with the following characteristics: Planning (Table A-1); Development (Table A-2); Verification of the high and low-level requirements and software architecture (Tables A-3, A-4, and A-5); Verification of source and executable codes (Tables A-5 and A-6); Testing and analysis (Table A-7); Configuration control (Table A-8); Quality assurance (Table A-9); and Certification (Table A-10). These processes are organized and presented in Fig. 3.

As part of the effort of the DO-178C release, other supplementary standards were also developed, including special recommendations, regarding tool qualification (RTCA DO-330 [11]), *model-based development and verification* (RTCA DO-331 [12]), object-oriented technology (RTCA DO-332 [13]), and formal methods (RTCA DO-333 [14]).

2.1.3 RTCA DO-254

According to Kounish *et al.* [15], the RTCA DO-254 [2] is the standard which is used for the airborne and safety-critical application providing a proper guidance to assure the design
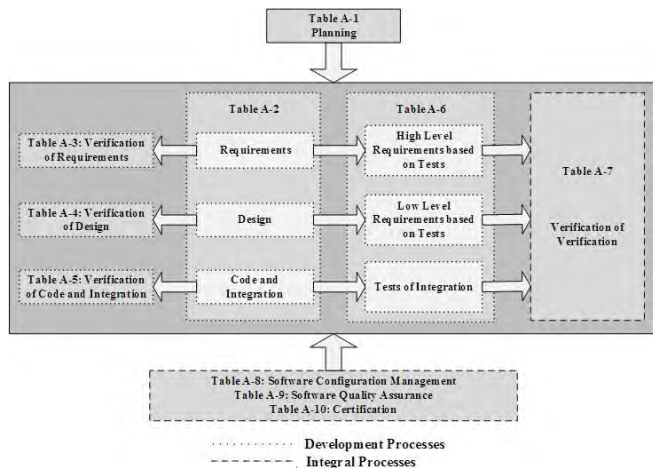
Figure 3: RTCA DO-178C processes organization [10]

of airborne electronic hardware. It has many similarities with RTCA DO-178C. As an example, the RTCA DO-254 has also five design assurance levels broken down into objectives to be satisfied.

The RTCA DO-254 objectives are presented in Table A-1 and published in Appendix A of the standard. It has 6 processes: Planning; Requirements; Conceptual design; Detailed design; Implementation; Validation; Verification; Configuration management; Assurance; and Certification. These processes are organized and presented in Fig. 4.
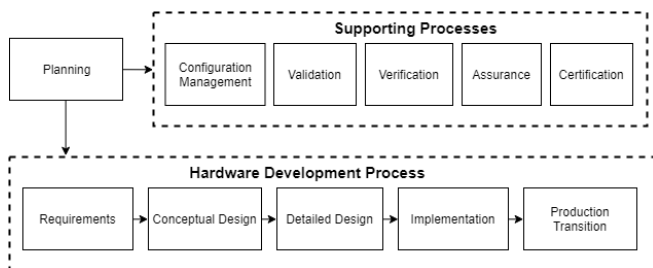

Figure 4: RTCA DO-254 processes organization [2]

2.1.4 RTCA DO-297

According to RTCA DO-297 [8] glossary, the integrated modular avionics (IMA) is

> "a shared set of flexible, reusable, and interoperable hardware and software resources that, when integrated, form a platform that provides services designed and verified to a defined set of safety and performance requirements, to host applications performing aircraft functions".

An IMA system architecture is composed of one or more platforms and includes interfaces to other aircraft systems and users.

The IMA typical approach consists of the following levels of acceptance: component; module; application; platform; IMA

system; and IMA aircraft. These levels of acceptance are organized as illustrated in Fig. 5.

A *component* is a self-contained hardware part, software part, database, or combination thereof that is configuration controlled. A component does not provide an aircraft function by itself.

A *module* may be software, hardware, or a combination of hardware and software, which provides resources to hosted applications. Modules may be distributed across the aircraft or may be co-located.

An *application* is a collection of software and/or hardware modules with a defined set of interfaces that, when integrated with a platform, performs a function.

At component, module, and application levels, the developments of such parts should follow the RTCA DO-178C and/or the RTCA DO-254.

A *platform* is a group of modules that establishes a computing environment, support services, and platform-related capabilities, such as health monitoring and fault management.

An *IMA system-level* consists of platform(s) and a defined set of hosted applications.

The IMA aircraft-level should demonstrate that each aircraft function and hosted application functions as intended, supports the aircraft safety objectives, and complies with the applicable regulations. However, during the installation activities, the interactions between hosted applications relative to the provided aircraft functions should be verified and validated during aircraft ground and flight testing.
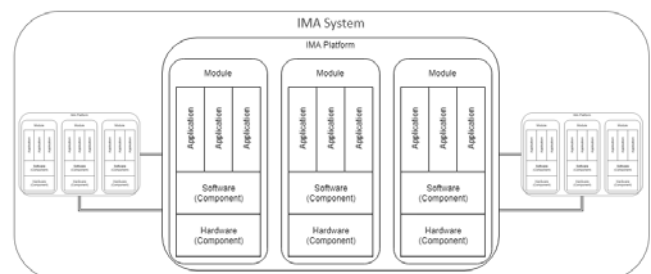
Figure 5: Levels of IMA approval

2.1.5 Wrap-Up

Together, these four standards are closely pursued in the development of avionics systems. From the system perspective, passing through modules, down to the software and the hardware.

The conservatism from the avionics system community is mainly due to the levels of impact a failure may cause, for example, level A in DO-178. Therefore, new trends are almost always seen as "dangerous" to be included or considered in avionics.

However, if those trends are backed-up by sound basis, they can be step by step introduced in future generation systems. We argue that it is the case regarding domain-specific architectures supported by reconfigurable computing and models

of computation.

2.2 Domain-Specific Architectures

During the semiconductor boom, supported by the Moore's Law, general-purpose code was accelerated in general-purpose cores, including techniques such as cache hierarchy; 512-bit single instruction, multiple data (SIMD) floating-point units; various pipeline stages; dynamic branch prediction; out-of-order execution; speculative execution; multithreading; and multiprocessing. Nowadays, it is needed a paradigm shift in computer architecture to achieve a new level of efficiency. This shift concerns moving from general-purpose hardware to domain-specific architectures [4].

Gupta [16] mentions three different architectures, as shown in Table 2.

Table 2: Architecture paradigm and domain [16]

Architecture	Domain
Scalar (*e.g.* CPU)	Complex algorithms with diverse decision tree. Limited in performance scaling.
Vector (*e.g.* GPU)	Efficient in reduced set of parallelizable functions. Suffers latency and penalties related to memory hierarchy.
Programmable logic (*e.g.* FPGA)	Customized to computer a particular function. May take hours to compile and synthetize.

Central Processing Unit (CPU); Graphics Processing Unit (GPU); Field Programmable Gate Array (FPGA).

Different from general-purpose architectures, the hardware can be specific and optimized for a particular domain in DSA. Moreover, this can be done in runtime, as discussed in the next section.

In this sense, computers will be more and more *heterogeneous*.

2.3 Reconfigurable Computing and Formal Models of Computation

Reconfigurable computing includes computation research towards the use of reconfigurable devices. In this case, for a specific application requirement, *i.e. domain-specific*, and at given time frame, the reconfigurable devices' spatial structure is changed to comply with a given objective [17, 18].

Reconfigurability is the "art of engineering degrees of freedom into embedded systems", as highlighted by [19]. Notice, that this concept is different from the *adaptiveness*. The latter stands for a possible behavioral ability a system can have that aids to decide how the system should be reconfigured in a optimized fashion. Adaptive computing is stated as one of the reconfigurable computing research areas.

A reconfiguration design is presented in [5]. The author introduces a hardware and software composition which shows

the application of both *partial* and *full* runtime hardware reconfiguration based on a heterogeneous computer. In other words, a SoC with *scalar* and *programmable logic* elements integrated in a single package. Then, that reconfiguration design, composed by hardware and software, can be classified as a domain-specific architecture.

Architectures like that can be applied to the avionics systems domains, as far as they are able to provide the safety inherent from this domain.

In this sense we can use formal *models of computation* (MoC) together with runtime reconfiguration to provide flexibility and performance along with safety to future avionics systems.

Different models types exist to address different purposes. *Functional modeling* is used to address the functional behavior of a system. On the other hand, design and synthesis refine into implementation details. A *model* is defined as an abstraction of an entity, *e.g.* a physical system or even another model, according to [20]. That author also states that abstraction is a method for choosing which aspects will be taken into account when modeling a system.

A MoC is an abstraction of a physical computing device, and so, different MoCs serve to different objectives.

MoCs are based on three main points: *processes*, *events*, and *signals*. The following definitions are derived from [20, 21].

Definition 1. *Event. Elementary information unit exchanged between processes.*

Definition 2. *Signal. Processes communicate to each other by writing to and reading from signals, which are a sequence of events. Signals preserve the order that events are entered. Each event has a tag and a value. Tags can be used to model physical time, events order, and other key properties of a MoC.*

Definition 3. *Process. Receives and send events. The process activity is comprised of evaluation cycles, i.e. an application of a function which maps inputs to outputs. Then, in each evaluation cycle the process receives inputs, computes, and sends outputs.*

Definition 4. *Process Constructor. Parameterizable templates for instantiating processes, i.e. a higher-order function.*

Processes constructors are used to create processes. New processes can be created to form a hierarchical concurrent process network.

Definition 5. *Model of Computation. Set of processes and process networks implemented by a well-defined set of processes constructors.*

The modeling task can be supported by a tool or framework. Horita *et al.* [22] introduced a framework comparison method with a number of desirable characteristics the tools should have to aid in the formal modeling and simulation of systems. ForSyDe was one of the frameworks analyzed.

DOI
10.3384/ecp19162018   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **159**

Formal System Design (ForSyDe) [23] is a transformational design methodology based on the *functional programming paradigm*. It targets heterogeneous embedded systems [20].

A system is modeled as a hierarchical concurrent process network in ForSyDe. Processes communicate with each other by signals.

In ForSyDe, we model a *signal* as a list of *events*, where the event *tag* is implicitly given by the event position in the list. The semantics of a tag is defined by the currently in-use MoC, *i.e.*, an identical tag of two events in different signals does not imply these events happened at the same time. All events in a signal must have values of the same type [21].

The synchronous (SY) MoC splits the time domain into slots. Everything inside a slot occurs at the same time. The evaluation cycle of processes lasts exactly one time slot in synchronous MoC [20], according to the *perfect synchrony hypothesis*, *i.e.* neither computation nor communication takes time.

Synchronous processes consume and produce exactly one event on each input and output in each evaluation cycle. This implies the *total order* of all events in any signal in the SY MoC. Events with the same tag appear at the same time instant.

Taking these fundamental basis explored so far, the runtime reconfiguration can be modeled as higher-order function [24], as follows.

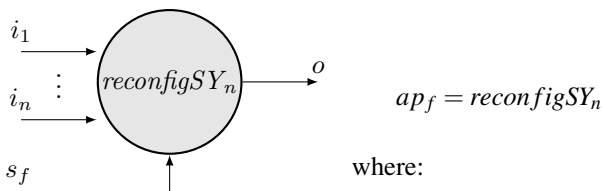The adaptive process *reconfigSY* definition is illustrated in Fig. 6.

$$ap_f = reconfigSY_n$$

where:

Figure 6: Process constructor for reconfigSY [21]

$$o = ap_f(s_f, (i_1, \ldots, i_n))$$
$$o[k] = s_f[k](i_1[k], \ldots, i_n[k])$$

reconfigSY synchronous reconfigurable process can be modeled by using the *mapSY* ForSyDe process constructor.

$$reconfigSY = mapSY(\$) \qquad (1)$$

Let's assume an N-modular redundancy [25]. Typically, this is applied to safety-critical system to provide dependability. Different modules may be implemented differently even using distinct hardware architectures, as long as their final output is the same. This is totally possible to be implemented using runtime reconfiguration, as shown in [26].

Concerning the aeronautics trend towards the second generation of integrated modular avionics adoption, DSA is certainly a very good candidate to be used in future avionics systems.

Examples of applications combining DSA and runtime reconfiguration are reduced instruction set computers (RISC) designed for a specific computation, and software-defined radio

(SDR). In the first case, it is possible to have many different RISC bitstreams in the memory and load or unload them in runtime using partial reconfiguration, according to the current need. Regarding SDR, it is possible to have different implementations and use them depending on the present context to address challenges such as size, weight, and power (SWaP).

Some benefits of this scheme are hardware reuse resulting in a reduced circuit area thus contributing to minimizing single event upset (SEU), and the full control of the trade-off involving power consumption and performance.

3 Next Challenges and Trends

Hennessy and Patterson [4] mention that a current DSA challenge is *how to port software*, generally written in programming languages such as C/C++ and Ada, to different architectures. This challenge is similar to the one faced back in the 80's: how to port code written in Assembly to different microcontrollers without the need to rewrite almost all of it. The answer for the latter was high level languages and hardware room. *i.e.* memory, to fit more and more code.

Another challenge to the integrated modular avionics is the *runtime hardware reconfiguration* application. The question is how to ensure the reconfiguration is safely performed in a deterministic way.

Confidently, DSA together with formal models of computation figure as a promising trend for the next years.

4 Summary

This paper presented some considerations, supported in the literature, on domain-specific architectures applicability to future avionics systems.

A research path is open concerning more and more application of formal models in real-time embedded safety-critical systems such as avionics. Besides, reconfigurable computing figure as one promising underlying element to achieve power-consumption and performance efficiency in domain-specific architectures.

5 Acknowledgments

References

[1] Radio Technical Commission for Aeronautics - RTCA. *DO-178C - Software Considerations in Airborne Systems and Equipment Certification*, 2012.

[2] Radio Technical Commission for Aeronautics - RTCA. *DO-254 - Design Assurance Guidance for Airborne Electronic Hardware*, 2000.

[3] Andrew A. Chien. Computer architecture: Disruption from above. *Commun. ACM*, 61(9):5–5, August 2018.

[4] J. L. Hennessy and D. A. Patterson. *Computer Architecture: A Quantitative Approach*. Morgan Kaufmann, 6th edition, 2017.

[5] Denis S. Loubach. A runtime reconfiguration design targeting avionics systems. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, pages 1–8, Sacramento, USA, September 2016. IEEE.

[6] Johnny Marques and Adilson Cunha. Verification scenarios of onboard databases under the rtca do-178c and the rtca do-200b. *36th IEEE/AIAA Digital Avionics Systems Conference*, 2017.

[7] Society of Automotive Engineers-SAE. *ARP-4754A Guidelines for Development of Civil Aircraft and Systems*, 2011.

[8] Radio Technical Commission for Aeronautics - RTCA. *DO-297 Integrated Modular Avionics (IMA) Development Guidance and Certification Considerations*, 2011.

[9] L.. Xiaoxun, Z. Yuanzhenb, F. Yichenb, and S. Duoa. A comparison of sae arp 4754a and arp 4754. *Procedia Engineering*, 17:400–416, 2011.

[10] Leanna Rierson. *Developing Safety-Critical Software: A Practical Guide for Aviation Software and DO-178C Compliance*. CRC Press, 2013.

[11] Radio Technical Commission for Aeronautics - RTCA. *DO-330 Software Tool Qualification Considerations*, 2011.

[12] Radio Technical Commission for Aeronautics - RTCA. *DO-331 Model Based Development and Verification Supplement to DO-178C and DO-278A*. RTCA, 2011.

[13] Radio Technical Commission for Aeronautics - RTCA. *DO-332 Object-oriented and Related Technologies Supplement to DO-178C and DO-278A*. RTCA, 2011.

[14] Radio Technical Commission for Aeronautics - RTCA. *DO-333 Formal Methods Supplement to DO-178C and DO-278A*. RTCA, 2011.

[15] R. Koushik, M. Anushree, B. J. Sowmya, and N. Geethanjali. Design of spi protocol with do-254 compliance for low power applications. *2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT)*, 2017.

[16] Amit Gupta. Advances in adaptable computing. In *Proceedings of the 2019 International Symposium on Physical Design*, ISPD '19, pages 37–38, New York, NY, USA, 2019. ACM.

[17] Christophe Bobda. *Introduction to Reconfigurable Computing: Architectures, Algorithms, and Applications*. Springer, 2007.

[18] Dirk Koch. *Partial Reconfiguration on FPGAs: Architectures, Tools and Applications*. Springer-Verlag New York, 1 edition, 2013.

[19] J. Lyke, C. G. Christodoulou, A. Vera, and A. H. Edwards. Reconfigurable systems: Advanced applications and technologies [scanning the issue]. *Proceedings of the IEEE*, 103(7):1000–1003, July 2015.

[20] Axel Jantsch. *Modeling Embedded Systems and SoC's: Concurrency and Time in Models of Computation*. Morgan Kaufmann, San Francisco, USA, 1st edition, 2003.

[21] Ingo Sander and Axel Jantsch. Modelling Adaptive Systems in ForSyDe. *Electronic Notes in Theoretical Computer Science*, 200(2):39 – 54, 2008. Proceedings of the First Workshop on Verification of Adaptive Systems (VerAS 2007).

[22] Augusto Y. Horita, Ricardo Bonna, and Denis S. Loubach. Analysis and comparison of frameworks supporting formal system development based on models of computation. *Advances in Intelligent Systems and Computing*, 800:161–167, 2019.

[23] I. Sander and A. Jantsch. System modeling and transformational design refinement in ForSyDe [formal system design]. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 23(1):17–32, Jan 2004.

[24] Denis S. Loubach, Eurípedes G. O. Nóbrega, Ingo Sander, Ingemar Söderquist, and Osamu Saotome. Towards runtime adaptivity by using models of computation for real-time embedded systems design. In *9th Aerospace Technology Congress*, Solna, Stockholm, October 2016. FTF - Swedish Society of Aeronautics and Astronautics.

[25] H. . Lo, L. . Ju, and C. . Su. General version of reconfiguration n modular redundancy system. *IEE Proceedings G - Circuits, Devices and Systems*, 137(1):1–4, Feb 1990.

[26] Ricardo Bonna, Denis S. Loubach, Ingo Sander, and Ingemar Söderquist. Triple modular redundancy based on runtime reconfiguration and formal models of computation. In *10th Aerospace Technology Congress*. Swedish Society of Aeronautics and Astronautics (FTF), 2019.

V Material and Structures

Post-buckling Analysis in Thin-web Laminated Composite Beams
Bruna Luiza Nolli, Carlos Cimini Jr. and Cristiano Pena

Failure Induced by Instability in Structural Composites under Longitudinal Compression
Alfredo Faria and Ragnar Larsson

Post-Buckling Analysis in Thin-web Laminated Composite Beams

Nolli, B. L., Cimini Jr, C. A., Pena, C. A.*

Federal University of Minas Gerais, Av Presidente Antônio Carlos, 6627, Belo Horizonte, Minas Gerais/Brazil
E-mail: nolli.bruna@gmail.com, carlos.cimini@gmail.com, cristiano.pena@embraer.com.br
*Embraer S.A, Rua Sete, 498, Belo Horizonte, Minas Gerais/Brazil

Abstract

Aeronautic industry is aiming to increase the efficiency and to reduce the costs of their aircrafts, in order to develop airplanes with better performance and lower fuel consumption. Researches have demonstrated that shear panels can carry a significant amount of load after reaching its initial buckling load. Consequently, exploring the post-buckling capacity of composite materials reinforced panels results in lighter and less expensive structures. For the metallic reinforced panels there is a consolidated methodology developed by NASA to calculate the diagonal tension, NACA TN2661. However, a theory for the panel's post-buckling behaviour in composite reinforced panels is yet to be developed. Therefore, the purpose of this study was to contribute to this new theory and focus on the influence of stacking sequence and fibre orientations on the post-buckling behaviour of composite reinforced panels. The main goals were to develop a method to build a FE model to represents the post-buckling behaviour of the composite reinforced panel in order to avoid having to use experimental results in future projects; study the variations in stacking sequence in post-buckling analysis; choose the material to build a reinforced panel that have the best behaviour during the post-buckling analysis: metal or composite. The results have shown that the FEM, considering the load, boundary conditions and materials described in this study, can represent the behaviour of the composite reinforced panel and its post-buckling behaviour. Also, after the comparison between six composite reinforced panels models with different stacking sequence, the layup that presented the lowest values for the failure indices was the one with +45 and -45 at the outside layers. This laminate was chosen to be compared with the metallic reinforced panel model, and this comparison has shown that the composite reinforced panel could withstand higher loads, so it is considered the best for the post-buckling behaviour analysis.

Keywords: Diagonal Tension, Post-Buckling, Finite Element Model, Composite Material

1 Introduction

The challenges of aeronautic industry are to save weight and to reduce the costs of their aircrafts through the combination of different materials and calculation methods. Applications of composite materials have grown, consequently, the challenges to find optimal designs and new developments emerged [1]. The use of composite materials to build panel's web at wings and fuselage where they are allowed to undergo an elastic buckling (diagonal tension state), is one way to achieve this goal [2].

It is possible to design thinner web structures made of composite material using the methodology that evaluates the panel post-buckling behaviour, comparing it to other methods, resulting in a lighter final product. This improvement in weight is feasible because the diagonal tension method allows the reinforced panel to buckle after a pre-defined load, but this is not the failure of the structure. At this moment, diagonal folds appear in the panel's web, and

the shear forces that caused the buckling are resisted by tension in the web by the diagonal folds, and by compression in the stiffeners [3]. Therefore, the final structure can be designed to support tension loads, and the tensions allowable are greater than the compressions one, so the complete structure withstand greater loads.

For the metallic reinforced panel there is a consolidated methodology developed by the *National Aeronautics and Space Administration* (NASA) to calculate the panel's diagonal tension, NACA TN2661 [4]. This semi-empiric method was developed based in several tests performed with aluminium panels using different geometries and loads, and is widely used by aircrafts manufactures. However, a theory for the panel's post-buckling behaviour in composite reinforced panels is yet to be developed. There are some researches ([5] and [2]) trying to adapt the NACA TN-2661 [4] method for composite materials, making it account for the anisotropy of the material and corroborate the results with tests. Most of them used *Glass Reinforced Aluminium* (GLARE) to build the panel. Some other studies [6] were

based in modelling the reinforced panel in finite elements and compared the results with tests data. Yet it was used a one bay panel with unidirectional layup.

To improve the understanding of post-buckling behaviour in composite reinforced panels, the studies should be focused in the variables that affect this behaviour, like stacking sequence, and orientation of the layers. And to overcome the expenses with tests, the reinforced panels modelled in finite element should have more than one bay, vertical and horizontal stiffeners and a complete layup using layer with different orientations. Therefore, the main goals of this study were to develop a method to build a *Finite Element Model* (FEM) to represents the post-buckling behaviour of the composite reinforced panel to avoid having to use experimental results in future projects; study the variations in stacking sequence in post-buckling analysis; and choose the material to build a reinforced panel that have the best behaviour during the post-buckling analysis: metal or composite.

The methodology developed in this study consist in five steps described hereafter. The first step was to build a FEM for an aluminium reinforced panel. One side of the panel is fixed, and a force is applied on the opposite side. A linear static analysis, a linear buckling analysis and a non-linear analysis were performed.

The second step was a comparison between results from the hand calculation for diagonal tension using NACA TN2661 [4] method and FEM analysis, in order to guarantee that the model is correct. Then a comparison between the buckling loads obtained from the linear and non-linear analyses was performed. If the differences are smaller than the error previous defined, it is possible to continue to the next step. Otherwise, the FEM must be modified in order to achieve that error. By the end of this step the model and mesh are considered validated.

Step three is a verification of the use of the validated mesh to build a FEM for the composite reinforced panel. In order to perform the verification, it was built two models: one with the equivalent properties for the layup chosen, and the other one modelling each layer and using the material properties of the tape. The models result from linear buckling analysis, and from the non-linear analysis, were compared. If the differences are smaller than the error previous defined, it is possible to continue to the next step. Otherwise, the FEM must be modified in order to achieve that error.

Step four consists in select the stacking sequence and the angle of fibre orientations for the layups, which have to be symmetric and balanced. For each layup, a FEM is created modelling each layer and using the material properties of the tape. Also, for each layup it will be applied several different loads values at the panel, in order to evaluate the reinforced panel behaviour for different loads magnitude. A non-linear analysis for each model is performed, and the failure index of the step when the diagonal tension is complete, and the panel's web redistribute the compressive load to the stiffeners for all of them is compared. The layup that presents the lowest values for the failure index for all the loads is chosen to be the laminate having the best post-buckling behaviour.

In step five a metallic reinforced panel FEM with the same weight that the composite reinforced panel layup chosen in step four is built. Then compare the buckling load from linear buckling analysis, and the load step when the diagonal tension is complete, and the panel's web redistribute the compressive load to the stiffeners from the non-linear analysis, in order to determine which one has the best behaviour for post-buckling analysis.

2 Results and discussion

2.1 Metallic reinforced panel FEM

The first step of this study was to build a FEM for the metallic reinforced panel with horizontal and vertical stiffeners. The materials and dimensions for the panel and stiffeners, and the load applied in the structure, were taken from an example of diagonal tension calculation presented in [7]. A schematic 3D model of the reinforced panel components and dimensions is presented in fig. 1.
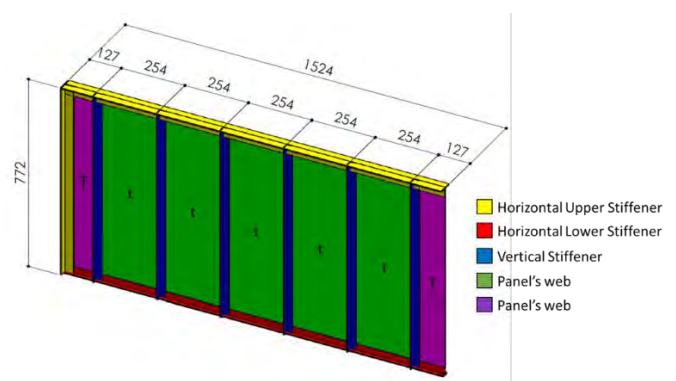

Figure 1: Metallic reinforced panel 3D model – geometry (all dimensions are in mm).

The panel's web was built with Aluminium 2024 T3, and the thickness for the panel shown in fig.1 as green is 0.635mm, and the purple one is 1.27mm. For the horizontal and vertical stiffeners, it was chosen Aluminium 7075-T3 Extruded.

The FEM was built using the software FEMAP 11.2® as pre- and post-processor and NASTRAN® as the solver. The panel's web was modelled using two-dimensions, four-nodes plate elements CQUAD4, which uses plane stress theory, capable of carrying in-plane forces, bending forces, and transverse shear forces [8]. The vertical and horizontal stiffeners were modelled using one-dimension elements that connects two grid points CBAR, which is a straight prismatic element with axial, bending, and torsional stiffness [8]. The CBAR element was chosen because the change in stiffener geometry using CBAR is easier to perform than it would be if it was modelled as plate elements. Consequently, the model is more versatile. Following the boundary conditions applied by [9], on the left-hand side of the panel, fixed boundary conditions were applied in the last column of nodes. Based on

an example presented in [7], a force of 60075 N was applied in the –Y direction on the right-hand side of the panel. On the right-hand side of the panel, a condition of displacement only in Y direction was applied in the first column of nodes. Also, the adjacent elements of the CBAR elements, representing the horizontal stiffeners, had their thickness increased. Therefore, these elements represent the web thickness and the stiffener thickness together. The FEM is shown in fig. 2.
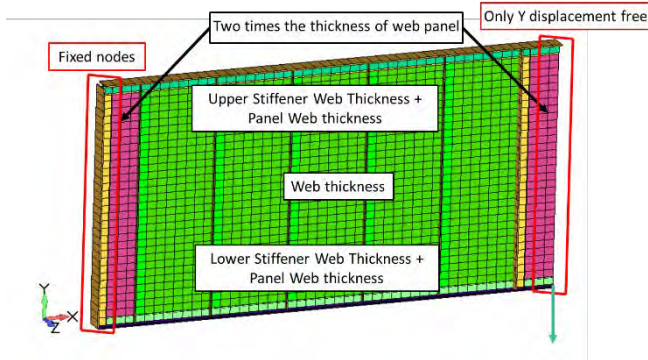

Figure 2: Metallic reinforced panel FEM final model.

It was performed a static analysis (SOL101) to evaluate the stress in the panel's web, a linear buckling analysis (SOL105) to calculate the buckling load, and a non-linear analysis including large displacements and elastic material properties (SOL106) to evaluate the post-buckling behaviour of the structure. The results obtained from the model and the hand calculation for diagonal tension using NACA TN2661 method [4] were compared.

The value of the shear stress from FEM was compared with the stress in the middle of panel (f_{max}) and the shear nominal stress in the web (f_S) in the region where appears the first eigenvalue (from linear buckling analysis SOL105). The comparisons are shown in fig. 3 and tab. 1.
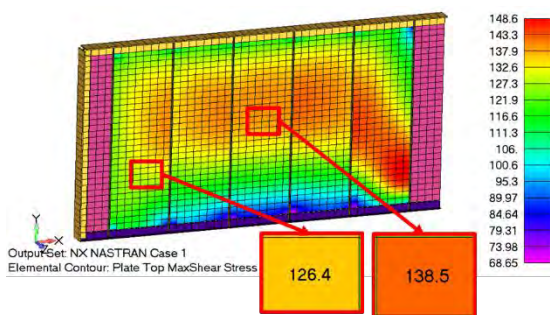

Figure 3: FEM stress (units in MPa).

Table 1: Comparison between shear stress from FEM and NACA TN2661 [4] method

FEM shear stress	NACA TN2661	Difference
138.5 MPa	134.2 MPa	3.2%
126.4 MPa	126.4 MPa	0.0%

Using the results from the linear buckling analysis (SOL105), it is possible to determine the buckling load (P_{cr_i}) for the metallic reinforced panel, shown in eq. (1).

$$P_{cr_i} = P_a\,\lambda_i = 60075N * 0.0227358 = 1365.85N \qquad (1)$$

Where P_a Is the load applied in the panel and λ_i is the eigenvalue calculated by the software FEMAP 11.2® in the linear buckling analysis (SOL105).

For the non-linear analysis (SOL106) the determination of the buckling load was performed plotting the non-linear load versus the total displacement of one node (it was chosen the one with greater displacement in the buckling region from the linear buckling analysis SOL105), as mentioned by [1]. On the plot load *vs.* displacement for the non-linear analysis (SOL106), a line A–A tangent to the first linear ramp was traced. Then a line B–B tangent to the second linear ramp was drawn where there is a significant change in the rigidity (fig. 4). Zooming into the non-linear region of the same plot, the buckling load was obtained by the intersection of these two lines, as showed in fig. 5.
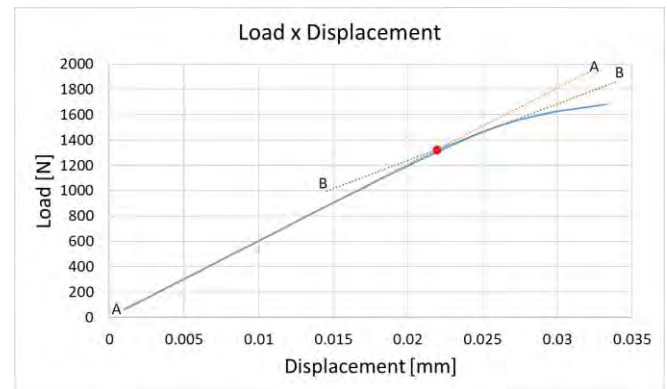

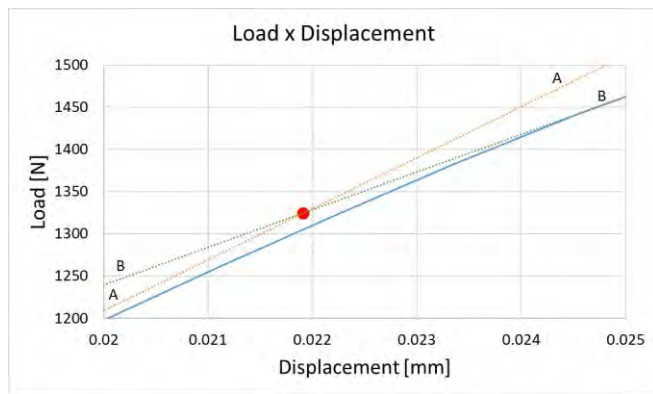*Figure 4: Non-linear analysis load-displacement graph*

Figure 5: Non-linear analysis load-displacement graph (zoom in the region where lines A-A and B-B cross)

Figure 6: Non-linear analysis – complete diagonal tension (deformations are plotted to scale at 150 times their actual value)

The comparison between buckling loads obtained from the linear buckling (SOL105) and non-linear analysis (SOL106) is presented in tab. 2.

Table 2: Comparison between buckling loads from linear buckling analysis and non-linear analysis

Linear buckling analysis	Non-linear analysis	Difference
1365.85 N	1324.36 N	3.13%

The results from comparison between FEM and hand calculation for diagonal tension using NACA TN2661 [4] method, (tab. 1), and the comparison between linear buckling and non-linear analysis, (tab. 2), have shown that the greatest difference was 3.2%. These encouraging results create the necessary confidence to use the FEM model in the following analyses.

The reinforced panel post-buckling behaviour was also analysed from the non-linear simulation (SOL106). In this analysis, the panel total load (60075 N) was divided into 100 increments of 600.75 N, that were successively applied in the model. Therefore, at each 0.01 increment that occurs in the model, a load of 600.75 N is added to the reinforced panel [10].

The complete diagonal tension phenomenon was considered to occur when diagonal folders were visually formed in the panel's web. Figure 6 shows the load step where it occurred.
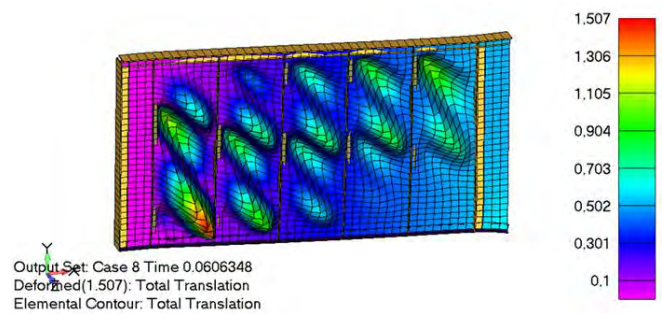
The complete diagonal tension load occurs when the panel's web redistributes the compressive load to the stiffeners. And this load is obtained by multiplying the step where first occurs diagonal tension by the total load applied in the reinforced panel. Consequently, this load is calculated in eq. (2):

$$0.0606348 * 60075N = 3642.64N \qquad (2)$$

In order to check if this result was correct, two other verifications were performed. The first one was an analysis of the stress state of one element in the buckling region, using the non-linear solution (SOL106). Minimum Principal Stress (Min Prc), Maximum Principal Stress (Max Prc), and Maximum Shear Stress (Max Shear) were considered. Figure 7 shows these stresses plotted as functions of the applied load.
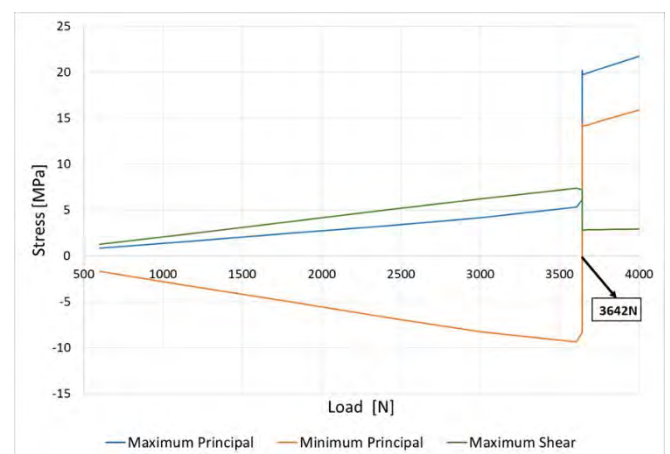

Figure 7: Stresses state during load increment

In fig. 7, all these stresses (Max Prc, Min Prc and Max Shear) experience sharp variations on their values at the load 3642 N. The explanation for these changes is the diagonal tension field. When the panel's web buckles and it cannot hold compression stress anymore, diagonal folds are formed

in the web, which is further loaded only in tension. The shear stress is distributed for the stiffeners [3].

The second verification was an evaluation of the load that the horizontal stiffeners were carrying during load application. It was chosen four elements, two from the upper horizontal stiffener and the others from the lower horizontal stiffener, and the results were plotted in a graph. For the linear static analysis, the axial stresses in the linear element representing the stiffeners in the end of the analysis were taken. These values were plotted in the graph and a line was traced until the zero (at the load step zero the axial stress in the stiffener is also zero). For the non-linear analysis the axial stresses in the linear element representing the stiffeners were taken for each step of the analysis and these values were plotted against the load. The final result can be seen in fig. 8.
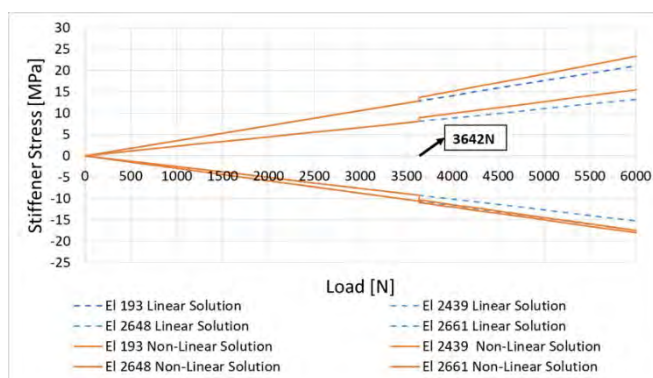

Figure 8: Stress on the stiffener with the increase of the load

The dashed line represents the axial stresses for the linear analysis, and the solid line represents the axial stresses for the non-linear analysis. It can be seen that for the load of 3642 N there is a discontinuity in the stress for the non-linear analysis, which means that at this load step the stiffeners start carrying more load than they would in the linear analysis. Therefore, this is the load step when the diagonal tension is complete, and the panel's web redistribute the compressive load to the stiffeners [3].

2.2 Composite Material Reinforced Panel versus Isotropic Reinforced Panel

The validated mesh aforementioned was used to build two FEM (with the same geometry) for the composite reinforced panel, model 1: modelling each layer and entering the **elastic** material properties of the tape, and model 2: with the equivalent properties for the layup selected. It was chosen a carbon/epoxy tape, [45/-45/0/90]$_S$, balanced and symmetric. The choice of using 0º, +45º, -45º and 90º fibre orientation is due to the recommendation made by [11]. They affirm that the laminate will be fibre dominant with the use of at least 10% of its plies in 0º, +45º, -45º and 90º. The balanced layups were chosen to remove the membrane coupling between in-plane normal and shear behaviour, and the choice of use symmetric laminate is to uncouple bending and membrane response, and to prevent warping under thermal loading [11].

They also affirm that laminates should be symmetric and balanced to maximize buckling strengths.

The mechanical and allowable properties of the carbon/epoxy tape used to build the layup came from [12]. The FEM material direction was aligned with the +X direction of global coordinate system, which means the tapes oriented in 0º were aligned with the +X direction.

The equivalent properties for the stacking sequence and tape material above-mentioned, is shown in tab. 3. Because the value of Ex is equal to Ey, an isotropic material was used in the FEM with the laminate equivalent properties.

Table 3: Equivalent property - Tape carbon/epoxy [45/-45/0/90]$_S$

Equivalent Properties	
Ex = Ey	56675.5 [MPa]
Gxy	22039.8 [MPa]
vxy = vyx	0.286

The materials for the reinforced panels can be seen in tab. 4. The material for the stiffeners were changed from aluminum to steel, and their thickness (web and flange) were doubled, and the load applied in the reinforced panel was altered from 60075 N to 143802 N. These modifications were necessary to guarantee that a post-buckling behavior would occur in the reinforced panel even with the increase in web's thickness from 0.635 mm to 1.52 mm. All the other panel's dimension (height, width, etc.) were kept the same.

Table 4: Reinforced panels materials

Material	
Web (model 1)	Tape carbon/epoxy [45/-45/0/90]$_S$
Web (model 2)	NA (it was used the Young's modulus equal to the composite model) E = 56675.5 MPa
Horizontal Upper Stiffener	Steel 4043
Horizontal Lower Stiffener	Steel 4043
Vertical Stiffener	Steel 4043

The results from the two models were compared, the first eigenvalue from the linear buckling analysis, and the load step when the diagonal tension is complete, and the panel's web redistribute the compressive load to the stiffeners from the non-linear analysis. The FEM plots are presented in figs. 8 and 9, and the comparison in tab. 5.
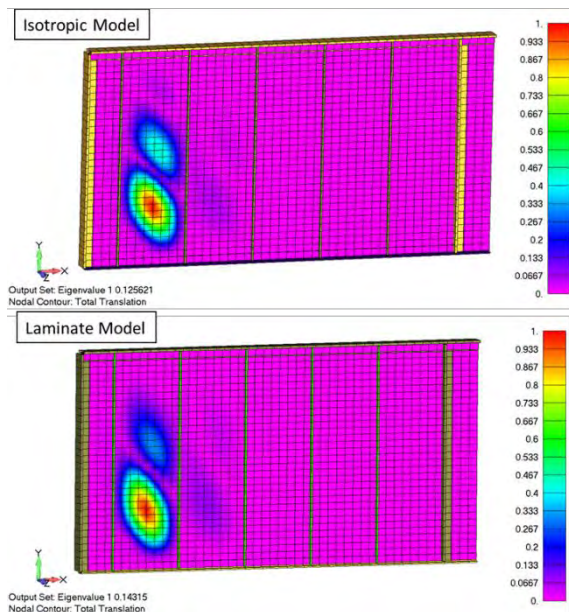
Figure 8: Isotropic and composite models: first eigenvalue (linear buckling analysis) comparison
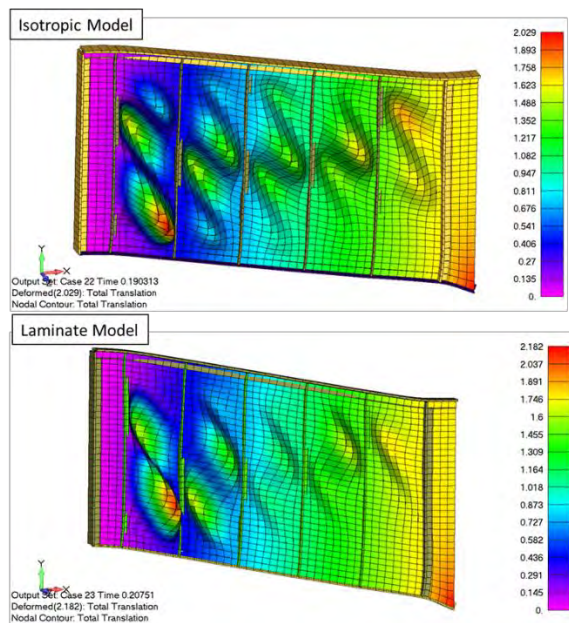

Figure 9: Isotropic and composite models: complete diagonal tension first occurrence (non-linear analysis) comparison (deformations are plotted to scale at 150 times their actual value)

Table 5: Comparison between Isotropic and Composite models.

	First Eigenvalue	First occurrence of Diagonal Tension
Isotropic Model	0.125621	0.190313
Composite Model	0.14315	0.20751
Difference	1.75%	1.72%

The results from comparison between Isotropic and Composite models, presented in tab. 5, show that the greatest difference is 1.75%. These encouraging results create the necessary confidence to use the FEM model in the following analyses.

2.3 Composite Reinforced Panel

Knowing that the FEM can be used to model the composite reinforced panel and to predict the post-buckling behavior, the model was used to choose the best laminate and compared it with the metallic reinforced panel. Six different laminates were selected to be compared, using carbon/epoxy (AS4/APC2) tape (same properties and allowable from [12]), all having 0º, +45º, -45º and 90º, being balanced and symmetric (the reasons for these choices were presented in Section 2.1). It was chosen to use the same material and fiber orientations and just change the stacking sequence because as mention by [13], the buckling modes are more sensitive to this type of variation. All the different stacking sequence are shown in tab. 6.

Table 6: Selected stacking sequence to be compared.

Model 1	[0/90/45/-45]$_s$
Model 2	[45/-45/0/90]$_s$
Model 3	[45/0/90/-45]$_s$
Model 4	[0/45/-45/90]$_s$
Model 5	[0/45/90/-45]$_s$
Model 6	[45/0/-45/90]$_s$

The reinforced panel dimensions and their vertical and horizontal stiffeners material and dimensions were the same as used in Section 2.2.

For each model it was applied seven loads, in order to verify the laminate behavior with different load magnitudes. For all

models, it was evaluated the Laminate Max Failure Index using the Tsai-Wu criterion and Maximum Strain criterion, calculated by the software FEMAP 11.2®. The failure index was used to classify the stress state of the laminate and not to determine the failure of a lamina, therefore the failure indices values presented are not close to one (that indicates failure).

Because of convergence problems in the models, not all the loads used for one criterion could be used for the other. The loads applied in the model for each criterion are presented in tab. 7.

Table 7: FEM applied loads for each criterion

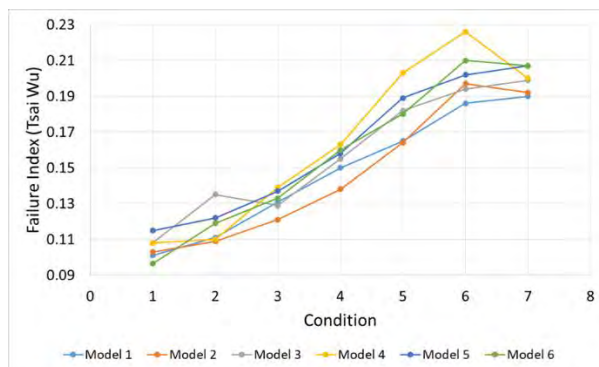|	**Tsai-Wu**	**Maximum Strain**
Condition 1	5500 N	5500 N
Condition 2	100000 N	100000 N
Condition 3	160000 N	160000 N
Condition 4	310000 N	310000 N
Condition 5	500000 N	480000 N
Condition 6	700000 N	700000 N
Condition 7	800000 N	900000 N

The results are presented in figs. 10 and 11.

Figure 10: Comparison of the Tsai-Wu Failure Index for all the models
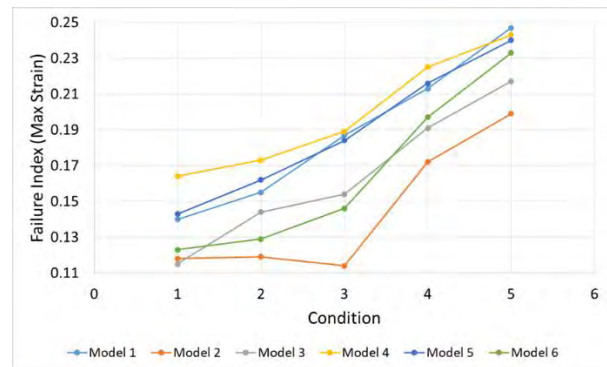

Figure 11: Comparison of the Max Strain Failure Index for all the models

Analyzing figs. 10 and 11 it is possible to conclude that the model that have the best behavior in the post-buckling is Model 2, because it presents the lowest failure index for the two criteria compared to all the other models. This result is corroborated by [11] that affirm for stability is better to use the +45º and -45º plies on the outer surfaces.

2.4 Composite Reinforced Panel versus Metallic Reinforced Panel

In order to verify which reinforced panel, metallic or composite, has the best behavior for the post-buckling analysis, the composite reinforced model chosen before (Model 2) was compared with a metallic reinforced model. Both models had the same mass and a load of 180000 N was applied in both panels.

The two reinforced panels use the same vertical and horizontal stiffeners, have the same boundary conditions applied, the web's thickness was the only parameter changed in the metallic one. Using the Aluminum 2524 T3 and Tape carbon/epoxy (AS4/APC2) densities, it was possible to calculate the web's thickness for the metallic reinforced panel which gives the same weight as the composite reinforced panel. Table 8 shows the thickness and mass for the two models.

Table 8: Thickness and mass for the composite and metallic models

	Composite Reinforced Panel	**Metallic Reinforced Panel**	**Difference**
Web's thickness [mm]	1.520	0.701	-53.87%
Mass [kg]	14.330	14.330	0.00%

The results from the two model were compared, the first eigenvalue from the linear buckling analysis, and the load step when the diagonal tension is complete, and the panel's web redistribute the compressive load to the stiffeners from the non-linear analysis. The FEM plots are presented in figs. 12 and 13, and the comparison in tab. 9.
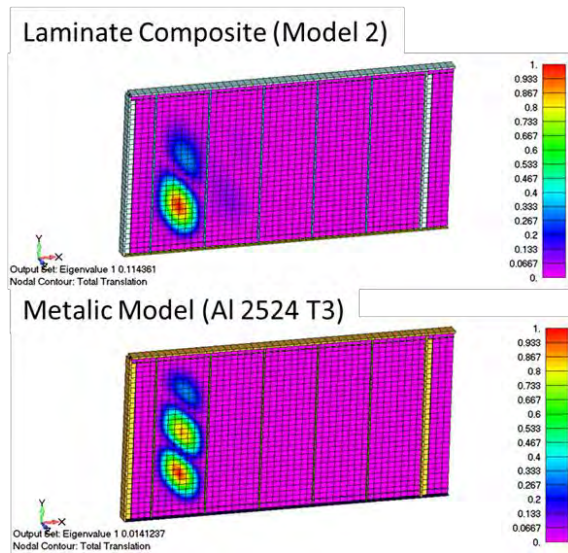

Figure 12: Composite and metallic models: first eigenvalue (linear buckling analysis) comparison
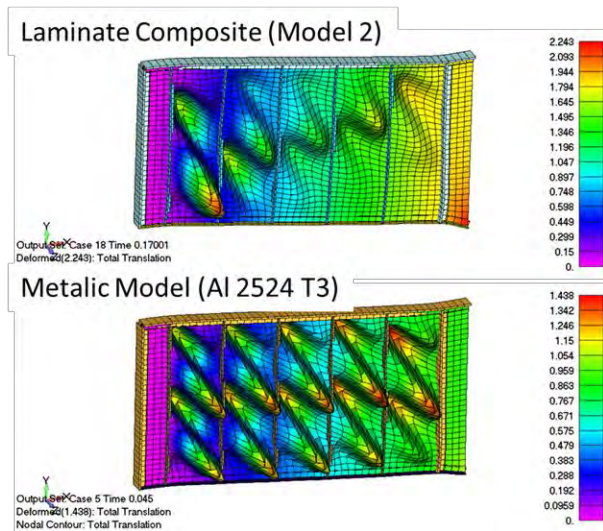

Figure 13: Composite and metallic models: complete diagonal tension first occurrence (non-linear analysis) comparison (deformation scaled in 150 times)

Table 9: Comparison between results from Composite and Metallic models

	First Eigenvalue	**First occurrence of complete Diagonal Tension**
Composite Model	0.114361	0.17001
Metallic Model	0.014124	0.04500

The comparison of the results of the two reinforced panels with the same mass, shows that the metallic reinforced panel supports lower loads before the buckling than the composite reinforced panel. This means that the web of the composite reinforced panel withstands to greater loads than the metallic one, consequently redistributes less load for the stiffeners. This result is corroborated by the analyses of fig. 14, which shows the comparison of stiffeners axial stress for both of the reinforced panels.
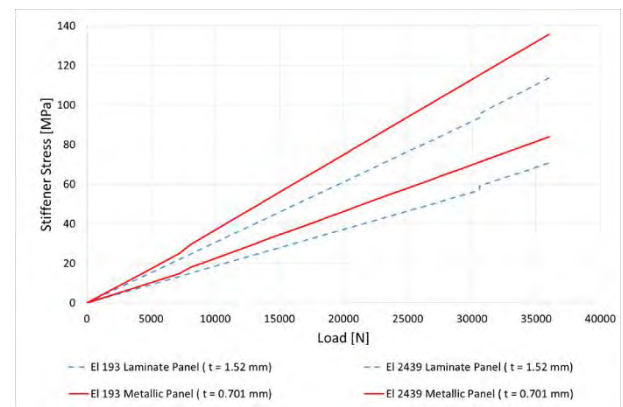
Figure 14: Stress on the stiffener for the composite and metallic models

Therefore, it is possible to conclude that the composite reinforced panel presents the best behavior for the post-buckling, in the conditions analyzed in this study (FEM boundary conditions, materials, etc.).

In order to determine the mass of the metallic reinforced panel for its behavior becomes similar to the composite reinforced panel, a new metallic reinforced panel was built. The web's thickness for the its panel was increased until its first eigenvalue and its load step when the diagonal tension is complete, and the panel's web redistribute the compressive load to the stiffeners got closer to the values for the composite reinforced panel. Table 10 shows the thickness and weight for the two models.

Table 10: Thickness and mass for the composite and new metallic models

	Composite Reinforced Panel	New Metallic Reinforced Panel	Difference
Web's thickness [mm]	1.52	1.51	-0.82%
Mass [kg]	14.330	17.368	21.20%

The FEM results for the two models are presented in figs. 15 and 16, and the comparison is shown tab. 11.

Figure 15: Composite and new metallic models: first eigenvalue (linear buckling analysis) comparison

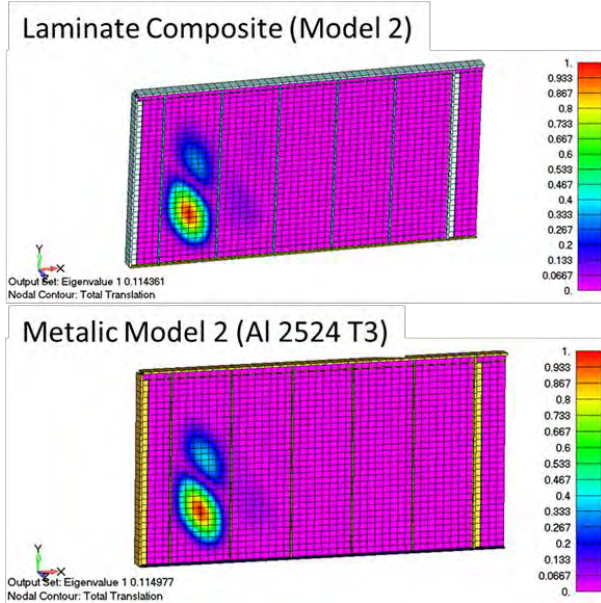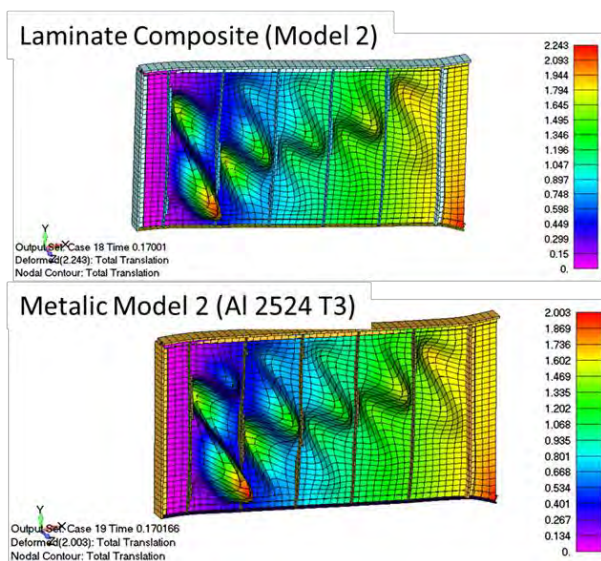*Figure 16: Composite and new metallic models: complete diagonal tension first occurrence (non-linear analysis) comparison (deformation scaled in 150 times)*

Table 11: Comparison between composite and new metallic models

	First Eigenvalue	First occurrence of complete Diagonal Tension
Composite Model	0.114361	0.170010
New Metallic Model	0.114977	0.170166

3 Conclusion

The purpose of this study was to improve the understanding of the post-buckling behaviour in composite reinforced panels, by verifying the influence of stacking sequence in this performance. Moreover, evaluate if the composite reinforced panel have a better behaviour in the post-buckling analysis than the metallic one was also a purposed.

A Literature review revealed that most studies are trying to adapt the method NACA TN2661 [4], used to design metallic reinforced panels, for the composite ones, taking into account the material's anisotropy. But they do not consider the fact that the anisotropy is not the only characteristic of the material that affects the post-buckling behaviour. Stacking sequence, and fibres orientation have influence in buckling analysis, therefore they have effects in the post-buckling behaviour.

Therefore, the main goals of this study were to develop a method to build a FEM to represents the post-buckling behaviour of the composite reinforced panel to avoid having to use experimental results in future projects; study the influence of stacking sequence in post-buckling behaviour; choose the material to build a reinforced panel that have the best behaviour during the post-buckling analysis: metal or composite.

The methodology adopted was to build a FEM for the metallic reinforced panel, and to compare its results with the method NACA TN2661 [4]. The model was modified until the differences reached a value smaller than the error previous defined. Using the validated model, two FEM were developed for the composite reinforced panel, an isotropic one, using equivalent properties, and the other modelling each layer. The results for both models were compared to verify if the model developed in this study could be used to model a composite reinforced panel modelling each layer. Then it was compared the failure index for six models of composite reinforced panel, having the same material, but varying the stacking sequence. The one having the lower value of the failure index was considered to have the best post-buckling behaviour. This model was then compared with the metallic reinforced panel, having the same mass, in order to verify which panel has the best behaviour during the post-buckling event.

The results from this study have shown that it was possible to represent the behaviour of the composite reinforced panel

during post-buckling using the FEM developed. For this FEM it was modelled each layer of the [45/-45/0/90]$_S$ laminate, which was balanced and symmetric. The boundary conditions and load used were described in this study.

Then six models with different layups were modelled, using the same material, the same total number of layers and the same number of layers for each orientation. The order of the layers was changed to evaluate the influence of the stacking sequence in the post-buckling behaviour. Comparing the failure indices for all the models in the moment when the diagonal tension is complete and the panel's web redistribute the compressive load to the stiffeners, it was concluded that the model that has better results, or lower failure index, is the one with +45 and -45 at the outside layers.

Two FEM were built for the metallic and composite reinforced panels, having the same mass, using the same stiffeners, only varying the web thickness. The comparison between buckling analysis and post-buckling behaviour shown the composite panel supports more load before the buckling than the metallic one. This means that the web of the composite reinforced panel withstands to greater loads than the metallic one, and consequently redistributes less load for the stiffeners. Therefore, it is possible to conclude that the composite reinforced panel presents the best behaviour during the post-buckling event.

References

[1] ARAKAKI, F. K.; FARIA, A. R. Composite-stiffened panel design under shear post-buckling behavior. Journal of Composite Materials, v. 50, n. 26, p. 3643-3662, 2016.

[2] JODOIN, A. et al. Diagonal tension in fibre-metal laminates. In: ICAS 2002, 23, 2002, Toronto. p. 7-8.

[3] NIU, M. C. Airframe Stress Analysis and Sizing. 2nd ed. Hong Kong: Conmilit Press Ltd., 2005. 810 p.

[4] KUHN, P; PETERSON, J. P.; LEVIN L. R. A Summary of Diagonal Tension: Part I – Methods of Analysis. NACA Tech. Note 2661: may, 1952.

[5] WITTENBERG, T. C.; VAN BATEN, T. J.; DE BOER, A. Design of fiber metal laminate shear panels for ultra-high capacity aircraft. Aircraft Design, v. 4, n. 2-3, p. 99-113, 2001.

[6] AGARWAL, B. L. Post-buckling behavior of composite shear webs. AIAA JOURNAL, Seattle, v. 19, n. 7, p. 933-939, 1981.

[7] BRUHN, E. F. Analysis and design of flight vehicle structures. U.S.A.: Tri-State Offset Co., 1973. 996 p.

[8] NASTRAN, M. S. C. Quick reference guide. MSC. SOFTWARE, v. 1, 2012a.

[9] BRAZ, T. B. C. Comportamento em flambagem e pós-flambagem de painel aeronáutico reforçado com laminado de camada fina versus camada grossa. 2018. 49 f. Monografia (graduação em engenharia mecânica) – Escola de engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

[10] NASTRAN, M. S. C. Implicit Nonlinear (SOL 600) User's Guide. MSC. SOFTWARE, v. 1, 2012b.

[11] BAILIE, J. A.; LEY, R. P.; PASRICHA, A. A summary and review of composite laminate design guidelines. Langley RC, Hampton, 1997.

[12] DANIEL, I. M.; ISHAI, O. Engineering mechanics of composite materials. New York: Oxford university press, 1994. 410 p.

[13] KASSAPOGLOU, C. Design and analysis of composite structures: with applications to aerospace structures. Chichester: John Wiley & Sons, 2013. 318 p.

Failure Induced by Instability in Structural Composites under Longitudinal Compression

Alfredo R. de Faria, Ragnar Larsson*

Department of Mechanical Engineering, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP/Brazil
E-mail: arfaria@ita.br, ragnar.larsson@chalmers.se
*Department of Industrial and Materials Science, Chalmers University of Technology, Göteborg, Sweden

Abstract

The objective of this work is to develop a damage propagation model applicable to the study of failure mechanism in composites induced by longitudinal compression. The damage propagation model shall be based on an energy principle, quantified by the critical energy release rate associated to matrix cracking, and on the study of the kinematics of the equilibrium of a representative volume of a ply material within the damaged region of the laminate. It is expected that the model proposed will be implemented within the context of the finite element technology in order to expand its usefulness to realistic situations where multiple ply laminates are used. The model to be proposed shall capture the onset of formation of kink-bands and subsequently its propagation. It is formulated within the context of an energy based approach that considers the critical energy release rate of the matrix in the failed elements of the mesh. A damage parameter that ranges from 0 (intact) to 1 (fully damaged) is used that relates relative shear strains in the matrix induced by loss of stability of a representative volume. The damage models developed will be useful to realistically predict failure due to kink-band formation and propagation mechanism. The results obtained will serve as the starting point for larger collaborative projects involving strain rate effects and damage tolerant composites.

Keywords: Composites, damage tolerance, stress analysis

1 Introduction

The most prominent failure mode in fiber reinforced composites subject to longitudinal compressive loadings is kink-band formation. Rosen [1] was the first researcher to propose a model to describe this type of failure in fiber reinforced composites. However, Rosen's model, and a series of subsequent models based on his model, predicted unrealistic high strengths. It was only a few years later that Argon [2] proposed an improved model that substantially reconciled experimental and theoretical values of longitudinal compressive strengths. The great contribution from Argon was the consideration of initial local fiber misalignment that facilitated initialization of fiber instability followed by matrix shearing. Although Argon's model did not consider additional fiber rotation due to the initial misalignment it paved the way for other researchers to propose more accurate models [3, 4].

Failure by kink-band formation is today recognized to be a consequence of microbuckling and kinking, which are induced by matrix yielding or cracking because of initial misalignment of fibers in the laminate manufactured [5-6]. The most successful models up to date are based on the investigation of the stability of representative volumes where initial misalignments are present and the equilibrium

equation of a fiber is imposed on its displaced configuration. Following that line of reasoning, Gutkin et al. [7] proposed a failure criterion to determine the onset of kink-band formation. All these earlier models, however, neglect damage propagation and lack therefore the ability to predict ultimate failure loads [8]. Moreover, they do not rely on or propose specially formulated finite elements that possess the capabilities to describe damage propagation in ways similar to those observed for instance in decohesive type elements. More recently, a number of published works included damage propagation modeling in longitudinal crushing of composites [9-10], but they do not explicitly consider microbuckling as the main damage mechanism. This work proposes to unite a microbuckling explicit model with matrix crack damage propagation resulting in a new alternative way to investigate failure under longitudinal compression.

2 Representative volume and matrix shear strain

The fibers are assumed to be cylindrical with diameter ϕ_f. In a 3D situation fibers are arranged in hexagonal packing with a fiber volume fraction v_f. Figure 1 presents the 3D and 2D configurations. The total area A of the hexagon in Fig. 1 is $A = 3\sqrt{3}(t_m + \phi_f)^2/2$ and the area covered by fibers is $A_f =$

$3\pi\phi_f^2/4$. Hence, the fiber volume fraction is $v_f = A_f/A$ leading to the conclusion that

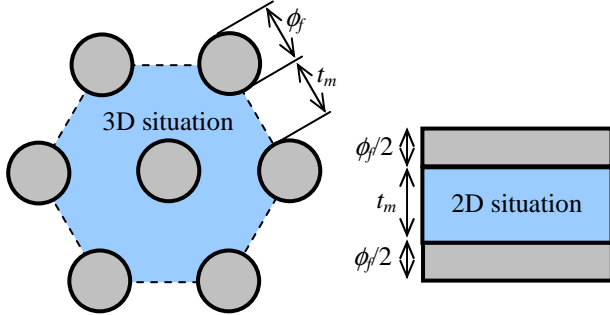$$t_m = \phi_f\left(\sqrt{\frac{\pi}{2\sqrt{3}v_f}} - 1\right) \tag{1}$$

Figure 1: representative volumes

whereas a 2D model implies

$$v_f^{2D} = \frac{\phi_f}{\phi_f + t_m} \tag{2}$$

The representative volume of the composite ply in the kink-band region is shown in Fig. 2, where the unloaded and loaded configurations are sketched. The initial geometric configuration of the representative volume is completely described by its length δx, the initial fiber misalignment angle θ_0 and the width of the matrix layers between fibers t_m.

(a) unloaded configuration
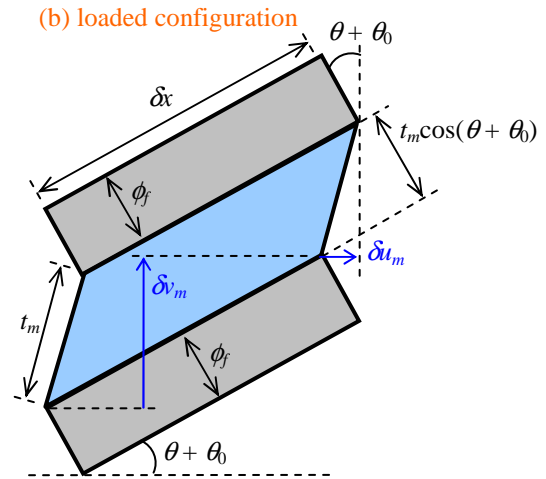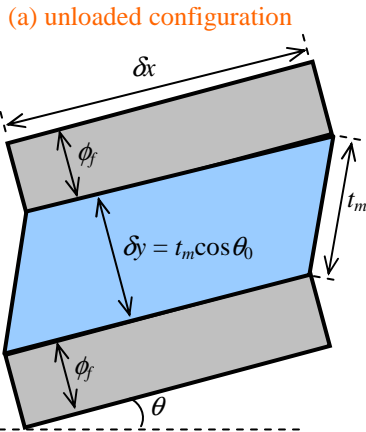

(b) loaded configuration

Figure 2: kink-band region: (a) unloaded and (b) loaded

The relevant geometric and kinematic information are displayed in Fig. 2. After loading the misalignment angle increases to $\theta + \theta_0$. It is assumed that the fibers are incompressible such that, initially, two neighboring fibers are separated by a distance $t_m\cos\theta_0$, and, after loading, this distance drops to $t_m\cos(\theta + \theta_0)$. Moreover, the length δx is fixed. The shear deformation in the matrix is given by

$$\gamma_m = \frac{\delta u_m}{\delta y} + \frac{\delta v_m}{\delta x} \tag{3}$$

where δu_m and δv_m, according to Fig. 2, can be expressed as

$$\begin{aligned}\delta u_m &= \phi_f[\sin(\theta + \theta_0) - \sin\theta_0] \\ \delta v_m &= \delta x[\sin(\theta + \theta_0) - \sin\theta_0]\end{aligned} \tag{4}$$

The value of δy can be computed as shown in Fig. 2: $\delta y = t_m\cos\theta_0$. Substitution of δy and Eq. (4) into Eq. (3) yields

$$\begin{aligned}\gamma_m &= \frac{\phi_f[\sin(\theta + \theta_0) - \sin\theta_0]}{t_m\cos\theta_0} + \frac{\delta x[\sin(\theta + \theta_0) - \sin\theta_0]}{\delta x} = \\ &\left(1 + \frac{\phi_f}{t_m\cos\theta_0}\right)[\sin(\theta + \theta_0) - \sin\theta_0]\end{aligned} \tag{5}$$

Since the initial misalignment angle θ_0 is usually small ($0^\circ < \theta_0 < 4^\circ$) a good approximation for Eq. (5) is

$$\gamma_m \approx \left(1 + \frac{\phi_f}{t_m}\right)\sin\theta \tag{6}$$

It is possible to go even further assuming that θ is small and, considering Eq. (2), write

$$\gamma_m \approx \left(1 + \frac{\phi_f}{t_m}\right)\theta = \frac{\theta}{1 - v_f^{2D}} \tag{7}$$

This approach linearizes the problem in the variable θ, what can sometimes be useful and may provide insight into the

constitutive relations involved. For a 2D model the shear modulus of the matrix is

$$G_m^{2D} = \frac{G_m}{1-v_f^{2D}} \qquad (8)$$

Thus, provided the deformation regime within the matrix remains elastic and angle θ remains small, one can write

$$\tau_m = G_m\gamma_m = (1-v_f^{2D})G_m^{2D}\frac{\theta}{1-v_f^{2D}} = G_m^{2D}\theta \qquad (9)$$

3 Equilibrium equation in the loaded configuration

When it comes to kink-band formation the relevant homogenized stresses acting on the representative volume are σ_1 and τ_{12}, which are depicted in Fig. 3. The fiber is subject to normal stress σ_f. Assuming that the fiber is much stiffer than the matrix, the classical theory of micromechanics states that

$$\sigma_f\phi_f = \sigma_1(\phi_f + t_m) \qquad (10)$$

The classical theory of micromechanics also indicates that the shear stress is the same throughout the representative volume. Hence, the shear stress in the fiber τ_f is equal to the homogenized shear stress τ_{12} i.e.,
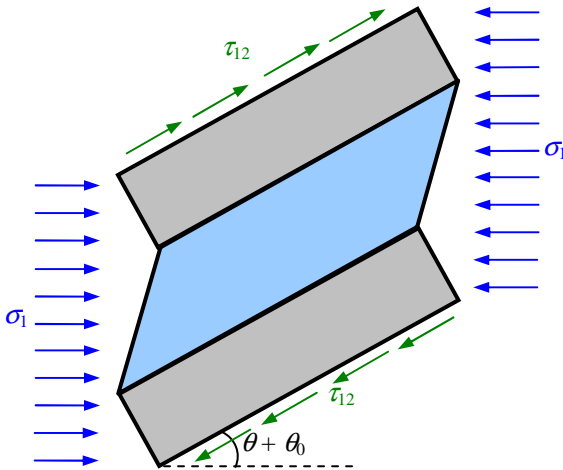
$$\tau_f = \tau_{12} \qquad (11)$$

Figure 3: Homogenized stresses applied to the representative volume

The free body diagram of a single fiber in the loaded configuration is shown in Fig. 4. The fiber orientation in the unloaded configuration is also shown with the initial misalignment angle θ_0. Force P is the net effect of the normal stresses in the fiber σ_f and force S is the net effect of the shear stresses in the fiber τ_f such that, with the aid of Eqs. (10) and (11),

$$P = \sigma_f\phi_f = \sigma_1(\phi_f + t_m) \quad , \quad S = \tau_f\phi_f = \tau_{12}\phi_f \qquad (12)$$
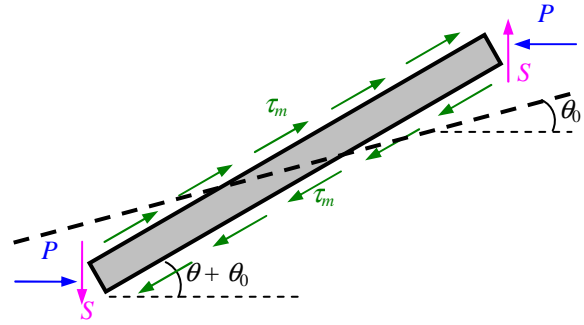

Figure 4: Free body diagram of fiber in the loaded configuration

The force equilibrium is clearly satisfied. Imposition of moment equilibrium yields

$$P\delta x\sin(\theta+\theta_0) + S\delta x\cos(\theta+\theta_0) = \tau_m\phi_f\delta x \qquad (13)$$

Substitution of Eq. (2) and (12) in (13) leads to

$$\frac{\sigma_1}{v_f^{2D}}\sin(\theta+\theta_0) + \tau_{12}\cos(\theta+\theta_0) = \tau_m \qquad (14)$$

Failure of the fiber is assumed to happen abruptly when the fiber stress σ_f in Eq. (12) reaches the fiber compressive strength X_f^c. Before reaching X_f^c it is assumed that the fibers behave linearly.

Equation (14) describes the buckling of an initially displaced rigid bar under the action of compressive and shear forces as shown in Fig. 5. If the applied forces P, S and the spring constant k are given by

$$P = \frac{\sigma_1}{Lv_f^{2D}} \quad , \quad S = \frac{\tau_{12}}{L} \quad , \quad k = \frac{G_m^{2D}}{L} \qquad (15)$$

the moment equilibrium equation matches exactly Eq. (14), provided τ_m is given by Eq. (9). Even if Eq. (9) is not applicable, the original relation $\tau_m = G_m\gamma_m$ can still be used and the spring constant k may be assumed to vary with θ.
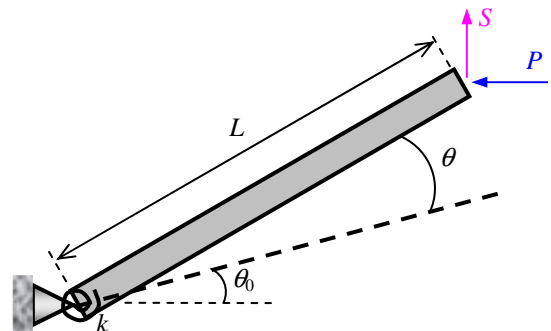

Figure 5: Initially displaced rigid bar

Notice that, if one assumes elastic regime and small angles, then substitution of Eq. (9) into Eq. (14) will result in a linear equation that can be solved for θ. Hence, under these simplifying assumptions, the moment equilibrium equation

can be explicitly solved for θ. A simple stress based criterion for fiber kinking can be subsequently proposed if energy balance considerations during crack formation come into play. This, however, is not the objective of this work. Instead, a model that goes beyond the onset of failure shall be proposed.

4 The matrix cohesive model

The proposed model behavior of the matrix in shear is shown in Fig. 6. The damage parameter starts at zero and reaches the maximum value 1 when the failure shear strain γ_{mf} is reached and, consequently, the matrix has no stiffness left, becoming completely degraded (see Fig. 6). S_L is a matrix property that determines the onset of matrix failure.
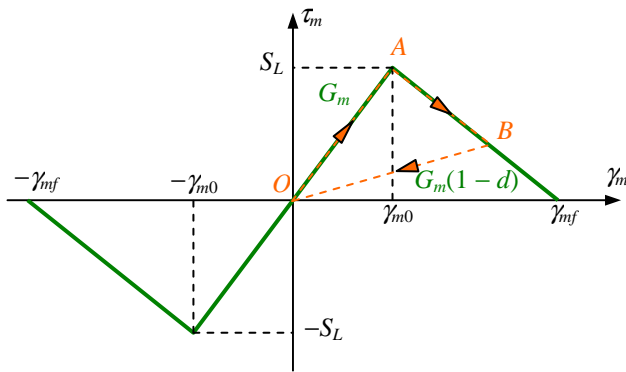

Figure 6: Model behavior in shear

Once the initiation (γ_{m0}) and failure (γ_{mf}) shear strains are available, the damage parameter can be expressed as

$$d = \frac{\gamma_{mf}}{\gamma_{mf} - \gamma_{m0}}\left(1 - \frac{\gamma_{m0}}{\gamma_m}\right) \tag{16}$$

Equation (16) is valid only for strain values that obey $|\gamma_m| > \gamma_{m0}$, otherwise $d = 0$. If the matrix becomes damaged then $d > 0$ and d never decreases, i.e., the model does not allow for restoration of mechanical properties once the matrix is damaged. Take for instance the loading path OAB in Fig. 6. If, after returning to point O the material is again loaded, it will follow path OB until it reaches point B. If further loading is applied beyond point B the original curve of the constitutive law is retaken until $d = 1$, when total failure is reached.

Figure 6 shows that the inclination of the OB path is $G_m(1 - d)$. Hence, generally the stress × strain relation can be written as

$$\tau_m = G_m(1-d)\gamma_m \tag{17}$$

which, in the incremental form reads

$$\Delta\tau_m = G_m(1-d)\Delta\gamma_m - G_m\gamma_m\Delta d \tag{18}$$

In Eq. (18) it is tacitly assumed that there is no plastic shear strain.

5 Incremental model and the finite element context

Two parameters are important to describe kink-band formation: θ that can be computed from Eq. (14) and d that can be computed from Eq. (16). It is also assumed that the fibers respond linearly up to catastrophic failure when $\sigma_f = X_f^c$. Thus, the constitutive relations for both fiber and matrix are

$$\sigma_f = E_f\varepsilon_f \quad , \quad \tau_f = G_f\gamma_f \quad , \quad \sigma_m = E_m\varepsilon_m \quad , \quad \tau_m = G_m\gamma_m \tag{19}$$

where E_f and G_f are assumed to be constant throughout. The third equation $\sigma_m = E_m\varepsilon_m$ may pose a problem since the matrix elastic modulus E_m is needed and it would clearly vary depending on the degree of degradation of the matrix. One possible way around this difficulty is to adopt a damage propagation model for the matrix under traction/compression similar to the one proposed in Fig. 6 for the matrix under shear. This would result in one additional damage parameter. Another option is to consider that, since kink-band formation is the relevant failure mechanism in this study, the matrix will be under compressive stresses just like the fibers. Hence, equation $\sigma_m = E_m\varepsilon_m$ can be used up to the point where σ_m is equal to the matrix strength under compression X_m^c. In this study the later option is considered.

The in-plane mechanical properties of a lamina can be computed using the Halpin-Tsai rules of mixture as

$$E_1 = E_f v_f^{2D} + E_m(1 - v_f^{2D}) \tag{20}$$

$$v_{12} = v_f v_f^{2D} + v_m(1 - v_f^{2D}) \tag{21}$$

$$\xi = 1 + 40(v_f^{2D})^{10} \tag{22}$$

$$\eta_E = \frac{(E_f / E_m) - 1}{(E_f / E_m) + \xi} \tag{23}$$

$$E_2 = E_m\frac{1 + \xi\eta_E v_f^{2D}}{1 - \eta_E v_f^{2D}} \tag{24}$$

$$\eta_G = \frac{(G_f / G_m) - 1}{(G_f / G_m) + \xi} \tag{25}$$

$$G_{12} = G_m\frac{1 + \xi\eta_G v_f^{2D}}{1 - \eta_G v_f^{2D}} \tag{26}$$

where v_f and v_m are the fiber and matrix Poisson ratios, respectively, assumed constant. The ply constitutive equation is

$$\begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{Bmatrix} = \frac{1}{1 - v_{12}v_{21}}\begin{bmatrix} E_1 & v_{12}E_2 & 0 \\ v_{21}E_1 & E_2 & 0 \\ 0 & 0 & (1 - v_{12}v_{21})G_{12} \end{bmatrix}\begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{Bmatrix} \tag{27}$$

which, in the incremental form reads

$$\begin{Bmatrix} \Delta\sigma_1 \\ \Delta\sigma_2 \\ \Delta\tau_{12} \end{Bmatrix} = \frac{1}{1-\nu_{12}\nu_{21}} \begin{bmatrix} E_1 & \nu_{12}E_2 & 0 \\ \nu_{21}E_1 & E_2 & 0 \\ 0 & 0 & (1-\nu_{12}\nu_{21})G_{12} \end{bmatrix} \begin{Bmatrix} \Delta\varepsilon_1 \\ \Delta\varepsilon_2 \\ \Delta\gamma_{12} \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ \gamma_{12}\Delta G_{12} \end{Bmatrix} \quad (28)$$

Notice that the matrix shear modulus G_m influences only G_{12} in Eq. (26). Hence, the term $\gamma_{12}\Delta G_{12}$ must be accounted for in Eq. (28). The other ply mechanical properties (E_1, E_2, ν_{12}) are assumed to be independent of G_m.

In the finite element context the problem must be solved incrementally. Initially consider that the equilibrium problem in the weak form has been solved for time step n such that

$$\sum_{e=1}^{N_e} \int_{V_e} \delta\mathbf{q}_e^T \mathbf{B}_e^T \boldsymbol{\sigma}_e^n dV_e = \sum_{e=1}^{N_e} \int_{V_e} \delta\mathbf{q}_e^T \mathbf{f}_e^n dV_e \quad (29)$$

where summation extends over all N_e elements in the mesh and $\mathbf{f}_e$ is the vector of external loads applied to element e. Matrix $\mathbf{B}_e$ is the discretization matrix that relates strains within a given element to the element degrees of freedom stored in $\mathbf{q}_e$: $\boldsymbol{\varepsilon}_e = \mathbf{B}_e\mathbf{q}_e$. At the subsequent time step $n + 1$ the new equilibrium equation in the weak form reads

$$\sum_{e=1}^{N_e} \int_{V_e} \delta\mathbf{q}_e^T \mathbf{B}_e^T \boldsymbol{\sigma}_e^{n+1} dV_e = \sum_{e=1}^{N_e} \int_{V_e} \delta\mathbf{q}_e^T \mathbf{f}_e^{n+1} dV_e \quad (30)$$

with $\boldsymbol{\sigma}_e^{n+1} = \boldsymbol{\sigma}_e^n + \Delta\boldsymbol{\sigma}_e^{n+1}$. Subtraction of Eq. (29) from Eq. (30), and recognizing that $\delta\mathbf{q}_e$ are arbitrary and geometrically admissible virtual displacements

$$\sum_{e=1}^{N_e} \int_{V_e} \mathbf{B}_e^T \Delta\boldsymbol{\sigma}_e^{n+1} dV_e = \mathbf{f}_{ext}^{n+1} - \mathbf{f}_{ext}^n = \Delta\mathbf{f}_{ext} \quad (31)$$

where the combined effects of all $\mathbf{f}_e$ have been grouped into $\mathbf{f}_{ext}$. Solution to Eq. (31) must be found iteratively. All the terms relating to $\Delta\boldsymbol{\sigma}_e^{n+1}$ in the left-hand side of Eq. (31) must be computed at the Gaussian stations used in the numerical integration.

Equations (5) and (14) in the incremental forms read

$$\Delta\gamma_m = \left(1 + \frac{\phi_f}{t_m \cos\theta_0}\right) \Delta\theta \cos(\theta + \theta_0) \quad (32)$$

$$\frac{\Delta\sigma_1}{\nu_f^{2D}} \sin(\theta + \theta_0) + \Delta\tau_{12} \cos(\theta + \theta_0) +$$
$$\Delta\theta \frac{\sigma_1}{\nu_f^{2D}} \cos(\theta + \theta_0) - \Delta\theta\tau_{12} \sin(\theta + \theta_0) = \Delta\tau_m \quad (33)$$

Substitution of Eqs. (5), (18) and (32) into (33) yields

$$\frac{\Delta\sigma_1}{\nu_f^{2D}} \sin(\theta + \theta_0) + \Delta\tau_{12} \cos(\theta + \theta_0) +$$
$$\Delta\theta \frac{\sigma_1}{\nu_f^{2D}} \cos(\theta + \theta_0) - \Delta\theta\tau_{12} \sin(\theta + \theta_0) =$$
$$G_m(1-d)\left(1 + \frac{\phi_f}{t_m \cos\theta_0}\right) \Delta\theta \cos(\theta + \theta_0) -$$
$$G_m\left(1 + \frac{\phi_f}{t_m \cos\theta_0}\right) [\sin(\theta + \theta_0) - \sin\theta_0]\Delta d \quad (34)$$

Therefore, given the stress and damage increments ($\Delta\sigma_1$, $\Delta\tau_{12}$, Δd), $\Delta\theta$ can be computed and the fiber misalignment is updated to $\theta + \Delta\theta$.

6 Numerical results

A 1 m × 1 m × 0.125 mm single layer specimen, with fibers oriented along the x axis is subject to compressive prescribed displacement also along x. Figure 7 depicts the situation. Only the center node is restricted to move along the y direction in order to avoid rigid body mode. A 10 × 10 mesh of bilinear elements is used.
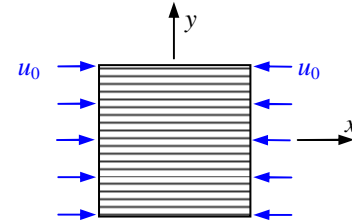
Figure 7: Single layer specimen under compression

The material properties used are reported in Table 1. Notice that the initiation (γ_{m0}) and failure (γ_{mf}) shear strains of the fibers are unnecessary for the present model. The fiber volume fraction is assumed to be 60%.

Table 1: Material properties used

property	matrix	fiber
E_{11} [GPa]	2.6	240
E_{22} [GPa]	2.6	240
ν_{12}	0.38	0.22
G_{12} [GPa]	0.94	98.36
γ_{m0}	0.002	–
γ_{mf}	0.005	–

The initial fiber misalignment θ_0 is defined (in degrees) according to Eq. (35)

$$\theta_0 = 3.5° \cos(\pi x / L) \quad (35)$$

where L is the edge length ($L = 1$ m) and $-L/2 \le x \le L/2$.

Figure 8 presents the result in terms of damage parameter d computed with Eq. (16). Figure 8 is obtained for the ultimate displacement $u_0 = 6$ mm as illustrated in Fig. 7.
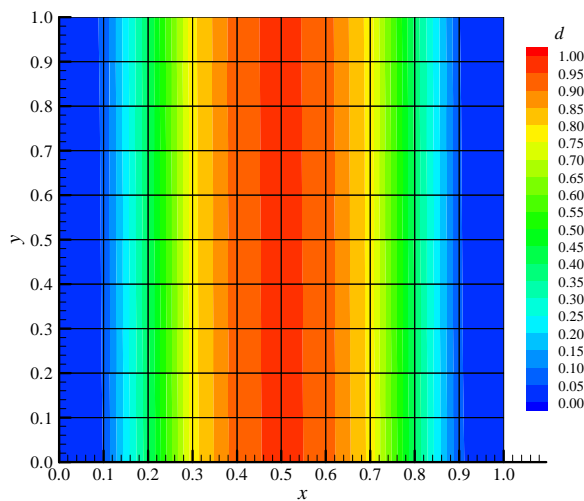
Figure 8: Damage parameter mapping

Figure 9 presents the evolution of the matrix shear stress τ_m and damage parameter d computed at the center node of the mesh. Observe that the τ_m behavior is consistent with Fig. 6 and the red curve for d follows the hyperbolic pattern described in Eq. (16).

Figure 9: Matrix shear stress and damage parameter evolution

7 Conclusions

A model is proposed to capture the onset of formation of kink-bands and subsequently its propagation. It is formulated within the context of an energy based approach that considers the critical energy release rate of the matrix in the failed elements of the mesh. A damage parameter that ranges from 0 (intact) to 1 (fully damaged) is used that
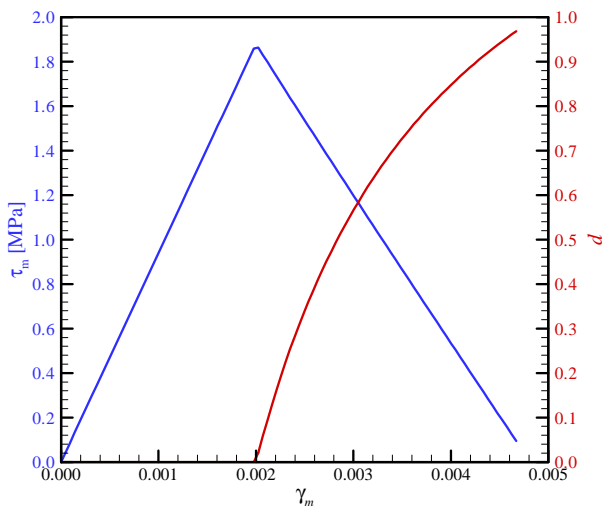
relates relative shear strains in the matrix induced by loss of stability of a representative volume.

Today commercially available finite element packages (e.g. Abaqus) possess damage modeling capabilites. However, they are not well suited for predictions of failure under longitudinal compression since they dot not inherently consider microbuckling and kink band formation, therefore delivering poor results whenever crushing is involved. The present model may be coupled with additional models for damage propagation and failure predictions [11], and latter incorporated into a commercial FE code. An element deletion criterion would then be required in order to simulate more realistic composite structures.

The model is based on physical principles and requires only two parameters whose extraction is not difficult: the initiation (γ_{m0}) and failure (γ_{mf}) shear strains of the matrix. The numerical result present shows that the model is simple to implement and captures the most essential behavior of structural failure under compressive loadings.

References

[1] Rosen BW. *Mechanics of composite strengthening*. Fibre composite materials. Metals Park, Ohio: American Society of Materials; 1965.

[2] Argon AS. *Fracture of composites*. Treatise on materials science and technology. New York: Academy Press; 1972.

[3] Steif PS. A model for kinking in fiber composites - I. fiber breakage via micro-buckling. *International Journal of Solids and Structures* 26(5/6): 549-561, 1990.

[4] Steif PS. A model for kinking in fiber composites - II. kink band formation. *International Journal of Solids and Structures* 26(5/6): 563-569, 1990.

[5] Pimenta S, Gutkin R, Pinho ST, Robinson P. A micromechanical model for kink-band formation: Part I - experimental study and numerical modelling. *Composites Science and Technology* 69: 948–955, 2009.

[6] Pimenta S, Gutkin R, Pinho ST, Robinson P. A micromechanical model for kink-band formation: Part II - analytical modelling. *Composites Science and Technology* 69: 956–964, 2009.

[7] Gutkin R, Pinho ST, Robinson P, Curtis PT. A finite fracture mechanics formulation to predict fibre kinking and splitting in CFRP under combined longitudinal compression and in-plane shear. *Mechanics of Materials* 43: 730–739, 2011.

[8] Dávila CG, Camanho PP. *Failure criteria for FRP laminates in plane stress*. NASA/TM-2003-212663, 2003.

[9] Gutkin R, Pinho ST. Combining damage and friction to model compressive damage growth in

fibre-reinforced composites. *Journal of Composite Materials* 49(20): 2483-2495, 2015.

[10] Gutkin R, Costa S, Olsson R. A physically based model for kink-band growth and logitudinal crushing of composites under 3D stress states accounting for friction. *Composites Science and Technology* 135: 39–45, 2016.

[11] Bürger D, de Faria AR, Almeida SFM, Melo FCL, Donadon MV. Ballistic impact simulation of an armour-piercing projectile on hybrid ceramic/fiber reinforced composite armours. *International Journal of Impact Engineering*, 43: 63-7, 2012.

VI Production and Manufacturing

Mathematical Optimization of the Tactical Allocation of Machining Resources for an Efficient Capacity Utilization in Aerospace Component Manufacturing
Sunney Fotedar, Torgny Almgren, Michael Patriksson, Ann-Brith Strömberg and Stefan Cedergren

Airframe Sealing Automation Using Snake-robot
Luís Gonzaga Trabasso, Doglas Negri, Walter Antonio Kapp, Magnus Engström, Guilherme Sartori Natal and Kerstin Johansen

Proposal of a Method for the Implementation of the Industry 4.0 - Aircraft Final Assembly Domain
Alexandre Leite Junior, Gléverson Lemos and Luís Gonzaga Trabasso

Additive Manufacture at Industrial, Aeronautical and Defence Area: How to Control the Production of a Good at an Embargo Scenario
Juan Manuel Iglesias Pascual and Jose Valentin Iglesias Pascual

Machining Distortion Analysis of Aerospace Components using the Contour Method
Mats Werke, Mosharraf Hossain, Daniel Semere, Anders Wretland

Mathematical Optimization of the Tactical Allocation of Machining Resources for an Efficient Capacity Utilization in Aerospace Component Manufacturing

Sunney Fotedar[1], Torgny Almgren[2], Stefan Cedergren[3], Ann-Brith Strömberg[1], and Michael Patriksson[1]

[1]Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden
E-mail: sunney@chalmers.se, torgny.almgren@gknaerospace.com, stefan.cedergren@gknaerospace.com,
anstr@chalmers.se, mipat@chalmers.se,

[2]Logistics, GKN Aerospace, Trollhättan, Sweden
[3]PTC, GKN Aerospace, Trollhättan, Sweden

Abstract

In the aerospace industry, with low volumes and many products, there is a critical need to efficiently use available manufacturing resources. Currently, at GKN Aerospace, resource allocation decisions that in many cases will last for several years are to some extent made with a short-term focus so as to minimize machining time, which results in a too high load on the most capable machines, and too low load on the less capable ones. This creates an imbalance in capacity utilization that leads to unnecessary queuing at some machines, resulting in long lead times and in an increase in tied-up capital. Tactical resource allocation on the medium to long-range planning horizon (six months to several years) aims to address this issue by allocating resources to meet the predicted future demand as effectively as possible, in order to ensure long range profitability. Our intent is to use mathematical optimization to find the best possible allocations.

Keywords: Tactical resource allocation, Capacity utilization, Mixed Integer Linear Programming, Optimization, Manufacturing, Aerospace, Resource loading, Logistics, Multi-objective optimization.

1 Introduction

GKN Aerospace Engine Systems (GKN for short) in Trollhättan, Sweden, is a leading supplier of aircraft engine components. GKN's components have presence in almost all major commercial aircraft. The products produced in Trollhättan include casings and rotating parts for both the compressor and the turbine section of the engine. Rotation and a high temperature differences between different parts of the engine puts high, in many cases extreme, quality demands (tight tolerance limits) on the components. The capital-intensive production at a large aerospace tier-1 supplier like GKN is generally influenced by expensive materials, long supply lead times, a large product mix and demand variations (see [1]). The potential reduction of tied-up capital that can be achieved by decreasing lead times is therefore substantial; in addition the project is expected to increase the overall resource utilization.

1.1 Production context

Manufacturing is performed in multiple steps, such as cutting (milling, turning, drilling, and grinding), welding, assembly, heat/surface treatments, and control/measurements. For cutting, GKN has a variety of production resources (machines) with different functions. The factory is organized in several functionally oriented production shops, and most of the production resources are shared by different products. Each production shop is organized as a job-shop, where similar types of machines are placed in proximity to each other. A complication is that it is, in practice, impossible to physically move machines, as they are bulky and fixed into the ground in 2–5 meter deep pits. Thus, the factory as such can only to a very limited extent be adapted to changes in the product mix; it is therefore not possible (highly unlikely) to maintain perfect flows of parts through the factory over time.

1.2 New product introduction

The process of introducing a new product is tied-up with tactical resource allocation decisions. These tactical resource allocations (which are referred to in this text) should not be mistaken for the short-term resource allocations done when choosing between several available machining resources while scheduling. The latter is commonly addressed in industry (see an example from GKN, [2]). The tactical resource allocations identify which production processes that need to be developed. The processes developed then constitute the solution space for the short-term allocations, by deciding which resources are available for allocation for the short-term planning. The tactical problem therefore differs significantly from the short-term planning problem.

The development of the required production processes consists of a number of activities that are needed to get the required approvals from customers and authorities in order to make a resource available for a specific operation on a product. The technical preparations (also called qualifications) often require several days of computer numeric control (CNC) programming, design and production of new fixtures, testing and validation, etc. These activities are the main parts of an iterative process to acquire the customer approval, and may take several months to complete. The time and costs associated with this process is often significant; the initial resource allocations are therefore seldom changed. However, if forced changes are made due to large variations in product demand and/or changes in available capacity of machines, it leads to delays and delivery disturbances. The tactical resource allocations will thus have a long-lasting effect on delivery times, product profitability and efficiency of the production system. Thus, it is of utmost importance that the resource allocations are made in an *optimal* way, so that changes in the later stages can be avoided.

The first step in the process to introduce new products is the mapping of the new product to a suitable value stream. Different products, or part types (sometimes a product is made of multiple part types), belonging to the same *value stream* are quite similar in terms of functionality. For example, all rotating parts should be in the same value stream. The reason is that it is easier to control geometrical tolerances by analyzing the data on machines producing similar products. This data is stored in the statistical process control (SPC) systems at GKN. Studying the previous SPC results has been one way of analyzing the suitability of a certain machine to handle new parts, by studying historical process outcome on similar features, such as a hole. Then experts apply the tolerance limits from the new part to get an understanding of how the machine would perform when introducing this part. This is easier if the newly introduced part is similar to the once already being machined. So, to get an overview on all the machining-related characteristics for different machines is possible. These issues are of the highest importance in the aerospace industry, as any regular occurrence of non-conformance (i.e., a slight deviation from a customer tolerance limit) results in halting of production.

Once a value stream is identified for the product, an operations list is created by the manufacturing experts. The operations list contains the operations needed to produce a certain product (or part), as well as the specific machine in which the respective operation must be performed. The capacity planners who are responsible for managing and determining production capacity are involved to check if the capacity is available in these machines. At this stage, however, no detailed assessment of the impact on resource allocation of varying demand for products is made. This is the reason that sometimes, when there are changes in the demand of some products or capacity of some machines, it may result in an additional effort to change some product routings which requires waiting for customer approval. Thus, the resource allocation plan should be robust such that the routing is still feasible, even when demand or capacity has varied slightly from their respective expected values.

1.3 The need for tactical resource allocation

Internal company analyses have identified queuing at machines as a main contributor to long production lead times (see [1, p. 66]). As the queues are often caused by uneven loads on the production resources (see [3]), the production system must be prepared for such unavoidable demand variations, i.e., the resources need to have some buffer capacity. Due to accounting principles, the preliminary resource allocations often have a misleading impact on the internal cost assessment (see [4]), which often results in sub-optimal decisions from an overall capacity planning perspective (i.e., *fast* resources are favored, leading to an unbalanced resource loading). Therefore, a long-range tactical resource allocation is essential to achieve the required flexibility to absorb these variations, and to avoid unnecessary queuing.

This leads us to the research questions to be discussed in the following sections. We need to identify an objective function which will result in product routings with desirable properties, and in a later phase, good resource allocations. These properties constitute a balanced distribution of work load between machines; the robustness of the final allocation (minimum reallocations between time-periods); high machining standards; and a minimum distance travelled by products/part types between different physical areas in the factory. The relative importance of these objectives may vary, and it is an area of policy discussions for GKN. However, a balanced capacity utilization and high machining standards are the two most important properties.

1.4 Research front

There are rather few studies focused on the tactical resource planning problem on the medium- to long-range planning horizon. Tactical planning in general is a tool that can be used for reducing the effect of uncertainties in the production environment, and assist in creating stability in the production system (see [5, 6]). One of the reasons why this type of optimization model has been developed with limited success in few cases is due to the computational complexity of the problem. From a combinatorial perspective, there are so many feasible solutions to evaluate that the computation times tend to be so long that the problem has historically been considered practically insoluble. Our research aims to utilize optimization theory to solve such models in a reasonable time frame. There are not many industrial scale tactical resource allocation problem solved from a mathematical optimization perspective. This type of approach provides, not only a feasible solution, but also a measure of *goodness* (*the optimality gap, i.e., the ratio of the difference between the best known solution and a value that bounds the best possible solution to the best known solution*) (see the full definition in [7]). Most of the commercial software used to our knowledge either use heuristics or some other inexact methods lacking this feature. However, there are commercial mathematical optimization solvers such as Gurobi [8] and CPLEX [9], which can be used to solve such problems, but it requires identifying the

right objective function and constraints, and also modelling them in an efficient way. This is also a part of what we do.

2 Scope and problem description

The scope of the proposed tactical resource allocation model is restricted to allocating products to machines which perform cutting operations. They includes milling, drilling, grinding, and turning machines. The model should ideally be used when a new product is introduced, or when there are significant changes in the machining capacity (e.g., decommissioning of machines or the introduction of new machines). The proposed model should be used by the capacity planner and by product introduction leaders at the company.

Each product has a *bill of material* containing details about the product structure, including all the *part types*, and their quantities used in the final product. However, it is quite common that products are composed of one part. Each product/part type has an operations list, i.e., a list of manufacturing operations, such as milling, drilling, turning, and grinding, required to manufacture this product/part. A combination of a product/part and such a specific *operation* constitutes a *job type*. The task of the resource allocation model is to assign each job type to resources (machines) such that the imbalance in capacity utilization is minimized. A simple example with five machines and a part type having three operations is illustrated in Fig. 1. We have a single part (Part 1) which requires three operations to be completed, and let us assume all five machines are capable of doing these three operations. We also define a user parameter τ which is the maximum number of machines which are allowed to do a specific job type, i.e., only machines 1 and 5 are allowed to first operation on the first part in Fig. 1. It is desired to identify only few machine alternatives for each job type or a user defined parameter ($1 \leq \tau \leq 5$). This restriction is to prevent too complicated flows and limit the number of qualifications or tests to be conducted by manufacturing experts to check the performance of a machine for a particular type of operation on a specific part. One possible solution (for $\tau = 2$) are qualifying machines 1 and 5 for the first operations, 3 and 4 for the second and 2 and 5 for the third operation. This results in the following eight possible routings $1-4-2$, $1-4-5$, $1-3-2$, $1-3-5$, $5-3-2$, $5-3-5$, $5-4-5$, and $5-4-2$. So, if we look at this small combinatorial problem, we started with $\left(\binom{5}{2}\right)^3 = 10^3$ alternative routings, and we selected only 2^3 which results in capacity balancing. Moreover, typically there are approximately 3–4 part types in a new product and each have at least 3–5 operations, and there are nearly 150 machines. Thus, a simple enumeration for such a large instance of the problem is too computationally time-consuming and we need to use mathematical optimization to find the best possible allocations. Thus, in a way optimization helps manufacturing experts to identify which machines to qualify, such that both qualification costs and capacity imbalance are minimized for the next 3–5 years.
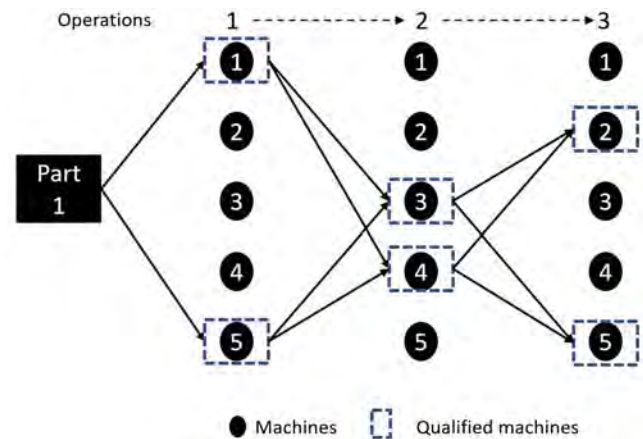

Figure 1: Part routes

2.1 Constraints

All product routings must satisfy some constraints. According to an internal company report, the constraints are classified into hard and soft constraints. Hard constraints are also referred to as level 1 constraints, while soft constraints are divided into level 2 and level 3 constraints. Going from level 1 to level 3, the importance of the constraints decreases (see Fig. 2). In Fig. 2, the base represents the size of the set of possible allocations of machines to orders of different job types; this set reduces as we add restrictions on, e.g., the number of qualifications (or qualification cost). The vertical position in Fig. 2 represents relative importance. In optimization parlance, the violation of soft constraints could be either combined and added as a penalty in the objective function (see [10]), which is to be minimized, or treated/modeled as different criteria in a multi-criteria optimization problem (see [11]).
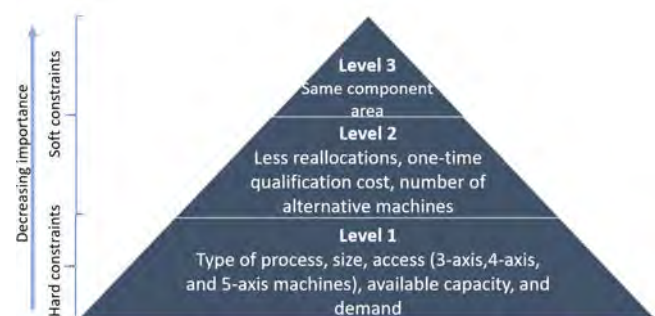
Figure 2: Constraints

Hard constraints   Hard constraints are those which must be satisfied during each time-period within the long-range planning horizon. These constraints include a required compatibility between job type and machine, based on the type of process (milling, turning, etc.) and whether a job type of a given dimension (e.g., diameter) can be processed on a given machine (i.e., size compatibility). Lastly, the projected workload on a machine should be less than or equal to its total

capacity, and the demand for each job type should be met in each time-period.

Soft constraints It is not mandatory to satisfy constraints which are associated with the preferences of the capacity planners. The constraints belonging to level 2 are mentioned below.

1. Reallocations: Preparing a machine for a job type is a resource-demanding activity; therefore, it is prudent to avoid changing the routing of orders of a given job type too often. For example, if we represent a job type as (i, j), where $i \in PartTypes$, and $j \in Operations$, then a binary decision variable s_{ijk}^t can be used to represent whether an order of job type (i, j) is processed in machine $k \in Machines$ at time-period $t \in TimePeriods$. The value of the variable s_{ijk}^t is 1, if job type (i, j) is performed in machine k, otherwise, it is equal to zero. Now, we use the introduced variables to clarify when a reallocation is considered. In Table 1, we highlight three alternative allocations (three rows) for machine 2 for the resource allocation sub-problem with one job type $(1, 1)$, and for five time-periods. Let us assume that all three allocations are feasible (that is, they satisfy the level 1 constraints). A reallocation of job type (i, j) to machine k at time t is determined by the variable values $s_{ijk}^{t-1} = 0$ and $s_{ijk}^t = 1$ (i.e., at the consecutive time periods $t - 1$ and t). The first solution in Table 1 comprises one reallocation at $t = 5$, while the second comprises two reallocations at $t = 3$ and $t = 5$, respectively. The third solution comprises no reallocations and it is therefore preferred. This reasoning is valid provided there are no reallocations at other machines.

Table 1: Reallocations between time-periods

s_{112}^1	s_{112}^2	s_{112}^3	s_{112}^4	s_{112}^5	Reallocations
0	0	0	0	1	1
0	0	1	0	1	2
1	1	1	0	0	0

2. One-time qualification: Generally, for any new product, there is a first-time qualification of each machine to which it is assigned. However, for orders of job types related to existing products, there exist machines which are already qualified or used historically. Nevertheless, whenever reallocating orders of an existing product to another machine, which is technically capable but not qualified before, a one-time cost for qualifying the other machine for that product arises. This means additional man-hours for programming CNC or developing new fixtures, and sometimes waiting for the customer approval. It is preferred to keep this *fixed cost* as low as possible. However, it is to be checked whether allowing qualifications results in an improved capacity balancing.

The level 3 constraints are the least important. One of the con-

straints included is the limitation of the total distance travelled by orders. Each machine is assigned to a physical space in the factory, so-called *component area* in the factory. It is desired to keep the number of times orders must flow in-between these component areas below a pre-defined number, or to keep it as low as possible. This helps to maintain an efficient flow and to reduce internal transportation.

2.2 Optimization model

The approach taken to solve this resource allocation problem is to utilize a mathematical optimization model. The output of the proposed optimization model is the assignment of the different job types to the machines. The model identifies a small number of different product routings for the capacity planners/short-term scheduling. We plan to obtain a resource allocation for a period of three to five years, with quarterly or monthly time-periods. The demand for products is steady, and there is a steady flow of raw materials through the factory. An optimization model requires an objective function which assesses the quality of the solutions, while a set of constraints to restrict the possible allocations, which in turn restricts the achievable objective value. We have mentioned all the constraints in the previous section. Thus, next we discuss how to incorporate the desired properties of the final resource allocation plan into the objective function.

1. Objective function: There are some articles which have discussed approaches to calculate internal production costs for products in similar production environments (see [4, 12, 13]). Most of these articles have considered product cost as a function of the total machining hours spent on a product. This is the legacy of the traditional accounting approach as mentioned in [4]. They use the product routing as a constant in their cost function, but with our approach we let the product routing be a variable, and instead of minimizing machining hours we minimize resource utilization imbalance. This is to our knowledge a novel approach and much more effective in improving flow efficiency in functional oriented production. We define *utilization* as the ratio of planned hours of machining to total available hours in a machine. Thus, to obtain a balanced utilization, we propose minimizing the (planned) utilization which is above a threshold value, that is, imposing a penalty for each time period when a machine's planned utilization is above a pre-defined threshold ζ, where $0 \leq \zeta \leq 1$.

2. Decision variables: To define the objective function, we introduce some notation. We represent the number of orders of type (i, j) produced in time-period t in machine k by x_{ijk}^t. The capacity (in hours) of machine k is denoted by using capacity $C_k \geq 0$ (Note: In case of planned decommissioning of machines capacities can be indexed over time). The number of hours required to process an order of job type (i, j) in machine k is represented by p_{ijk} (also including average set-up time as we do not assume batch production). In (1), the inner bracket represents the total resource loading of machine k during time-period t. Thus, dividing it by the total capacity of

DOI
10.3384/ecp19162021   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **186**

machine k yields the relative utilization at time-period t. The plus-operator, denoted $[\cdot]_+$ sets any negative value to zero. Our first proposed objective function to be minimized over the set X of feasible solutions satisfying the level 1 constraints is defined as to

$$\text{minimize } z_1(x \in X) := \sum_t \sum_k \left[\frac{1}{C_k} \left(\sum_i \sum_j p_{ijk} x_{ijk}^t \right) - \zeta \right]_+ \tag{1}$$

Our second alternative for an objective function is defined as to

$$\text{minimize } z_2(x \in X) := \sum_t \max_k \left[\frac{1}{C_k} \left(\sum_i \sum_j p_{ijk} x_{ijk}^t \right) - \zeta \right]_+ \tag{2}$$

The function z_1 represents the average of non-negative excess capacity utilization (summed over the machines and the time-periods). The function z_2 represents the maximum (over the machines) excess utilization (summed over the time-periods). Since the average can be composed by both small and (very) large values, the function z_2 represents over-utilization more adequately than z_1. A review of such inequity-averse optimization approaches is described in detail in [14]. The benefits of using the (2) are also highlighted in the work done on resource allocation problems in [15].

3   Expected results and future research work

The proposed optimization model is a mixed-integer linear programming model. The objective function and constraints are linear or made linear using integer/binary variables. The goal of the optimization model is to produce optimal or near-optimal solutions in a reasonable time frame. Since we are not aware of the lower limit (best possible feasible values) of some of the soft constraints, we intend to use one of the iterative frameworks as mentioned in [11].

3.1   Expected results

The proposed decision model should enable a faster and better adaptation of the production system to changes in the market and production conditions. Using our model, the organization's aim is to increase the available machining capacity by at least 2%, which could mean a 4-digit number of additional hours of more machining time available each year. These additional machining time could be made available when new products are introduced, or when the demand for some of the existing products increases. It is expected to reduce the number of over-utilized machines, which will result in reduced queues and, subsequently, reduced lead times.

3.2   Future research work

We intend to use real production data extracted from the Enterprise resource planning (ERP) system at GKN. We will identify an existing product and consider it as a new product

to be introduced in the factory to test the model. We will utilize the historical resource allocation to find out which machines are qualified to perform operations on the existing products. We also plan to take input from manufacturing experts from GKN on identifying other similar machine alternatives which can perform some of these operations. This will increase the size of the feasible set of possible allocations, which is important for improving the optimal value. We will perform computational experiments to benchmark different model formulations with respect to computation time and solution quality.

References

[1] A.J. Lewestam and H.S.U. Mäki. Increasing throughput potential in functional oriented machining plants. *Master Thesis, Division of Logistics and Transportation, Chalmers University of Technology, Göteborg, Sweden*, 2015.

[2] K. Thörnblad, A.-B. Strömberg, M. Patriksson, and T. Almgren. Scheduling optimisation of a real flexible job shop including fixture availability and preventive maintenance. *European Journal of Industrial Engineering*, 9(1):126–145, 2015.

[3] E.M. Goldratt. Computerized shop floor scheduling. *International Journal of Production Research*, 26(3):443–455, 1988.

[4] A. Myrelid and J. Olhager. Hybrid manufacturing accounting in mixed process environments: A methodology and a case study. *International Journal of Production Economics*, 210:137–144, 2019.

[5] J. Olhager and E. Selldin. Manufacturing planning and control approaches: market alignment and performance. *International Journal of Production Research*, 45(6):1469–1484, 2007.

[6] A.M.T. Thomé, L.F. Scavarda, N.S. Fernandez, and A.J. Scavarda. Sales and operations planning: A research synthesis. *International Journal of Production Economics*, 138(1):1–13, jul 2012.

[7] Optimality gap: `https://glossary.informs.org/ver2/mpgwiki/index.php/Optimality_gap`.

[8] Gurobi's website: `https://www.gurobi.com/`.

[9] IBM's CPLEX optimizer: `https://www.ibm.com/analytics/cplex-optimizer`.

[10] M.L. Fisher. The Lagrangian relaxation method for solving integer programming problems. *Management Science*, 50(12):1861–1871, 2004.

[11] R.T. Marler and J.S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, 2004.

[12] C. Windmark and C. Andersson. Cost assessment of a production system—a method targeting a product's aggregated value stream costs. *Procedia Manufacturing*, 25:231–238, 2018.

[13] E.W. Hans, W. Herroelen, R. Leus, and G. Wullink. A hierarchical approach to multi-project planning under uncertainty. *Omega*, 35(5):563–577, 2007.

[14] Ö. Karsu and A. Morton. Inequity averse optimization in operational research. *European Journal of Operational Research*, 245:343–359, 2015.

[15] H. Luss and D.R. Smith. Resource allocation among competing activities: a lexicographic minimax approach. *Operations Research Letters*, 5(5):227–231, 1986.

Airframe Sealing Automation Using Snake-robot

Luís Gonzaga Trabasso[*1,2], Doglas Negri[*2], Walter Antonio Kapp[*2], Magnus Engström[*3], Guilherme Sartori Natal[*1]
and Kerstin Johansen[*4]

*1 Department of Mechanical Engineering, Aeronautics Institute of Technology, SJCampos, SP/Brazil
E-mail: gonzaga@ita.br, guilesn@gmail.com
*2 SENAI Institute of Innovation, Joinville, SC/Brazil
E-mail: luis.gonzaga@sc.senai.br, doglas.negri@sc.senai.br, walter.kapp@sc.senai.br
*3 SAAB AB, Aeronautics, Linköping, Sweden
E-mail: magnus.engstrom@saabgroup.com
*4 Department of Management and Engineering, Linköping University, Linköping, Sweden
E-mail: kerstin.johansen@liu.se

Abstract

The work described herein have had the objective of carrying out a technical-economic feasibility study of a snake-robot for sealant application in an aircraft fuselage. The execution of the project brought knowledge related to the design and development of robots, constrained Cartesian trajectory generation, optimization methods for inverse kinematics computation, collaborative robotics and 3D simulations besides experiments related to aircraft components and application of sealants in the fuselage.

Concerning the technical feasibility of the project, the solution proposed proved to be an adequate solution for the proposed task. The sealant application trajectories in fillets (union of two plates of the fuselage) and fasteners were tested, being all well executed by the robot in a virtual environment, without collisions and without relevant complications. All the trajectories used for the evaluation were taken from the test framework provided by Saab and respecting the process specifications described for this task.

Economical evaluation of ROM (Rough Order of Magnitude) gives a result showing fairly good pay-back time although the potentials would be even better with a lower cost.

A list of commercial components for the manufacturing of the snake robot was specified, such as: motors, reducers, vision system, components for programming and controlling the robot (drivers, software and robot interface with the user) etc. It is important to mention that, despite the design of the robot drivers, the CAD model developed here needs further detailing for fabrication purposes. Structural analysis must be carried out in the structure and, after any changes are made, evaluations and corrections must be done. A study about the dimension and mass reductions of an existing sealant dispenser was carried out, where the cartridge and the sealant application valve were chosen to be coupled directly to the robotic arm.

In general, all activities were executed according to the proposed schedule. The project presented a very high potential and could be applied not only to the aeronautical industry (which was the main objective of the present work), but also to other fields of the industry. Therefore, a continuation of this project is highly recommended, such that the proposed snake robot can be manufactured and physically tested.

Keywords: SNAKE-Robot, Aircraft Industry, Sealing Automation, Collaborative Robotics.

1 Introduction

In the aircraft industry normally most part of the airframe must be sealed. For example, the internal structure of the wing defines the tank and needs to be sealed to prevent any leakage. Due to the weight optimization, there is no internal tank. The fuel is in direct contact with the inner side of the wing skin.

To seal a wing a small and thin worker have to go inside of the airframe through inspection openings. For a big commercial aircraft, this task is easy, but for an executive jet, or combat jet the sealing task could be very hard, or even impossible, putting constrains to the designer when optimizing the structure for aerodynamics and loads.

The internal fuselage structure is also sealed in order to provide a pressurized cabin. In addition sealing is often used for corrosion protection, especially when mixing airframe part materials and in areas with high level of moisture.

The standard anthropomorphic robots cannot fit to this task due the lack of degrees of mobility necessary to avoid collisions in the restricted workspace. This article has proposed a study of a robot design similar to a snake morphology.

The kinematic chain of a snake robot has several redundant degrees of freedom. This makes it possible to control a path and a shape of a robot, making it able to insert the robot and work inside the airframe, avoiding collisions.

The kinematic solver and the trajectory planner give some new ideas of how to realize a feasible robot design.

2 Project Objective

2.1 General Objective

The objective of this project was a feasibility study of a concept of a virtual snake-robot for automatic application of sealant in aeronautical structures.

The developed robot has been tested (virtually) through the generation of trajectories, kinematics solution strategies, accessibility simulations and the execution of sealing tasks. Subsequently it has been evaluated for technical and economic feasibility for implementation at Saab Aeronautics, a Swedish industrial partner in this project.

2.2 Specific Objectives and Established Goals

Specific Objectives	Established Goals
Definition of an application and requirements for automated sealing that can be used for additional evaluation of the snake robot and the sealing agent.	Define a representative application for sealing, which includes detailed requirements regarding product aspects: sealing; and of production: interface man / machine. Delivery includes CAD application models that will be used for further system evaluation.
Design the conceptual 3D model of the snake robot based on industrial components, including the kinematic model.	Design a low-TRL snake robot for an industrial application of sealant in aircraft fuselage using CAD modeling, CAE and simulation tools.
Study to reduce mass and dimensions of an existing sealant application tool (used in larger robots), to be coupled to the snake robot.	Considering the sealant applicator developed by ITA, assess possibilities of reduction of dimensions and mass of the device.
Technical and economic criteria used for evaluating the feasibility of the proposed solution. Assess the	Considering the constraints of the process described above, assess the performance of the robot from the
potential for establishing an international network among robot integrators Swedish, Brazilians and component suppliers.	point of view of implementing the task of sealing. Still, evaluate the manufacturing costs of the robot, in order to observe the economic viability of the project.

Table 1: Specific objectives

3 Scope of the project execution

3.1 Definition of an application and requirements

The application is defined by Saab and consists of a virtual DMU – Digital Mock-Up, of a section of an aircraft fuselage with the overall dimension 1 m³. The DMU contains the main requirements for the robot geometrical accessibility, quality requirements of sealing of fillets and rivets including process cycle times, see Fig. 1.

3.2 Automated sealing process

The sealing processes that are of interest for automation are the fillet sealing within the structure (not the aerodynamic sealing) and the coating of the fasteners (photos below).

The ASASR project uses the virtual test structure (DMU) to simulate the sealing process using the snake robot and the end effector. Specific parts of interest in the test are displayed on the following pages.
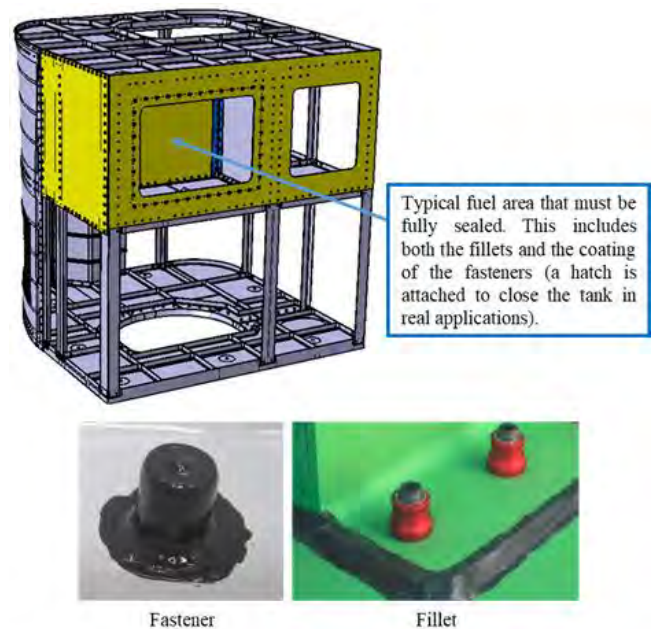

Typical fuel area that must be fully sealed. This includes both the fillets and the coating of the fasteners (a hatch is attached to close the tank in real applications).

Fastener   Fillet

Fig. 1: Test airframe and project requirements

3.3 Product requirements

As described above, the ASASR project focuses on the fillet sealing and the coating of fasteners of the airframe. The PR1776 B2 (PPG) sealant was chosen for this project. It is important to mention that, when mixing the two materials that form the sealant, air is introduced to the mixture, which will

result in quality problems (with a risk of rework). In this way, it is preferable to use the existing cartridge system for the sealing end-effector.

3.4 Production requirements

The following production requirements are defined for the use of a snake robot with sealant application system for fastener coating and fillet sealing:

- The application cycle time, including the positioning between two neighboring rivets, must be less than 10 s/rivet;

- The fillet sealing must reach a minimum speed of 50 mm/s;

- The installation should not require physical safety barrier and must still conform to the guidelines of the relevant European collaborative robots machines;

- The reach position is found up to 1 m, from the robot's entry into the aircraft structure;

- The error of product position, both for rivets and for fillets has a maximum tolerance of +/- 2 mm. Therefore, it is necessary to use a calibration system for accurately positioning the robot

- Snake robot programming should be done offline, using commercially available software, such as CATIA V5 Robotics, with appropriate software.

4 Snake robot project

The kinematic chain model, which is the basis for the development of this project, is a highly flexible concept when it comes to performing tasks in confined environments (limited accessibility and restricted workspaces). This high flexibility is guaranteed mainly by the hyper redundant robot. A first analysis is related to kinematics modeling of this robot, where the models of direct and reverse kinematics are considered. Then there are studies to generate the path to be followed by the snake robot while performing tasks in a 3D environment. Finally, the dimensioning of the elements and the CAD modeling of snake robot.

4.1 Kinematic model of the snake robot

For robotic chains of any kind, the creation of kinematic models is of paramount importance to the knowledge of the dynamic behavior of the robot, or for the characterization of the stresses suffered by certain joints. Skeleton models, associated with the Denavit-Hartenberg parameters [1] (serial robots) for the translations and rotations of coordinate systems of each joint, are commonly used in these robotic cinematic models, as showed in Fig. 2. In this study, this application added to some techniques of numerical optimization is used in the models of direct and inverse kinematics of the robot-snake. Following are the kinematic models employed for this solution.

4.1.1 Direct kinematics

Initially, considering the process requirements presented by company Saab, it was proposed a distribution of joints that made up the kinematic chain of the snake robot. This kinematic chain represented schematically in Fig. 2, consists of a pedestal (J_1), which provides the robot rotation and inversion complete. Followed by five joints snake type (J_2-J_6). These snake joints have the particularity of maintaining the alignment of the links, however, they rotate the axis in 90° joints in consecutive joints. Finally, like joints 7 and 8, they are part of the semi-spherical robot of the snake that assists in the execution of the sealing tasks.
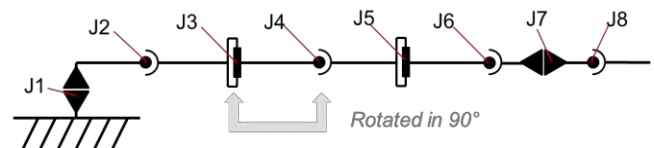

Fig. 2: Kinematic Model

Then, an 8-degree freedom skeleton model for the direct kinematics of the snake robot is proposed and developed in numerical calculations that solve in Scilab 6.0.1 (Fig. 3). The main objective of this model is the understanding of the distribution of joints and their axes of local coordinate referred to in the workspace of the robot.
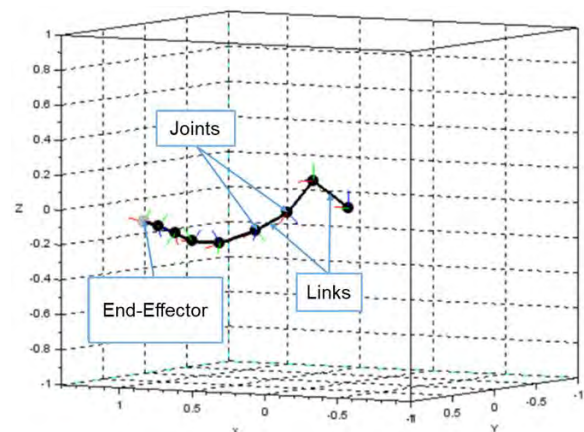

Fig. 3: Skeleton model for direct kinematics solution

Table 2 presents the D-H parameters used in the design of the kinematic model.

The same skeletal model using Denavit-Hartenberg parameters, as well as optimization techniques (PSO - Particle Swarm Optimization), is used for the solution of the inverse kinematics of the robot.

4.1.2 Inverse kinematics

Inverse kinematics of a robot, nothing more, than from a three-dimensional cartesian coordinate (X, Y, Z), directed (V_X, V_Y, V_Z) or not, the effector, through algebraic solutions (numerical or analytical), a set of angular amplitudes joints (in this case, $\theta_1, ..., \theta_8$) [2], which meets this solution.

*1 – All inout information of angles must be in radians [rad].
*2 – All input information of length must be in meters [m].

Link	θ^{*1}	d^{*2}	a^{*2}	α^{*1}
#1	θ_1	0,022	0,277	$\pi/2$
#2	θ_2	0	0,249	$\pi/2$
#3	θ_3	0	0,212	$\pi/2$
#4	θ_4	0	0,196	$\pi/2$
#5	θ_5	0	0,196	$\pi/2$
#6	$\theta_6; -\pi/2; 0$	0; 0; 0,060	0; 0; 0	0; $-\pi/2$; 0
#7	$\theta_7; 0; \pi/2; 0$	0; 0; 0; 0	0; 0; 0; 0,165	0;$\pi/2$; 0; 0
#8	θ_8	0	0,084	0

Table 2: Denavit-Hartenberg parameters of snake robot

In this study, an iterative numerical solver was proposed in Scilab 6.0.1 to solve the trajectories (oriented three-dimensional Cartesian coordinate sequence), which uses concepts of the optimization method PSO - Particle Swarm Optimization in its solution. This solver receives the cartesian trajectories oriented in the form of a ".txt" file, performs the solution, and returns the angular amplitude of each of the 8 joints $\theta_1, \dots, \theta_8$ in a file of the same format for each discrete point of the trajectory, as can be seen in Fig. 4.
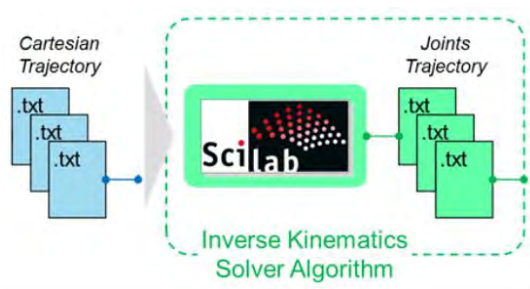

Fig. 4: Inputs and outputs of the inverse kinematics solver

As it is a redundant robot, that is, it has more degrees of freedom (number of θ = 8) than is necessary to execute a given task (X, Y, Z) and oriented (V_X, V_Y, V_Z), a strategy of divide the solution of the inverse kinematics of the robot into two parts is adopted. This strategy consists of, starting from the proposed chain, to create two virtual robots. The first robot consists of the three initial joints of the snake robot, making it possible to position the joint 4 in a known Cartesian coordinate (X_4, Y_4, Z_4) within a working space. For a better positional control of the robot, this known coordinate is strategically positioned in the trajectory of the robot, as can be seen in Fig. 5. After the inverse kinematics proposed for the first robot, find X_4, Y_4 and Z_4 through the PSO, the solution of the second virtual robot containing 5 degrees of freedom is performed. In this case, besides the positioning of the process effector (X_{EE}, Y_{EE}, Z_{EE}), the orientation of the vector V_X must be guaranteed (Fig. 5). For this, two (α, β) of the three Euler angles are required for the orientation to be maintained. These are calculated from the V_X vector itself, such as:

$$V_X = [a \quad b \quad c],$$

$$\beta = sin^{-1}(-c),$$

$$\alpha = cos^{-1}(b, a).$$

Thus, the effector orientation matrix of the process can be calculated as:

$$EE_{orientação} = \begin{bmatrix} cos\,\alpha * cos\,\beta & -sin\,\alpha & cos\,\alpha * sin\,\beta \\ sin\,\alpha * cos\,\beta & cos\,\alpha & sin\,\alpha * sin\,\beta \\ -sin\,\beta & 0 & cos\,\beta \end{bmatrix}$$

With the remaining vectors of orientation V_Y and V_Z being the result of this matrix of rotation. With this, again the optimization algorithm can perform the search of the angles $\theta_4, \dots, \theta_8$ that satisfy the considered tolerance.
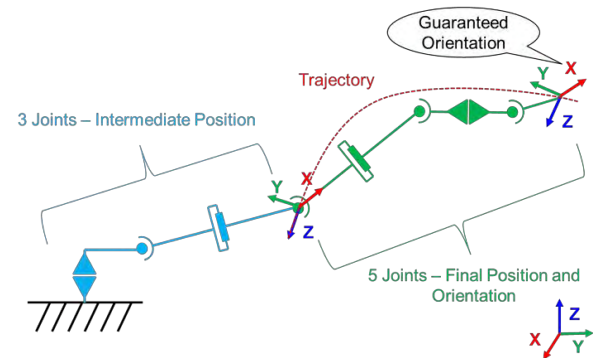

Fig. 5: Proposed of strategy to solve the inverse kinematics.

The way the optimization algorithm finds a solution that satisfies the required tolerance is identical for the two steps of the solution strategy (intermediate position and final position and orientation). However, the difference between the two solutions is the objective function to be optimized. For the intermediate positioning, the objective function is defined as:

Minimize $\qquad f(\theta_1, \theta_2, \theta_3)$

$$f = \sqrt{\left(X - \overline{X}\right)^2 + \left(Y - \overline{Y}\right)^2 + \left(Z - \overline{Z}\right)^2}$$

where   (X, Y, Z) the desired Cartesian coordinate,

$(\overline{X}, \overline{Y}, \overline{Z})$ the optimized Cartesian coordinates.

and for the positioning and orientation of the effector of process defined as:

Minimize $\qquad f(\theta_4, \theta_5, \theta_6, \theta_7, \theta_8)$

$$f = \sqrt{\left(\sum V_x - \sum \overline{V_x}\right)^2 + \left(X - \overline{X}\right)^2 + \left(Y - \overline{Y}\right)^2 + \left(Z - \overline{Z}\right)^2}$$

where V_x is the desired orientation vector,

(X, Y, Z) the desired Cartesian coordinate,

$\overline{V_x}$ is the optimized orientation vector

($\overline{X}$, $\overline{Y}$, $\overline{Z}$) the optimized Cartesian coordinates.

An important definition is the working limits of each of the snake robot's joints. These limits are used in the optimization algorithm to restrict the search field of the PSO particles. The limits for each of the joints are shown in Table 3.

θ_1	θ_2	θ_3	θ_4
±180°	-90°~+135°	±120°	±120°
θ_5	θ_6	θ_7	θ_8
±120°	±120°	±180°	±90°

Table 3: Angle limits for the joints

In this way, the optimization algorithm seeks to minimize the objective functions. The PSO, an optimization algorithm originally developed by Eberchart and Kennedy [3], consists of, from an initial population of random particles whose members interact locally with each other and are governed by global rules, perform the minimization of any previously defined objective function. The present work uses the fully connected topology (g_{best}), which means that a particle takes the entire population as its topological neighbors.

Initially, each particle has a random position and velocity. The particles interact with each other, informing the best position. With the data, velocities and positions are adjusted for each particle i. The velocity of a particle for the next increment (k + 1) is given as:

$$v_{k+1}^i = \omega v_k^i + C_1 R_1 \left(p_{best}^i - x_k^i \right) + C_2 R_2 \left(g_{best}^i - x_k^i \right);$$

$$i = 1, 2, \ldots, N$$

where, N is the total number of particles, R_1 and R_2 are random values from 0 to 1, p_{best} is the best position for that particle and g_{best} is the best overall position. ω, C_1 and C_2 are the chosen parameters: the first is the inertial component of the particle, and the latter are terms of "reliability" among the particles of the group [4]. The position in the next iteration can be determined as:

$$x_{k+1}^i = x_k^i + v_{k+1}^i$$

where, x_k^i is the position of the particle in the current iteration. Fig. 6 shows the position and velocity of a particle in one step and how the best overall position and the best particle position will influence the next particle position.
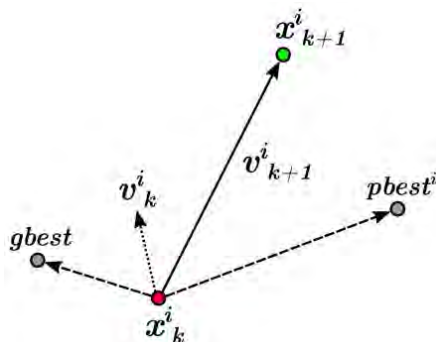

Fig. 6: Present and future position of the particle and speed

The entire optimization process requires a stop condition. For this problem, the set of stop tolerance is 1E-4, or even the maximum number of iterations is reached 150. This procedure is triggered for each discrete coordinate that makes up a path to be followed by the robot.

4.2 Generation of trajectories

A great challenge for the snake robot (hyper redundant serial manipulator) is to develop its interface to the programmer to be able to make use of all the geometric performance that this mechanical system model can offer.

In this feasibility study on the use of a snake robot for the application of aircraft fuselage seals, a versatile graphical interface was adopted to simulate a simplified CAD/CAM system for specific off-line programming application of the virtual robot prototype. This section describes the routines developed with the Rhinoceros 5 software and its Grasshopper 0.9.0076 plug-in for the generation of Cartesian trajectories in different types of tasks. Fig. 7 shows the user input workflow to view the programmed path.
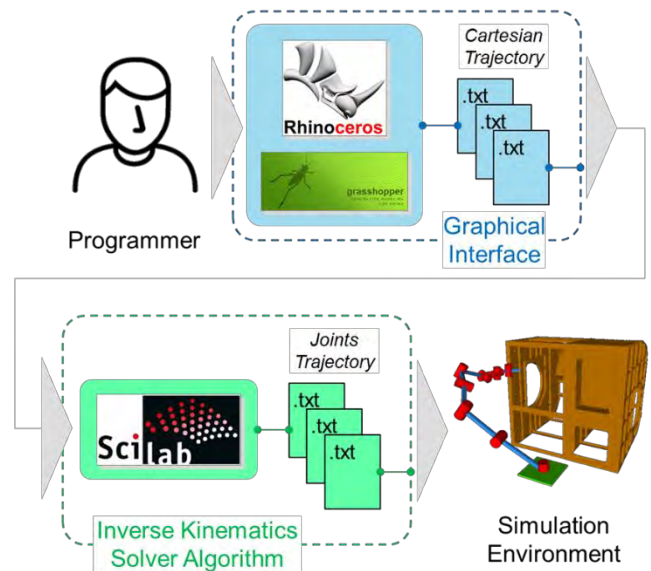

Fig. 7: Overview of workflow for trajectory generation.

To generate a trajectory for the snake robot, steps are followed, consisting of a list of coordinates of sequential points and their respective orientations of the effector; therefore, the programmer must first use the Rhinoceros 5 CAD module to: (i) draw free curves or (ii) extract curves from an existing CAD model (eg, fuselage). In both cases, the curves serve as the basis for the complete definition of the trajectory. Different systematic algorithms developed with Grasshopper especially for this project, segment the curves selected by the programmer and also, in most cases, give orientations to the effector at each segmentation point, generating a text file (Cartesian trajectory). In this step, the text files of the Cartesian trajectory are given as input in Scilab 6.0.1, which implements the inverse kinematics solution algorithm implemented. This procedure is described in Section 4.1, along with a detailed explanation on the kinematic modeling of the snake robot mechanism. The last

stage of the trajectory generation workflow is motion simulation. In a dedicated simulation environment generated in LabView 2017 with the help of the Robotics supplement. Joints trajectory text files are used to represent the synchronized angular position of each motor / joint for each trajectory segmentation point.

4.2.1   Types of trajectories

Two categories of main trajectories were designed for this feasibility study: chain insertion and tasks. The task execution trajectories for the application of sealing in aeronautical structures comprise three subcategories: general trajectory, fillet sealing and rivet covers, the latter being divided into two subclasses as to the geometric strategy appropriate for the application (propeller and cone).

Fig. 8: Types of trajectories for the snake robot.schematically shows the types of trajectory categories for the snake robot and its subdivisions. The following sections briefly describe when and how the programmer should apply each trajectory to make use of the mechanical flexibility offered by the snake robot.
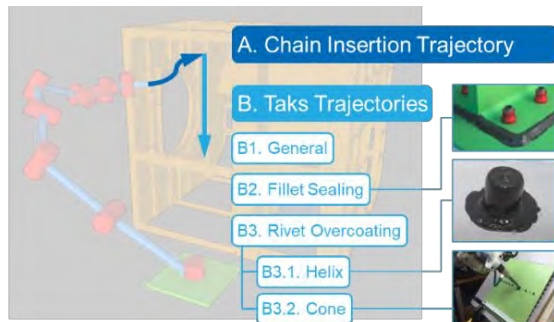

Fig. 8: Types of trajectories for the snake robot.

4.2.1.1   Chain insertion path

This type of trajectory is used for the insertion of the kinematic chain within a restricted working space (obstacles). The objective is to position the final effector ("snake head") as close as possible to the region where a specific task is to be conducted, with a certain amount of control over the path through which the entire kinematic chain ("snake body") moves along this trajectory. This is obtained by controlling the spatial trajectory of the fourth joint of the robot (J4) and the effector (EE) separately, but synchronously. In this sense, the simplest way for the programmer to generate an insertion trajectory is to draw a curve in space to serve as the basis for the movement of the entire snake robot. This curve is called the complete path and is the first input of the algorithm (Fig. 9). Throughout the insertion, the fourth joint and the effector follow different sections of this curve.
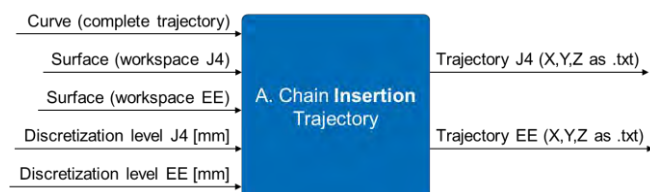

Fig. 9: Inputs and outputs of the algorithm for the generation of insertion trajectory of kinematic chain

The algorithm basically divides the entire trajectory with the use of the surfaces that define the boundaries of the workspace for the fourth joint and the final effector. These surfaces were previously generated and are saved as global variables in the Grasshopper application. The last entries required in this case are the desired discretization levels for each segment of the trajectory. All this information is processed and two text files containing each trajectory (fourth joint and effector) are created. It is important to note that both trajectories have the same number of points. The inverse kinematic solver will have a pair of points (one from each trajectory) to find the joint locations values. Note that output text files don't contain orientation information (such as vectors or Euler angles). The tangent vector of the curve is calculated in Scilab and the solver will align the X axis of the robot to this vector. Fig. 10 shows an example of the use of this type of trajectory.
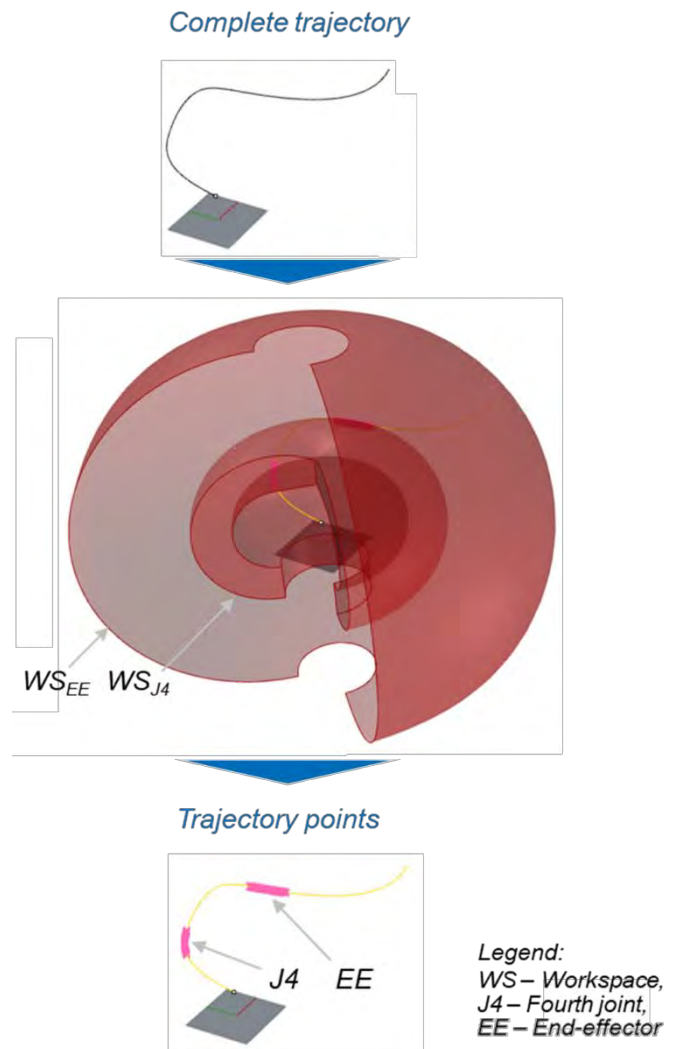

Fig. 10: Example of the use of the chain insertion trajectory

4.2.1.2   Task Execution Trajectory – General

General trajectories are the most flexible type of programming that can be applied to the execution of a task in this feasibility study. In this case, as in all paths of the task, it

is assumed that the snake robot is already positioned with its effector near the area where the fence task will be conducted. The first three joints are kept fixed (i.e., the center of the fourth joint does not move in space) and the remaining five joints are responsible for positioning and orientation of the effector center and the X-axis, respectively.

For this type of trajectory, the algorithm takes an entire path given as input by the user, as well as its level of discretization and returns a list of segmentation points and tangent vectors (V_X), with respect to which the orientation of the final effector will be solved. As shown in Fig. 11, the last input required for the definition of the remaining vector (V_Y and V_Z) is the reference rotation angle (α).

Fig. 11: Inputs and outputs of the algorithm for generating a history of general tasks.

The reference rotation angle is used to control the angular position of the orthogonal pair V_Y-V_Z in the normal plane at each segmentation point. This is the last parameter needed to fully define the trajectory in terms of position and orientation. The cross product $V_X \times \hat{u}_Z$, where $\hat{u}_Z$ represents the overall Z direction, generates an auxiliary vector (V_{Z_AUX}) which is normal for V_X. This result is then rotated around V_X from an angle α to create V_Z. Finally, the orientation vector of the last trajectory is given by $V_Y = V_Z \times V_X$. The influence of the reference rotation angle can be seen in Fig. 12.
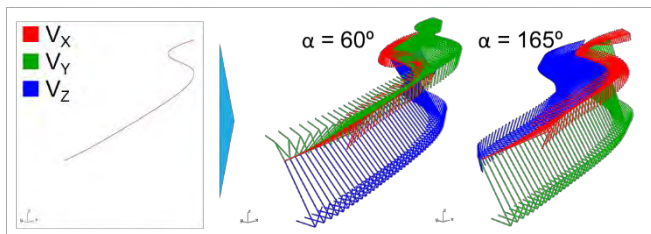

Fig. 12: Example of using the trajectory of general tasks and influence of the angle of rotation

4.2.1.3   Task execution path – Fillet sealing

This type of path is used to facilitate the scheduling of tasks in which the Effector of the robot should follow a corner (for example, fillet from the union of two aluminum plates) with a given orientation to the homogeneous application of the sealant. Unlike the general trajectory, there is no need to use the global Z direction to generate of an auxiliary vector. Once the plane curve representing the bottom of the fillet and its discretization level are given, the orientation of the final effector is totally defined by three other entries (see Fig. 13).

Fig. 13: Inputs and outputs of the algorithm for generation of fillet in sealing task trajectories.

First, the upper fillet curve, which consists of the lower curve folded in an orthogonal direction, guides the definition of local coordinate systems for each segmentation point. The tangent vectors of the lower curve, together with the vectors that join the segmentation points of each inserted curve, form the basis for these coordinate systems. Finally, the two orientation angles in relation to the previously formed local coordinate systems, namely the y-lead angle and the ω-tilt angle, define the inclination of the sealing applicator to the surface of the fillet. Fig. 14 shows the influence of each of these angles on the effector orientation along the programmed path.
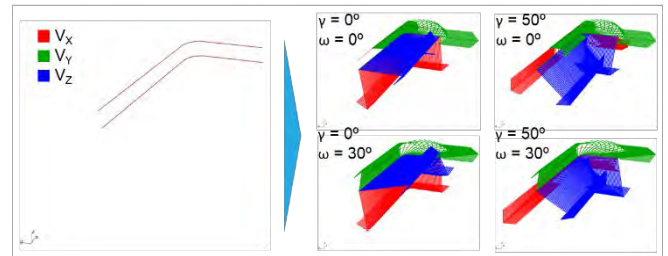

Fig. 14: Example of using of sealing trajectory of fillets and influence of rotation angle of reference

4.2.1.4   Task execution path – Rivet coating – Helical

Although this was not the main focus of this feasibility study, the SENAI Institute for Innovation in Manufacturing and Laser Processing Systems studied, as requested by Saab, to what extent the snake robot would also be able to conduct the auxiliary tools (such as sealant covers). The inputs required to generate the overlap helical rivet trajectory, are listed in *Fig. 15*.
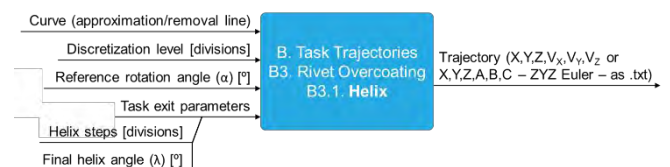

Fig. 15: Inputs and outputs of the algorithm for generating coating for a helical trajectory of rivets

In this type of task, the effector must approach the rivet location orthogonally in a linear trajectory along which its orientation is kept constant. The trajectory of exit / removal is also a line with the same direction and level of discretization as the approach trajectory. The main difference here is the helical behavior of the final effector orientation along the tool removal. The propeller pitch parameters and the final helix angle (λ) determine how the vectors V_Y and V_Z should rotate to the application flow of the sealant to be cut and the surface finish within the acceptable sealing patterns. The final angle of the propeller provides the angular displacement between the set of vectors V_Y-V_Z for the first segment of the approach path and for the last segmentation point of the trajectory of removal. The steps of the propeller, on the other hand, define the extent to which reorientation / transition regime should last (from the beginning of the trajectory of removal). Fig. 16 shows the influence of each of these parameters on final effector orientation throughout the propeller overcoating task.
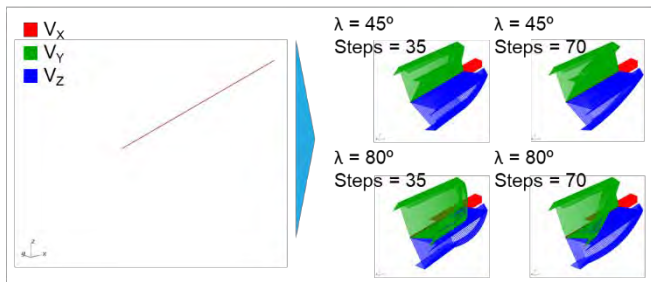
Fig. 16: Example of the use of helical trajectory for rivet and the influence of the parameters of the task's output

It is important to note that, as in the general path of the task, the general Z direction and a reference rotation angle (α) are necessary to define the approach orientation of the effector in the helical rivet liner. In addition, in this case of trajectory, 6-DOF are necessary to solve the inverse kinematics, since each parameter of orientation of the final effector must be combined so that the trajectory can be totally followed. The tests were performed with the freezing of the first two joints of the robot, but the stability of the solution found by the optimization algorithm still needs to be improved. If necessary and approved by all partners, the scope of a future project will meet this challenge.

4.2.1.5   Task execution path – Rivet coating – Cone

This is the latest trajectory concept designed for the feasibility study and provides a way for the application of sealant around a rivet, driven without any additional tools, but with the extruder's own nozzle. Similar to the concept of fillet-sealing trajectory, the trajectory of the rivet cone also needs two base curves (bottom edge and top row) for the creation of auxiliary coordinate systems, in relation to which the orientation angles of the sealing application process will be configured. The main input parameters that differ for these types of trajectory are the task output parameters present in the conical edge overlap (Fig. 17).
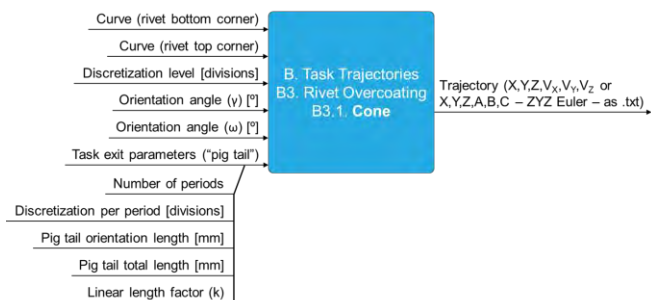

Fig. 17: Inputs and outputs of the algorithm for trajectory generation facing cone in rivets

An exit path in the form of a "pigtail" was defined at project meetings as ideal for cutting the flow of sealant in this type of application. Thereafter, after the sealant feed is interrupted, the "pigtail" trajectory takes care of the remaining sealant volume in the extruder and accommodates it so that the finish of the rivet seal is not compromised. There are a large number of parameters used for the complete definition of this output trajectory, but the influence of the main parameters can be seen in Fig. 18. The number of periods determines how many complete revolutions the effector center will be triggered

until it reaches a linear output trajectory. The total length of the "pigtail" defines how many millimeters of trajectory will be required for the number of defined periods to complete. In other words, it has influence on the "density" of the conical spiral. Finally, the linear length factor (k) indirectly defines how long the last (linear) part of the trajectory should be. The higher this factor, the smaller the linear part of the output path. If k = 1, the linear part of the output trajectory will have exactly the same length as the "pigtail". If k = 2, the length of the linear part will be reduced by half.
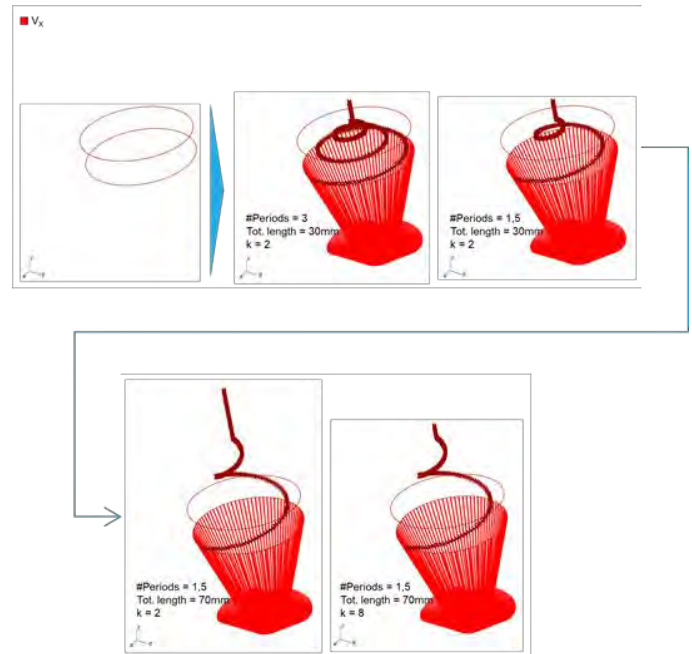

Fig. 18: Example of using cone trajectory for rivet and the influence of the parameters of the task's output.

4.3   Dimensioning and CAD model of the snake robot

For the snake robot design, some operational specifications are necessary. These specifications were basically taken from the process requirements defined by Saab (Section 3).

These specifications are used to dimensioning actuation drives for robot joints. Thus, a selection of some commercial components can be performed to compose this study. For gearboxes of main torque, the KHK Gears and Harmonic Drive® Gearboxes suppliers have been selected.

In the same way, from efforts to drive de gearboxes, the motors and control drivers of the joints of the snake robot have been defined. Available components from the Maxon Motor manufacturer have been selected.

The modeling of the snake robot was performed in the SolidWorks 2018 graphic environment. For this feasibility study, the dimensions related to the bearings and other structural components of the robot were not considered, since it was expected to evaluate its accessibility capacity and not its structural resistance. Fig. 19 show the CAD model of the proposed snake robot. A folder packed with the robot CAD file will be attached to this report.
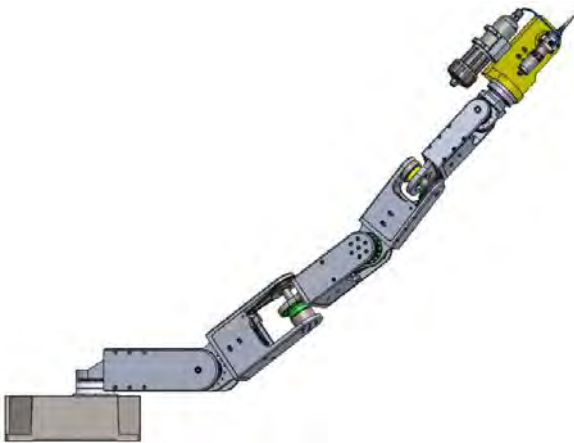
Fig. 19: Side view of the snake robot

4.4 Vision system

The proposed vision system for the project will serve to determine a local coordinate system in the sealing task workspace. The proposal is to use a measuring head with two integrated measuring systems: a LASER profilometer with three light sheets for 3D measurements and a 2D shape measuring camera on a surface.

The principle of measurement of the profilometer for 3D measurements is that of triangulation, with the measurement direction aligned with the axes of the LASERs and the camera positioned obliquely observing the line projected by the LASERs. The image shift of the line along the image at 2D sensor array lines determines the Z coordinate of the image on the LASERs axis.

The LASER profilometer with three light sheets, serves to accurately measure the normal vector of the surface to be machined. This could be done by using a LASER light sheet profilometer integrated into the robot movement, but which would carry the errors of the robot's movement. Like three sheets of light, the normal vector is measured with the robot in static pose, in a stable condition, and measurement is made only with recourse in the measurement head. Another function is to locate the robot at reference poles in the workspace and use local coordinates to minimize the errors of displacement of the kinematic chain.

Because of the complexity of the geometries found in the airframe, the measurement of the triple profilometer would require complex algorithms to identify the lines projected by each profilometer in the same image. To solve this problem, the LASERs are pulsed alternately, and the monochrome 3D camera captures images synchronized with each of them.

The camera that makes the 3D measurement is monochromatic and in its field of view there is a bandpass interference filter tuned to the LASER wavelength of the projectors. This feature, in addition the adjustment of the numerical aperture of the lens, set to the minimum value for the required depth of focus, causes the image captured by the camera to record only the line projected by the LASER without the sketch texture in the image, the 3D geometry

processing is also analytical, with basic image processing. Due to the possibility of addressing the LASER, it is possible to activate only one LASER to record the shape of the sealant deposited on the joints and on the fasteners (rivets and screws).

In the head there is even a second camera, colored, unfiltered passing through its field of vision. There is also a set of white LED illuminators to produce a precise episcopic image with appropriate contrast. This image is used to locate reference points for the task on the surface to be machined. Fig. 20 shows in detail a measuring head.
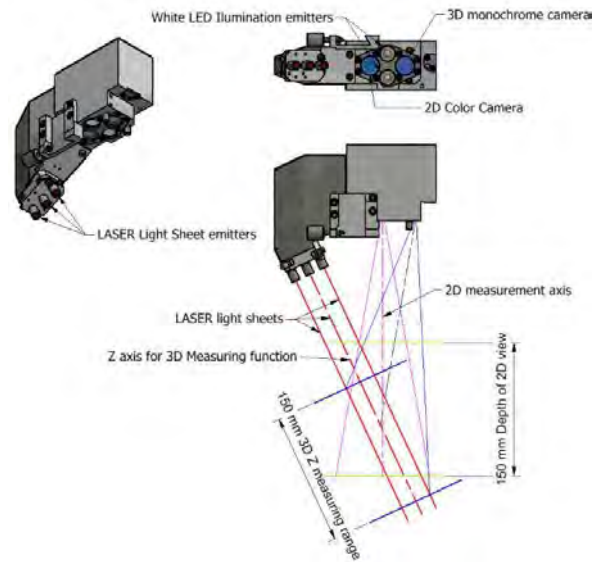

Fig. 20: LASER head for 3D / 2D measurements

4.5 Study of weight reduction and dimensions of the sealant dispenser 2K

A study for the development of a sealant dispenser for airframe fuselage has already been developed by the project partner Aeronautics Institute of Technology - ITA, where size and mass reductions have made on this development, resulting in a complete new design of this dispenser has proposed, where the coupling of the reservoir cartridge and the seal dispensing valve to the robot structure was chosen. For this purpose, the same cartridge and valve used in the dispenser developed by ITA were used in this proposal (Fig. 21), but now the dispenser is integrated on link 7 and valve is on the flange after the Joint 8.
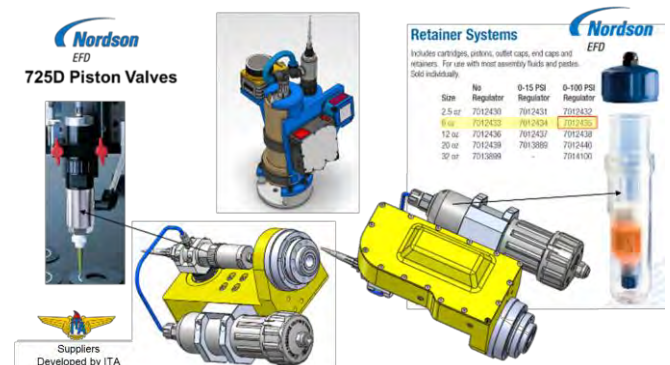

Fig. 21: Snake robot process effector

Thus, an even greater mass reduction was achieved for the sealant dispenser, reaching 2.7 kg considering the transmission components of joints 7 and 8, however, without considering the vision system in the dispenser developed by the ITA.

5 Conclusions

The present project, whose objective was to carry out a technical feasibility study on the use of a snake robot for the application of sealant in aircraft fuselage, had its scope and schedule fully executed. The execution of the project brought knowledge related to the design and development of robots, collaborative robotics, besides experiments related to aircraft components and application of sealant in fuselage.

In terms of the technical feasibility of the project, the solution proposed here, proved to be an excellent solution for the proposed operation. The trajectories of application of sealant in fillets (union of two plates of the fuselage) and fasteners were tested, being all very well executed by the robot in a virtual environment, without collision and without many complications. All the trajectories used for the evaluation were taken from the test framework provided by Saab, and respecting the process specifications described for this task. For the economic feasibility, although the estimated cost of the snake robot is higher when compared to the IIWA and the UR10e, the project is still feasible, since for the snake robot, after the complete development of the project, its cost of will be reduced.

A series of commercial components for the manufacture of the snake robot specified, from components for the manufacture and construction of the robot, such as: motors, reducers and vision system, to components for programming and control like drivers, software and robot interface with the user. It is important to point out that, despite the design of the robot drives, the CAD model developed here is not suitable for fabrication, structural analysis must be carried out in the structure and, after these changes have been made, evaluations and corrections are necessary robot drives must be performed. A study of the reduction of dimensions and mass in the sealant dispenser was carried out, where the coupling of the cartridge and sealant application valve was chosen directly to the robotic arm, causing an initial mass reduction of the applicator of 1.1 kg.

References

[1] L-W. TSAI, Robot Analysis: The Mechanics of Serial and Parallel Manipulators. Institute for Systems Research University of Maryland. Joh Wiley & Sons, Inc, 1999. ISBN 0-471-32593-7;

[2] B. Siciliano, L. Sciavicco, L. Villani, G. Oriolo, Robotics: Modeling, Planning and Control. Springer, 2009. ISBN 978-1-84628-641-4 / e-ISBN 978-1-84628-642-1

[3] Eberchart, R. and Kennedy, J. (1995). Particle swarm optimization. IEEE International Conference on Neural Networks.

[4] Perez, R. E. and Behdinan, K. (2006). Particle swarm approach for structural design optimization. Computers and Structures 19-20: 1579-1588.

Proposal of a Method for the Implementation of Industry 4.0 - Aircraft Final Assembly Domain

Alexandre Leite Junior, Gléverson Fabner Condé Lemos, Luís Gonzaga Trabasso

Department of Aeronautical and Mechanical Engineering, Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, São Paulo/Brazil
E-mail: alexandreleitejr@gmail.com, gleverson.lemos@engenharia.ufjf.br, gonzaga@ita.br

Abstract

The present work proposes a method to guide companies in a gradual implementation of Industry 4.0 based on the best practices identified in the scientific literature and customised to their contexts. In this sense, nine methods developed by other authors in the literature (ACATECH, PwC, Capgemini etc.) were analysed comparatively in order to identify opportunities to build an improved solution. Thus, the best practices identified were brought together and integrated on a new method which had each of its stages systematised. Furthermore, as a way of exemplifying its usage, the proposed method has been represented in the Aircraft Final Assembly domain – based on information collected from the literature and through interviews conducted with professionals of this area. At the end, the novel method has been evaluated in two ways: through a comparative analysis with the proposals of other authors and through a value perception survey applied to professionals from the industry and academy, related to the Industry 4.0 and innovation field within the aeronautical domain. Accordingly, this paper presents a comparative analysis of nine methods for the implementation of Industry 4.0 and provides a more complete solution based on the best practices identified in the scientific literature. In addition, it systematises a procedure for the elaboration of a customised Industry 4.0 Maturity Scale, presents the macro process mapping of a general Aircraft Final Assembly area and an Industry 4.0 Maturity Scale designed and customised for the Aircraft Final Assembly domain.

Keywords: Aircraft Manufacturing, Advanced Manufacturing, Industry 4.0 Maturity Scale

1 Introduction

According to Schuh et al. [1] Industry 4.0 (I4.0) is a new concept of the market characterised by the wide integration of information, communication and Cyber-Physical Systems (CPS) in the industrial environment. In general, it denotes an unprecedented transformation in terms of productivity, flexibility and agility; stimulating the efficient and sustainable use of limited resources, profitable production of highly customised products and the emergence of new business models in the market. In addition, some authors [1] [2] agree that Industry 4.0's most significant contribution is to transform companies into fast-learning organisations with agile interactions, capable of continuous adaptation to changing environments.

However, Schuh et al. [1], Porter and Heppelmann [3] emphasise that the phenomenon of Industry 4.0 is still a novelty without clear boundaries because it is not standardised and well disseminated in the market yet. Pessl et al. [4] consider that this is mainly due to the following obstacles: high investment costs, lack of transparency regarding potential benefits and fears related to changes in the organisational structure and the Information Technology security.

Thus, to obtain the benefits of I4.0, companies have to adopt a systematic and gradual plan of implementation, based on its market strategy and its structure as a whole. [1] [5] [6] For this purpose, some authors have proposed Methods for the Implementation of Industry 4.0 (MII4.0), in order to guide companies in this endeavor.

However, none of the methods found in the literature summarises all the best practices presented by them individually. Moreover, the related authors do not systematise how companies with specific characteristics – like from the aeronautical industry – could create the vision of I4.0 projected to their particular context, so that the deployment project is properly customised.

In this sense, the present work analyses, proposes and evaluates a more complete, flexible and customised MII4.0, in order to provide a scientifically based solution with a practical approach. The proposal is presented and has its structure detailed (with the description of its stages and decision points).

As a way of complementing the work, it was represented in the aeronautical context (more specifically, in the final assembly area), which is a very fertile environment for this

subject, due to its demanding levels of quality, flexibility, reliability, performance and cutting-edge technologies. [7]

Considering the subject exposed, the objective of this work was to propose a Method for the Implementation of Industry 4.0 (MII4.0) that met the best practices identified in the literature and could be customised to its application context, representing it in the Aircraft Final Assembly area.

2 Industry 4.0

The fourth industrial revolution refers to a new global wave that aims to seamlessly combine manufacturing, automation and Information & Communication Technology (ICT) into a vertical network within an organisation by connecting two or more of these organisations in a horizontal chain. These connected systems (also defined as CPS) are fused – blurring the lines that divide the physical and digital domains – being able to interact with each other using standard Internet-based Protocols, allowing data analysis, failure prediction and self-configuration in order to adapt to changes. [5] [6] [8] [9] [10] [11]

The concept of CPS indicates that physical production plants are connected to each other in an open network. This enables an autonomous and integrated value chain, where data and information travel without barriers among the different internal sectors of the company, as well as among the external stakeholders involved in the process. The goal is to create an integrated, flexible, efficient, transparent and optimised production flow. [5] [6] [8] [9] [10] [11]

The great advantage of this revolution, for Schuh et al. [1], is how quickly the companies will be able to react to variations in the value chain by: (1) having an instant insight of any new circumstance (real-time capability and systems integration); (2) accurately analysing information, based on reliable data (Big Data Analytics for known hypotheses and Machine Learning/Artificial Intelligence for new contexts); (3) making decisions with precision and speed (decision support systems / automated decision making); and finally (4) acting seamlessly and optimally (vertical / horizontal process / systems integration and CPS).

In addition, Rüßmann et al. [10] and Bloem et al. [12] highlight gains in productivity, cost decrease, inventory reduction and increased customisation due to increased agility and intelligence generated thanks to improvement in Machine-to-Machine (M2M) communication, Human-Machine Interaction (HMI) and predictive maintenance.

Specifically in the Aerospace, Defense and Security sector, the expected cost reduction to be obtained from I4.0 deployment is 3.7% per year. This might explain why executives of this market intend to invest 5% of its annual revenue in digital operations solutions, over the next five years. [13]

2.1 Industry 4.0 in aeronautics

The aeronautical industry is characterised by low volumes of products, with a restricted and specific market and a complex manufacturing process. [14]

The latter can be broadly divided into five distinct stages: (1) fabrication of primary parts, through mechanical forming, machining and composite lamination processes. (2) Fabrication of elementary subsets of wiring, pipes and welded parts. (3) Assembly of parts by the grouping of primary parts, producing panels, fairing, stringers, ribs etc. (4) Assembly of segments by joining the fabricated / assembled parts (from the initial three phases), resulting in the half-wings, empennages and fuselage sections. (5) Last stage, where the aircraft parts are painted, pre-equipped / equipped with aircraft systems (Air-Management-Systems, hydraulics, landing gear, flight controls etc.), finishing when the aircraft passes to the flight preparation and delivery to the client. [15]

The complex manufacturing in addition to the low production cadence results in weak automation and significant labour along the processes, achieving up to 80% of total costs. For this reason, the key contribution of I4.0 in the aeronautical context is supporting human workers with digital enhanced solutions to improve company results. [8] [14]

In this work, a focus was given to the final assembly area, which is located in the last stages (after structural assembly and before flight preparation), being characterised by activities of high diversity, complexity, customisation and with the use of specific technologies, showing a high potential for progress with the Advanced Manufacturing.

In the literature, several authors [9] [10] [15] [16] have listed nine key technologies that are enabling the introduction of I4.0 (even though there are others being explored). Many of them are already being used in some contexts, but integrated, they will be able to transform aeronautics around the world.

One of those technologies, capable of significantly transforming aircraft manufacturing, is the Internet of Things (IoT). Using an IoT platform to connect workers and other I4.0 solutions (like Augmented Reality) could make processes much quicker and reliable, by allowing critical information to flow seamlessly. For instance, an employee could use a tablet or smart glasses to scan a specific part of the airplane and automatically receive (from a local Cloud) the precise data related to installation procedure, needed parts and other technical information. Still, this information could be sent to a robotic tool, which could perform the task. [8] [14] [17]

On another situation, IoT could be used to locate tools on the shop floor and track its usage. This could validate if tools are only being handled by authorised and trained people (increasing people safety and product security) and monitor usage data, based on assembly specifications (reducing mistakes, avoiding the burden of regulation compliance and reducing costs with quality inspections). [8] [17]

Further contributions of I4.0 can come from Horizontal and Vertical System Integration (HVSI), where Information Technology/Operational Technology (IT/OT) infrastructure will have a much stronger integration among internal departments and beyond the company boundaries (suppliers, clients, third parties etc.). This is specifically beneficial for the aeronautical sector, since its supply chain is based on a complex network (with a large number of actors, specific

technologies providers, traffic of ultra-confidential data etc.). [8] [14]

So that the HVSI is enabled, Cloud solutions (that are already present in several enterprises) will be significantly improved to allow a greater data traffic (inside and outside the company), with a minimal reaction time and reliable transmission. [10]

Due to its low volume of production and high number of different tasks, the aircraft industry has been limited in the application of conventional automation. With substantial enhancements in terms of safety, interaction and autonomy; Advanced Robotics intends to enable a collaborative environment where humans and robots can work together, sharing tasks in an open and fenceless shop floor. This new advent allows processes do be more flexible and optimised, by delegating potentially risky, repetitive, of difficult access, not ergonomic or precise activities to robots; allowing human operators to work on those with highly added value or that would be too complex/expensive to be automated. [18] [19]

Considering 3-D simulations, the design of materials, products and processes are already a common reality in the aeronautical companies. However, new technologies (called Digital Twins) will be able to simulate plant operations in real-time; monitoring machines, products and even humans. This will allow more precise analysis, tests of possible scenarios and processes optimisation in digital world before physical implementation (reducing mistakes, identifying risks, driving down machine setup times and reducing costs). [10] [17]

With the implantation of sensors – in manufacturing systems, enterprise- and customer-management systems –, there will be an unavoidable data explosion generated from many different sources. In this context, Big Data Analytics solutions will enable aircraft companies to properly treat and use data to have precise information about key aspects, relate different variables, identify patterns and predict likely future scenarios. This new acquisition will support timely decision-making, enable a self-learning and autonomous value chain, guarantee synergy along the supply chain, allow proactive action, increase productivity and avoid process interruptions. [6] [10] [17]

Because of the increased connectivity and use of standard communications protocols, brought by the digital revolution, the application of advanced Cybersecurity will be mandatory in the aeronautical context. Enterprises will have to place technologies to ensure data security, sophisticated access management (of machines and humans), reliable communication and protection of the IT/OT infrastructure. [10]

Another technology that is already a reality in the aeronautical environment is the additive manufacturing, such as 3-D Printing (3DP). It has been used in the fabrication of prototypes, tooling and even parts to compose the airplane. Companies are now investing in Research & Development to assure reliability on 3DP products, so that they can benefit from the production of complex shapes parts, in a short time, with different types of materials (sometimes combining

them), reducing the part weight and waste of raw material. [6] [10] [17]

Finally, a popular solution that is already being applied to industrial applications (and has been improved in terms of precision, reliability and scanning) is the Augmented Reality. This versatile tool may help aircraft assembly lines in: providing work instructions in mobile devices (as previously mentioned), indicating parts in a warehouse, provide real-time information to support decision-making and train employees. [8] [10] [17]

From the description above, it might be tempting for the aircraft companies to apply some of those solutions to their organisations. However, pushing new technologies does not necessarily bring results and can even be risky (in terms of financial losses, under-exploitations of solutions, damages to IT/OT infrastructure, conflicts of previous and new IT/OT, difficulty in staff adaptation etc.). [20] Considering that, it is essential that companies have a clear vision of how I4.0 is projected to its specific context, understand its actual reality, align I4.0 deployment goals with the corporate strategy, identify the solutions that best fit their needs, plan realistic projects and manage its progress to achieve the desired results. With this mission, the present work intends to propose a complete method to guide companies in this endeavor.

3 Analysis of the literature MII4.0

Initially, in order to identify opportunities in the elaboration of a new method, an exploratory research on MII4.0 proposed by other authors was structured.

There were searched publications in the databases: *Scielo*, *IEEE Xplore*, *Conference on Manufacturing Systems*, *Internet of Things World Forum*, *ABEPRO*, *Portal Capes*, *Simpósio de Eng. de Produção* and *Google Scholar*. The period considered was from 2014 to 06/2018. The used keyword were (in English and Portuguese): *Method*, *Model*, *Guideline*, *Roadmap*, *Deployment*, *Maturity Model* and *Steps* combined with *Industry 4.0*, *Advanced Manufacturing* and *Digital Manufacturing*. As a result, nine publications were found (ordered by date and alphabetically): Cabral et al. [21]; Cordeiro et al. [22]; Pessl et al. [4]; Schuh et al. (ACATECH) [1]; Erol et al. [23]; Ganzarain e Errasti [24]; Geissbauer et al. (PwC) [13]; Bechtold et al. (Capgemini) [25]; Rockwell Automation [2].

Subsequently, these methods were studied, analysed comparatively and evaluated based on the good practices identified in this work (understood as those not common to all methods, but that make the deployment process more grounded in data, less subjective and more customised to the reality of each company).

Still within those good practices, there were highlighted some best practices that presented prominent points within the analysed criteria. Good and best practices are described as follows:

Projection of Vision 4.0: this criterion intended to verify if the methods presented a phase to create a Vision 4.0 in order to ensure the leveling of knowledge (among all those

involved in the implementation process) about the concept of I4.0. Pessl et al. [4] define a detailed step for understanding the concepts of I4.0, while Erol et al. [23], Ganzarain and Errasti [24] and Bechtold et al. [25] go a step further and emphasise the need to bring this vision into the specific context of the organisation (Customised Vision) – although they do not systematise a procedure for that.

Self-assessment: the self-assessment is an important step where the organisation identifies its current maturity position and stablishes a realistic basis for its decision-making process. Although this step is present in almost all methods, Schuh et al. [1] and Pessl et al. [4] stand out for establishing a logical procedure, where the current status is calculated from a multiple choice questionnaire.

Strategy definition: a phase to define the strategy of the implementation project, ensuring that the project goal is aligned with the objectives of the company. Another emphasis is given to Schuh et al. [1] and Pessl et al. [4], since those authors present a procedure (based on the calculation of current maturity stage) where the goals are set and represented by numerical levels to be reached with the evolution of the current levels.

Definition of 4.0 solutions: there is a need for a step that systematises the selection of 4.0 solutions to achieve the defined goals. Ganzarain and Errasti [24] were the only authors to mention this step; however, none of the methods has detailed this step, which is why no best practice was identified.

Evaluation of results: to collect and evaluate the achieved results, in order to evidence gains and / or losses of the project. Schuh et al. [1] emphasise in their work the need for the application of quantitative indicators for the monitoring and evaluation of the results obtained with the implementation of I4.0.

Gradual deployment: this aspect verifies if the method allows the gradual implementation of Industry 4.0. Although most methods allow this type of implementation, the method of Bechtold et al. [25] stands out for having a logical and clear cyclical structure, which visually highlight the need for continuous and incremental I4.0 deployment.

Real application: to check if the methods have already been applied in a real environment, in order to validate it. Few authors presented this differential and specially Schuh et al. [1] and Pessl et al. [4] have evidenced the applicability of the method with examples from Harting AG & Co. and an unidentified Austrian company, respectively.

Levels of maturity: to verify if the authors propose a maturity scale (in order to allow the diagnosis of the current situation and the definition of the goal) and how many levels of gradation it presents. In this regard, Schuh et al. [1] presented the most complete scale (6 degrees), the only one containing the predecessor stages of Industry 4.0 (computerisation and connectivity).

Origin: to verify if the methods were originated from an Academic (A), Corporate (C) or both sources. While an academic source is distinguished for the scientific

background, the second one values a more practical and applicable approach to the industrial context. Schuh et al. [1] and Cabral et al. [21] presented the best practices for having their MII4.0 elaborated from the collaboration of academic and corporate sources.

In this sense, fig. 1 summarises the comparative analysis, indicating the works that meet the good practices and those with the best practices (highlighted in blue).

	Cabral et al. [21]	Cordeiro et al. [22]	Pessl et al. [4]	Schuh et al. [1]	Erol et al. [23]	Ganzarain e Errasti [24]	Geissbauer et al. [13]	Bechtold et al. [25]	Rockwell Automation [2]
Projection of vision 4.0	-	-	Yes	-	Yes	Yes	-	Yes	-
Self-assessment	Yes	Yes	Yes	Yes	-	-	-	Yes	Yes
Strategy definition	-	Yes	Yes	Yes	Yes	-	Yes	Yes	-
Definition of 4.0 solutions	-	-	-	-	-	Yes	-	-	-
Evaluation of results	-	-	-	Yes	-	-	-	-	-
Gradual deployment	Yes	Yes	Yes	Yes	-	-	Yes	Yes	-
Real application	-	-	Yes	Yes	-	-	-	-	Yes
Levels of maturity	3	4	4	6	-	4	4	-	-
Origin	A/C	A	A	A/C	A	A	C	C	C

Figure 1: Comparative analysis of the MII4.0 of the literature

Through this analysis, it was concluded that none of the methods meet all the identified best practices, evidencing an opportunity to create a more complete method. In addition, the combination of these best practices allowed the creation of an ideal model that would serve as a reference for the proposal of this work, which is presented in the following section.

4 Proposal of a MII4.0 and representation in the Aircraft Final Assembly

Based on the previous analysis and the information collected from the literature, the 8-step method was proposed and represented as in fig. 2.

As shown, the method was strategically built over the PDCA structure, due to the three motivations: (1) it is widely known in the market, generating greater receptivity among users; (2) it summarises the basic concepts of management, facilitating the understanding of the sequential logic of the method; and (3) it has a cyclical representation, inducing the culture that the implementation of Industry 4.0 should occur progressively since, as observed by Schuh et al. [1] and Chukwuekwe et al. [6], a systemic and robust transition must be made in a gradual way, respecting the organisational structure.
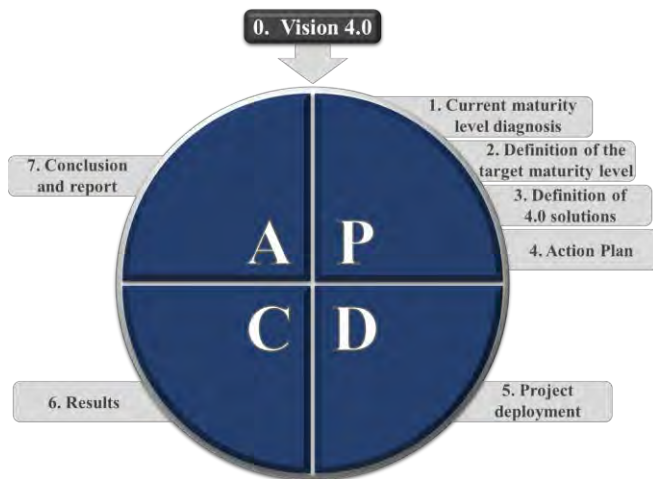
Figure 2: Proposed MII4.0

In this way, each of the stages is described in the following topics and then exemplified by a representation performed in the Aircraft Final Assembly scenario. It is important to emphasise that this representation was made based on the literature study and through interviews conducted with 22 professionals from this area. In addition, the data used is from a fictional example, without referring to any specific company, due to confidentiality issues.

4.1 Vision 4.0

Considered the most important step of the method, it represents the formalisation of the Industry 4.0 concept, specifically translated to the context where the method will be deployed - which is achieved through a customised maturity scale.

The I4.0 maturity scale can be understood as a reference metric that guides the company in its classification regarding the different levels of deepening in this universe. [4]

For the construction of the customised maturity scale, the present work developed an innovative procedure based on three dimensions: Internal, Structural and Conceptual 4.0 - according to fig. 3. They are explained and exemplified in the Aircraft Final Assembly context, as it follows.
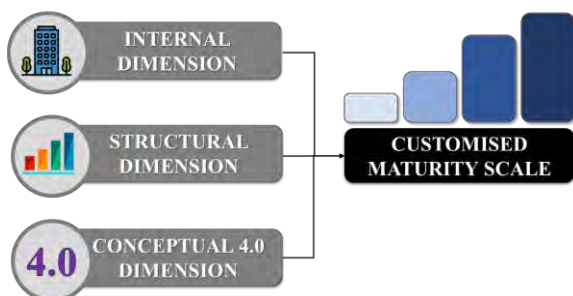

Figure 3: Dimensions used in the construction of the Customised Maturity Scale

The Internal Dimension refers to the nature of activities that are present in the company/areas and is responsible for the characterisation of the maturity scale to the specific context of the company.

The nature of activities list is obtained from the macro process mapping, exemplified in four steps in fig. 4, for the Aircraft Final Assembly context.
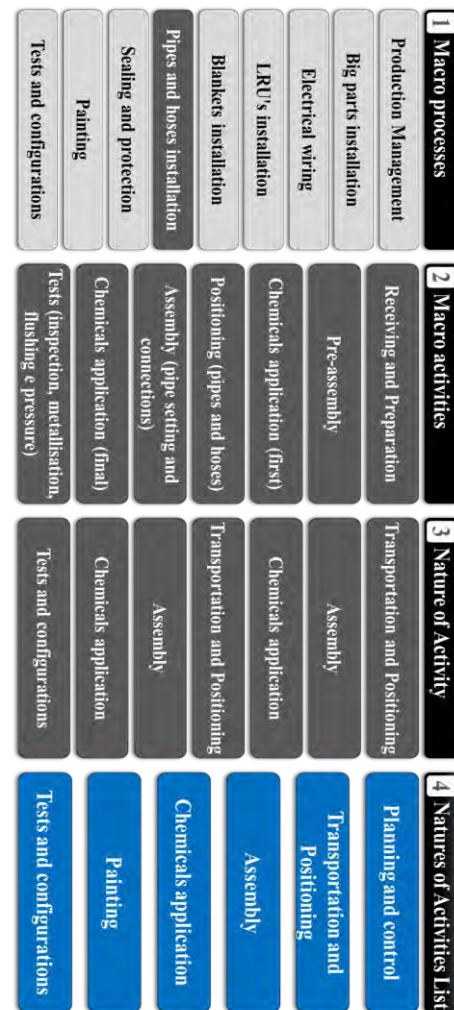

Figure 4: Procedure for the identification of the Activity Nature List (Internal Dimension)

(1) Macro processes are mapped – in fig. 4, nine macro processes were mapped in the Aircraft Final Assembly.

(2) Macro activities from each macro process must be identified – in fig. 4, the "Pipes and hoses installation" macro process was explored as an example.

(3) Macro activities with similar characteristics are grouped – forming "Nature of Activity" clusters.

(4) All the different "Natures of Activities" clusters are identified and listed – for the Aircraft Final Assembly context, the six clusters indicated in blue were identified. These clusters represent the internal dimension and shall be projected on the customised maturity scale.

Note: the full macro process mapping and the Internal Dimension definition for a general Aircraft Final Assembly area is presented in the **Attachment 1.**

The Structural Dimension corresponds to the structure of the maturity scale itself (number of levels, nomenclatures and main characteristics of each level).

It consists of a generic maturity scale and can be obtained from the literature (developed by other authors) or conceived by specialists of the company (able to build one).

It is suggested in this work the use of the structure of the ACATECH maturity scale [1] – fig. 5, since it summarises concepts of I4.0 that are frequently cited in the literature, has a logical transition between levels and is the most complete amongst the authors studied in this work (6 maturity stages).
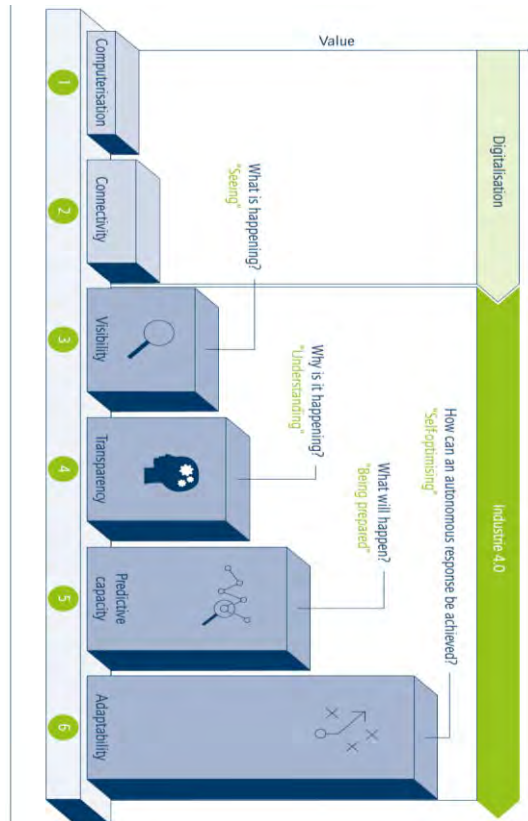

Figure 5: ACATECH maturity scale [1]

In this scale, the authors classify the first two levels (Digitalisation) as a prerequisite for entry into Industry 4.0. Thus, level (1) *Computerisation* represents the stage where the company starts to use automated and disconnected equipment. At level (2) *Connectivity*, automated equipment is connected, with automatic data and information traffic. In level (3) *Visibility*, in addition to previous levels, the company has monitoring sensors throughout the process, capable of providing real-time data on production status. In level (4) *Transparency*, the organisation acquires intelligent manufacturing systems, able to relate different data and identify the probable causes for different behaviours of the process. In level (5) *Predictive Capacity*, IT systems are able, based on current and historical data, to predict likely future scenarios, assisting in the preventive decision-making. Finally, in level (6) Adaptability, the manufacturing systems acquire the intelligence to make decisions and act autonomously in predefined situations to optimise the production process, without human interference.

Finally, ***Conceptual 4.0 Dimension*** consists of the Industry 4.0 technologies and systems that are applicable to the processes of the organisation.

For this purpose, the company should work on researches, benchmarking, contact with commercial representatives etc. Examples of 4.0 terms: Big-Data, Additive Manufacturing, Cloud Computing, Internet of Things, Augmented Reality etc. (More examples are given in Section 2.1).

With the three dimensions identified, the natures of activities (*Internal Dimension*) have to be individually described according to the characteristics it should present in each of the different levels of maturity (*Structural Dimension*). Furthermore, the I4.0 terms (*Conceptual 4.0 Dimension*) should complement the descriptions, detailing the characteristics of the maturity stage to which they belong.

Figure 6 represents one of the lines of the Aircraft Final Assembly maturity scale (related to the Transport and Positioning cluster), in order to exemplify it.

Note: the full Aircraft Final Assembly maturity scale (created based on [1] [6] [8] [10] [14] [17] [18] [19]) is presented in the **Attachment 2.**

4.2 Current maturity level diagnosis

Based on the customised maturity scale, the company should classify each of its macro activities (already mapped in the first stage) in one of the maturity levels (1-6), according to the characteristics of the cluster to which they belong.

At the end, the grades are summarised in a table, indicating the number (No) of macro activities classified on each level within each Nature of Activity cluster. Thus, the resulted Maturity Level (ML) of each cluster (column) is given by the weighted average of the grades as indicated in eq. (1) – adapted from [1] [4]. A theoretical example is given in tab. 1.

$$ML_{cluster} = \frac{\begin{array}{c}(No\ of\ macroactivities\ level\ 1) \times 1 + \\ (No\ of\ macroactivities\ level\ 2) \times 2 + \\ (...) + \\ (No\ of\ macroactivities\ level\ 6) \times 6\end{array}}{Total\ No\ of\ macroactivities} \quad (1)$$

Table 1: Current Maturity Level diagnosis table.

Classification	Planning and control	Transportation and Positioning	Assembly	Chemicals application	Painting	Tests and configurations	**Final Assembly (total)**
0	0	5	0	0	0	0	
Level 1	1	1	1	1	0	1	
Level 2	2	8	4	2	1	2	
Level 3	2	1	3	1	1	5	
Level 4	0	0	0	0	0	0	
Level 5	0	0	0	0	0	0	
Level 6	0	0	0	0	0	0	
Current Maturity Level	**2.20**	**1.33**	**2.25**	**2.00**	**2.50**	**2.50**	**2.13**

As shown in tab. 1, from fifteen macro activities of the *Transportation and Positioning* cluster, five were below level

1 (then, they were classified at level 0), one was at level 1, eight were classified at level 2 and one was classified at level 3. From the eq. (1), the resulted current ML for this cluster was 1.33. This procedure was made for all the clusters, resulting in the global current ML (2.13), which is given by the arithmetical average of the clusters.

Figure 6: Maturity levels for the Transportation and Positioning cluster

Note: in this example, it was considered that all clusters had the same degree of importance, so a simple arithmetical average was used for the global ML. If, in another context, a company consider the clusters to have different degrees of importance, a weighted arithmetic average can be used.

4.3 Definition of the target maturity level

After calculating its current ML, the company is able to set a goal of progress on the Industry 4.0 path. This definition should involve the managers of the related areas and take into account: the area / company strategy, the current ML and the customised maturity scale (which provides future state projections as a benchmark). [1] [4]

In defining its goal, the company must formalise it by means of a table similar to the Current Maturity Level diagnosis table. Thus, in the example of tab. 2, the company chose to increase the ML of the *Transportation and Positioning* cluster, migrating four positioning macro activities from level 0 to level 3. As a result, the Target ML of this cluster was set to 2.13 (being elevated from 1.33) and for the Final Assembly area (as a whole), it was set to 2.26 (moving from 2.13).

Table 2: Target Maturity Level definition table.

Classification	Planning and control	Transportation and Positioning	Assembly	Chemicals application	Painting	Tests and configurations	Final Assembly (total)
0	0	1	0	0	0	0	
Level 1	1	1	1	1	0	1	
Level 2	2	8	4	2	1	2	
Level 3	2	5	3	1	1	5	
Level 4	0	0	0	0	0	0	
Level 5	0	0	0	0	0	0	
Level 6	0	0	0	0	0	0	
Target Maturity Level	2.20	2.13	2.25	2.00	2.50	2.50	2.26

4.4 Definition of 4.0 solutions

From the goal definition, the company must then identify the 4.0 solutions that will allow its achievement.

This procedure occurs in three stages: (1) initially, the company conducts a market research for possible solutions; (2) these solutions are filtered through mandatory requirements (local legislation, mandatory process requirements, internal standards etc.); (3) Finally, candidate solutions that meet the mandatory requirements should be evaluated through a Decision Making Method (DMM).

For the example used in this work, the progress of the ML of the positioning macro activities was simulated, using collaborative robotics solutions (ensemble with the systems required for their integration with the company IT/OT infrastructure, allowing them to be digitally visible and interact with other systems/humans in the company).

Thus, as represented in fig. 7, an initial research resulted in 15 available solutions, which were filtered by requirements of payload, system weight and standard Protocol. As a result, five candidate solutions were selected to be evaluated through a Decision Making Method (DMM). Although the proposed MII4.0 does not delineate which one shall be used, it is suggested the use of AHP, due to its high versatility, the

DOI
10.3384/ecp19162023   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **205**

possibility of using qualitative / quantitative variables and ease of application.
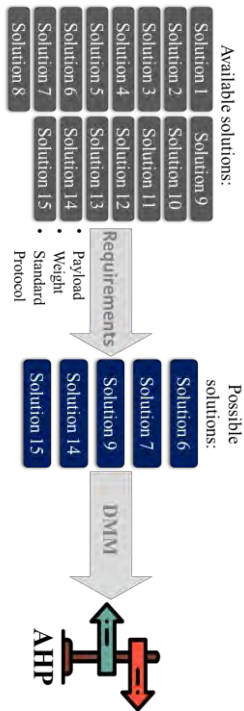

Figure 7: Procedure for the definition of 4.0 solutions

At the end of the decision-making analysis, the company is able to compare the evaluation grade and the cost of each solution, so that the most suitable one(s) can be selected for implementation. Note: the simulated solutions had their details hidden for brand exposure reasons.

4.5   Action plan

Similar to the same name step of the traditional PDCA tool, in this stage the actions required for the implementation of the selected solution(s) are identified and formalised.

It is important to notice that before the solution implementation itself, several predecessor actions might be required in order to enable the proper integration with the company IT/OT infrastructure.

According to Andrade [26], this step is the product resulted from all the activities contained in the P phase (Plan) and must be detailed in order to achieve the goal. This detailing can be done through the 5W1H: What, When, Who, Where, Why and How.

Note: because this work describes a representation (and not a real application), this and the following steps do not have a content for the Aircraft Final Assembly example.

4.6   Results

At the end of the project deployment, the company must collect and analyse the results in terms of two aspects:

Project – verify if the deployment was within the defined scope, finished on time and did not exceed the budget.

Goal – confirm if the target maturity level was achieved.

This information should then be registered and serve as input for the conclusion of the project, in the last step.

4.7   Conclusion and report

Finally, the achievement of all expected results in terms of project and goal, indicates the success of the project. In this case, the company should: (1) disseminate the results obtained, (2) record the lessons learnt, (3) formalise the new process and (4) monitor the new process indicators to ensure robustness.

The non-achievement of at least one of the expected results represents a failure (partial or complete) of the project. In this case, the responsible team should: (1) carry out a cause analysis, (2) decide on correction (through a new action plan) or project disruption, (3) execute the decision and (4) register the lessons learnt, so as not to make the same mistakes on future projects.

Finally, the conclusion indicates a new iteration of the cycle, following the philosophy of continuous improvement.

Note: the full method summary flowchart is detailed in the **Attachment 3**.

5   Evaluation of the proposed MII4.0

The proposed method was evaluated in two ways: (A) comparative evaluation and (B) value perception research.

(A) By comparative evaluation, the proposed MII4.0 was compared with the work of other authors through the nine best practices identified in the literature, as represented in fig. 8.

	Proposed MII4.0	Cabral et al. [21]	Cordeiro et al. [22]	Pessl et al. [4]	Schuh et al. [1]	Erol et al. [23]	Ganzarain e Errasti [24]	Geissbauer et al. [13]	Bechtold et al. [25]	Rockwell Automation [2]
Projection of vision 4.0	Yes	-	-	Yes	-	Yes	Yes	-	Yes	-
Self-assessment	Yes	Yes	Yes	Yes	Yes	-	-	-	Yes	Yes
Strategy definition	Yes	-	Yes	Yes	Yes	Yes	-	Yes	Yes	-
Definition of 4.0 solutions	Yes	-	-	-	-	-	Yes	-	-	-
Evaluation of results	Yes	-	-	-	Yes	-	-	-	-	-
Gradual deployment	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes	-
Real application	-	-	-	Yes	Yes	-	-	-	-	Yes
Levels of maturity	6	3	4	4	6	-	4	4	-	-
Origin	A/C	A/C	A	A	A/C	A	A	C	C	C

Figure 8: Comparative evaluation of the proposed MII4.0

It was verified that the proposed MII4.0 has 8 of the 9 best practices analysed, being the most complete one:

- It has a Vision 4.0 stage with an innovative procedure for the construction of a customised maturity scale;

- It systematises a procedure to diagnose the current maturity level of companies – adapted from [1] and [4];
- It has a stage to define the target maturity level, based on the company's internal strategy, current maturity level and the built customised maturity scale – adapted from [1] and [4];
- It introduces (in an innovative way among the existing MII4.0) a systematised procedure to define 4.0 solutions;
- It presents a stage for the collection and analysis of the project results;
- It has a cyclical structure to induce the gradual implementation of I4.0 – based on the PDCA tool;
- It has its deployment based on a six-levels maturity scale (comprising the two prerequisite steps for the I4.0 deployment – presented by ACATECH [1])
- It was built from an academic work and corporate contributions (feedback from professionals from industry, after the method was simulated and presented).
- The negative aspect raised by this analysis concerns the application of the method in a real environment. If this criterion had been met, a more concrete evaluation about the method quality would be possible. However, it was not possible to do so within the term of this work, due to corporate unavailability. Nevertheless, the present work is an initial study towards a field that is yet little explored and, as such, it is understood that the knowledge generated here may contribute to future implementations or even later works.

(B) For the perception survey, thirty-three interviewees were selected based on their experience and expertise in innovation projects in the aeronautical industry – considering they had to be qualified enough to evaluate the applicability of the proposed method.

From this group, four people worked on *Research and Development of New Technologies*, being responsible for searching market-incoming technologies (from low Technology-Readiness-Levels, TRL) and developing them, until they are ready to be applied in the industrial environment. Those professionals had an average of 18.6 years of experience in the aircraft industry.

Three other respondents were specialists in *Manufacturing Digital Solutions*, with an average of 27.7 years of experience in the implementation of CAD (Computer Aided Design), CAE (Computer Aided Engineering), CAM (Computer Aided Manufacturing), MES (Manufacturing Execution System), PLM (Product Lifecycle Management), SPDM (Simulation Process & Data Management) and systems integration in the aviation sector.

Moreover, three participants worked on the *Manufacturing Strategy* department, responsible for coordinating processes, methods, indicators and resources, ensuring the competitiveness of the manufacturing sector. This group counted with an average of 13.7 years of experience in the aeronautics field.

The greater group (twenty respondents with an average of 12.6 years of experience) were from the Final Assembly area, working directly with the application of manufacturing technologies in the productive process.

Finally, the last three interviewees came from the academic area. Two of them were PhD's in *Aeronautical and Mechanical Engineering* (working as professors and researchers in the manufacturing field) and one was MSc. in *Aeronautical and Mechanical Engineering* (working as a researcher in the manufacturing field).

In this sense, the proposed MII4.0 was presented to all participants, who evaluated it in six aspects:

- If it was relevant (for the industrial and literature field), fig. 9 – item evaluated by unanimity (100%) as very relevant or relevant, evidencing the importance of the theme explored in this work.
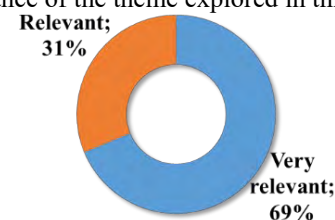

Figure 9: perception survey – relevance of the method

- If it was visually friendly, fig. 10 – receiving a very positive or positive perception by 97% of the respondents (3% in the neutral zone), which is probably due to its structuring according to the PDCA tool, which generates greater familiarity, understanding and acceptance on the part of users.
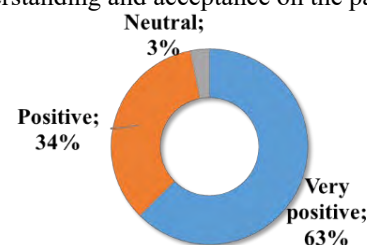

Figure 10: perception survey – visual friendliness

- If it had a logical structure, fig. 11 – also understood by unanimity (100%) as very positive or positive, indicating that the interviewees were able to assimilate the logic of the stages of the customised model and the way in which it leads to the proposed objective. Again, the use of the PDCA method may have contributed in this aspect, by guiding the steps according to the fundamentals of management.
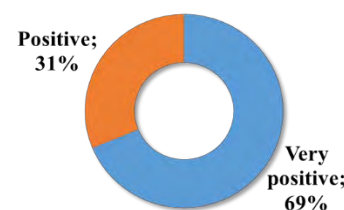

Figure 11: perception survey – logic of the proposed structure

- If it was complete (i.e., it did not lack any relevant step for a project of this nature), fig. 12 – in this question, all the interviewees classified the method as complete or had the main steps, without pointing out any missing steps.
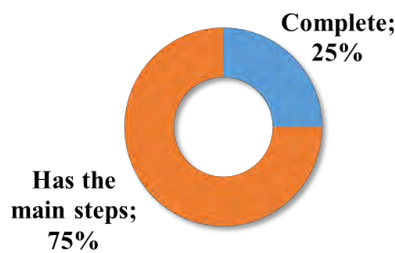

Figure 12: perception survey – completeness of the proposed method

- If it was practical, fig. 13 – 91% of the respondents considered it to be practical or very practical, 6% neutral and 3% complex. The criticisms were based on the diagnosis (which was previously done by questionnaire as proposed by [1] and [4]). As a solution, it was simplified to a direct classification of the mapped macro activities by nature of activity clusters in the maturity scale, without the need for questionnaires, allowing a complete, practical and yet representative evaluation.
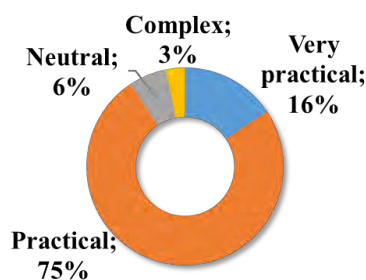

Figure 13: perception survey – practicality of the proposed MII4.0

- If it was versatile (that is, applicable to different industrial contexts), fig. 14 – in the last topic, 97% of respondents pointed the MII4.0 as very versatile or versatile and 3% as neutral (due to doubt of the robustness of the proposed method, that should be applied on a real context to be confirmed). This evaluation is very favourable to the elaborated MII4.0, indicating its potential of application in companies and areas of different sectors, even for very specific business as the aircraft industry.
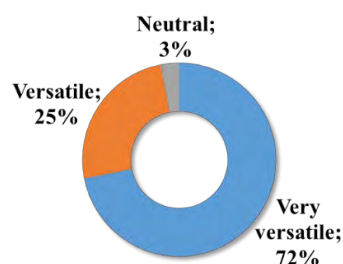

Figure 14: perception survey – versatility of the method

6 Results and conclusions

From the previous study, it is concluded that the main objective of this work was reached, with the proposal of a method based on the best practices identified in the literature and based on a customised I4.0 maturity scale, so that it could be applied to specific contexts (as the aeronautical industry).

In addition, the work had the following results:

- Presented a comparative analysis of nine methods in the literature, through nine good practices identified;
- Developed a more complete method, among those analysed - presenting 8 of the 9 best practices identified;
- Created an innovative procedure for creating the customised maturity scale;
- Carried out the mapping of macro processes and macro activities of a general Aircraft Final Assembly context;
- Elaborated a customised maturity scale for the Aircraft Final Assembly context.

Finally, the evaluations by comparative analysis and perception survey presented very favourable results to the method, indicating its applicability potential and its value of contribution in the industrial and academic field.

References

[1] G. Schuh, R. Anderl, J. Gausemeier, M. T. Hompe, W. Wahlster, Industrie 4.0 Maturity Index. Managing the Digital Transformation of Companies, Munich: Herbert Utz. Acatech study, 2017.

[2] Rockwell Automation, "The Connected Enterprise Maturity Model", 2014. [Online]. Available: http://literature.rockwellautomation.com/idc/groups/ literature/documents/wp/cie-wp002_-en-p.pdf. [Accessed 10-05-2018].

[3] M. E. Porter, J. E. Heppelmann, "How Smart, Connected Products are Transforming Competition", *Harvard Business Review*, 2015.

[4] E. Pessl, S. R. Sorko, B. Mayer, "Roadmap Industry 4.0 – Implementation Guideline for Enterprises. International Journal of Science", *Technology and Society*, vol. 5, nr. 6, pp. 193, 2017.

[5] M. Damm, Fraunhofer, "Industrie 4.0 – an overview", 2017. [Online] Available: https://www.ipa.go.jp/sec/old/users/seminar/seminar _yokohama_20170227-03.pdf. [Accessed 18-02-2019]

[6] D. O. Chukwuekwe, P. Schjolberg, H. Rodseth, A. Stuber, "Reliable, Robust and Resilient Systems: Towards Development of a Predictive Maintenance Concept within the Industry 4.0 Environment", *EFNMS Euro Maintenance Conference*, 2016.

[7] G. F. Barbosa, "Aplicação da metodologia DFMA - Design for Manufacturing and Assembly - no projeto e fabricação de aeronaves", São Carlos, 165 p., Masters dissertation, Universidade de São Paulo, 2007.

[8] European Comission, "Industry 4.0 in Aeronautics: IoT applications", 2017. [Online] Available: https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/DTM_Aeronautics%20-%20IoT%20Applications%20v1.pdf. [Accessed: 14-09-2019]

[9] I. Colotla, I.; A. Fæste, A. Heidemann, A.Winther, P. H. Andersen; T. Duvold, M. Hansen,Winning the Industry 4.0 race. How ready are Danish manufacturers?, Copenhagen: Boston Consulting Group & Innovationsfonden, 2016.

[10] M. Rüßmann; M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, M. Harnisch, " Industry 4.0: the future of productivity and growth in manufacturing industries", *Boston Consulting Group*, vol. 9, nr. 1, pp. 54-89, Apr. 2015.

[11] R. C. Schläpfer, M. Koch, P. Merkofer, Industry 4.0: challenges and solutions for the digital transformation and use of exponential technologies, Zurich: Deloitte, 2015.

[12] J. Bloem, M. V. Doorn, S. Duivestein, D. Excoffier, R. Maas, E. V. Ommeren, The fourth industrial revolution - things to tighten the link between IT and OT, Sogeti VINT, 2014.

[13] R. Geissbauer, J. Vedso, S. Schrauf, "Industry 4.0: Building the digital enterprise", 2016. [Online]. Available: https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2016.pdf. [Accessed 16-05-2018].

[14] I. Guyon, R. Amine, S. Tamayo, F. Fontane, "Analysis of the opportunities of industry 4.0 in the aeronautical sector". *10th International Multi-Conference on Complexity, Informatics and Cybernetics – IMCIC*, 2019.

[15] G. F. Barbosa, R. V. Aroca, "Advances of Industry 4.0 concepts on aircraft construction: An Overview of Trends". *Journal of Steel Structures & Construction*, DOI: 10.4172/2472-0437.1000125, vol. 3, nr. 1, 2017.

[16] M. A. K. Bahrin, M. F. Othman, N. N. Azli, M. F. Talib, "Industry 4.0: a review on industrial automation and robotic", *Jurnal Teknologi*, vol. 78, nr. 6-13, pp. 137-143, Jan. 2016.

[17] Deloitte, "Exponential technologies in manufacturing – Transforming the future of manufacturing through technology, talent, and the innovation ecosystem", 2018. [Online] Available: https://www.compete.org/storage/reports/exponential_technologies_2018_study.pdf. [Accessed 11-09-2019]

[18] I. Maurtua, A. Ibarguren, J. Kildal, L. Susperregi, B. Sierra, "Human–robot collaboration in industrial applications: Safety, interaction and trust", *International Journal of Advanced Robotic Systems*, vol. 14, nr. 4, pp. 1-10, Dec. 2017.

[19] A. L. E. S. Vysocky, P. E. T. R. Novak, "Human-Robot Collaboration in Industry", *MM Science Journal*, vol. 9, nr. 2, pp. 903-906, Jun. 2016.

[20] A. Moeuf, S. Lamouri, R. Pellerin, R. Eburdy, S. Tamayo, "Industry 4.0 and the SME: a technology-focused review of the empirical literature", *7th IESM Conference*, 2017.

[21] A. C. D. Cabral, A. N. R. Costa, C. J. Gamarra, "Micro, Pequenas e Médias Empresas a Caminho da Manufatura Avançada: uma Proposta de Roteiro", *XXXVII ENEGEP*, Joinville, 2017.

[22] G. A. Cordeiro, R. E. C. Ordonez, R. Ferro, P. B. Nicolela, B. C. Stella, "Etapas para Implantação da Indústria 4.0: Uma Visão Sob Aspectos Estratégicos e Operacionais", *Encontro Nacional de Engenharia de Produção,* Joinville, 2017.

[23] S. Erol, A. Schumacher, W. Sihn, "Strategic guidance towards Industry 4.0–a three-stage process model", *International Conference on Competitive Manufacturing*, pp. 495-501, 2016.

[24] J. Ganzarain, N. Errasti, "Three stage maturity model in SME's toward industry 4.0", *Journal of Industrial Engineering and Management (JIEM)*, vol. 9, nr. 5, pp. 1119-1128, 2016.

[25] J. Bechtold, C. Lauenstein, A. Kern, L. Bernhofer, "Industry 4.0 - The Capgemini Consulting View", *Capgemnini Consulting*, vol. 31, Pags. 32-33, 2014.

[26] F. F. de Andrade, "O método de melhorias PDCA", São Paulo, 169 p., Masters dissertation, Universidade de São Paulo, 2003.

Additive Manufacture at Industrial Aeronautical and Defence Area: How to control the production of a good and some questions related copyright intellectual property performance at an embargo scenario

Juan Manuel Iglesias Pascual* and José Valentin Iglesias Pascual**

*E-mail: unifjpascual@fei.edu.br jose.pascual@metodista.br
*R&D, SaveInProcess SBC SP/Brazil **Professor Doctor UMES, SBC SP/Brazil

Abstract

One question raised in this exploratory work with focus on how assure that the Additive Manufacture, AM, let's say 3D printing, produced part fits all the requirements, to fulfil the demands of warranty and performance to assure the proper operation of the system in that this part is included and how to perform the traceability of them. Some other important points in this discussion, regarding the digital files and the parts produced from those files or by files produced from scanned 3D parts, include among others: Copyright and Patent issues, Licensing private or Public (GPL), Creative Commons (CC). Other important topics are about customer adapted production, support licenses i.e. the AM techniques recommend by the supplier of the system to produce a specific part to fulfil the requirements of the system. Traceability of this production and Digital Right Management (DRM) since suppliers are demanded regarding the International Traffic in Arms Regulations (ITAR) and Export Administration Regulations (EAR). Embargo statements frequently are needed for humanitarian and geopolitical reasons and ensure the fulfilling dimension and scope and demands of an embargo sanction are a trick challenge. The AM could be an opportunity or a threat in this issue. The point surrounding the 4.0 industry is how to define and control the permission to copy and some kind traceability of the produced goods e.g. IoT and "IP or DOI mark", but the ontology logical data and right technological and legal model must be defined. However, if those goods, the spare parts or additional ones, will not be controlled, they potentially could be produced by the entity or nation under embargoes sanctions or some undisclosed ally. The other concerns regarding AM solutions are about software to control in some extent the production of some good and with the possibility that the supplier of the product restrict the production at the buyer as an undeclared embargo, one option is private blockchain key shared by the supplier of the technology able to be AM, the customer and a trusted third part.

Keywords:Defence-aerospace Additive Manufacture Intellectual Property ontology

1 Introduction

The history of the Additive Manufacturing, mainly regarding the 3D printing, has more than a century and the solutions been becoming more complex [1]. One of the very first references and patents was done in the end of the XIX century by BLANTHER [2]. The current possibilities for several different applications are enormous with different technologies, niches and materials commercially available, since the first commercial stereolithography in 1980s [3] [4] [5]. The efforts to do this technology closer to students occurs at different levels and sometimes include the participation of the community [6] or develop the improve professional capabilities by online training. [7]

The standard file exchange format is stl, Stereo Lithographic or Standard Tessellation Language [8] and is usually described as a succession of tiny layers [9]. Other files formats used are: 3MF, AMF, SPE-NC, STEP, each one has some specificities. [10].

Additive Manufacture (AM) is changing the production standards and the value chains in an unprecedented way [11] [12] [13]. Its distribution sometimes is geographically dispersed, in some scenarios it could be in clusters and then 4.0 industry pops up some considerations about regional ecosystem influence [14] and also RISE Research Institutes of Sweden [15], there are also discussion about frameworks and threats and opportunities. [16]

2 Acceptance of AM

The field of AM is becoming more professional and the normalization is a good metrics of this. In Brazil one of the

groups is the ABNT/CCE 261, the author is currently working in this group, with ABIMAQ/FIESP that works in the normalization in this field in groups with ISO and ASTM. Some standards try to develop an unambiguous and specific the vocabulary on this field, at least for its current situation [17] [18], others the documentation [19] and several other subjects important for the development technology and diffusion of this manufacture. This effort to making clear the concepts and the basic validations makes the traceability and quality something more real since the concepts and definitions will be at list very similar. The market presents several suppliers of powder solutions for AI [20]

The impact in the economy is diversified and improving the logistic tools, supporting of the maintenance by Additive Manufacture could be made ease by some kind of parts' library, e.g. gaskets and O-rings [21] [22]

Evolution of the basic paradigms of intelligent manufacturing and some questions regarding Human – Cyber-Physical Systems [23]. As it is very active area and in continuous evolution its ontology is not fully defined and perhaps increases will be possible, e.g. what could be the difference between a printer and a robot. If we accept that they tend to be very similar, perhaps concepts as *cobots* will appear [24]. Industry 4.0 is a wide concept and include several prominent technologies and the research in the subject is a hot topic with risks and opportunities [25] [26].

3 Security and Intellectual Property

Some concepts about intellectual property (IP) looks diffuse and the legal concepts even in daily subjects [27]. Otherwise some items must be subject of some control, like the *Liberator* printed gun [28] include in printers [29] [30]. Oher considerations are the safety or integrity of the produced part [31] [32], the ideal mark to track a good must be invisible at the produced part [33]. Some very specific characteristics of one equipment, let's say some trend some noise, distortion could be used to link the produced part with the printer, that trace the authenticity and quality of individual parts [34]. A different approach could be print a QR code, with some tracer chemical with or not blockchain connection [35], some discussions about the custody chain of the blockchain information was also supplied. There are considerations to keep safe the file 3D part and some encryption technics as discrete cosines transforms was proposed [36], the labelling or marking the parts to improve the traceability of the SS316L produced parts is reported elsewhere [37]. This tracking or identification tag is not a consensus but in a Scenario of embargo or of restriction use of Defence/aeronautical parts perhaps labelling could become a normal practice.

Forensics of AM equipment is possible and is becoming a trend [38] [39].

Vulnerability assessment of the manufacturing enterprise process monitoring in manufacturing systems. AM process is raising several questions about the safety in the physical and

net chain and attack vectors, <.stl> extension files, are mentioned in the literature [40] [41]. The monitoring for parts and process authentication and verification of the design integrity could be performed by several different ways, destructive or not [42].

Security features for additive manufacturing in electronic files is one approach to keep some additional control of the AM process [43] [44].

To track the integrity of the production system several possibilities are idealized e.g. by the monitoring sound and noise at the manufacture island or by tracking the pattern of energy consumption of the motors and actuators in the shop [45] [25]

4 Aerospace and Defence industries and BDS questions

Systems and Software for Supporting Decisions based on AM Technologies in the Context of Defence are discussed elsewhere [46] [47] [48] [49] and a review of Laser Engineered Net Shaping (LENS) with pictures of some parts produced by this technology. [50]

Boycott, Divestment, and Sanctions (BDS) and its power of refusal could be a weapon or one of the tools to control or at least try to positioning about an issue [51]. The big challenge is more political than technical, the logistics to execute this process is difficult and several times involves crossing gray areas. To make it effective, it is necessary to ensure that both the scope and the objectives are clear, regular checks must be carried out to achieve goals and objectives. The purpose of this paper is to discuss and propose means to implement the sanctions envisaged, when proposing or implementing an embargo or restriction of productive capacity, it should be clear that the credibility of supervision and sanctions is the key to success [52]. With a technology that eventually facilitates the movement of manufacturing clusters the game of cat and mouse does not get simpler the surveillance and sanctions at an BDS scenario.

5 Proposal of this paper

The authors' proposal on intellectual property deals with safety issues in industry 4.0 so as not to limit its potential but aiming to improve quality and seek to ensure quantification and traceability in the production of a good in a third-party manufacture island / client. In our view the issue to be addressed in this commercial, political and legal arena, is how to manage and if appropriate block the production of a manufacturing cluster.

Obviously, this premise is made considering that the part is made with a protocol and / or "recipe" defined by the technology supplier and thus satisfactorily and consistently meets the safety and reliability requirements of the system where this part is inserted. If the customer, or final user, try

by its own risk or some illegal way produce the part will be more challenging to implementation any BDS. Technically, two broad classes of solutions can be considered, which ideally could not be done without tools like digital right management (DRM) and / or blockchains solutions and Internet of the Things (IoT):

Hardware: propose a topology in an IoT environment with a firewall designed to maintain communication in only 3 clouds (end user / client "C", Supplier of the technology "S" and trust one "T" chosen by "C" and "S") to track the production.

At Figure 1 presents a adapted Swedish tapestry, with a similar BDS clouds schema.
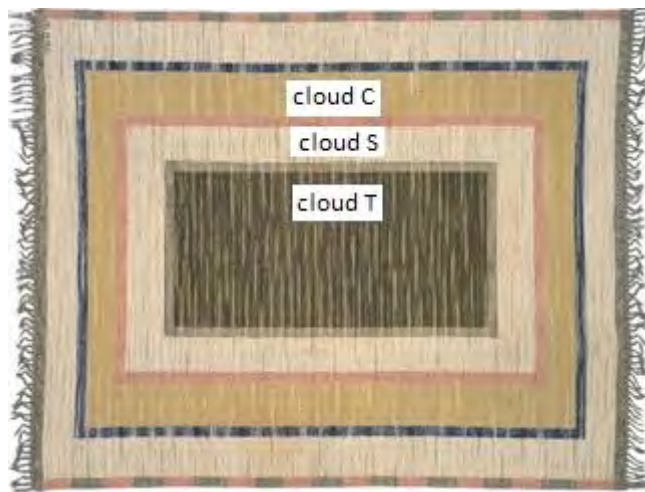
Figure 1: adapted Swedish tapestry represents similar BDS clouds.

Source: Nazmiyal Antique Rug Gallery, NYC

Software: The manufacturing files and instructions would be stored with the Trust one "T" and would be decoded and processed by a trusted channel in the manufacturing island for a defined amount of selected parts. So, this approach, would be, at least in ideal world, able to run in an untrusted environment to achieve a security goal.

Filev [53], schematics at Figure 2 and Figure 3, presents the innovation solutions to isolate the industrial area by firewall and or co-supervision at cloud computing and used at EXPOMAFE 2017 and 2019. FEI's University also participate of the exposition of Cluster with multiprotocol communication, with emphasis on OPC-UA and MTConect at **EXPOMAFE 2019 (https://www.expomafe.com.br/en/Home.html)** an initiative of The Brazilian Machinery and Equipment Builders' Association (ABIMAQ http://www.abimaq.org.br/site.aspx/abimaq-en).
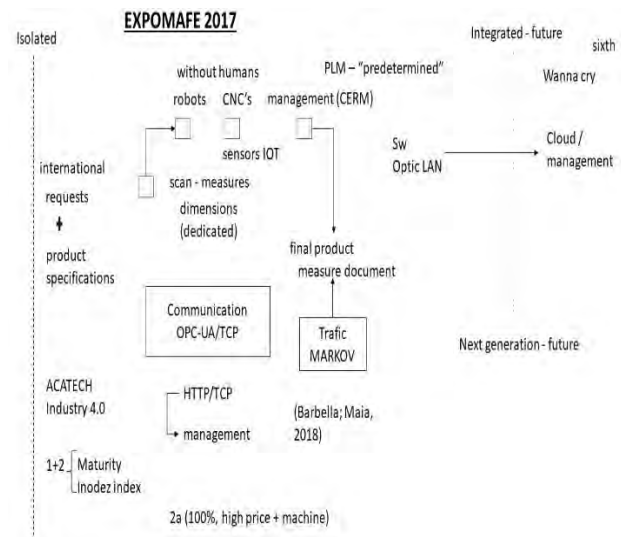

Figure 1: Conceptual architecture for an AM unit.
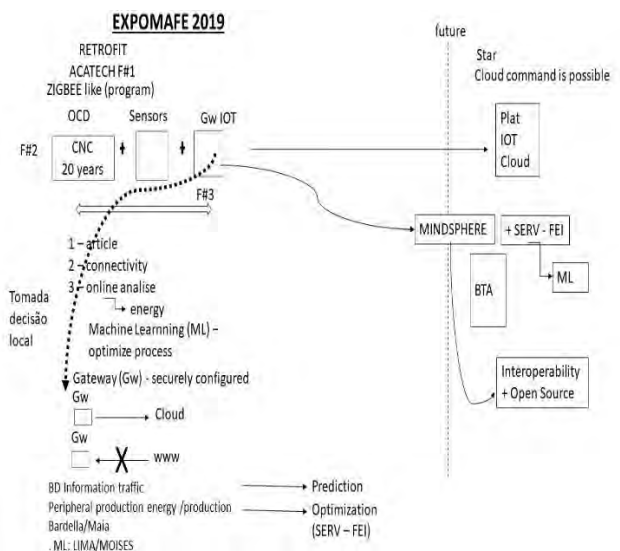
Source: Authors.

Figure 3: Conceptual architecture for an AM unit.

Source: Authors.

Conceptual architecture for AM unit, at Figure 2 and Figure 3, shows the importance integrating concepts about, open IOT, BDS clouds, machine learning and interoperability BTA.

The question to handle is what is, the more interesting architecture for intellectual protection and for a more restrictive scenario like as Defence industry or other specific ones that could be subject to embargo discussion, and sanctions that could be implemented after. Both approaches have their pros and cons, while in the hardware-focused option the manufacturing island would open the external inspection? Would it be technically feasible?

The archive virtual warehouse solution, with assurance of its integrity, would probably be more effective in avoiding "unwanted mirror" production units, since the production will be dependent of a file and instructions that will be managed by the three parts agreement. The basic idea is the files won't be able to duplicate and each batch need a new set of instructions and files to produce a very specific amount of parts, some newspapers are already produced in a similar way [54].

The point is that under embargo or conflict between the parts the solution the production's island architecture will be designed to not allow the operation, that will be stopped becoming inoperative if 2 of 3 clouds does not authorize the production.

On the other side of this solution, the part "C" cluster would be sure that the manufacturing process would not be altered without its knowledge, this is an inherent risk to current systems since the STUXNET and incidents where one relied on a security that proved to be fallacious. [55] [56]

Put that the recipe/files for production protocol will be shared by the three clouds, and just work together. The target is how to implement this at the hardware level and the net level and discuss this approach from the supply chain perspective.

Having made these considerations the safest, and most certainly the most complex, would be that the files could be segmented among the three participants of this agreement. In this way with the simultaneous input of the 3 participants the material would be produced. In case of a sanction and / or boycott, client "C" would only have his share of the files and thus should not be able to continue to produce the item in a clear formal and transparent sanction and / or boycott condition.

6 Final Conclusions

The AM becomes an important player at the logistic and supply chain due it's flexibility, capability for handle several materials, for production line its implementation is not so ease due questions related as price mainly due the raw material and for the printer and accessory units. For special parts or prototypes and small quantities the acceptation of this technics is bigger every day.

The question about how to guarantee the intellectual property and copyright is a fuzzy are with several questions raising tighter with the dissemination of the Additive Manufacturing / 3D printers. The perception is that some items area not challenging to by produced in this way, let's say use to copy and then how to control the copyright? The other side are the critical and complex parts that are difficult to produce and to copy and they could be represented for parts for aeronautic Defence and aerospace, those critical items that represent

high tech and could be subject of BDS are the object of this paper and how to control its production. The technology is under continuous evolution and the supply chain, legislation must address those questions.

References

[1]   E. OZTEMEL and G. SAMET, "Literature review of Industry 4.0 and related technologies," *Journal of Intelligent Manufacturing,* pp. 1-56, 24 Jul 2018.

[2]   D. L. BOURELL, J. J. BEAMAN, M. C. LEU and D. W. ROSEN, "A Brief History of Additive Manufacturing and the 2009 Roadmap for Additive Manufacturing: Looking Back and Looking Ahead," in *US – TURKEY Workshop On Rapid Technologies, September 24 – 24, 2009*, 2009.

[3]   ,. C. LEINENBACH, "Material Aspects in Metal Additive Manufacturing Challenges, Opportunities, Visions," in *LANL Workshop*, Santa Fé, 2015.

[4]   M. N. ISLAN, H. GOMER and S. SACKS, "Comparison of dimensional accuracies of stereolithography and powder binder printing," *Int J Adv Manuf Technol,* pp. 3077 - 3087, 2017.

[5]   ASTM, "ASTM F42/ISO TC 261 Develops Additive Manufacturing Standards," ASTM, [Online]. Available: https://bit.ly/2LsaUhg. [Accessed 05 Feb 2019].

[6]   D. DUMOND, S. GLASSNER, A. HOLMES, D. C. PETTY, T. AWISZUS, W. BICKS and R. MONAGLE, "Pay it forward: Getting 3D printers into schools," in *IEEE Integrated STEM Education Conference*, PRINCENTON, 2014.

[7]   MITxPRO, "Additive Manufacturing for Innovative Design and Production," 2019. [Online]. Available: https://additivemanufacturing.mit.edu/.

[8]   J. GARDAN, "Additive Manufacturing technologies: state of art and trends," *Internation Journal of Production Research,* vol. 54, no. 10, pp. 3118-3132, 2016.

[9]   A. Gebhardt and J.-S. Hotter, "Additive Manufacturing : 3D Printing for Prototyping and Manufacturing," 2016.

[10] E. PEI, M. RESSIN, R. CAMPBELL, B. EYNARD och J. XIAO, "Investigating the impact of additive manufacturing data exchange standards for re-distributed manufacturing," *Progress in Additive Manufacturing,* pp. 1-14, 12 Jun 2019.

[11] E. BLACKWELL, T. GAMBELLl, V. MARYA och C. SCHMITZ, "The great re-make: Manufacturing for modern times," 2017.

[12] M. HANNIBAL and G. KNIGHT, "Additive manufacturing and the global factory: Disruptive technologies and the location of international business," *International Business Review,* pp. 1116 - 1127, 2018.

[13] T. JOHNSTON, T. D. SMITH and J. L. IRWIN, Additive Manufacturing in 2040 Powerful Enabler, Disruptive Threat, RAND Corporation, 2018, p. 31.

[14] M. GÖTZ and B. JANKOWSKA, "Clusters and Industry 4.0 – do they fit together?," *European Planning Studies,* vol. 25 (9), pp. 1633 - 1654, 2017.

[15] RISE Research Institutes of Sweden, "ADDITIVE MANUFACTURING AT RISE IVF," [Online]. Available: https://bit.ly/2KMu8il.

[16] M. STEHN, I. WING, T. CARLILE, J. DICHAIRO and J. MARIANI, "3D opportunity for adversaries Additive manufacturing considerations for national security," Deloitte Development LLC, 2017.

[17] ISO/ASTM, *52921 (ASTM F2921) under revalidation,* 2013.

[18] ABNT/ISO, *Manufatura aditiva - Principios Gerais - Terminologia,* CEE 261/ Manufatura Aditiva, 2018, p. 25.

[19] ASTM, "The Global Leader in Additive Manufacturing Standards," ASTM, West Consohocken, 2017.

[20] SCHMOLZ + BICKENBACH Group, *Printdur® Metal powder for Additive Manufacturing,* Witten, 2018-016.

[21] H. KIM, M. CHA, B. C. KIM och D. MUN, "Part library-based information retrieval and inspection framework to support part maintenance using 3Dprinting technology," *Rapid Prototyping Journa,* vol. 25, pp. 630-644, 2019.

[22] R. AUGUSTSSON and D. BECEVIC, "Implementing Additive Manufacturing for Spare Parts in the Automotive Industry A case study of the use of additive manufacturing for spare parts," CHALMERS UNIVERSITY OF TECHNOLOGY, Gothenburg, 2015.

[23] Z. JI, L. PEIGEN, Z. YANHONG, W. BAICUN, Z. JIYUAN och M. LIU, "Toward New-Generation Intelligent Manufacturing," *ENGINEERING,* vol. 4, pp. 11-20, 2018.

[24] A. R. SADIK and B. URBAN, "An Ontology-Based Approach to Enable Knowledge Representation and Reasoning in Worker–Cobot Agile Manufacturing," *Future Internet,* vol. 9, no. 90, 2017.

[25] C. BAYENS, T. LE, L. GARCIA, R. BEYAH, M. JAVANMARD and S. ZONOUZ, "See No Evil, Hear No Evil, Feel No Evil, Print No Evil? Malicious Fill Patterns Detection in Additive Manufacturing," in *Proceedings of the 26th USENIX Security Symposium,* Vancouver, 2017.

[26] M. K. THOMPSON, G. MORONI, T. VANEKER, G. FADEL, I. CAMPBEL, I. GIBSON, A. BERNARD, . J. SCHULZ, P. GRAF, B. AHUJA and F. MARTINA, "Design for Additive Manufacturing: Trends, Opportunities, Considerations and Constraints," in *CIRP Annals Manufacturing Technology*, 2016.

[27] M. WEINBERG, "3 Steps for Licensing Your 3D Printed Stuff," 2015.

[28] W. A. WINDLE, "ADDITIVE MANUFACTURING: PREPARING FOR THE REALITY OF SCIENCE FICTION," Homeland Security Digital Library, Monterey, 2015.

[29] ALA American Library Association, "3-D Printing in Libraries: Policies & Best Practices," Chicago, 2018.

[30] T. MODEGI, "Proposal for 3D-Printing Regulation Technique in Fabricating Illegal Objects Using Feature-vector Based Matching Algorithm of 3D Shapes," in *SICE Annual Conference*, Tsukuba, 2016.

[31] M. YAMPOLSKI, T. R. ANDEL, J. T. McDONALD, W. B. GLISSON and A. YASINSAC, "Intellectual Property Protection in Additive Layer Manufacturing: Requirements for Secure Outsourcing," in *PPREW-4 Proceedings of the 4th Program Protection and Reverse Engineering Workshop*, New Orleans, 2014.

[32] M. YAMPOLSKIY, A. SKJELLUM, M. KRETZSCHMAR, R. A. OVERFELT, K. R. SLOAN och A. YASINSAC, "Using 3D printers as weapons," *International Journal of Critical Infrastructure Protection,* pp. 58-71, Sept 2016.

[33] V. ITIER, W. PUECH och A. BORS, "CRYPTANALYSIS ASPECTS IN 3-D WATERMARKING," i *IEEE International Conference on Image Processing (ICIP)*, Paris, 2014.

[34] F. PENG, J. YANG, Z.-C. LIN and M. LONG, "Source identification of 3D printed objects based on inherent equipment distortion," *Computers & Security,* vol. 82, May 2019.

[35] Z. C. Kennedy, D. E. Stephenson, J. F. Christ, T. R. Pope, B. W. Arey, C. A. Barretta and M. G. Warner, "Enhanced anti-counterfeiting measures for additive manufacturing: coupling lanthanide nanomaterial chemical signatures with blockchain technology," *Journal of Materials Chemistry C,* no. 37, pp. 9570-9578, October 2017.

[36] G. N. PHAM, J.-H. PARK, O.-H. KWON, H.-J. SONG, S.-H. LEE, K.-S. MOON, S.-T. KIM, Y.-R. CHOI och K.-R. KWON, "Selective Encryption for 3D Printing Model in DCT Domain," i *ICUFN 2018*, 2018.

[37] T. NIENDORF, F. BRENNE, M. SCHAPER, A. RIEMER, S. LEUDERS, W. REIMCHE, D. SCHWARZE and H. J. MAIER, "Labelling additively manufactured parts bymicrostructural gradation – advancedcopy-proof design," *Rapid Prototyping Journal,* vol. 22, no. 4, pp. 630-635, 22 May 2015.

[38] V. GARCIA och C. VAROL, "Digital forensics of 3D printers," i *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*, Antalya, Turkey, 2018.

[39] D. BRADFORD MILLER, W. B. GLISSON, M. YAMPOLSKIY, Kim-Kwang and R. CHOO, "Identifying 3D printer residual data via open-source documentation," *Computers & Security,* vol. 75, pp. 10-23, Jun 2018.

[40] L. D. STURM, C. B. WILLIANS, J. A. CAMELIO, J. WHITE och R. PARKER, "Cyber-physical vulnerabilities in additive manufacturing systems : A case study attack on the .STL file with human subjects," *Journal of Manufcturing Systemas,* pp. 154-164, Jul 2017.

[41] Z. Xu and Q. ZHU, "Cross-Layer Secure Cyber-Physical Control System Design for Networked 3D Printers," in *American Control Conference (ACC)*, Bonston, 2016.

[42] T. KOMOLAFE, W. TIAN, G. T. PURDY, M. ALBAKRI, P. TARAZAGA and J. CAMELIO, "Repeatable part authentication using impedance based analysis for side-channel monitoring," *Journal of Manufacturing Systems,* vol. 51, pp. 42-51, 2nd Apr 2019.

[43] F. CHEN, G. MAC och N. GUPTA, "Security features embedded in computer aided design (CAD) solid models for additive manufacturing," *Materials & Design,* pp. 182 - 194, 2017.

[44] M. DAWSON, "Cyber Security in industry 4.0: The Pitfalls of Having Hyperconnected Systems," *Journal of Strategic Management Studies,* vol. 10, pp. 19-28, 22 Oct 2018.

[45] S. B. MOORE, J. GATLIN, S. BELIKOVETSKY, M. YAMPOLSKIY, W. E. KING and Y. ELOVICI, "Power Consumption-based Detection of Sabotage Attacks in Additive Manufacturing," 2017.

[46] D. S. GONZÁLEZ and A. G. ÁLVAREZ, "AM Manufacturing Feasibility Study & Technology Demonstration EDA AM State of the Art & Strategic Report," Fundación Prodintec, 2018.

[47] J. A. SAUCEDO-MARTíNEZ, M. PÉREZ-LARA, J. A. MARMOLEJO-SAUCEDO, T. E. SALAIS-FIERRO and P. VASANT, "Industry 4.0 framework for management and operations: a review," *Journal of Ambient Intelligence and Humanized Computing,* vol. 9, pp. 789 - 801, Jun 2018.

[48] E. N. ROTHMAN and A. ZIMMERMAN, "Editors' Introduction," *RADICAL HYSTORY Review,* pp. 1-23, 1 May 2019.

[49] A. BUSACHI, J. ERKOYUNCU, P. COLEGROVE, R. DRAKE, C. WATTS, F. MARTINA, N. TAPAGLOU and H. LOCKETT, "A system approach for modelling additive manufacturing in defence acquisition program," in *11th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME '17*, 2018.

[50] A. B. BUSACHI, J. E. ERKOJUNCU och COLEGROVE, "Modelling Applications of Additive Manufacturing in Defence Support Services," Cranfield University, 2017.

[51] M. HEDGES och N. CALDER, "Near Net Shape Rapid Manufacture & Repair by LENS," NATO, Neuilly-sur-Seine, 2006.

[52] D. FRUCHART, P. HOLTON, S. T. WEZEMAN, D. STRANDOW and P. WALLENSTEEN, "United Nations Arms Embargoes - Their Impact on Arms Flow and Target Behaviour," SIPRI Arms Transfers Project, Solna, 2007.

[53] R. FILEV MAIA, Interviewee, *Impressions on information security at manufacturing islands in the 4.0 industry.* [Intervju]. 12 Jun 2019.

[54] K. Van SYCKLE, "See how the Times gets printed and delivered," 2018.

[55] Ernst & Young LLP, "Cybersecurity for Industry 4.0 Cybersecurity implications for government, industry and homeland security," Kolkota, 2018.

[56] INCIBE, "Emerging Threats to industrial Control Systems," 23 Aug 2018. [Online]. Available: https://bit.ly/32eeVMy.

Machining Distortion Analysis of Aerospace Components using the Contour Method

Mats Werke[a]*, Mosharraf Hossain[b], Daniel Semere[b], Anders Wretland[c]

[a]RISE IVF AB, Box 104, 431 22 Mölndal, Sweden
[b]KTH Industrial Production, Brinellvägen 68, 100 44 Stockholm
[c]GKN Aerospace AB, Flygmotorvägen 1, 461 81 Trollhättan, Sweden
*Corresponding author, mats.werke@ri.se

Abstract

During machining the residual bulk stresses induced by previous process steps like forging and heat treatment will be released. This may cause undesirable geometric distortions of the final component and thereby high rejection rates and costs. This problem can be reduced by adjusting process– and design parameters. This paper presents a methodology for minimizing machining distortions. The methodology is based on a combination of procedures for prediction of machining distortions, using the Contour method and procedures for adjustment of machining distortions. Practical experiences are discussed and demonstrated using an aerospace component. The methodology should be executed in close cooperation between several actors in the value chain and best results may be achieved by combining several concepts for adjustment of machining distortions. Further research in conjunction with the Contour method, adaptive fixturing and toolpath adjustment is recommended.

Keywords: Machining distortions, Contour method, Adaptive Fixtures, NC-Code optimization

1 Introduction

During machining the accumulated bulk stresses induced by previous process steps, such as forging and heat treatment, will be released. This may cause undesirable residual deformations and geometry errors of the final component. The distortions are difficult to predict and in practice, therefore wide tolerances and several machining steps are applied to avoid reprocessing or, at worst, scrap. By taking into account the residual stresses during process planning a significant improvement in dimensional accuracy can be achieved. This may save millions of Euros in scrap costs and correction procedures for the manufacturing industry [1].

This paper proposes a methodology for minimizing such distortions that emerges during machining, to be used in order to facilitate a minimum of deviations from the nominal shape of the final product. The methodology is based on a combination of procedures for prediction of residual stresses in the work piece, prediction of deformations in the component during material removal and procedures for minimizing the distortions.

The overall methodology is described in Section 3 whereas procedures for residual stress prediction and mitigation during subsequent material removal operations is described in Section 4, supported by a test case, a forged aerospace component. Different concepts for adjusting machining distortions are described in Section 5 and conclusions are summarized in Section 6 together with suggested issues for further research.

2 Related Work

Several investigations have been reported in the literature on machining and residual stresses. The majority of these are focusing on the residual stresses induced from the machining operations. However, relatively few papers have addressed distortions coming from the redistribution of the initial residual stresses after the material removal. Distortions in machining of aerospace parts are often attributed to the redistribution of the residual stresses given all other operation parameters are optimized for adequacy [2]. The existing research literatures refer to the prediction and measurements of the initial residual stress field as well as estimating the distortions and proposing solutions to counter these distortions to attain the tolerances.

Earlier investigations concerning how to measure residual stresses using both non-destructive or destructive techniques such as X-ray diffraction and hole drilling have been reported. The limitations to these methods in terms of depth of measurement is well known. In this paper, the Contour Method [3][4], which is an accurate method that overcomes these limitations has been used. The method has been validated by a combined experimental and computational analysis as reported in an earlier work by Werke et al in [5]. This paper is based on these findings.

Redesign of parts and considerations during process planning, particularly machining sequence and fixture configurations, are proposed in some papers to comply the desired tolerance [6][7]. Cerutti et. al. discussed the effect of positioning the workpiece with some offset, a parameter numerically optimized to find the position where the stress

distribution assumes its minimum [8]. Chantzis et.al. presented a work with similar approaches and also provide a simulation based work flow to find the optimized offset for a minimum distortions in [9]. These works specifically refer to conditions where the stress profile fits to the machining task. Therefore, it is difficult to generalize for other stress profiles.

This paper proposes a new method for adjustment of the NC-code in order to compensate for the distortions during the manufacturing process.

3 Methodology

The first step is to predict the work piece bulk stresses induced prior to machining, see fig. 1.
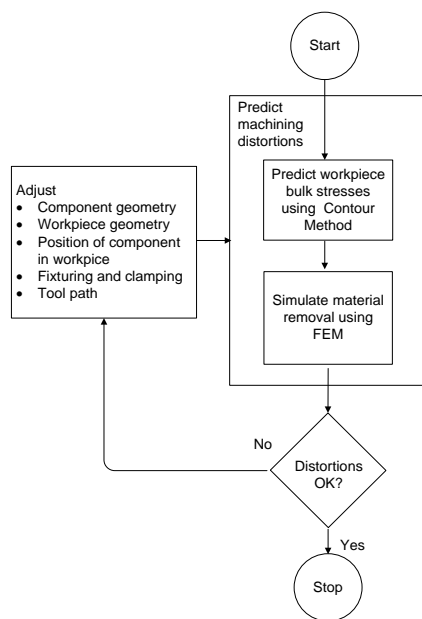
Figure 1 Overall methodology

This is performed using the Contour method. This method uses the fact that a body that contains residual stresses will deform when the body is cut into sections. The tractions required to restore the deformed section to its original shape are equivalent to the out of plane residual stress released by sectioning, see fig. 2.
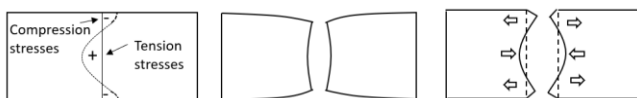
Figure 2 Schematic illustration of Contour method [3][4]

This information is used as input data to the second step, a material removal simulation for calculation of the distortions. If the distortions exceed the requirements from the customer, then adjustments should be performed until the level of distortions are within tolerances. The methodology is described below in conjunction to a test case, an aerospace component.

4 Prediction of machining distortions

An aerospace component in a nickel based alloy was used as test case, see fig. 3. The component was analyzed in its net shape condition before machining where after the deviations from its nominal geometry was verified. The work piece was hot forged in 1100 C with a screw press followed by solution annealing and then machining in several setups to final form. The geometrical shape had a tendency to drift out of tolerance and twist around the axial direction. The twist was max 0.51 mm and min -0.69 mm. The ambition was to set up a model for analysis and adjustments of the drift.
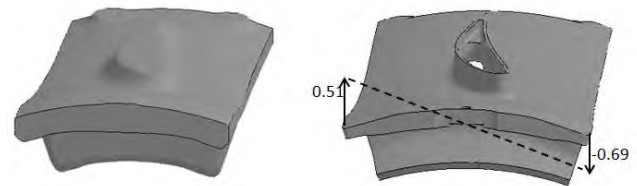

Figure 3 Workpiece before and after machining and illustrated twist

4.1 Prediction of workpiece bulk stresses

Predict stresses in cut sections: The work piece was cut in 5 sections (0, 50, 90, 120, 240 mm) according to fig. 4, using Wire Electro Discharge Machining (W-EDM). The cut section displacements were then measured using Geometric Optical Measurement (GOM) technique. The GOM measurement requires the cut surfaces to be made with lowest possible surface roughness, to make it possible to detect the subtle deviations from an ideal flat shape after the parting is performed. Thus, the choice of a precision machining method that is able to cut along a straight line while securely clamped is crucial.
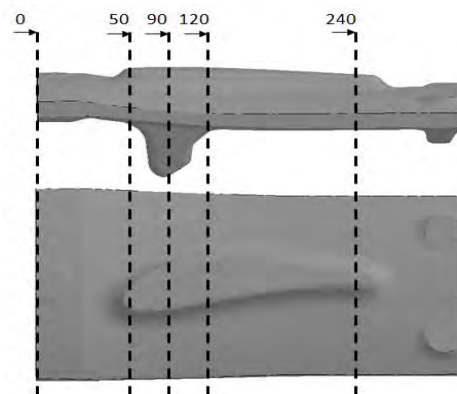

Figure 4 Workpiece cut in 5 sections

In the test case this procedure was carefully planned where the work piece was prepared with start and finish holes for the EDM wire, to enable lowest possible deflection of the component during each parting operation, see fig. 5. Thus, the component was actually functioning as its own fixture during the major part of each parting operation.
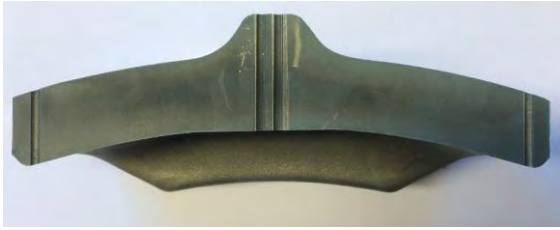
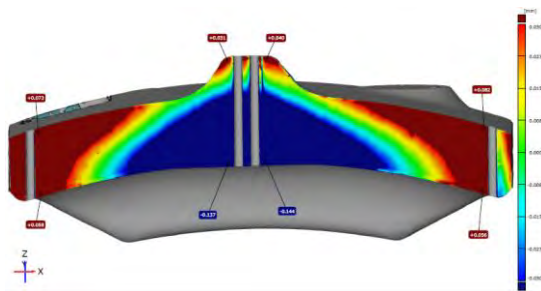*Figure 5 Cut section with four drilled holes for guidance of EDM*

Figure 6 GOM measurement of cut section

The results from GOM according to fig 6, were translated to Finite element models of the sectioned parts. The out of plane distortions in the cut sections are directly related to the out of plane stresses before sectioning and were applied as boundary conditions to finite element models of the parts after sectioning. The displacement surfaces from opposite sides of the cut were averaged in order to improve the accuracy of the results. The out of plane stresses (x-stresses) were then calculated according to fig. 7.
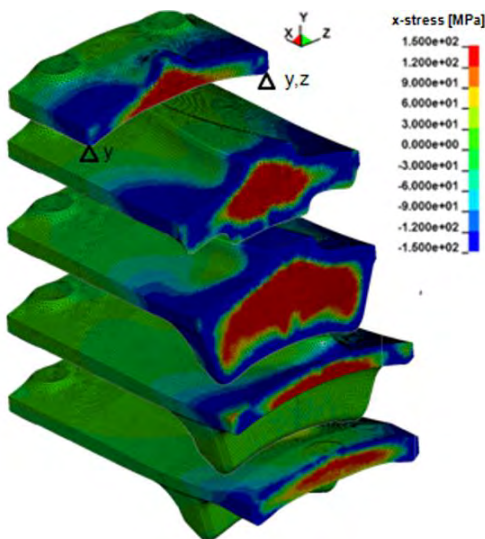
Figure 7 FE-models of the sectioned workpieces with applied x-displacements and resulting x-stresses

Calculate work piece stresses: One limitation, using the Contour method, is that the calculated stresses are limited to the out of plane stresses from the cut sections. The method may be expanded to a 3-dimensional stress state using multiple cuts at 90 degrees [10] or 45 degrees from the first direction [11]. Thus, substantial work is needed in order to

predict the complete stress field on a component, using multiple cuts. In order to simplify the stress analysis a numerical iterative approach using FEM for prediction of 3-dimensional stresses based on single cut measurements was developed and tested according to the following.

The calculated out of plane stresses (x-stresses) in the cut sections were mapped onto the work piece mesh and interpolated between the sections. After mapping and interpolation, a spring back calculation was performed which, due to equilibrium, established a 3-dimensional stress condition. This redistribution of stresses caused the originally mapped out of plane stresses to decrease. To overcome this problem an iterative procedure with remapping of the out of plane stresses on the work piece with subsequent spring back calculations was performed. The procedure was repeated until the out of plane stresses reached its original level and resulted in a more correct 3-dimensional stress state [5].

4.2 Simulation of material removal

A simplified approach for simulation material removal was utilized, where material was instantaneously removed using linear FEA. The mesh of the removed material was deleted and a spring back analysis of the remaining component was performed according to fig. 8.
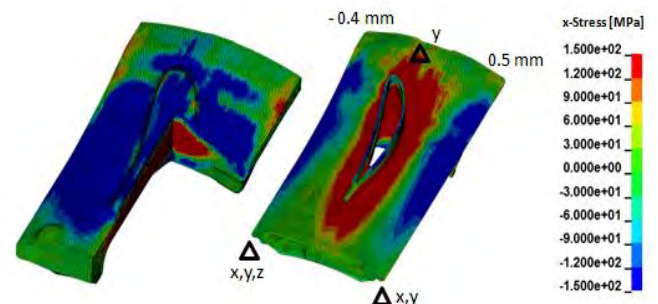

Figure 8 Stresses and distortions after material removal

Correlation between the calculated results and measurements of the top surface of the component using a conventional Coordinate Measurement Machine (CMM) show that the calculation is in good agreement with the appearance of the real part, see fig. 9.
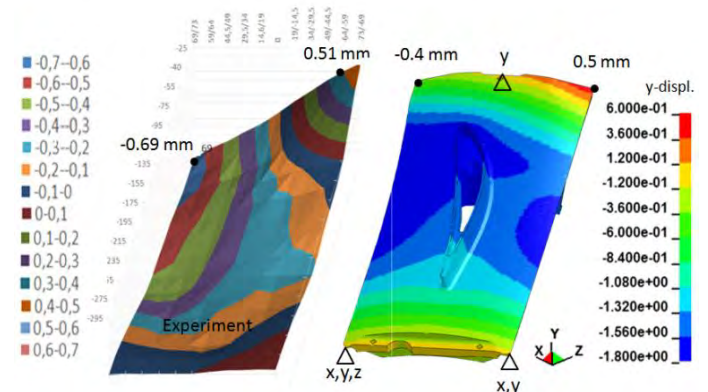

Figure 9 CMM measurement at the top surface (left) compared to calculated deformations (right).

5 Adjusting for distortions

Several concepts for how to handle geometry deviations caused when inherent bulk stresses are released during subsequent material removal are described below. They all assume the existence of a billet, in which the inherent stress field from previous forming and heat treatment operations is well known. Thus, the accurate stress field will act as the foundation for further advanced analysis of how the billet behaves during the following steps of material removal. Further, additional benefit can be achieved from accurate simulations of the intended actions of how to handle distortions most efficiently.

5.1 Adjust component position in work piece

The level and distribution of the residual stresses that are released in the component during the material removal process depends on where the final component is positioned in the billet. To move the part in the billet can be most effective if the blank and the final component have relatively simple geometries. Thus, one possibility to minimize for distortion is to virtually investigate the effects on part geometry while changing the position of the final part relative to the blank. Thereby the volume in the billet with the least effect on geometry deviations of the final part can be determined, see fig. 10.

Figure 10 Adjustment of position of the component relative to workpiece

5.2 Adjust work piece and component geometry

Another approach is to adjust the design of the component and redistribute the process induced stress field. This is illustrated in fig. 11 where a stiffness bar is included in the component in order to minimize deflections due to internal stresses. Even though this approach, that from practical reasons is a less likely choice for an already existing design, it is an approach much similar to topological design optimization including residual stresses. As such it might be able to not only change the geometry of the work piece and thereby redistribute and reduce the stress field but also be able to contribute to weight optimization.
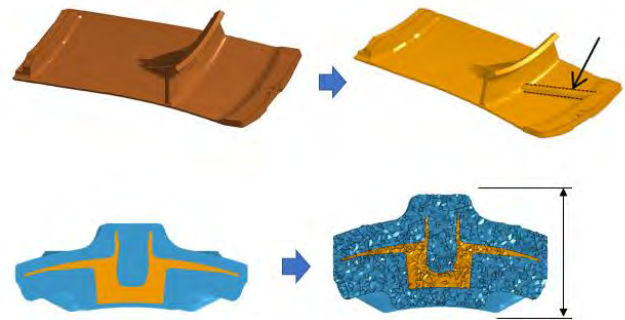

Figure 11 Adjustment of component- and workpiece geometry

5.3 Adaptive fixtures

During machining, the work piece must be held and locked to a fixed position in the workspace of the machine tool. This is normally done by use of some kind of fixture or clamping device. The upside of this concept is that the position of the part is well known. The downside on the other hand is that the clamps at the same time may prevent the part from needed relaxation. This may introduce new stresses due to the changed stress state which may affect the geometrical stability of the part and cause unexpected distortions during material removal operation. The preferred solution would therefore be a clamping device that could be adaptive and provide some dimensional flexibility.

The general requirements for an adaptive and reconfigurable fixture set-up have been discussed by several researchers. A comprehensive list of work performed in this area is compiled by Borboni et al [12] whereas more specific applications with respect to aero engines are discussed by Papastathis et al [13].

Distortions during material removal are largely dependent on the continuous change of inherent phenomena in the work piece, such as changes in the residual stress state due to material removal. It is therefore important to enable the part to relax, i.e. regain a state free from the distortions imposed because of the relieved internal stresses. This is ideally done continuously, but in practice at reasonably frequent instances, most often related to by "best practice".

Thus, the interface between the part and the clamping device should, to some degree, be flexible in order to accommodate for distortions that inevitably will be the result of released stresses. However, this flexibility must be exercised such that the dynamic stiffness of the ongoing material removal process can be maintained. This is of course of particular importance when high and intermittent process forces are applied, such as during conventional milling and even more so during milling with super abrasives such as ceramic or Cubic Boron Nitride (CBN) tools. Flexibility to enable a rigid enough clamping of the part must therefore be exercised in steps - between the material removal operations, rather than continuously during the operations.

The most direct and simple form of adaptive clamping is therefore to just release and immediately clamp the part again.

This concept requires that the interface elements are "floating" to regain the stiffness requirements as just discussed. A number of practical solutions to this problem has been discussed and also practically tested in the European collaborative project INTEFIX [14].

Because of the many parameters to set during the design of a fixture, the release and clamp concept ha to be validated first using simulation. Furthermore, simulation enables the optimal amount of material removed in the release and clamp operation and also offers support when defining the most effective position of each clamping device.

Thus, with support from a FE-model, it is possible to calculate the required clamping position and the corresponding force, introduced at each clamping interface. In the following, a calculation procedure using several steps of material removal is demonstrated in conjunction to the test case according to fig. 12. The calculation indicated the need for high clamping forces during the first material removal steps (node 3, 4, 7 – 10), whereas the required clamping forces during the second steps were low, see fig. 13. The proposed calculation procedures may be used for virtual test of various clamping options. Further, to adapt the clamping to the geometry variation during machining and in addition minimize the stresses introduced by clamping.
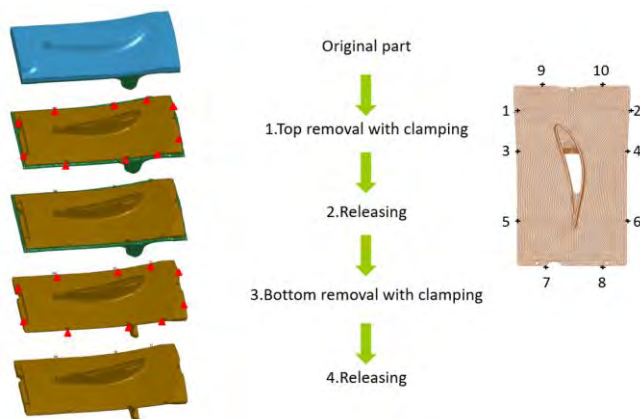

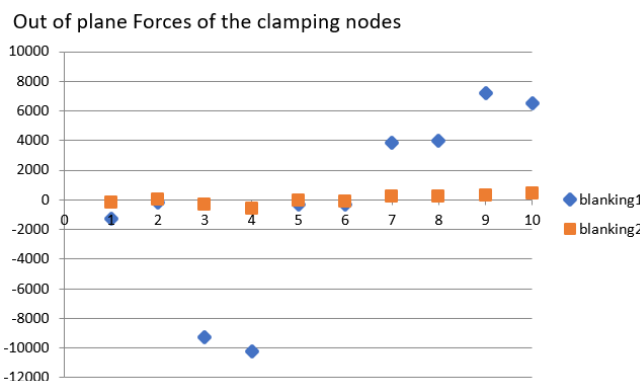*Figure 12 Material removal in two steps including clamping*

Figure 13 Calculation of out of plane clamping forces at node 1 - 10

A more profound way to arrange for adaptive holding of the part that is subject for material removal would be to use a multi-axis and adaptively programmable machine tool. This application offers the possibility to be able to reach several sides of the billet within the same set-up without any intermediate re-clamping.

The deflections imposed by continuous release of residual stresses due to the ongoing material removal can under such conditions be handled to a very high degree of accuracy.

Figure 14 Multi axis machine tool configuration, Courtesy of Precise Tool and Manufacturing; Rochester, NY 14624

5.4 Adjust tool path

By taking into account the residual stresses during the tool path generation a significant improvement of geometrical accuracy can be achieved. A schematic overview of the steps required for adjusting the tool path using simulation is shown in fig. 15.
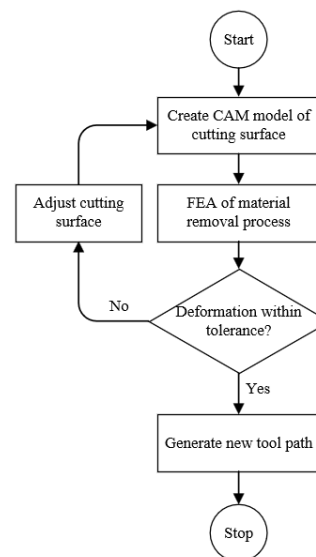

Figure 15 Schematic illustration of tool path optimization

The intended CAM cutting surfaces can be modeled with the help of CAD/CAM software. The generated 3D model will be used as input to the material removal simulation to predict the distortions. The input to the material removal simulation can be residual stresses generated with support from the Contour method. If the deformations are out of tolerances the cutting surface is adjusted. The loop is repeated until the

deformations are within tolerances. A tool path that is adjusted due to the release of residual stresses is then generated. The concept is explained according to fig. 16 using an example with stress-free block exposed to clamping. The NC-code is prepared to machine a planar surface but after release of the clamping deviations in planarity will occur. By calculation of the deviations it is possible to adjust the NC-Code by e.g. mirroring the deviations.
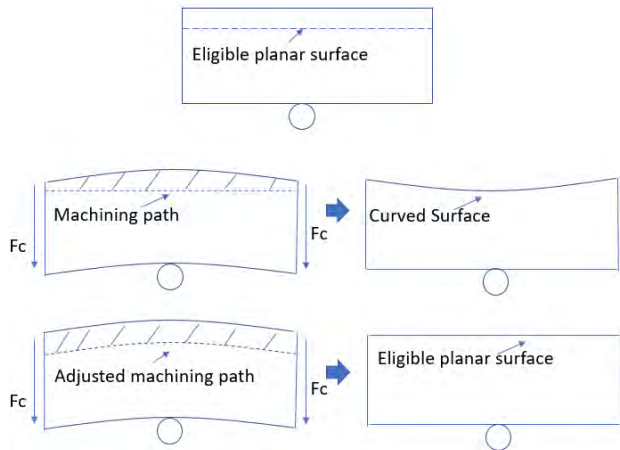

Figure 16 Schematic illustration of tool path compensation in a stress-free block exposed to clamping

However, the adjustment of one specific dimension may have consequences on other dimensions and tolerances. This is schematically illustrated with the example in fig. 17 using a block with internal residual stresses. Here the eligible planar machining top surface will be distorted due to the release of stresses during machining. This effect may be adjusted by preparing a modified machining path for the top side which adjust for the top surface deformations. However, the top side adjustment may also have effects on other parts of the geometry such as the bottom side geometry which in turn then must be adjusted. This may cause the top side to deform and the coupled problem may be solved using an iterative FEA adjustment strategy in order to create the two planar surfaces according to fig 18.
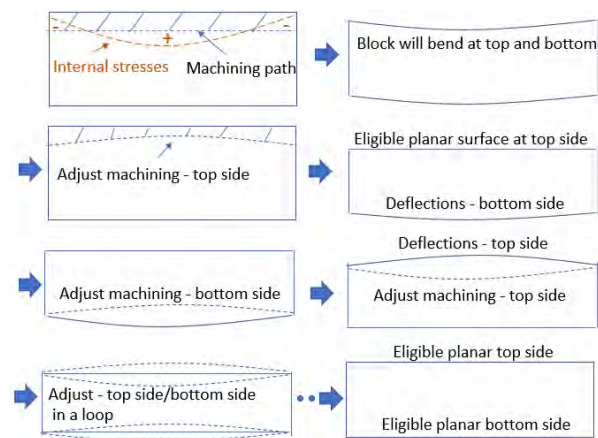

Figure 17 Schematic illustration of iterative tool path compensation in a block with internal stresses.
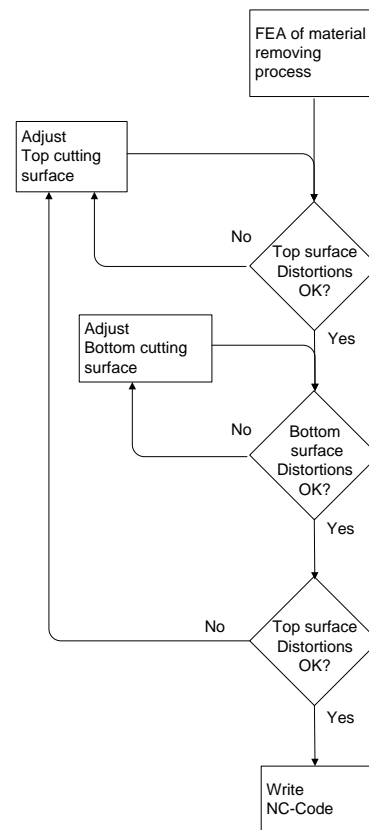

Figure 18 Adjustment strategy

Comparison logic have to be built to compare the expected surface with simulated surface. If the distortions reside within the customer defined specifications, then adjusted NC code should be generated using CAM software.

The tool path optimization methodology requires successive selection of software applications, boundary conditions and application of initial stresses, thereby estimating the deformation arises during material removal. This will be run in an iterative manner involving CAD, CAM and FEA tools in order to minimize the distortion. The suggested adjustment should be evaluated carefully because the magnitude of distortion is a function of depth of cut, spring-back and cutting profile.

6 Conclusions and future research

This paper presents a methodology for minimizing distortions. The concept is a combination of procedures for prediction of machining distortions with procedures for adjustment of machining distortions. Machining distortion is a complex task and a close cooperation between several actors in the value chain may improve the possibilities to achieve good results. Issues for further research are discussed according to below.

Contour method - Sectioning: The sectioning must be done with high precision using e.g. EDM. However, EDM generates heat and leaves a thin re-cast layer on the cut surface. This effects the geometry and topology of the cut

surface negatively and reduce the accuracy of the proceeding GOM- and FE analysis. Thus, the possibilities to increase the accuracy by replacing EDM with e.g. abrasive micro waterjet cutting should be investigated. The process is based on particle erosion without thermal impact on the sectioned surface. It shares the benefits of traditional waterjet cutting, whereas the smaller cutting system increase the precision 10 times cutting with +/-0.01 mm in precision and 1 μm Ra-value. The impact on surface topology is thereby minimal.

Contour method – 3D stress prediction: In order to reduce the need for multiple cuts a numerical iterative approach for prediction of 3-dimensional stresses based on single cut measurements and FEM is proposed according to Section 3. The results for the test case indicate very were good results but further investigations are recommended.

Adaptive fixtures: During machining, the workpiece must be held and locked to a fixture which must be able to handle the clamping- and geometry changes that occur during the machining. Thus, further research is required in order to develop flexible adaptive clamping and support of the workpiece to the fixture in order to handle springback effects during machining.

Adjusting Tool Path: This paper presents a new concept for tool path optimization. Further research is required in order to analyse parts with complex surfaces and residual stress distributions and where several demands concerning geometry dimensions and tolerances from the customer have to be fulfilled.

7 Acknowledgements

References

[1] Chatelain, Lalonde, Tahan, "A Comparison of the Distortion of Machined Parts Resulting From Residual Stresses Within Workpieces", Recent Advances in Manufacturing Engineering, ISBN: 978-1-61804-031-J2

[2] Lequeu, Lassince, Warner, Raynaud, "Engineering for the future: Weight saving and cost reduction initiatives," Aircr. Eng. Aerosp. Technol., vol. 73, no. 2, pp. 147–159, 2001.

[3] Prime, Dewald, "The Contour Method" Chapter 5 in Practical Residual Stress Measurement Methods, Wiley-Blackwell, pp. 109-138Chapter," Wiley-Blackwell, Las Alamos, 2001.

[4] Johnsson, "Residual stress measurements using the Contour method", PhD thesis, University of Manchester School of materials, 2008 Manchester UK.

[5] Werke, Wretland, Ottosson, Holmberg, Machens, Semere, "Geometric distortion analysis using a combination of the contour method and machining simulation", 51st CIRP conference of Manufacturing Systems, 2018 Stockholm

[6] Mocellin, Cerutti, "Numerical prediction of distortions during machining of large aluminium aeronautical parts. Materwiss. Werksttech., vol. 47, no. 8, pp. 699–709, 2016.

[7] Guo, Zuo, Wang, Xu, Wang, Hu, "The application of FEM technology on the deformation analysis of the aero thin-walled frame shape workpiece," Key Eng. Mater., vol. 315–316, pp. 174–179, 2006.

[8] Cerutti, Mocellin, "Influence of the machining sequence on the residual stress redistribution and machining quality: Analysis and improvement using numerical simulations," Int. J. Adv. Manuf. Technol., vol. 83, no. 1–4, pp. 489–503, 2016.

[9] Chantzis, Van-Der-Veen, Zettler, Sim, "An industrial workflow to minimise part distortion for machining of large monolithic components in aerospace industry," in *Procedia CIRP*, 2013, vol. 8, pp. 281–286.

[10] Prime, Newborn, Balog, Quenching and cold work residual stresses in aluminum hand forgings: Contour method measurement and FEM prediction. Materials Science Forum 426-432 (2003) 435-440.

[11] DeWald, Hill, Multi-axial contour method for mapping residual Stresses in continuously processed bodies. Experimental Mechanics 46 (2006) 473–490.

[12] Borboni, Aggogeri, Merlo, Pellegrini, Amici, PKM Mechatronic Clamping Adaptive Device, International Journal of Advanced Robotic Systems, April 2015.

[13] Papastathis, Ryll, Bone, Ratchev, Development of a Reconfigurable Fixture for the Automated Assembly and Disassembly of High Pressure Rotors for Rolls-Royce Aero Engines, IFIP AICT 315, pp. 283–289, 2010.

[14] Intelligent fixtures for the machining of low rigidity components, INTEFIX Report June 2016, EU Seventh Framework Programme (grant agreement no 609306).

VII Maintenance, Failure Mitigation and Operational Aspects

Phase Out Maintenance Optimization
Olle Wijk, Patric Andersson, Jan Block and Thord Righard

Prescriptive Maintenance: Building Alternative Plans for Smart Operations
Henrique Marques and Alessandro Giacotto

Fault Detection and Isolation Based on Bond Graph Models:
Application to an Electromechanical Actuator
Gabriel Dos Santos Sobral and Luiz Carlos Sandoval Góes

Model-based Sensor Fault Detection in an Autonomous Solar-powered Aircraft
Paulo Victor Padrao Lopes, Liu Hsu, Michael Vilzmann and Konstantin Kondak

Phase Out Maintenance Optimization

Olle Wijk[a], Patric Andersson[b], Jan Block[c] and Thord Righard[a]

[a] Systecon, [b] P2J Solutions AB, [c] Saab Support and Services
E-mail: olle.wijk@systecon.se, patric.andersson@p2j.se, jan.block@saabgroup.com, thord.righard@systecon.se

Abstract

This research studies the performance of an optimization technique for phase out optimization of technical systems. The optimization search is done by a genetic algorithm where good candidate samples can be drawn and evaluated in a fast manner using a novel concept denoted matrix simulation. The method is illustrated by applying it to a phase-out scenario of an aircraft fleet, where the optimal stop-maintenance strategy is determined for a set of repairable items (rotables). The main contribution of this paper is a parametric study of how the optimization search is expected to perform when varying the sampleset and problem size. In the parametric study the genetic algorithm result is compared against the real optimum which was found in a separate time-consuming brute force search, going through all possible solutions.

Keywords: phase-out, maintenance optimization, genetic algorithm, matrix simulations

1 Introduction

For capital-intensive systems, such as aircraft, the retirement process will often stretch over a decade or more, during which the operational fleet gradually contracts. This requires managing the resources for the remaining fleet at acceptable levels of cost, availability and risk. Examples of resources include maintenance personnel, tools, workshops and spares. The improvement of key profits on logistics and maintenance performance can be achieved by inventory management of costly components, which is important for equipment-intensive industries [1]. In a phase-out scenario, the number of aircraft in operation will decrease with time, and the demand for spares will also normally diminish. This combination may increase the relative quantity of spares in stock during the phase-out, which can represent an undesired capital investment.

Several studies have been done to manage spares during the operation phase of a product. In terms of maintenance policy, many researchers studied the joint optimization of maintenance and stock provisioning policy for spare part logistics, see [2], [3], [4] and [5]. In addition, Ferreira and Wang in [6] proposed a hybrid of simulation and analytical models for spare parts optimization, taking into account the residual life of equipment.

During a phase-out period, retired aircraft are usually dismantled to support the remaining operational fleet; called "parting-out" process, see the right-hand side of fig. 1. The "parting-out" refers to the collection of spares from retired aircrafts to be used in remaining operational fleet, i.e. the retired fleet is a source for increasing the spares stock.

In [7] it is suggested from a top-down overview perspective, how a Part-out Based Spare Provisioning (PBSP) management programme can be structured for managing a phase-out scenario in an effective manner. Block in [7] introduces the parting-out approach for spares provisioning management of a fleet and describes the prerequisites of such a method, furthermore, associated key decision criteria are discussed and a framework is presented for the phase-out management process.

This paper focuses on analyzing a method for optimizing the maintenance strategy decisions during a phase-out scenario [8], which is one out of several factors necessary to establish a cost-effective PSBP management programme. More specifically, the paper focuses on finding cost efficient points in time to; stop corrective maintenance (CM), stop preventive maintenance (PM), stop the parting-out process from disposed aircraft, see fig. 1.
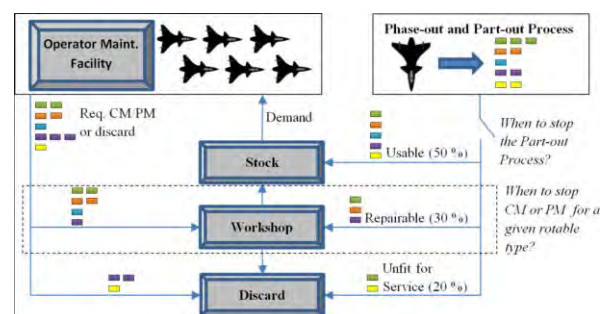

Figure 1: Schematic of the rotable flow in a PBSP management program [7]. The rotables are here represented by rectangles and the rotable type (see tab. 1) is represented by the rectangle color.

2 Scenario

The optimization technique applied in this paper is illustrated using field data gathered during the operation life of a military aircraft system. The data cover $N = 10$ rotables and have been slightly modified from actual values to allow publication.

In total the data set covers 90 aircraft that are phased out during a $T = 10$ year period; each aircraft in operation accumulates 20 flight hours/month. The aircraft phase-out scheme is shown in fig. 2, at the end of the phase-out period there are still 30 aircraft remaining in operation.
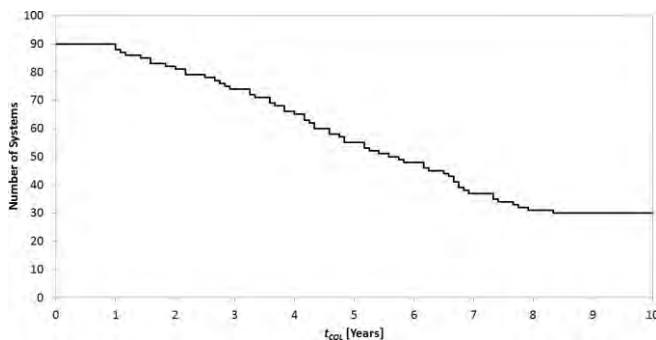

Figure 2: The phase-out scheme for the 90 aircraft system. After 10 years only 30 aircraft remain in operation.

The recovered rotables from a retired aircraft are classified in three categories:

- Usable (serviceable, sent directly into storage).
- Repairable (unserviceable, but can be reused after a repair action, sent to the workshop).
- Unfit for service (neither reusable nor worth repairing and should be scrapped).

On average 50 % of the recovered items are classified as usable, 30 % as repairable and 20 % as unfit for service.

Data for all maintenance significant events on the 10 rotables is listed in tab. 1. The reason for studying the particular rotables given in tab. 1 is that they comprise a relatively large maintenance volume and that the maintenance cost for these rotables is relatively high. In tab. 1 it is seen that there is a cost associated with recovering an item from a phased-out aircraft, which means that it can be beneficial to stop recovering items at some point in time. Furthermore, there is a cost for performing CM/PM on an item.

As long as maintenance activities are performed there is also an associated maintenance capability cost

$$c_K = 500\ 000 \text{ SEK/year}$$

i.e. costs for keeping necessary maintenance resources, facilities etc. available. The higher the maintenance capability c_K cost is, the more cost efficient it becomes to stop maintenance on all items as soon as possible and preferably at the same time, i.e. due to the c_K parameter it is not possible to find an optimal solution with an item-by-item approach.

3 Optimization problem

The phase-out maintenance optimization problem studied in [8] can be formulated as

$$min\ C(\boldsymbol{t}_{CP}, t_{COL})$$

$$(P)$$

$$\text{s.t. } B_i(t) = 0 \text{ for } i = 1, 2, \dots, N.$$

where $C(\boldsymbol{t}_{CP}, t_{COL})$ is the objective cost function that can be chosen freely to map reality, t_{COL} represents the time when we stop recovering items from phased out aircraft and $\boldsymbol{t}_{CP}$ is a time vector

$$\boldsymbol{t}_{CP} = (t_{CM_1}, t_{PM_1}, t_{CM_2}, t_{PM_2}, \dots, t_{CM_N}, t_{PM_N})$$

where the two first time elements t_{CM_1} and t_{PM_1} is the time to stop CM/PM respectively for the first rotable item, the third and fourth element t_{CM_2} and t_{PM_2} is the time to stop CM/PM respectively for the second rotable item and so on.

For the specific scenario studied in this article, a backorder constraint of zero is used for each rotable, i.e. no backorders allowed. A more detailed description of the optimization problem and the terminology used is found in [8], where it is also outlined how one can handle non-zero backorder constraints.

4 Solving the optimization problem by brute force

An optimization problem can always be solved by going through all possible solution, one obstacle though is that this might be just too time consuming. In [8] a novel concept denoted matrix simulation is developed in order to narrow down the solution space of the optimization problem (P). For the scenario described in section 2 the matrix simulations

Table 1: Empirical data used for modelling.

Item ID	Quantity per aircraft	Initial stock	Fail rate [1/10³ fh]	CM TAT [Days]	CM cost [SEK]	PM Interval	PM TAT [Days]	PM cost [SEK]	Recover cost [SEK]
Warning Beacon	1	11	0.13	20.1	1500	1600 fh	23.6	2000	560
Relief Valve	1	11	0.07	40.3	10000	1010 fh	91.2	14000	3750
Transducer	1	15	0.20	23.1	12000	510 fh	35.3	10000	4500
Cut-Off Valve	4	23	0.20	30.2	6000	2010 fh	90.3	18000	2250
A/C - Start Generator	2	18	0.16	89.6	20000	800 fh	69.6	20000	7500
Hydraulic Generator	1	10	0.13	30.2	6000	3600 days	71.5	8000	3000
Electric Motor	2	25	0.72	41.3	6000	2600 fh	42.6	6000	2250
Electric Jack	1	12	0.88	19.1	12000	1300 fh	41.9	20000	4500
Oxygen Hose	1	18	0.34	1.0	1000	510 fh	21.0	400	500
Cooling Turbine	1	20	0.27	115.3	30000	1010 fh	132.5	43000	15000

narrows down the solution space to 43 866 471 255 357 candidates, which are tested one by one in a time-consuming 4000 hour brute force search, see [8]. The minimum cost found is

$$C_{min} = 24\ 046\ 241 \text{ SEK}$$

which in the subsequent sections is used for comparison when evaluating the ability of a genetic algorithm to solve the optimization problem (P).

From [8], the optimal time to stop collecting items from phased out aircraft is

$$t_{COL}^{opt} = 7.92 \text{ years}$$

and the optimal time to stop CM/PM on the respective item is listed in tab. 2.

Table 2: Optimal time instances to stop CM and PM.

Item ID	t_{CM}^{opt}	t_{PM}^{opt}
Warning Beacon	7.65 years	3.20 years
Relief Valve	6.62 years	4.11 years
Transducer	1.37 years	7.53 years
Cut-Off Valve	7.53 years	3.88 years
A/C - Start Generator	7.65 years	5.82 years
Hydraulic Generator	5.71 years	1.26 years
Electric Motor	4.79 years	6.74 years
Electric Jack	7.65 years	4.68 years
Oxygen Hose	2.74 years	7.88 years
Cooling Turbine	7.53 years	4.00 years

5 Attacking the optimization problem with a genetic algorithm

The optimization problem (P) in section 3 can be solved using a genetic algorithm as described in fig. 3
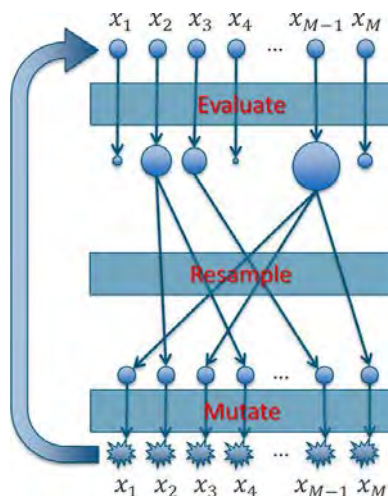

Figure 3: Genetic algorithm: 1) Evaluate the samples; 2) Survival of the fittest samples by resampling; 3) Mutate the samples.

The samples used in the genetic algorithm can be described as

$$x_m = (\boldsymbol{t}_{CP}, t_{COL}), \quad m = 1, 2, \dots, M$$

When ***initializing*** the genetic algorithm, the M samples are drawn randomly. In [8] it is described how good sample candidates can be randomly drawn from a narrowed solution space satisfying the optimization problem backorder constraints. This is achieved by a novel concept denoted matrix simulation, furthermore as a byproduct of the matrix simulations, fast ***evaluation*** of samples becomes possible. For the specific scenario described in this article, samples were evaluated close to a speed of 3 000 000 samples/sec on a laptop with a 2.7 GHz processor. In the evaluation process samples with low cost $C(\boldsymbol{t}_{CP}, t_{COL})$ are given more support at the expense of samples with higher cost, see [8] for details.

In the ***resampling*** step of the genetic algorithm the next sample population is drawn from the support distribution obtained in the evaluation step. The new sample population then undergoes a ***mutation*** process where the samples are randomly modified slightly within the narrowed solution space obtained by the matrix simulations, see [8] for details.

The evaluation, resampling and mutation steps of the genetic algorithm are repeated for a number of iterations until the best solution found stabilizes. In each iteration the best sample, i.e. the sample with lowest cost $C(\boldsymbol{t}_{CP}, t_{COL})$, is treated separately and is always guaranteed to survive to the next iteration (without mutation) as long as not a new better sample is found taking its place.

In [8], the proposed genetic algorithm is run $I = 100$ iterations using two different sampleset sizes $M = 100$ and $M = 2000$. When comparing with the brute force optimum (see section 4) it is concluded that with $M = 100$ the genetic algorithm comes very close to the optimum (only 0.05 % higher cost) while with $M = 2000$ the optimum is found.

To estimate the probability of the genetic algorithm producing a solution with a cost within a given distance from the minimum cost (found by brute force), it is necessary to run the genetic algorithm multiple times to get sufficient statistical confidence. The next section, which is the main contribution of this article, shows the results from such parametric study when varying the samples set size M.

6 Parametric study of the genetic algorithm when varying the sampleset size M

In [8] it is shown that it is possible to find the optimal phase-out strategy for a field case with 10 items by running a genetic algorithm $I = 200$ iterations with a sampleset size $M = 2\ 000$. However, the fact is that the result of the genetic algorithm depends on random numbers drawn during the optimization search (resample and mutate steps). Therefore, changing the random initializer seed of the computer and re-running the genetic algorithm will generate a different search trajectory that may or may not end up in the optimum. Therefore, an interesting question that this article tries to answer is:

Q: What is the probability of finding the minimum cost of problem (P), or being a certain distance from it, after running the genetic algorithm a certain number of iterations I using a sampleset size M?

To answer this question one can run the genetic algorithm many times, each time using a different initializer seed to the random generator in order to get a different optimization path. This type of experiment is done 1000 times for the sampleset sizes

$$M = 10, 20, 50, 100, 200, 500, 1\,000, 2\,000, 5\,000, 10\,000$$

and the results are shown in fig. 4 which represents the probabilistic state of the best sample for the genetic algorithm after $I = 200$ iterations.
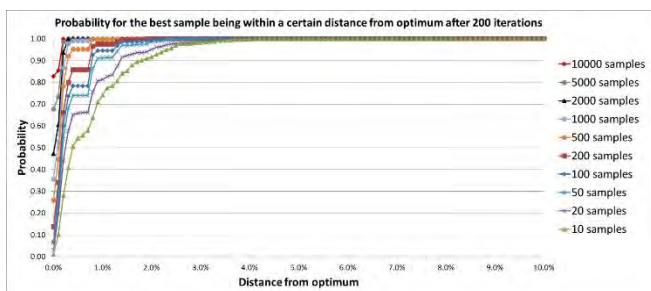

Figure 4: Probability of being a certain distance from the optimum after running the optimization I = 200 iterations.

Although running the genetic algorithm once is fast, 2 seconds when $M = 100$ and 1 minute when $M = 2000$, doing it repeatedly 1 000 times for the different sampleset sizes M takes a long time. Therefore, in order to speed up the the parameter study the genetic algorithm searches were spread out on a number of computers.

In the graph of fig. 4 the distance to the minimum cost (found by brute force search) is given in percent on the x-axis. The statistically estimated probability of being a certain percentage distance from the minimum cost after 200 iterations (based upon 1 000 trials) is shown on the y-axis. As expected, the probability of finding the optimum increases as the sampleset size M increases. If using $M = 10\,000$ the probability of finding the optimum after $I = 200$ iterations is greater than 0.8. Furthermore, it is seen that all 1 000 trials for a sampleset size of $M = 10\,000$ manages to get within a distance of 0.2 % of the optimum.

7 Parametric study repeated with a modified mutation step

In [8], the mutation step of the genetic algorithm is achieved by modifying the stop time for CM/PM for *one* of the rotables (randomly selected) to a neighbor candidate in the solution space produced by the matrix simulations. There is however a risk of getting stuck in local minima when mutating the solution in small steps, and therefore an alternative mutation process is tested where the stop time for CM/PM (for the randomly selected rotable) was randomly selected among *all* possible solutions provided by the matrix simulations (equal chance of being drawn). When re-running the parametric

study of sampleset sizes M with this slight modification of the genetic algorithm the results of fig. 5 are obtained.
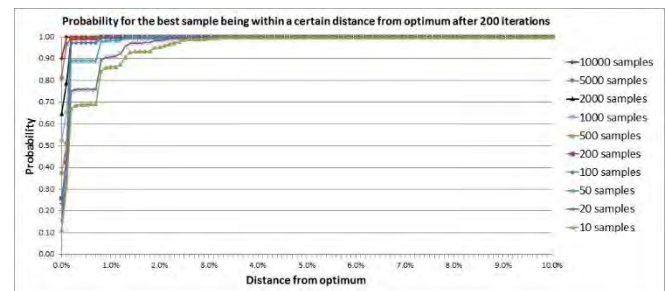

Figure 5: Probability of being a certain distance from the optimum after running the optimization I = 200 iterations after modifying the mutation process.

Comparing the results of fig. 4 with fig. 5 the convergence to the optimum is faster since every sampleset size M in fig. 5 has a higher probability of finding the optimum compared to fig. 4. In particular, the probability of finding the minimum cost with $M = 10$ samples after $I = 200$ iterations has increased from 0 % to 10 %, and the probability of finding the optimum with $M = 10\,000$ samples after $I = 200$ iterations has increased from the 80 % range to the 90 % range. Furthermore, it is seen in fig. 5 that all 1000 trials for a sampleset size of $M = 10\,000$ manages to get within a distance of 0.1 % of the minimum cost (compared to 0.2 % in fig. 4)

8 Parametric study on a large scale problem

In a real-world application there will be more than $N = 10$ items involved in the phase-out optimization problem described in this article. Therefore, the parametric study of section 7 was repeated for a problem involving $N = 100$ items. In this case it was not possible to find the true optimum by a brute force search (too many solution candidates), but it was still possible to compare the genetic algorithm performance against the minimum cost value found in *all* genetic algorithm searches

$$MINVALUE = 174\,550\,398 \text{ SEK}.$$

The result of this study is found in fig. 6 which shows the probabilistic state the genetic algorithm after $I = 4000$ iterations compared to $MINVALUE$ for varying sampleset sizes $M = 10, 20, 50, 100$ and 200. For $I = 4000$ and $M = 200$ it is seen that there is a 40 % probability to find $MINVALUE$. An increased amount of iterations or samples is expected to increase the probability to find the $MINVALUE$ but unfortunately there was no time to finish such additional parametric variations within the given project time frame.
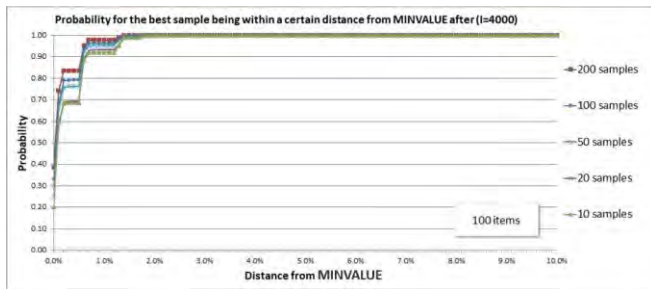
Figure 6: Probability of being a certain distance from the MINVALUE after running the optimization $I = 4000$ iterations.

9 Conclusions

This article analyzes the capability of a genetic algorithm to solve a phase-out optimization problem for a fleet of aircraft, but the ideas are general and can be applied to other types of technical systems.

For a problem size of $N = 10$ items the genetic algorithm, after running $I = 200$ iterations, has a high probability (0.7) of getting within a 1 % range of the optimum already with a small sampleset size of $M = 10$, with higher sampleset sizes, $M = 10\,000$, there is a high probability (0.9) of finding the optimum (fig. 5).

When increasing the problem size to $N = 100$ items, more iterations are required to get close to the optimum. After running the genetic algorithm $I = 4000$ iterations with a sample set size of $M = 200$, there is a high probability (above 0.9) to be within a 1 % range of what could be an optimum (fig. 6).

Looking closer at the results from fig. 5 it is noted that with

$$N = 10, I = 200 \text{ and } M = 200$$

there is a probability of 0.25 to find the optimum. An interesting question is if N and I are increased with a factor 10, i.e.

$$N = 100, I = 2000 \text{ and } M = 200$$

is there then still a probability of 0.25 to find the optimum? More generally, are the results of the analysis done in fig. 5 ($N = 10$) and fig. 6 ($N = 100$) similar after I and $10I$ iterations, respectively? The answer to this question is shown in fig. 7 (the 200 samples probability curve from fig. 5) and fig. 8 (the 200 samples probability curve from fig. 6, but after $I = 2000$ iterations instead of $I = 4000$ iterations). Although the curves of fig. 7 and fig. 8 have some resemblances, it is clear that the probability curve of $N = 100$ items (fig. 8) is mostly below the probability curve of $N = 10$ items (fig. 7). This is an indication that if the problem size N increases with a factor k then the number of iterations I needs to be increased more than a factor k to achieve similar results, further research is required to establish a more detailed relationship between I and N.
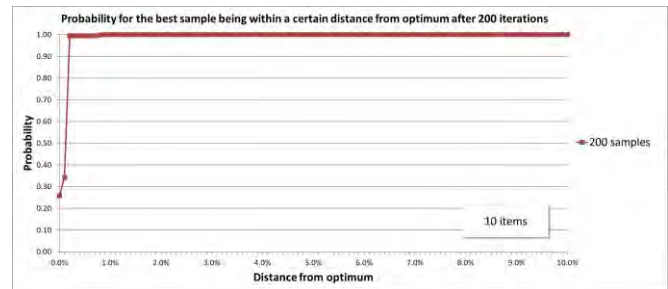

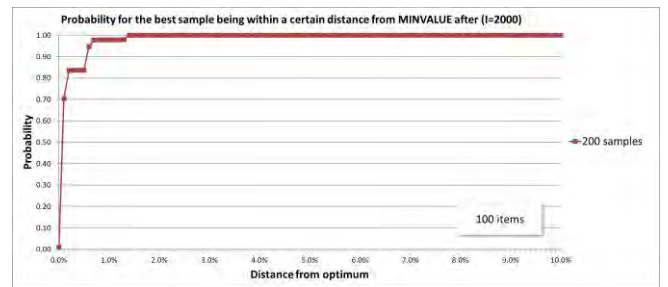*Figure 7: Probability of being a certain distance from the optimum when N=10, I=200 and M= 200.*

Figure 8: Probability of being a certain distance from MINVALUE when $N = 100, I = 2000$ and $M = 200$.

Producing a genetic algorithm parameter study, varying the number of items N, the number of iterations I and the sampleset size M is a time consuming process and was enabled by a fast sample evaluation, through matrix simulations described in [8].

Acknowledgments

References

[1] Braglia, M. & Frosolini, M. (2013), "*Virtual pooled inventories for equipment-intensive industries. An implementation in a paper district*", Reliability Engineering & System Safety, 112, 26-37

[2] Chen, M. C., Hsu, C. M. & Chen, S. W. (2006), "*Optimizing joint maintenance and stock provisioning policy for a multi-echelon spare part logistics network*", Journal of the Chinese Institute of Industrial Engineers, 23, 289-302

[3] Geiger, C. D., Martinez, O. E. & Lodree Jr, E. J. (2007), "*A multiobjective modeling approach for joint maintenance and spare parts inventory policy optimization*", IIE Annual Conference. Proceedings. pp. 1740-1745

[4] Scarf, P. A. & Cavalante, C. A. V. (2012), "*Spare parts provision for a maintained system with a heterogeneous*

lifetime", Advances in Safety, Reliability and Risk Management pp. 983-988

[5] Ilgin, M. A. & Tunalis S. (2007), "*Joint optimization of spare parts inventory and maintenance policies using genetic algorithms*", International Journal of Advanced Manufacturing Technology, 34, pp. 594-604

[6] Ferreria, R. J. P. & Wang, W. (2012), "*Spare parts optimisation subject to condition monitoring*", 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL, Vol 1, pp. 163-170

[7] Block, J., Ahmadi, A., Tyrberg, T., Peter Söderholm (2014), "*Part-out-based spares provisioning management: A military aviation maintenance case study*", Journal of Quality in Maintenance Engineering, Vol. 20 Iss: 1, pp.76 – 95

[8] Wijk, O., Andersson P., Block J. & Righard T. (2017), "*Phase-out maintenance optimization for an aircraft fleet*", International Journal of Production Economics, Vol. 188, pp 105-115

Prescriptive Maintenance: Building Alternative Plans for Smart Operations

Henrique Marques and Alessandro Giacotto

AerologLab-ITA, Aeronautics Institute of Technology, São José dos Campos, São Paulo/Brazil
E-mail: hmarques@ita.br, agiacott@ita.br

Abstract

Smart operations require the ability to generate alternative plans whenever a change in operations occurs in an unplanned manner. Alternate maintenance plans, in a highly dynamic context such as airline operations, require the ability to foresee small developments in terms of labor allocation, repairable items, and downtime, when and where they were not previously scheduled. In addition to being able to cause the disruption of the air transport network and consequent financial losses, it causes loss of trust in the company brand. Prescriptive maintenance is a potential technological response when using Artificial Intelligence to suggest alternative plans promptly so that decision-makers can reduce the impact on air operations. This paper proposes a framework for the construction of an integrated prescriptive maintenance solution that is certifiable by using auditable methods and extensible to complex systems of other industries. The adoption of prescriptive maintenance not only enhances the use of health management systems, widely available in modern aircraft fleets that have the potential to predict the remaining useful life of items of interest, but also allows identifying more than one response alternative to conflicts of interest in the conduction of the smart operations of air transport companies.

Keywords: Prescriptive Maintenance, Artificial Intelligence, e-Maintenance, Smart Operations, Information Fusion, Diagnosis, Prognosis

1 Introduction

The aircraft maintenance schedule is among the major decisions an airline has to make during its operation since it directly affects maintenance costs and aircraft availability. The development of aircraft maintenance schedule is a complicated task involving the combination of a variety of parameters such as mechanic and engineering maintenance labor, consumable materials and equipment spare logistics, and, from the operator point of view, the demand for air transportation [1]. On the other hand, more than 80% of passengers are now traveling on tickets priced at less than base fare and this downward pressure on revenues is leading many carriers to focus their attention on controlling maintenance costs while ensuring safety and airworthiness, complying thus with regulators requirements.

The predictive maintenance philosophy uses sensors and other monitoring strategies to gauge when complex systems actually will require maintenance. It is condition-based and relies on real-time health monitoring data available embedded in the system of interest. Data analytics procedures calculate the remaining useful life (RUL) of the equipment, allowing the operator to perform the maintenance before the faults happen, as the same philosophy of preventive maintenance, but not in a predetermined period (or cycles). The idea is to be able to establish the opportunistic moment to make the maintenance, postponing, or anticipating the moment, causing less downtime.

The prescriptive maintenance philosophy is a step ahead providing real-time adaptive recommendations, using artificial intelligence, about the tasks that should be done and keep updating as the operation continues. In other words, the prescriptive maintenance not only is based on the failures' prediction accordingly to the analysis of data patterns and trends, but also taking the specific company's maintenance process into consideration to provide detailed recommendations, and supports the solution-finding process.

This paper proposes a framework to tackle the scheduling maintenance problem adopting prescriptive maintenance strategy, which has to be certifiable, auditable, customized at the aircraft level [2], comprised of key indicators to measure its efficiency, capable of an adaptive context-awareness system and extensible to other industries.

2 Literature review

According to [3], key performance indicators are initially considered during the systematic inspection of failure patterns to understand their technical characteristics, criticality (downtime, costs, and occurrence) and their effects

on maintenance indicators. Besides the distinct task of timely executing a maintenance order, the measurement and control of the success of predictions and useful recommendations are not only used for the refinement of a machine learning method and a decision model, but also for the calculation of maintenance key performance indicators. By quantifying the defined key indicators, the current state maturity level is assessed, and weaknesses are identified. These results serve as a basis for the definition of company-specific target maturity state based on its strategic priorities.

As stated in [4], a feedback mechanism has to be added so that the recommendations are continuously improved based on cost-related sensor-generated data during action implementation. In this way, the phases of diagnostics, prognostics, and recommendations can be further investigated and enriched with new and improved algorithms.

Prescriptive maintenance is context-aware by nature. According to [5], in the prescriptive maintenance solutions proposed so far, the maintenance operators have to accept or reject the suggested tasks since their implicit experiential and "tribal" knowledge is necessary to a context-dependent maintenance decision, in a highly flexible operational environment. The information regarding the acceptance or rejection of the suggested tasks is then processed and analyzed as feedback information to improve the defined rules and measures continuously. In this sense, work is required to automatically adapt the proposed prescriptive maintenance set of rules to a dynamic operational environment, allowing the entire maintenance process to be less reliant on human knowledge and experience through machine learning approaches.

The prescriptive maintenance looks to the maintenance process powered by the Internet of Things (IoT), and its value is on the non-routine maintenance tasks when talking about heavy maintenance and in the challenge coming from the fleet schedule evolution in terms of line maintenance tasks. In terms of maintenance planning, while the Condition Based Maintenance (CBM) approach looks to the "when" and "why" to maintain, the prescriptive maintenance looks to the "what" and "where", defining the place and order of the maintenance tasks to accomplish and "who" should be engaged to execute each one [6]. To prescribe the way the tasks should be done is necessary to have inputs such as the MRO processes, available GSEs, spares, manpower and the condition of the system which is receiving the maintenance – resources availability status should be provided in real-time through the IoT. In this sense, the prescriptive maintenance is not a substitute or option to the Reliability-Centered Maintenance (RCM) paradigm, but the next step to its complete implementation on the field.

As mentioned by [7], nowadays, preventive maintenance does not consider operational contexts, such as environmental parameters, and these strongly affect components' lifetime. The ever-growing dynamism of the operational environment has also spurred the exploration of flexible on-line optimization approaches capable of reworking maintenance schedules incrementally (also known as rescheduling) to accommodate the contextual variability of the operation in different aspects.

When turning the focus to prescriptive maintenance, the most challenging paradigm encountered in practice remains tightly coupled to the match between the formulated optimization problem and the decision-making process that such a problem aims to model. Industries are complex environments where humans and machinery coexist and interact, often without holistically centralized management. It is often the case that actions triggered by a predictive model do not conform to the practical criteria or constraints under which such actions would be manually enforced, like time and cost. The developed models would fail to apply when deployed.

New working methodologies are needed to ensure that the hypothesis is aligned with the real operational requirements. Besides, such methods should also account for other practical aspects including the variability of metrics and constraints along time, cost implications of decisions made by the model, or the presence of conflicting objectives in the criteria guiding such choices.

In addition to it, as mentioned by [8], although the prescriptive maintenance concepts deliver promising results, the generalizability of the approach has not been tested yet and, as future work, the concept shall be validated in different industries to test its generalizability.

3 Prescriptive maintenance framework

In this section, the authors propose the framework, from now on called **S**mart **P**rescriptive **M**aintenance **F**ramework (SPMF) that can be used to build an effective prescriptive maintenance program for a fleet of commercial aircraft. The section starts with a description of the framework structure, then goes on with the description of the key enablers' elements and technologies requirements, continues with the description of key indicators to measure the efficiency of the framework, and describes the requirements of the adaptiveness and generalizability.

3.1 The SPMF framework

The framework is built on three domains of interest: The system's reliability, availability, maintainability, and safety (RAMS) factors, the operating environment in which the system is operated, and the maintenance environment that will perform the maintenance tasks required to restore the system to the necessary operating conditions, within its intended reliability characteristics (see fig. 1). Time and cost are treated as constraints in all domains.
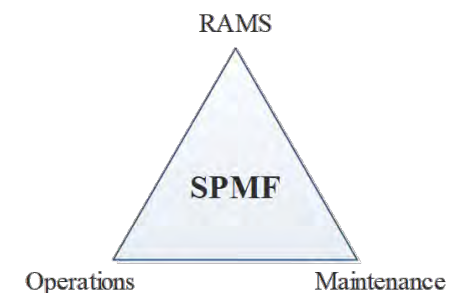

Figure 1: Framework domains of interest.

Each domain has essential information that will feed the framework through the various phases of information fusion until a group of maintenance actions is defined. The SPMF fusion flow diagram describes how the framework works as well as the inputs needed and the outputs expected in each phase.

As shown in fig. 2, the flow of information fusion is based on the capability to make the diagnosis, identifying the next operational demands to achieve the prognosis and based on

the maintenance capabilities, the prescription of the set of maintenance tasks.

From another perspective, the framework is constituted by five building blocks: inputs, fusion algorithm, output, the supported system, and the efficiency check process, as shown in fig. 3.
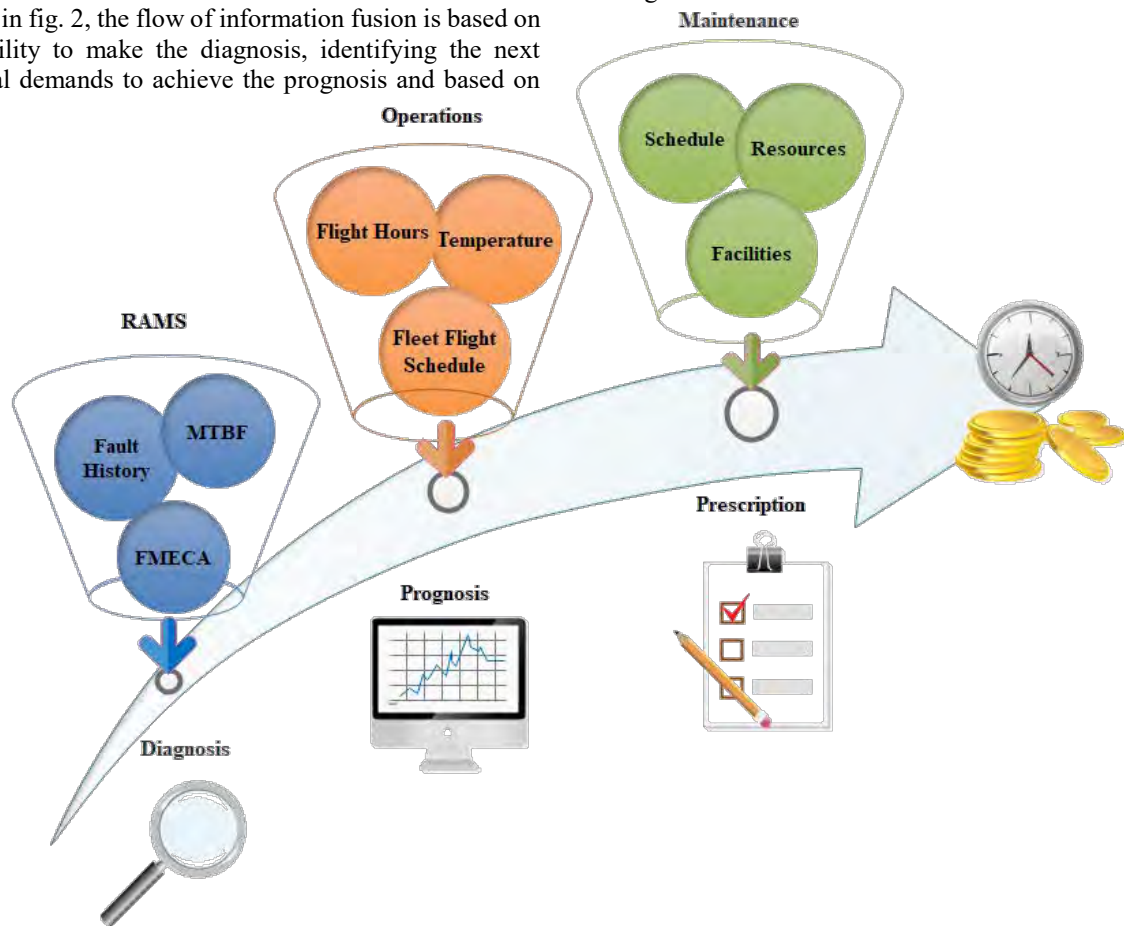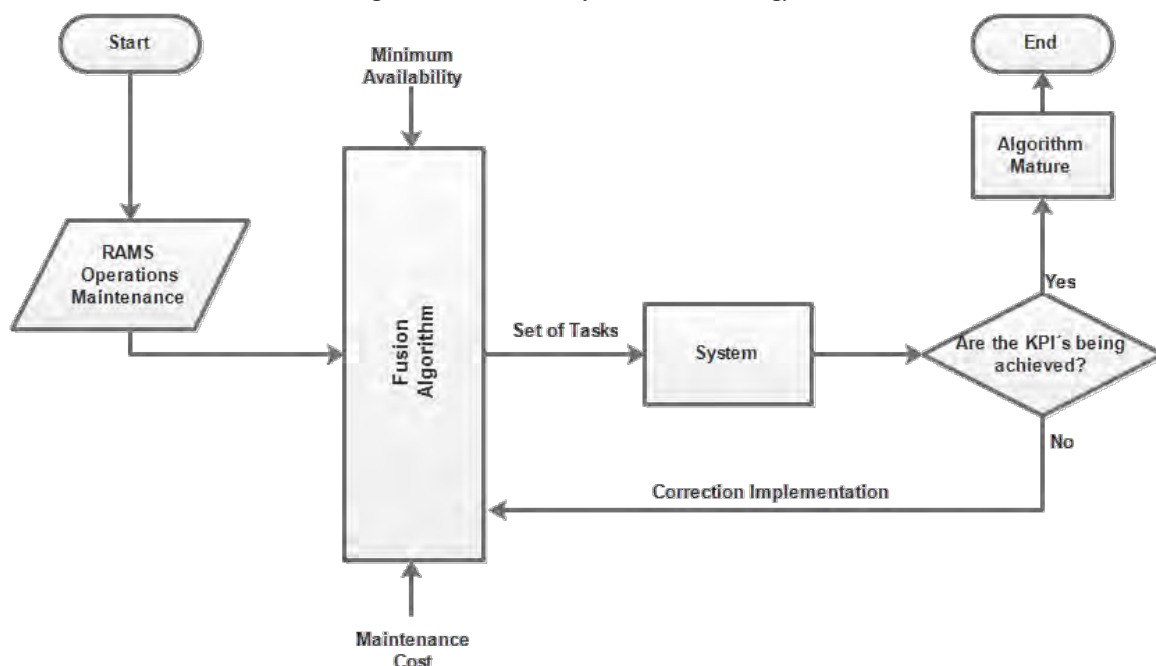

Figure 2: Framework fusion methodology.

Figure 3: Framework workflow.

3.2 Inputs

The inputs building block comprises the information needed to feed the fusion algorithm. Table 1 describes a simplified list of information from the RAMS, Operations, and Maintenance domains.

3.3 Fusion algorithm and its output

The fusion algorithm proposed is an Artificial Intelligence (AI) problem-solving agent [9]. Once all the inputs are informed, the algorithm starts to search for the most probable effective action or sequence of actions to solve the maintenance problem. Once the actions are selected, the algorithm recommends the best scheduling according to the minimum fleet availability requirement and the constraint of the direct (DMC) and indirect maintenance cost (IMC), providing, thus, the output of the fusion step.

3.4 Efficiency check

After the implementation of the tasks over the system, the algorithm itself verifies, during the efficiency check, if the maintenance actions are practical, the fleet availability is according to customer's requirement and maintenance costs minimized. If all these requirements are satisfied, then the selection of the same actions is reinforced, and the algorithm is considered mature. If at least one of the above conditions is not satisfied, then the probability attached to each task is redistributed, and a new set of actions is proposed in the next iteration.

3.5 Key indicators

As previously mentioned, the performance of the system is checked against two requirements: fleet availability and minimization of the system's maintenance cost. Both parameters are also key indicators of the algorithm performance: is the fleet availability above the threshold aligned with the operator`s intent? Are the DMC and IMC contained in comparison to historical maintenance cost or operator estimates?

3.6 Adaptiveness and context awareness

Adaptiveness and context awareness are essential characteristics of the SPMF. These characteristics allow prescriptive maintenance generation specifically for each "tail number". Some inputs, for example, the number of flight hours (FH) or flight cycles (FC), temperature || humidity, tribal knowledge, RUL and fault history, are a function of each operational environment, that is, one aircraft presents different values from another, and the same happens to maintenance teams and MROs. The prescriptive approach favors the utilization of the capability to identify the individual characteristics of each system, instead of using their average performance.

Thus, this approach makes use of some enabling technologies present in the industry 4.0 such as networking, availability of extensive data collected by sensors and the use of efficient algorithms necessary to support real-time maintenance execution monitoring [10].

Table 1: Framework inputs

Input	Description
Manufacturer Maintenance Plan	Initial maintenance plan provided by the manufacturer at the aircraft entry into service. It is developed according to MSG-3 principles, and it is updated yearly depending on the commercial agreement between the airliner and the operator
Flight Hour || Flight Cycle	The number of flight hours (FH) or flight cycles (FC) represent the usage of the aircraft, system or specific equipment. These parameters are useful since, very often, the reliability of the equipment is a function of the system usage
Temperature || Humidity	Environmental parameters have to be considered since operational conditions influence the failure rate of the equipment in use
Remaining Useful Life	The remaining useful life (RUL) curve of each aircraft of the fleet should be predicted, based on the expected fleet flight schedule. Thus, some prognostics and health management (PHM) system has to be present in the aircraft to help plan the maintenance schedule. If not present, the fusion algorithm would rely on fault history, making forecastings
"Tribal" Knowledge	"Tribal" knowledge is any information gathered from experience and lessons learned by the team who is responsible for the maintenance execution. This subset of data is often unstructured
Fault History	Historical data of all the equipment faults
Failure Mode and Criticality Analysis	Failure mode and criticality analysis (FMECA) particularly helpful to support inductive AI approaches
Mean Time Between Failures	The mean time between failures (MTBF) represents the expected time between failures of a system, during normal operation
Minimum Equipment List	It details which equipment is allowed to be inoperative without grounding the aircraft
Maintenance Cost	The maintenance cost, which includes man-hour and material related to all aircraft maintenance activities
Fleet Flight Schedule	The schedule of all the company´s flights to be accomplished

3.7 Extensibility to other industries

Ideally, the SPMF algorithm can be used in those fields where a complex system is in play. The critical point, in this case, is the adoption of parameters that can describe the specific complex system accurately. The current development proposes to extend the SPMF to the medical field (Medical-SPMF), being the human body the complex system to be "supported". Table 2 suggests, for each input listed in the tab. 1, the respective information applicable for the Medical-SPMF.

Table 2 – Medical SPMF inputs

Aircraft complex system input	Human body complex system input
Manufacturer Maintenance Plan	Formally recommended checks and their frequency
Flight Hour || Flight Cycle	Age
Temperature || Humidity	Body temperature
Remaining Useful Life	Prognostics
"Tribal" Knowledge	Informal lessons learned of the medical team
Fault History	Patient health history
Failure Mode and Criticality Analysis	Medical knowledge related to the specific organs failures being treated
Mean Time Between Failures	Not applicable
Minimum Equipment List	Medical statistics knowledge related to the organs failures being treated
Direct Maintenance Cost	Cost of treatment

The prescriptive maintenance has been deployed in industrial plants [11] and elevators and escalators maintenance [12] with improved results, but still in its early stages of development.

4 Conclusion and future work

This work proposes the Smart Prescriptive Maintenance Framework (SPMF) as a structure to implement an AI, context-aware, adaptive algorithm to support a fleet of commercial jets. The framework includes efficiency checks against operator requirements such as the fleet availability and the direct maintenance cost, and depends on specific operational, engineering and economic data to be assertive such as: manufacturer maintenance plan, FH, FC, temperature, humidity, RUL, "tribal" knowledge, fault history, FMECA, MTBF, MEL and DMC, among others.

The SPMF concepts apply to all industries where a complex system has to be supported, and in this paper is mentioned the extension to the medical field, as the human body the system to be supported. In this sense, it is recommended the adoption of suitable inputs and requirements for the treatment of a human body.

Some essential questions and open technical challenges have been identified, related to the AI modeling and the right metaheuristics to be used. It also remains to be clarified if there are key performance indicators other than cost and fleet availability to drive the prescriptive maintenance and check its efficiency. Regarding the inputs data, it is also unquestionable that such information will have to be consolidated in an interoperable data standard to be determined. Referring to the medical field still has to be addressed the best inputs to be adopted.

The next steps refer to the determination of the best metaheuristics to be used by the SPMF, identification of the most suitable KPI's and test of the algorithm framework in a real fleet. Successively, once the algorithm is mature and the right inputs are identified, tests will be conducted in the medical field to verify the extensibility of the SPMF concepts.

References

[1] C. Sriram and A. Haghani, "An optimization model for aircraft maintenance scheduling and re-assignment," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 1, pp. 29–48, 2003.

[2] Federal Aviation Administration, "Advisory Circular 43-218 Integrated Aircraft Health Management (Draft)," 09-Jul-2019. [Online]. Available: https://webcache.googleusercontent.com/search?q=cache:jOGamJpbqB4J:https://www.faa.gov/aircraft/draft_docs/media/afx/AC_43-218_Coord_Copy.pdf+&cd=1&hl=pt-BR&ct=clnk&gl=br. [Accessed: 19-Sep-2019].

[3] T. Nemeth, F. Ansari, W. Sihn, B. Haslhofer, and A. Schindler, "PriMa-X: A reference model for realizing prescriptive maintenance and assessing its maturity enhanced by machine learning," *Procedia CIRP*, vol. 72, pp. 1039–1044, 2018.

[4] A. Bousdekis, N. Papageorgiou, B. Magoutas, D. Apostolou, and G. Mentzas, "Enabling condition-based

maintenance decisions with proactive event-driven computing," *Computers in Industry*, vol. 100, pp. 173–183, 2018.

[5] K. Matyas, T. Nemeth, K. Kovacs, and R. Glawar, "A procedural approach for realizing prescriptive maintenance planning in manufacturing industries," *CIRP Annals*, vol. 66, no. 1, pp. 461–464, 2017.

[6] S. K. editor contributing, "RxM: What is prescriptive maintenance, and how soon will you need it?," *Plant Services*. [Online]. Available: https://www.plantservices.com/articles/2017/rxm-what-is-prescriptive-maintenance-and-how-soon-will-you-need-it/. [Accessed: 23-Sep-2019].

[7] A. Diez-Olivan, J. Del Ser, D. Galar, and B. Sierra, "Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0," *Information Fusion*, vol. 50, pp. 92–111, 2019.

[8] M. Abramovici, P. Gebus, J. C. Göbel, and P. Savarino, "Provider-Driven Knowledge Allocation Concept for Improving Technical Repair Tasks in IPS2 Networks," *Procedia CIRP*, vol. 64, pp. 381–386, 2017.

[9] S. Russell and P. Norvig, "AI a modern approach," *Learning*, vol. 2, no. 3, p. 4, 2005.

[10] A. Ceruti, P. Marzocca, A. Liverani, and C. Bil, "Maintenance in Aeronautics in an Industry 4.0 Context: The Role of Augmented Reality and Additive Manufacturing," *Journal of Computational Design and Engineering*, 2019.

[11] P. Deka, "Predictive and Prescriptive maintenance of manufacturing industry with machine learning," *Medium*, 25-Dec-2018. [Online]. Available: https://towardsdatascience.com/predictive-and-prescriptive-maintenance-for-manufacturing-industry-with-machine-learning-2078afa76bfb. [Accessed: 23-Sep-2019].

[12] ThyssenKrupp Elevator, "MAX - Predictive maintenance solution - ThyssenKrupp Elevator." [Online]. Available: https://max.thyssenkrupp-elevator.com/en/. [Accessed: 23-Sep-2019].

Fault Detection and Isolation Based on Bond Graph Models: Application to an Electromechanical Actuator

Gabriel dos Santos Sobral and Luiz Carlos Sandoval Góes

Department of Mechanical Engineering, Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, SP/Brazil
E-mail: sobral.gbrl@gmail.com; goes@ita.br

Abstract

Driven by the rapid development towards the concept of more electric aircraft in recent years, the power-by-wire technology, especially *electromechanical actuators* (EMA), is being progressively adopted with the promise of bringing performance improvements for future aircraft in different areas such as weight, maintainability and fuel consumption. However, for safety-critical applications like primary flight controls, the use of EMA introduces specific issues related to the actuator's response to failure. To that end, focusing on EMA fault diagnosis, this work proposes the development of a model-based quantitative *Fault Detection and Isolation* (FDI) methodology based on bond graph models. Firstly, an EMA behavioural model is created in bond graph, considering the actuator's most common failure modes: stator winding faults, backlash, jamming, and mechanical disconnection. Next, this behavioural model is applied to create a *Diagnostic Bond Graph* (DBG) model, which is based on the use of virtual residual detectors, or residual sinks, for numerical residuals generation, without the need for deriving symbolic *Analytical Redundancy Relations* (ARRs). These residuals are then evaluated with experimentally defined thresholds for fault detection, indicating if the system behaviour has departed from the acceptable operating range. Finally, the causal and structural properties of the bond graph are used to derive a *Fault Signature Matrix* (FSM) through the analysis of the DBG model causal paths. The FSM is employed for fault isolation, reducing the number of fault candidates, in order to isolate the component responsible for the faulty behaviour. Simulation results, using bond graph models implemented in the *20-sim* environment, show the successful detection and isolation of all EMA modelled faults using the proposed method.

Keywords: Bond Graph, Electromechanical Actuator, Fault Detection, Fault Isolation

1 Introduction

Although not new in the industry, the *more electric aircraft* (MEA) concept is receiving a growing interest in recent years, aiming at the development of safer, lower-cost, and greener technologies for the next generation aircraft. One of the main trends towards increasing the use of electrical power in aircraft systems resides in the adoption of *power-by-wire* (PBW) actuators. The PBW technology seeks the replacement of conventional *hydraulic servo actuators* (HSA) by electrically powered systems, with a series of advantages due to the exclusion of the heavy and bulky hydraulic system. With the recent engineering progresses in PBW actuation systems, these actuators are already found in service, with *electric-hydrostatic actuators* (EHA) and *electromechanical actuators* (EMA) becoming sufficiently mature to be introduced in the latest commercial and military aircraft [1–3]. Among the various kinds of actuators, EMAs are chosen for this study because of their growing role in aerospace applications.

Despite the potential benefits from EMA adoption such as the possible weight reduction, lower fuel consumption, and reduced maintenance costs, EMAs are not considered mature enough to completely replace conventional HSAs in safety-critical functions such as primary flight controls, where safety concerns due to the probability of jamming and backlash mechanical faults prevents the industry to embrace this technology. To that end, targeting the improvement on safety, reliability, availability and maintainability of complex engineering systems such as EMAs, it is crucial to develop autonomous supervisory systems that are able to quickly detect abnormal behaviours and isolate the causes of any fault which could seriously degrade the performance of the system.

EMA fault diagnosis poses an interesting research problem as the system is composed of several interacting components in different physical domains, resulting in intricate failure modes. Modelling a system with such characteristics can be accomplished by the use of the bond graph methodology, which is particularly suitable for modelling multidisciplinary engineering systems, and can also be used for the develop-

ment of powerful *Fault Detection and Isolation* (FDI) algorithms. From this perspective, the use of model-based and simulation-driven techniques can unquestionably provide engineers with efficient means to address these critical issues. Therefore, modelling and off-line simulation becomes a focus, not only to support conceptual design but also to the implementation of FDI systems before it is built and introduced in the industry.

The present work aims to develop a model-based quantitative FDI procedure, based on bond graph models, with application to electromechanical actuators. The proposed method will provide product development engineers with new tools to address EMA fault diagnosis, anticipating FDI since the conceptual design of the system. To accomplish that, an EMA bond graph model, representative of the main dynamic effects of the system, will be developed. That model shall also contemplate the most common failure modes of an electromechanical actuator, making it possible to simulate the system behaviour in different faulty scenarios. From this model, a FDI system based on the technique of *Diagnostic Bond Graph* (DBG) will be designed. The DBG will then be coupled with the behavioural model for off-line simulation, validating the applicability of the proposed method, revealing what faults can be detected and whether they can be isolated or not.

This work is structured as follows: Section 2 provides a review on FDI techniques and details the proposed bond graph model-based quantitative FDI, defining the steps for implementation of the DBG and construction of the *Fault Signature Matrix* (FSM). Section 3 presents the bond graph modelling of the EMA using an incremental approach, modelling the electric motor, mechanical power transmission and load coupling subsystems, which are connected to create the EMA behavioural model; the main EMA failure modes are also modelled and incorporated to the bond graph model; next, the DBG and FSM are generated according to the proposed method. Section 4 provides the implementation of the bond graph models and the FDI framework in the *20-sim* simulation environment, giving results from the actuator model simulation response and the residuals evaluation for EMA failure modes. The results are discussed in Section 5, which also proposes the topics for future research. Finally, Section 6 presents the conclusions of this work.

2　Fault Detection and Isolation

Before discussing any diagnostic system details, it is necessary to clearly define what is understood as a *fault*. In the aeronautical industry a standard definition is given by *Aerospace Recommended Practice* (ARP) 4761, *Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment*, from SAE International [4]:

> *Fault*: an undesired anomaly in an item or system that may lead to a failure.
>
> *Failure*: a loss of function or a malfunction of a system or a part thereof.

In this work's context, a more precise interpretation is necessary and the *fault* definition commonly accepted in the academic literature will be used. This definition is given by Venkatasubramanian [5], as follows:

> The term *fault* is generally defined as a departure from an acceptable range of an observed variable or a calculated parameter associated with a process. This defines a *fault* as a process abnormality or symptom, such as high temperature in a reactor or low product quality and so on.

Once a fault is clearly defined, the diagnostic system can be divided in four distinct phases [6] as follows: fault detection, fault isolation, fault identification, and fault accommodation. This work will focus on FDI algorithms, comprising the two first steps of this process.

The *fault detection* stage is concerned in determining if the system behaviour has departed from the acceptable range of operation, raising a fault alarm in case of unacceptable behaviour. The second stage, *fault isolation*, has the objective of reducing the number of fault candidates, using one or more decision procedures to isolate the component responsible for the fault behaviour.
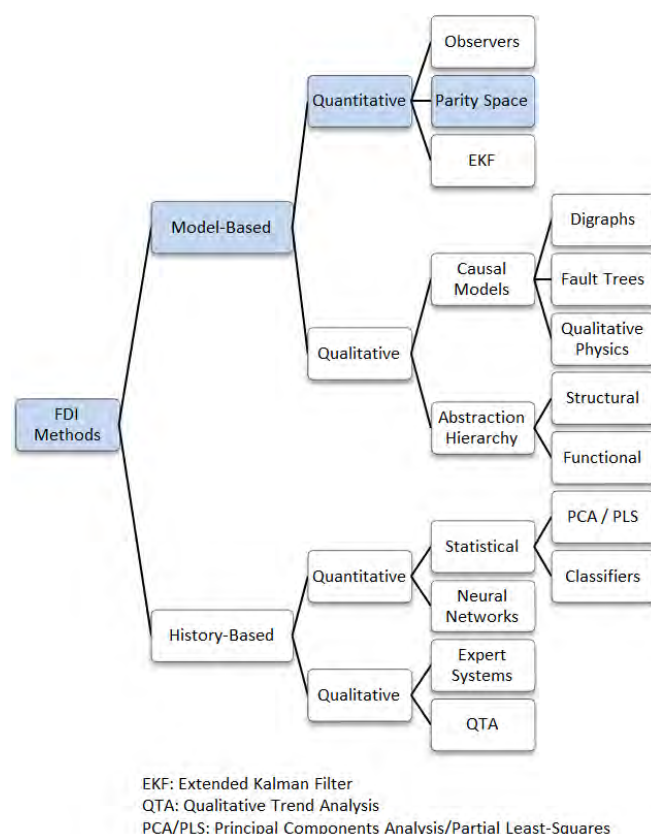

EKF: Extended Kalman Filter
QTA: Qualitative Trend Analysis
PCA/PLS: Principal Components Analysis/Partial Least-Squares

Figure 1: Classification of FDI algorithms. Adapted from [5]
·

In order to perform these tasks, any FDI algorithm will need *a priori* knowledge about the system, regarding the set of possible failures and the relationship between the observed variables and these failures, i.e. a set of causal relationships. The

FDI algorithms may be broadly classified in two classes, depending on how it develops the *a priori* knowledge about the system. Model-based algorithms derive its knowledge from a fundamental understanding of the physics of the process and its causal relationships. On the other hand, history-based algorithms will rely on past experience with the process, requiring the availability of a large amount of historical data. Regardless of the origin of that knowledge, it can be analysed on a quantitative or qualitative approach. Different FDI methods have been developed in the literature, these methods are presented in fig. 1, following the discussed classification.

This work explores a model-based quantitative FDI method, using the parity space approach. These techniques are based on a comparison of outputs from the real process with those from a behavioural model that describes the non-faulty process, creating redundancy relations. Any discrepancies between the two are described by parity vectors, or residuals. Every residual is expected to be zero when no fault is present and distinguishably different from zero when a fault occurs.

In effect, a perfectly accurate mathematical model of the system is never available. Therefore, there will always be a mismatch between the measured system variables and its mathematical model, even under normal operation, i.e. residuals will always be different from zero, even in the absence of faults. Because of that, it is necessary to establish thresholds for the acceptable levels of each residual. Once the residuals are generated, they are monitored for exceedances over these thresholds, in which case a fault alarm is declared. The subsequent analysis of the residuals leads to fault isolation. The basic scheme for model-based quantitative FDI is given in fig. 2 showing the two basic steps of residual generation and residual evaluation, to be detailed in the following sections.
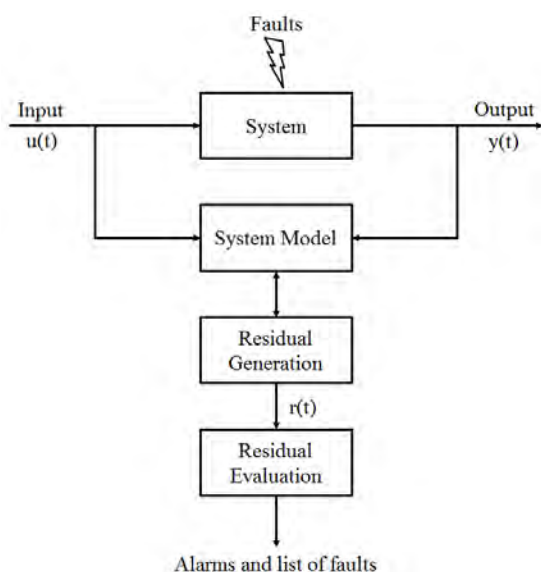

Figure 2: A general scheme for quantitative model-based FDI.

2.1 Residual Generation: Diagnostic Bond Graph

A proper residual design method is the best way to achieve the requirements of isolability and robustness on the quantitative model-based FDI. Appropriate residual generation, with robust and structured residuals, will make the residual evaluation step relatively easy.

The bond graph formalism is a knowledge representation form suitable for developing models of multi-energy domain engineering systems, being capable of holding a large amount of structural, functional and behavioural information about the system and describing essential causal relationships. These are reasons that make bond graph models a powerful tool for the construction of model-based FDI systems.

The most common method for generating residuals from bond graph models is through the analytical evaluation of the model physical constraint laws. These constraint laws are written symbolically, eliminating the unknown variables in a way that they will only depend on the system inputs, the parameter values and the measurements from sensors. Such equations are termed *Analytical Redundancy Relations* (ARR) and the numerical evaluation of ARRs give the residuals. Obtaining ARRs from bond graph models can be a straightforward and structured process, following the causality inversion approach, well documented in the literature [6]. However, a closed symbolic expression cannot be derived for all kinds of processes with all kinds of instrumentation because it may not be possible to, in certain cases, eliminate all unknown variables from the model through symbolic algebra, e.g. when the bond graph model contains algebraic loops that cannot be eliminated.

Samantaray [6] proposes a modified method for direct numerical evaluation of the residuals, introducing the DBG model, with no need for symbolically derivation of ARRs, being applicable to all situations whether the set of constraints can be symbolically resolved or not. The proposed method leads to the same set of residuals which can be obtained through classical means, if the equations are symbolically resolvable. For the construction of the DBG model, preferred differential causality is assigned to the storage elements of the bond graph model, along with the inversion of sensor causalities. This means sensors will impose the measured variable on the system, i.e. sensors become sources.

For instance, given an effort detector, **De**, the flow in the bond connected to this sensor is equal zero. When the causality is inverted for this sensor it becomes an effort source, imposing the measurement from the real process, and the expression for the source loading, the flow variable, is equated to zero. This expression is indeed, a residual, and its output is measured by a virtual flow sensor, which has only computational existence. All virtual sensors measuring residuals are presented with a superscript '*': **Df*** or **De***. The bond graph for substitution of the inverted causality effort sensor, containing the virtual residual sensor, is presented in fig. 3.

Similarly, the substitution proposed for the inverted causality flow sensor is dual to the one discussed for the effort sensor. Using this method there will be a virtual residual sensor for
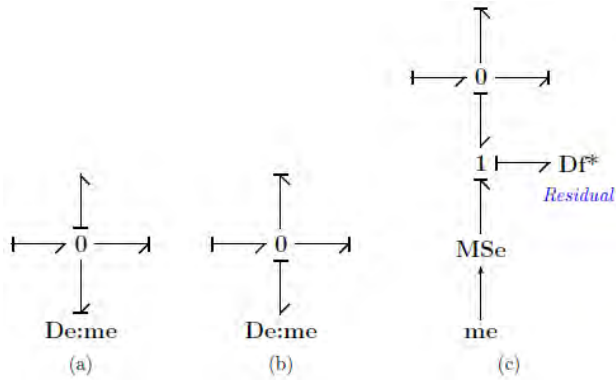
Figure 3: (a) Effort sensor; (b) Effort sensor in inverted causality; (c) Corresponding substitution and residual virtual sensor. Note: me *represents a generic measured effort variable.*

each real sensor on the system model. Furthermore, for a sensor in inverted causality, there exists at least one causal path to it which is different or does not exist when compared to the causal paths to any other sensor in the inverted causality. This proves that all the virtual residual sensors have distinct signatures [6]. A similar approach, which refers to these virtual sensors as residual sinks, has been developed by Borutzky [7].

2.2   Residual Evaluation: Fault Signature Matrix

Once the residuals are generated, the next step is residual evaluation, which aims to isolate the faulty components in the system's model. The proposed method uses a FSM for that. Each residual is sensitive to some faults and insensitive to some others. For good isolability it is important that different residuals become sensitive to different faults. If a residual fault signature, i.e. the faults to which this residual is sensitive, is different from the fault signatures of any other residuals, the residual is termed *structurally independent* of the others. The DBG approach guarantees that the number of structurally independent residuals obtained is equal to the number of sensors in the system [6].

The FSM, *S*, will describe the structural sensitivity of each residual to possible faults in the system's model, forming a structure that links the residuals to discrepancies in each component of the system. The elements in the *S* matrix are determined from the analysis presented in eq. 1.

$$S_{ij} = \begin{cases} 1, & \text{if the } i\text{-th residual is sensitive to} \\ & \text{faults in the } j\text{-th component;} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Additionally, two indexes are assigned to each component on the FSM. The monitorability index M_j will be 1 if at least one residual is sensitive to it, in other words, the component's fault is monitorable. The isolability index I_j is set to 1 when the component's fault signature is different from fault signatures of all other components, that is, the fault can be isolated. Since the DBG approach does not necessarily have closed form ARRs, the fault signatures are obtained by analysis of

the bond graph causal paths to each residual.

Once the residuals are obtained from the DBG model and the FSM is constructed, a decision procedure shall be applied to the set of evaluated residuals in order to generate the alarms. A threshold ε_i has to be defined for each residual r_i. The necessity for this non-zero threshold bounding the residual is due to process uncertainties, modelling errors and measurement noises.

3   Application to an Electromechanical Actuator

EMAs can have different configurations, adapted to the requirements of particular operations. As a more general application, the linear EMA architecture is mainly considered and discussed in this study. The linear EMA can have a gearbox or not, in order to reduce the electric motor speed before the power is transmitted to the load. The gear-drive EMA is more susceptible to certain single-point failures that would generate an undesired mechanical jamming, bringing complications to critical applications. On the other hand, direct-drive EMAs are less prone to mechanical jamming, have less probability of actuator damping loss, increased efficiency and reliability, and reduced system inertia due to the elimination of the in-between gearbox [1]. However, these advantages do not come for free, requiring a high-performance servo motor with high torque density and reliability, and in the absence of the gearbox buffering effect, any load disturbance can be directly reflected to the motor and the control system. For these reasons the direct-drive EMA will be studied in the present work.

The schematic for a direct-drive linear EMA is presented in fig. 4. Electric power is provided by the aircraft electrical buses and delivered to the actuator by the *Power Drive Electronics* (PDE) according to the command signal from the *Actuator Control Electronics* (ACE). The electric power is transformed into rotational mechanical power by the *Electric Motor* (EM), and then, converted into translational mechanical power by a *Mechanical Power Transmission* (MPT) mechanism, transferring power to the control surface.
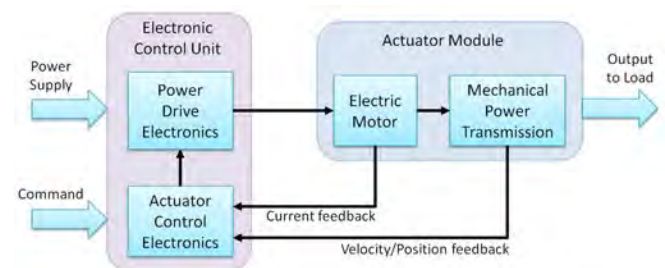

Figure 4: Schematic of a direct-drive linear EMA actuation system.

This work will concentrate on fault diagnosis of the EMA actuator module, centred on the MPT but also considering a simplified EM and load coupling models. The actuator model is proposed in the next section.

3.1 EMA Behavioural Model

Since the EMA can be divided into different interconnected subsystems, the bond graph model can be constructed by an incremental approach, modelling each of the subsystems and then coupling these sub-models into the complete EMA model. The non-causal word bond graph model for the EMA, neglecting the control signals, is presented in fig. 5.
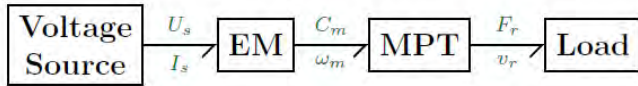

Figure 5: Word bond graph for a direct-drive EMA actuation system.

The EM performs electromechanical power conversion, being the key enabler for all kinds of PBW systems in the MEA concept. This fundamental device in currently implemented on EMAs using a high-performance *permanent magnet synchronous motor* (PMSM), which reduces the complexity of actuator control design [8]. A simplified model of the EM is based on a perfect DC motor, enabling perfect power transformation between the electrical and mechanical energy domains. On the bond graph formalism it is represented by a power gyrator (**GY**), which relates the flow variable of electric current (I_m) on one port to the effort variable of electromagnetic torque (C_{em}) on the other port through the motor torque constant (K_m). The major effects in the electric domain of the EM are the cooper loss, which can be defined as a voltage drop due to the winding resistance (R_w) and modelled by a **R** element in bond graph, and the electrical storage effect due to the winding inductance (L_w) and modelled by an **I** element. On the mechanical side of the EM the more important effects are the rotor inertia (J_m) modelled by an **I** element and the mechanical friction (R_{fm}) modelled by a **R** element

MPT is responsible to transform rotational into translational mechanical power, which is achieved using a roller-screw device where the torque (C_m) and rotational velocity (ω_m) from the EM are converted into force (F_r) and translational velocity (v_r) of the actuator rod. A perfect roller-screw mechanism can be modelled by a bond graph power transformer element (**TF**) that relates the rotational power variables (C_m, ω_m) to the translational ones (F_{ns}, v_{ns}) according to the transformation ratio ($2p$) of the screw pitch. Three additional significant effects are considered in the basic MPT model: roller-screw inertia, mechanical friction loss and compliance effects. For simplification the inertia is assumed to be concentrated on the translational side of the MPT, represented by the rod mass (m_r) and modelled by a bond graph **I** element. The power loss is modelled as a viscous friction force (F_{fr}), proportional to the operational velocity and the viscous coefficient (R_{fr}), being represented by a bond graph **R** element. The compliance effect is due to the deformation of MPT mechanical components under stress and can be modelled by associated spring and damping effects. The linear spring effect is expressed by an elastic force (F_e) proportional to the relative displacement and the compliance stiffness (k_{cr}), while the damping force (F_d), also a linear function, is proportional to the relative

velocity and the damping coefficient (R_{cr}). The spring and damping effects are respectively associated to a bond graph capacitance (**C**) and resistance (**R**).

The actuator load, composed of the flight control surface, is simply modelled as an equivalent translating mass (m_L) represented by a bond graph **I** element. An effort source (**Se**) is also included to represent external disturbance aerodynamic forces (F_{aer}) applied over the control surface. In addition, a compliance effect similar to the one included on the MPT model is considered here, representing the spring and damping effects associated to the mechanical link between the control surface and the actuator rod.

The complete EMA behavioural bond graph model, which is presented in fig. 6, is created by combining the EM, MPT, and load coupling sub-models. The model also includes four sensors, which are normally used in EMA systems: EM current (i_m), EM rotational velocity (ω_m), MPT rod translational velocity (v_r) and load translational velocity (v_L).
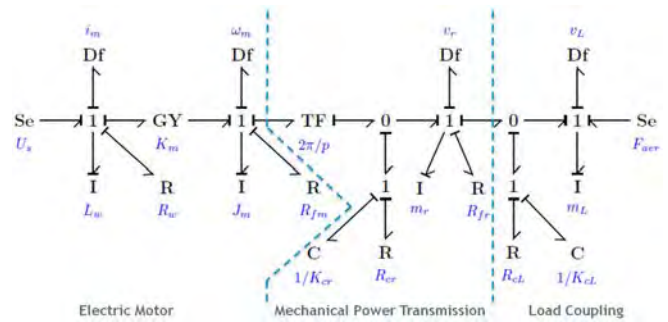

Figure 6: EMA behavioural bond graph model (integral causality).

The next step is to identify the most common failure modes that may affect the system and include these faults into the EMA behavioural model. Mechanical faults are credited to be the main source of concern for EMA systems. The rigorous conditions of aerospace applications lead to backlash, vibration and mechanical jamming on the roller-screw mechanism of the MPT, often caused by excessive loads, environmental factors, lubrication issues, and manufacturing defects. The second source of concern is the EM, where the high operational demands often leads to increased temperature within the motor housing, which may contribute to winding circuit faults. This work will consider the most probable failure modes [9] presented and described in the sequence:

Stator winding failure: higher temperatures will mostly affect the EM, leading to insulation deterioration on the stator winding and possibly resulting in short or open circuits which will directly affect the EM performance. The bond graph model already has a component representing the stator winding resistance (R_w), what makes it easier to insert a short or open circuit fault by simply changing the resistance value.

Backlash: the excessive wear of the mechanical parts on the MPT, specially in the contact areas between the mov-

ing parts will contribute to increased backlash, generating the so-called lost-motion and significantly affecting the dynamic performance of the actuator. When it exists, backlash can be considered as a compliance effect on the MPT model, an elastic force proportional to relative motion. Given that, a more realistic compliance model is necessary in order to represent not only the pure spring and damping effects but also the possible backlash and preload. Fu [8] proposes such a model, presented in fig. 7, which is incorporated in the bond graph EMA behavioural model.

Jamming: results in the complete blockage of the actuator movement and can be modelled on the MPT by increasing the mechanical friction parameters in order to force stiction. On the developed MPT bond graph model this can be simply implemented by increasing the viscous friction coefficient (R_{fr}).

Mechanical disconnection: this fault is not related to EMA internal components but to the mechanical linkage between the actuator and the flight control surface, modelled here as a compliance effect on the load coupling. The mechanical disconnection, also called free-run, occurs in the case of a complete structural failure on the mechanical linkage, decoupling the actuator load.
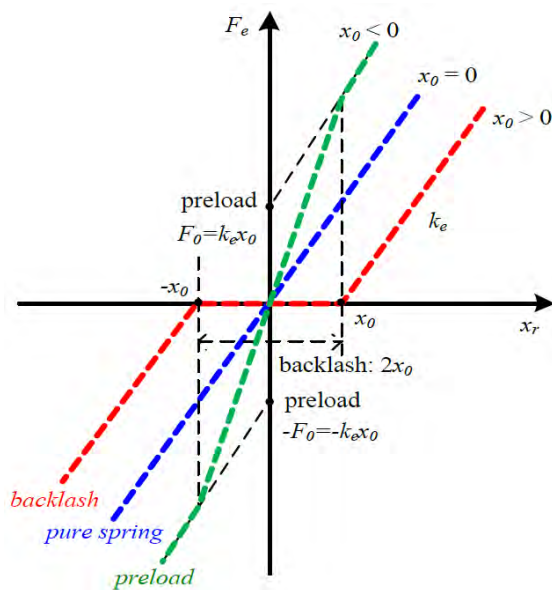

Figure 7: Compliance model with backlash and preload effects [8].

3.2 EMA Diagnostic Bond Graph Model

The DBG model is directly derived from the behavioural model performing the causality inversion of detectors, i.e. putting the model in preferred derivative causality and replacing the sensors by corresponding sources with their virtual residual sensors, or residual sinks. The diagnostic model will receive the signals coming from the behavioural model flow sensors and give the residuals in the virtual effort detectors

(**De***). This approach enables the numerical residual evaluation without the need of symbolically deriving ARRs. It also guarantees that all virtual residual sensors will have different sets of causal paths, meaning that residuals obtained are always structurally independent, as described in Section 2.1.

The framework for FDI simulation involves coupling the behavioural and the diagnostic models. The resulting coupled model is presented in fig. 8. To simulate real instrumentation all sensor outputs from the behavioural model are contaminated with noise (W_1 to W_4) before going to the diagnostic model. The noise model is based on white Gaussian noise with magnitude chosen at 1% of each signal measurement. Finally, the diagnostic model evaluates the signals coming from the behavioural model and gives the residuals (res_1 to res_4).

Once we have the residuals available from the DBG model, the next step for FDI is to obtain the fault signature matrix (FSM). This is accomplished through the analysis of the causal paths leading to each residual detector, resulting in the FSM presented in tab. 1, where the additional columns D and I indicates the detectability and isolability indexes, respectively. It is seen that faults in all of the components are detectable, i.e. at least one of the residuals is sensitive to that component, however, only fault in the rotational velocity sensor ω_m can be isolated, the other components do not have distinct fault signatures. Although most of the faults cannot be isolated to a single component, their signatures will indicate well defined segments of the system with a few possible faulty components.

Table 1: Fault Signature Matrix for the EMA system.

Subsystem	Component	res_1	res_2	res_3	res_4	**D**	**I**
EM electrical	U_s	1	0	0	0	1	0
	L_w	1	0	0	0	1	0
	R_w	1	0	0	0	1	0
EM mechanical	K_m	1	1	0	0	1	0
	J_m	0	1	0	0	1	0
	R_{fm}	0	1	0	0	1	0
MPT roller-screw	p	0	1	1	0	1	0
	R_{cr}	0	1	1	0	1	0
	k_{cr}	0	1	1	0	1	0
	M_r	0	0	1	0	1	0
	R_{fr}	0	0	1	0	1	0
Load coupling	R_{cL}	0	0	1	1	1	0
	k_{cL}	0	0	1	1	1	0
Load	M_L	0	0	0	1	1	0
	F_{aer}	0	0	0	1	1	0
Sensors	i_m	1	1	0	0	1	0
	ω_m	1	1	1	0	1	1
	v_r	0	1	1	0	1	0
	v_L	0	0	1	1	1	0

4 Simulation and Results

The behavioural and diagnostic bond graph models developed in Section 3 are implemented and simulated using the *20-sim* software, by *Controllab Products* [10]. In order to evaluate the aspects of the main dynamic effects of the proposed model, the EMA is simulated with real system parameters, gathered from industry product specifications and from the literature. These parameters are listed in tab. 2.
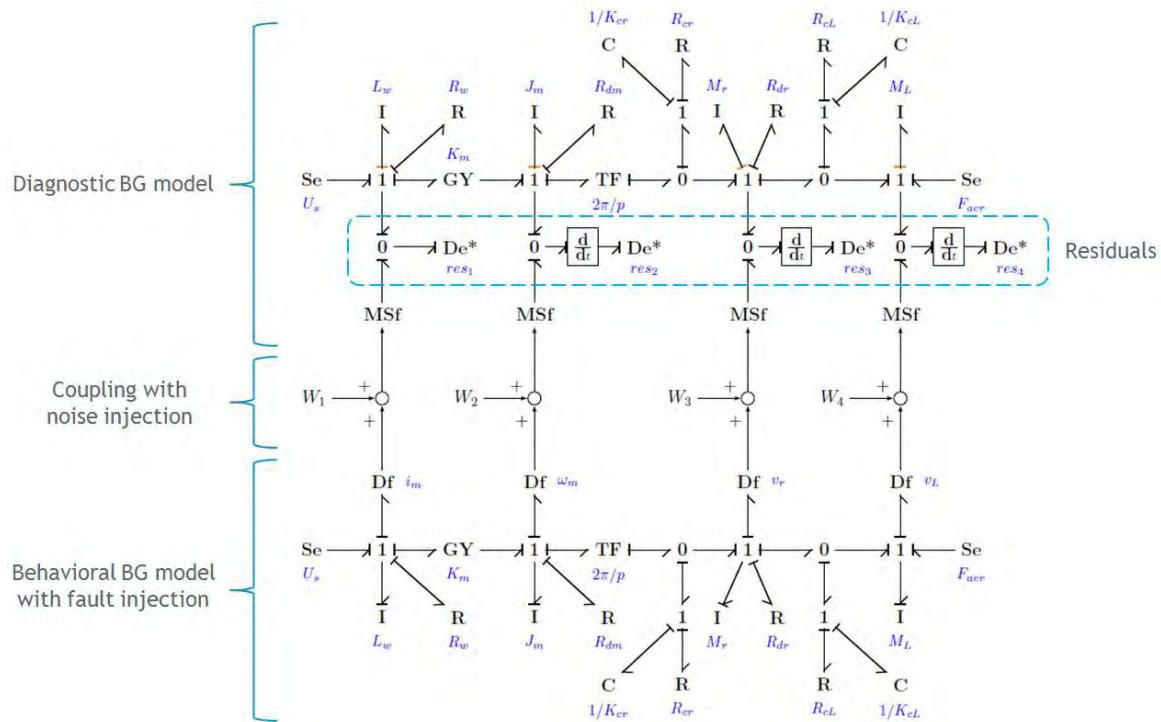
Figure 8: Coupling of EMA behavioural and DBG models.

Table 2: EMA model parameters.

Parameter	Value	Unity	Description
J_m	0.001 279	$\mathrm{kg\,m^2\,rad^{-1}}$	EM rotor inertia
K_{cL}	3×10^8	$\mathrm{N\,m^{-1}}$	Load compliance stiffness
K_{cr}	3×10^8	$\mathrm{N\,m^{-1}}$	Roller-Screw compliance stiffness
L_w	3	mH	EM stator winding inductance
M_L	600	kg	Load reflected mass
M_r	1	kg	Roller-screw rod mass
p	2.54	mm	Lead of roller-screw
R_{cL}	1×10^4	$\mathrm{N\,s\,m^{-1}}$	Load compliance damping
R_{cr}	1×10^4	$\mathrm{N\,s\,m^{-1}}$	Roller-screw compliance damping
R_{fm}	1×10^{-3}	$\mathrm{N\,s\,m^{-1}}$	EM friction coefficient
R_{fr}	1×10^4	$\mathrm{N\,s\,m^{-1}}$	Roller-screw friction coefficient
K_m	0.46	$\mathrm{N\,m\,A^{-1}}$	EM torque constant
R_w	1.5	Ω	EM stator winding resistance
U_s	115	V	Bus voltage

Source: [8, 11, 12].

First, the behavioural model for the EMA is implemented and simulated, validating the system open-loop response and its behaviour in all modelled failure modes: backlash, jamming, mechanical disconnection, stator open-circuit and stator short-circuit. The diagnostic model is then implemented and coupled with the behavioural model, following the FDI framework presented in fig. 8. The coupled model is simulated in all failure modes, generating the residuals, which are evaluated for fault detection and isolation.

For normal operation, the system is simulated in open-loop, with a 115 V step input on the voltage source (U_s) and a step aerodynamic force disturbance (F_{aer}) of 3 kN. Figure 9 gives the EMA simulated open-loop response for normal operation, which is as expected. The load displacement velocity (v_L) stabilizes at $0.1\,\mathrm{m\,s^{-1}}$ after a fast transient oscillation. It is also observed that the stiffness and damping compliance effects of the MPT and load coupling contribute significantly to increasing the load velocity oscillations. The electric motor stator winding current (i_m) presents a transient peak of about 60 A, which shall be controlled in a closed-loop implementation. Neither of the residuals will exceed the experimental thresholds, showed in fig. 10, indicating the system is operating without failures.

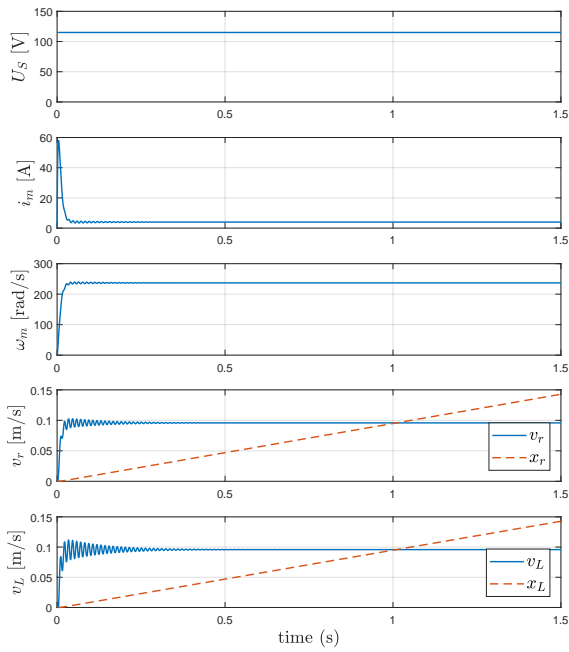The EMA response to failure is evaluated by injecting the

Figure 9: Simulation results for the EMA behavioural model in normal operation.
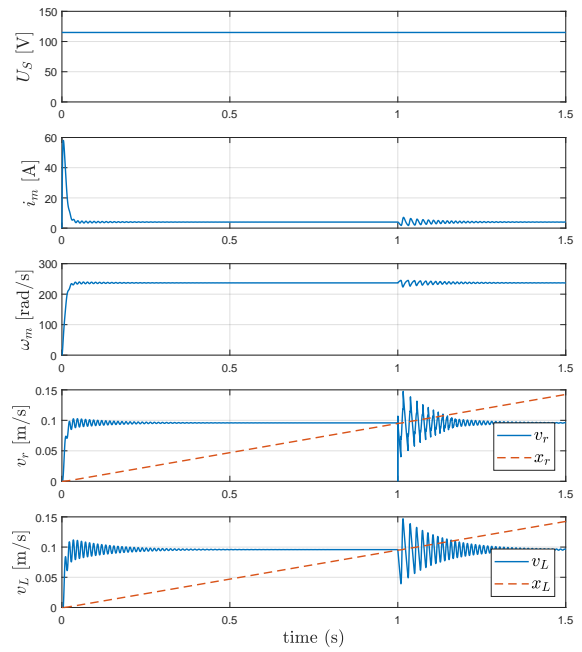

Figure 11: Simulation results for the EMA behavioural model with 0.3 mm *backlash.*
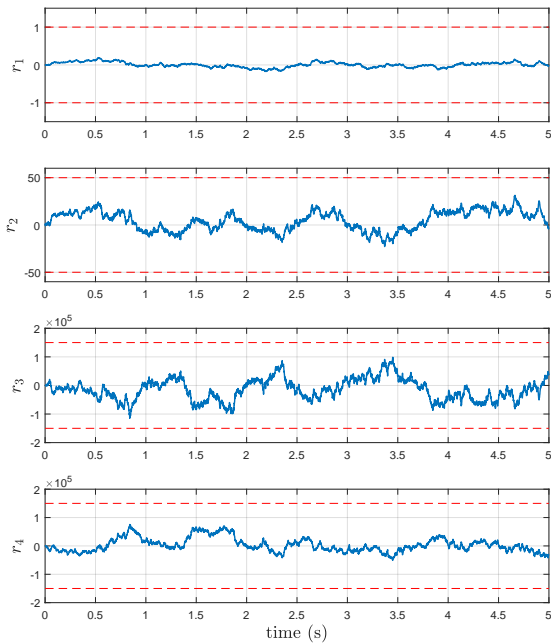

Figure 10: Residuals for the simulation of EMA coupled models in normal operation. Residual thresholds are displayed in red dashed lines.
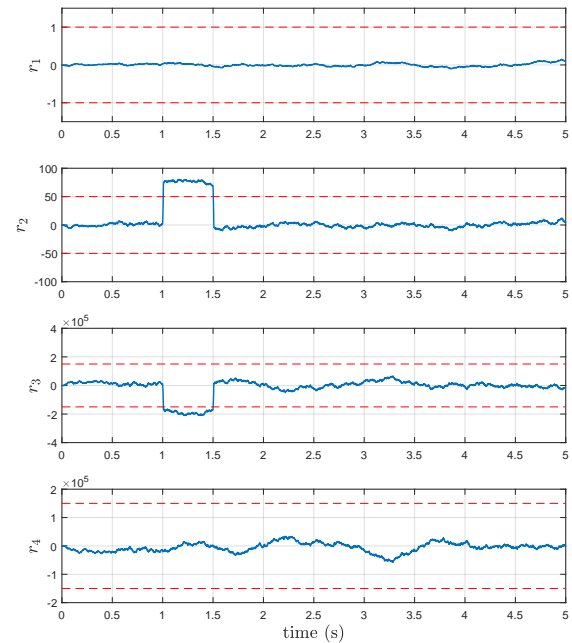

Figure 12: Residuals for the simulation of EMA coupled models with 0.3 mm *backlash. Residual thresholds are displayed in red dashed lines.*

faults modelled in Section 3.1. The model is simulated with the same parameters as before and the faults are injected, one at a time, as a step input always at 1 s of simulation.

As an example, the simulation results for the backlash fault are presented in fig. 11. The compliance effects on the mechanical power transmission are implemented in *20-sim* ac-

Table 3: FDI results summary.

Fault	Fault Injection	Coherence Vector	Fault Candidates
Normal operation	-	$[0,0,0,0]$	-
Open-circuit	R_w	$[1,0,0,0]$	U_s, L_w, R_w
Short-circuit	R_w	$[1,0,0,0]$	U_s, L_w, R_w
Backlash	R_{cr}, K_{cr}	$[0,1,1,0]$	p, R_{cr}, K_{cr}, v_r
Jamming	R_{fr}	$[0,0,1,0]$	M_r, R_{fr}
Mechanical disconnection	R_{cL}, K_{cL}	$[0,0,1,1]$	R_{cL}, K_{cL}, v_L

cording to the backlash model given in fig. 7, the results show the actuator response for a 0.3 mm backlash applied at 1 s. Backlash will occur due to excessive wear in the contact surfaces of the roller-screw mechanical system, creating a lost-motion effect and introducing oscillations that will propagate throughout the entire system, affecting the EMA dynamic performance.

The residuals evaluation, presented in fig. 12 show that the defined threshold is crossed for residuals r_2 and r_3. Giving the coherence vector $C = [0,1,1,0]$, which according to the FSM, means there is a faulty component in the MPT mechanism (parameters p, R_{cr} or K_{cr}) or in the actuator velocity sensor (v_r), which represents a correct diagnostic.

All failure modes were simulated and the DBG residuals evaluation pointed to the correct diagnostics, according to the fault signature matrix. Table 3 presents a summary of the simulation results.

5   Discussion

The recent progresses in PBW technology is driving the EMA as the next generation of actuation systems, already replacing conventional hydraulic actuators in commercial and military aerospace applications. However, their use in safety critical systems, such as primary flight controls, still faces some obstacles until it proves to be reliable and to behave safely, especially in fault scenarios. One of the research topics that can help on the evolution and acceptance of EMA in future aircraft architecture is the development of condition monitoring and diagnostic systems, such as the proposed FDI methodology.

The developed bond graph models allowed the simulation of an EMA system in different faulty conditions, providing an easy to implement diagnostic system which can be really helpful in the early diagnosis of EMA faults, improving its acceptance in aerospace applications. These models can be used for evaluation during the early process of EMA design, without the need of any real prototypes and expensive test sets.

The models were implemented in the *20-sim* simulation environment, as presented in Section 4. First, the open-loop behavioural model was simulated to verify its response. The use of a behavioural model as a tool for fault simulation brings the advantage of introducing any kinds of faults without the risks associated with the real process. The DBG model was then coupled with the behavioural model, generating numeri-

cal residuals which were processed and evaluated for the definition of experimental thresholds. The diagnostic result was correct to all simulated failure modes, confirming the feasibility of the DBG method proposed in the literature [6] and its application to EMA development.

Opportunities derived from this work, aiming the further development of the proposed method are listed below, as suggestions for future research topics:

1. Refinement of the EMA model, including the power drive electronics and actuator control electronics models, with its respective failure modes;

2. Implementation of the complete EMA controller and analysis of the system closed-loop response, specially on the impacts of the control loop over the response to different failure modes;

3. Analysis of multiple-fault scenarios, applying multiple-fault isolation techniques based on parameter estimation by Gauss-Newton least squares output error minimization [13, 14];

4. Development of *Bicausal Bond Graph* (BBG) models for analysis and improvement of sensor placement and residuals generation [15];

5. Use of uncertain bond graph models for robust FDI residuals generation, and sensitivity analysis of the residuals in the presence of parameter uncertainties [16];

6. Study of reconfigurable *Fault Tolerant Control* (FTC) techniques for fault accommodation / passivation [17];

7. DBG online implementation experiment, the outputs of the behavioural model are to be replaced by measured outputs from the real EMA system.

6   Conclusions

This paper presented a bond graph model-based approach to FDI, showing its application to an electromechanical actuator, relevant to MEA actuation systems. The bond graph modelling formalism is used as the base of the DBG FDI method being proposed. The DBG makes use of virtual residual detectors, also called residual sinks, enhancing the availability of information for the diagnostic system by generating all the structurally independent residuals in numerical form, as the numerical evaluation of ARRs. This is important in situations

where symbolic ARRs cannot be obtained due to the presence of algebraic loops in the system bond graph model. The DBG is also used for the construction of a FSM by analysis of the causal paths leading to the residuals.

The DBG approach was proven as a powerful tool, suitable for the implementation of FDI on complex multidisciplinary systems, such as the EMA. It is concluded that there are significant advantages in using the bond graph formalism, when compared to other modelling methods, as a tool for developing model-based quantitative FDI, confirming its capacity on the diagnostic of EMA faults. Still, the FDI approach proposed here is not restricted to this use, being applicable on the diagnostic of any complex interdisciplinary system such as those commonly found on the aerospace industry.

References

[1] Guan Qiao, Geng Liu, Zhenghong Shi, Yawen Wang, Shangjun Ma, and Teik C Lim. A review of electromechanical actuators for More/All Electric aircraft systems. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2017.

[2] Michel Todeschi and Laura Baxerres. Health monitoring for the flight control EMAs. *IFAC-PapersOnLine*, 28(21):186–193, 2015.

[3] Jean Charles Maré. *Aerospace actuators 2 : signal-by-wire and power-by-wire*. John Wiley & Sons, 2017.

[4] SAE. *ARP4761, Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment*. SAE International, 1996.

[5] Venkat Venkatasubramanian, Raghunathan Rengaswamy, Kewen Yin, and Surya N. Kavuri. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Computers and Chemical Engineering*, 27(3):293–311, 2003.

[6] Arun K. Samantaray and Belkacem Ould Bouamama. *Model-based process supervision : a bond graph approach*. Springer, London, 2008.

[7] Wolfgang Borutzky. Residual Bond Graph Sinks for Numerical Evaluation of Analytical Redundancy Relations in Model Based Single Fault Detection and Isolation. *Proc. 20th European Conference on Modelling and Simulation (ECMS'06)*, 5(Cd):166–172, 2006.

[8] Jian Fu, Jean Charles Maré, and Yongling Fu. Incremental modeling and simulation of mechanical power transmission for more electric aircraft flight control electromechanical actuation system application. *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)*, 1, 2016.

[9] Edward Balaban, Abhinav Saxena, Prasun Bansal, Kai F. Goebel, Paul Stoelting, and Simon Curran. A diagnostic approach for electro-mechanical actuators in aerospace systems. *IEEE Aerospace Conference Proceedings*, pages 1–13, 2009.

[10] C. Kleijn, M. A. Groothuis, and H. G. Differ. *20-sim 4.6 Reference Manual*. Controllab Products B.V., Enschede, 2017.

[11] EXLAR. Product Specifications. *GSX Series Integrated Motor Actuator*, 2018.

[12] Lijian Wang and Jean Charles Maré. A force equalization controller for active/active redundant actuation system involving servo-hydraulic and electro-mechanical technologies. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 228(10):1768–1787, 2014.

[13] Arun K. Samantaray and Sanjoy K. Ghoshal. Sensitivity bond graph approach to multiple fault isolation through parameter estimation. *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, 221(4):577–587, 2007.

[14] Wolfgang Borutzky. *Bond graph methodology: Development and analysis of multidisciplinary dynamic system models*. Springer, London, 2010.

[15] Arun K. Samantaray and Sanjoy K. Ghoshal. Bicausal bond graphs for supervision: From fault detection and isolation to fault accommodation. *Journal of the Franklin Institute*, 345(1):1–28, 2008.

[16] Mohand A. Djeziri, Belkacem Ould Bouamama, and Rochdi Merzouki. Modelling and robust FDI of steam generator using uncertain bond graph model. *Journal of Process Control*, 19(1):149–162, 2009.

[17] Youmin Zhang and Jin Jiang. Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control*, 32(2):229–252, 2008.

Model-based Sensor Fault Detection in an Autonomous Solar-powered Aircraft

Paulo Victor Padrao Lopes , Liu Hsu , Michael Vilzmann[*], and Konstantin Kondak[*]

Departament of Electrical Engineering, COPPE, Federal University of Rio de Janeiro, Brazil
E-mail: paulo.padrao@iff.edu.br
[*]Department of Robotics and Mechatronics, German Aerospace Center, Germany
[*] E-mail: michael.vilzmann@dlr.de

Abstract

Developed by Elektra Solar, an official spin-off of the DLR Institute for Robotics and Mechatronics (DLR-RMC), the Elektra 2 is a solar-powered autonomous aircraft designed to endure long distances as well as high altitudes. The main motivation of this work is to develop a sensor fault detection and reconfiguration (FDR) approach to be applied to the Elektra 2 Aircraft. Currently, the Elektra 2 Solar aircraft provides a simple limit-checking of certain IMU measurements such as aircraft angular velocities and pitch, roll and yaw angles. In the proposed FDR approach, residual generation for both longitudinal and lateral dynamics of the aircraft is achieved based on the design of Kalman filters. Combined with the sequential probability ratio test (SPRT) as a decision function, selected aircraft sensor residuals are evaluated using an adaptive threshold with low sensitivity to control signal variation and noise and high sensitivity to fault symptoms. Due to the critical aspects of IMU faults, the reconfiguration action is defined as switching from primary IMU to backup IMU in case of fault occurrence.

Keywords: Sensor Fault Detection, Fault Reconfiguration, Autonomous Aircraft Systems

1 Introduction

Sensors are the most critical components for aircraft safety and it is well known that hardware redundancy plays a fundamental role in fault tolerant systems [1], [2]. In order to reduce cost limitations and system complexity, analytical redundancy approach arises as a suitable option for navigation sensor fault detection of such systems [3], [4]. Analytical redundancy uses a mathematical model of the monitored system and it is here achieved by properly designed estimators that provide robust residual generation for fault detection framework [5]. These residuals are then evaluated using an adaptive threshold with low sensitivity to control signal variation and noise and high sensitivity to fault symptoms. The main motivation of this work is to develop a fault detection and reconfiguration approach to be applied to the Elektra 2 Aircraft (fig. 1). Currently, the aircraft provides a simple limit-checking of certain measurements such as aircraft angular velocities and pitch, roll and yaw angles. Developed by Elektra Solar, an official spin-off of the Institute for Robotics and Mechatronics at the German Aerospace Center, the Elektra 2 is a solar-powered autonomous aircraft designed to endure long distances as well as high altitudes.

2 Proposed Fault Detection Approach

The proposed fault detection and reconfiguration approach for the Elektra 2 aircraft is illustrated in figure 2. Residual gen-

Figure 1: Autonomous solar-powered aircraft: Elektra 2.

eration for both longitudinal and lateral dynamics of the aircraft is achieved based on the design of Kalman filters [6]. The sequential probability ratio test (SPRT) is then used as a decision function to be evaluated with adaptive thresholds (ATLMS) for each of the aircraft measurements. Due to the critical aspects of IMU faults, the reconfiguration action is defined as switching from primary IMU to backup IMU in case of fault occurrence. Therefore, a hot standby reconfiguration scheme is used for sensor fault tolerant purposes.

2.1 Aircraft Model and Flight Conditions

The six-degree-of-freedom, linearized model of the aircraft is decoupled into longitudinal and lateral-directional dynamics as follows
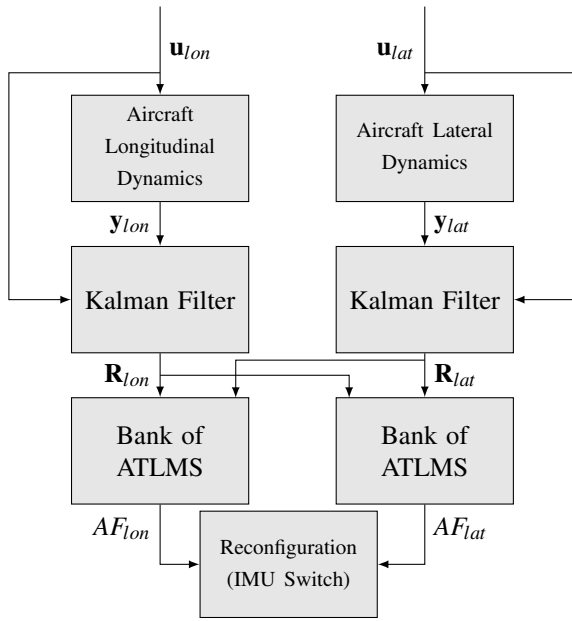
Figure 2: $\mathbf{u}_{lon}$ *and* $\mathbf{u}_{lat}$ *are the control signals for the decoupled longitudinal and lateral aircraft dynamics, respectively.* $\mathbf{y}_{lon}$ *and* $\mathbf{y}_{lat}$ *are the aircraft measurements from longitudinal and lateral aircraft dynamics, respectively.* $\mathbf{R}_{lon}$ *and* $\mathbf{R}_{lat}$ *are the aircraft residuals from longitudinal and lateral aircraft dynamics, respectively.* AF_{lon} *and* AF_{lat} *refer to the longitudinal and lateral alarm flags, respectively.*

$$\begin{bmatrix} \dot{\mathbf{x}}_{lon} \\ \dot{\mathbf{x}}_{lat} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{lon} & 0 \\ 0 & \mathbf{A}_{lat} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{lon} \\ \mathbf{x}_{lat} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_{lon} & 0 \\ 0 & \mathbf{B}_{lat} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{lon} \\ \mathbf{u}_{lat} \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \mathbf{y}_{lon} \\ \mathbf{y}_{lat} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{lon} & 0 \\ 0 & \mathbf{C}_{lat} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{lon} \\ \mathbf{x}_{lat} \end{bmatrix} \quad (2)$$

The longitudinal dynamics describes the aircraft motion in the (x, z) body plane and it has 3 degrees of freedom: pitch, longitudinal motion and vertical motion.

The state vector for longitudinal motion is given by

$$\mathbf{x}_{lon} \triangleq (u, w, q, \theta, h)^T \quad (3)$$

where u and w are the linear velocities along the x and z axes respectively, in body frame; q corresponds to the pitch rate; θ indicates the pitch angle; h indicates the altitude above ground the aircraft. The respective input vector is defined as

$$\mathbf{u}_{lon} \triangleq (\delta_e, \delta_t)^T. \quad (4)$$

where δ_e and δ_t are the elevator and throttle commands, respectively.

The lateral dynamics describes the aircraft motion around the z-axis and consists of roll, yaw and lateral motion. The state vector for lateral motion is given by

$$\mathbf{x}_{lat} \triangleq (v, p, r, \phi, \psi)^T \quad (5)$$

where v is the linear velocities along the y axis, in body frame; p and r correspond to the roll and yaw rates, respectively; ϕ and ψ the roll angle and yaw angle, respectively;

and the respective input vector is defined as

$$\mathbf{u}_{lat} \triangleq (\delta_a, \delta_r)^T. \quad (6)$$

where δ_a and δ_r are the aileron and rudder commands, respectively. All states and measurements are provided in the International System of Units (SI). The measured output vectors used here are rewritten in terms of V_a, α, and β as follows

$$\mathbf{y}_{lon_m} \triangleq (V_a, \alpha, q, \theta, h)^T \quad (7)$$

$$\mathbf{y}_{lat_m} \triangleq (\beta, p, r, \phi, \psi)^T. \quad (8)$$

where V_a is the true airspeed, α is the angle of attack and β the sideslip angle, which are measurable by air data sensors on the aircraft as in equation (9)

$$\begin{aligned} V_a &= \sqrt{u^2 + v^2 + w^2} \\ \alpha &= arctan\left(\frac{w}{u}\right) \\ \beta &= arcsin\left(\frac{v}{V_a}\right) \end{aligned} \quad (9)$$

2.2 Proposed IMU Faults

In this work, three different types of additive IMU faults are presented in table 1:

- *Abrupt Fault*: modelled as a step-wise function. It represents bias in the monitored signal;

- *Incipient Fault*: modelled by ramp signals. It represents drift of the monitored signal;

- *Extra Noise Fault*: modelled by an abrupt change of the signal standard deviation.

Table 1: Proposed IMU Faults

Fault	Measurement	Amplitude
Incipient	θ	0.2 rad/s
Extra Noise	q	$\mathcal{N}(\bar{\mu} = 0, \sigma = 0.01)$
Abrupt	ϕ	0.2 rad

2.3 Adaptive Threshold based on the LMS Filter (ATLMS)

The adaptive threshold (ATLMS) presented here is based on the technique exposed in [7]. The ATLMS technique allows the threshold to be tuned by changing well known parameters independently of the application and is based on SPRT algorithm, the LMS (Least-Mean-Square) algorithm, and residual sensitivity. The residual is said to be fault-sensitive if

it is highly influenced by fault occurrence and consequently deviates significantly from non-faulty mode. In contrast, a residual is said to be fault-insensitive if it is not influenced by fault occurrence or such influence is negligible in comparison with the sensitive residual. Differently from the application proposed by [7] and taking advantage of the decoupling nature of the system, the design of sensitive and insensitive residuals was obtained by the decoupling of the lateral and longitudinal aircraft linear models. Once this is not always possible for other systems, sensitive and insensitive residuals can also be accomplished by structured residual technique as presented in [8] or by a bank of observers/estimators as in [9], for example.

The basic idea behind the proposed adaptive threshold is to use the adaptability of the LMS algorithm to reinitialize the SPRT algorithm automatically (fig. 3) and the requirements to the dynamic behavior of such threshold are (i) low sensitivity to control signal variation; (ii) low sensitivity to noise and (iii) high sensitivity to faulty residuals.
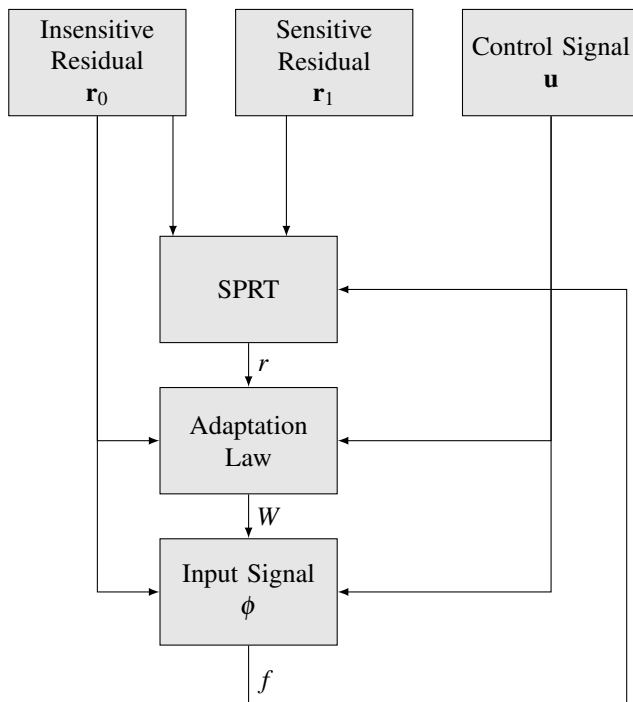
Figure 3: ATLMS Scheme: $\mathbf{u}$ is the control signal vector, r_0 is the insensitive residual, r_1 is the sensitive residual, r is the instantaneous SPRT response, $\mathbf{W}$ is the weight vector of the LMS algorithm, and ϕ is the input signal vector composed of $\mathbf{u}$ and recursions of the insensitive residual. A fault alarm flag is raised if the adaptive threshold f crosses the insensitive residual r_0.

The adaptive threshold f is computed at each sample period as:

$$f_k = \phi_k^T \mathbf{W}_k \tag{10}$$

where $\mathbf{W}$ is the weight vector of the LMS algorithm, and ϕ is the input signal vector composed of the control signal $\mathbf{u}$ and

recursions of the insensitive residual r_0. The SPRT can be calculated as

$$L_k = L_{k-1} + \ln \frac{p_0(r_{0k}|r_{0k-1},\ldots,r_{0k-d})}{p_1(r_{1k}|r_{1k-1},\ldots,r_{1k-d})} \tag{11}$$

where p_0 and p_1 are the probability density function of the insensitive and sensitive residuals r_0 and r_1, respectively. Considering that r_0 and r_1 present normal distribution with zero mean and same variance σ_0^2, the ratio $\ln\left(\frac{p_0}{p_1}\right)$ can be simplified by $\frac{r_{0k}^2 - r_{1k}^2}{2\sigma_0^2}$. For each data sample, three decisions can be made:

- In favor of a faulty mode (Hypothesis $\mathcal{H}_1$);
- In favor of a normal mode (Hypothesis $\mathcal{H}_0$);
- There is insufficient information to decide in favor of faulty or normal mode.

For this reason, under different hypotheses, the expected value of the SPRT response presents trends in opposite directions as follows

$$\begin{aligned} E\left\{\frac{r_{0k}^2 - r_{1k}^2}{2\sigma_0^2}|\mathcal{H}_0\right\} &< 0 \rightarrow \dot{L}_k < 0 \\ E\left\{\frac{r_{0k}^2 - r_{1k}^2}{2\sigma_0^2}|\mathcal{H}_1\right\} &> 0 \rightarrow \dot{L}_k > 0 \end{aligned} \tag{12}$$

where $\dot{L}_k = \frac{dE\{L_k\}}{dt}$.

The SPRT decision update is based on the difference between the adaptive threshold f and the insensitive residual r_0 as follows

$$\gamma = f - r_0 \tag{13}$$

The update term γ should also be affected by a varying term e as

$$e = e_0 - a\dot{L}_k \tag{14}$$

where $\dot{L}_k$ is the expected trend of the SPRT, e_0 is the safe distance between the adaptive threshold f and insensitive residual r_0, and a is the sensitivity due to changes in the trend of the SPRT at instant k. More details about these tuning parameters are shown in section 2.4. Without loss of generality and for the sake of implementation procedures, the derivative of the expected value of L_k, $\dot{L}_k$, is simplified by the approximation $\dot{L}_k = \frac{L_k - L_{k-1}}{T_s}$, where the additional parameter T_s is the sampling period. Thus, the objective function to the synthesis of the adaptive threshold is defined as

$$\min(\gamma_k - e_k)^2 \tag{15}$$

and the LMS algorithm is used to find the optimal weight vector $\mathbf{W}_k^*$ that minimizes such objective function.

The quadratic performance index can be rewritten as

$$(\gamma_k - e_k)^2 = \left(\phi_k^T \mathbf{W}_k - r_{0k} - e_0 + \underbrace{\left(\frac{a}{2T_s \sigma_0^2} \right)(r_{0k}^2 - r_{1k}^2)}_{s_k} \right)^2$$
$$= \mathbf{W}_k^T \phi_k \phi_k^T \mathbf{W}_k - 2\phi_k^T \mathbf{W}_k s_k + s_k^2 \tag{16}$$

The weight vector $\mathbf{W}_k$ that minimizes the quadratic performance index can be found by the gradient method as in the LMS algorithm as follows

$$\mathbf{W}_k = \mathbf{W}_{k-1} + \mu \left[-\frac{\partial (\gamma_k - e_k)^2}{\partial \mathbf{W}} \right]_{\mathbf{W} = \mathbf{W}_{k-1}} \tag{17}$$

.

An approximate estimate of the gradient at each sampling instant is given by

$$\frac{\partial (\gamma_k - e_k)^2}{\partial \mathbf{W}} = 2(\gamma_k - e_k)\frac{\partial (\gamma_k - e_k)}{\partial \mathbf{W}} = 2(\gamma_k - e_k)\phi_k \tag{18}$$

and its proof is provided below.

Proof. The expected value of equation (16) is

$$E\left\{ (\gamma_k - e_k)^2 \right\} = E\left\{ \mathbf{W}_k^T \phi_k \phi_k^T \mathbf{W}_k \right\} - E\left\{ 2\phi_k^T \mathbf{W}_k s_k \right\} + E\{s_k^2\} \tag{19}$$

and its partial derivative regarding the weight vector $\mathbf{W}$ is

$$\frac{\partial E\left\{ (\gamma_k - e_k)^2 \right\}}{\partial \mathbf{W}} = 2E\{\phi_k \mathbf{W}_k \phi_k^T\} - 2\underbrace{E\{\phi_k^T s_k\}}_{\Theta} \tag{20}$$
$$= 2\left[E\{\phi_k \mathbf{W}_k \phi_k^T\} - \Theta \right]$$

.

The result presented in eq. (20) must be the same as the expected value of eq. (18), as follows

$$E\{2(\gamma_k - e_k)\phi_k\} = 2\left[E\{\gamma_k \phi_k\} - E\{e_k \phi_k\} \right]$$
$$= 2\left[E\{f_k \phi_k - r_{0k}\phi_k\} - E\{(e_0 - a\dot{L}_k)\phi_k\} \right]$$
$$= 2\left[E\{\phi_k^T \mathbf{W}_k \phi_k - r_{0k}\phi_k\} - E\{(e_0 - a\dot{L}_k)\phi_k\} \right]$$
$$= 2\left[E\{\phi_k^T \mathbf{W}_k \phi_k\} - E\{\underbrace{(r_{0k} + e_0 - a\dot{L}_k)}_{s_k}\phi_k\} \right]$$
$$= 2\left[E\{\phi_k^T \mathbf{W}_k \phi_k\} - \underbrace{E\{\phi_k s_k\}}_{\Theta} \right]$$
$$= 2\left[E\{\phi_k^T \mathbf{W}_k \phi_k\} - \Theta \right] \tag{21}$$

This proves that

$$\frac{\partial E\{(\gamma_k - e_k)^2\}}{\partial \mathbf{W}} = E\{2(\gamma_k - e_k)\phi_k\} = 2[E\{\phi_k \phi_k^T\}\mathbf{W}_k - \Theta] \tag{22}$$

and therefore $\mathbf{W}_k$ gets closer to the optimal weight vector $\mathbf{W}^*$.
□

The final form of the adaptive threshold is defined as

$$f_k = \phi_k^T \mathbf{W}_k$$
$$\mathbf{W}_k = \mathbf{W}_{k-1} - 2\mu(\gamma_k - e_k)\phi_k \tag{23}$$

2.4 ATLMS Tuning Parameters

A fault flag is raised if the ATLMS crosses the insensitive residual $r_0(t)$. Therefore, the dynamics of the insensitive residual $r_0(t)$ must be taken into account in order to minimize false/missed alarms. For this reason, the parameter *Safety Offset* (e_0) is set to prevent undesired false alarms during the transient response of $r_0(t)$ or to decrease the distance between the adaptive threshold and $r_0(t)$ in case of $r_1(t)$ signal is not very sensitive to the applied fault.

The *Sensitivity Rate* (a) describes the sensitivity of the adaptive threshold due to changes in the trend of the residuals as determined by the SPRT. In case of fault occurrence and for $a < 0$, the adaptive threshold increases in the direction of the insensitive residual $r_0(t)$. Otherwise, the adaptive threshold increases in the opposite direction of the insensitive residual. This is a useful feature of ATLMS to avoid erroneous interpretations of adaptive threshold responses due to possible mistakes in the design of residual generators (when sensitive and insensitive residuals are incorrectly interchanged, for example), or by unexpected behavior of the residuals.

The *ATLMS Convergence Rate* (μ) is originally from the LMS algorithm and it is responsible for the adaptation stability and convergence speed of the ATLMS.

3 Simulation Results and Discussion

3.1 Simulation Methodology

The simulation methodology consists of three different flight scenarios with each of the proposed IMU faults presented in table 1 with following assumptions:

- Occurrence of single faults;

- Backup IMU does not fail;

- Reconfiguration procedure is defined as switching from primary IMU to standby/backup IMU once a fault is detected.

For the sake of simplification, all faults are introduced at 50 s. For incipient faults, the amplitude values described in table 1 refer to the slope of the faulty signals. For the extra noise faults, additive white noise is introduced with the indicated distribution. These faults cover both longitudinal and lateral aircraft dynamics and were chosen based on *a priori* knowledge and discussions with Elektra 2 Solar team. Due to the critical aspect of IMU faults, the fault detection flag is activated - and kept active - if any of the dedicated adaptive thresholds indicates a fault. Figure 4 illustrates the scheme of the bank of ATLMS for longidutinal dynamics. A similar

scheme is used for the lateral dynamics. For the longitudinal measurements, the bank of ATLMS uses the yaw angle ψ as the insensitive residual. For the lateral measurements, the bank of ATLMS uses the pitch angle θ as the insensitive residual.
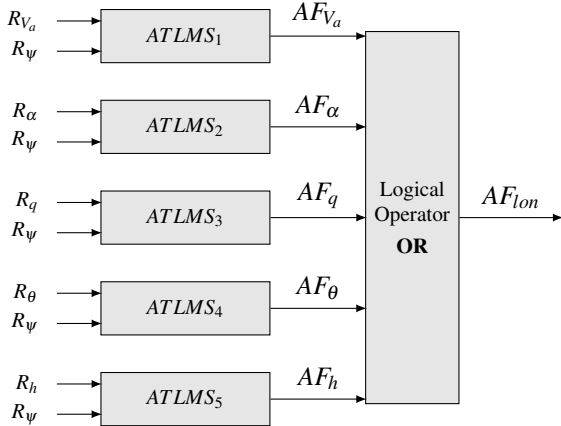

Figure 4: *Bank of ATLMS for Longitudinal Dynamics. R_i is the residual of measurement i. V_a is the aircraft airspeed, α is the angle of attack, q is the pitch rate, θ is the pitch angle, and h is the aircraft altitude. The bank of ATLMS uses the yaw angle ψ residual as the insensitive residual. The label AF_i refers to the alarm flag of measurement i and AF_{lon} is the combined alarm flag for longitudinal dynamics used for sensor switching reconfiguration in case of fault occurrence.*

3.2   Incipient fault in pitch angle θ

The first simulation scenario illustrates an incipient fault in pitch angle θ at 50s. As expected, a fault in pitch angle affects the longitudinal alarm flag *Lon Flag* but not the lateral alarm flag *Lat Flag* (fig.5 and fig. 6). The lateral adaptive thresholds present responses in the opposite direction of the insensitive residual. The overall aircraft performance would not remain in an acceptable flight condition if the reconfiguration action does not take place. As illustrated in figure 8, the aircraft airspeed/altitude suddenly increases/decreases in case of the proposed incipient fault occurs. This behavior could lead to overall system failure. However, due to the fault detection promptness and reconfiguration, the aircraft was slightly affected. Figure 7 presents the faulty and reconfigured aircraft roll, pitch and yaw angles.

3.3   Extra noise fault in pitch rate q

The second simulation scenario illustrates extra noise fault in angular rate q at 50s. As expected, the proposed fault affects the longitudinal alarm flag *Lon Flag* but not the lateral alarm flag *Lat Flag* (fig.9 and fig. 10). Although the overall aircraft performance remains in an acceptable flight condition, the elevator control signal δ_e shows an undesired intermittent behavior (fig. 12). In the worst case scenario, this behavior could damage the elevator control surface and lead to overall system failure. However, due to the fault detection promptness and reconfiguration, the unpermitted behavior of the elevator surface is minimized. Figure 11 presents the faulty and
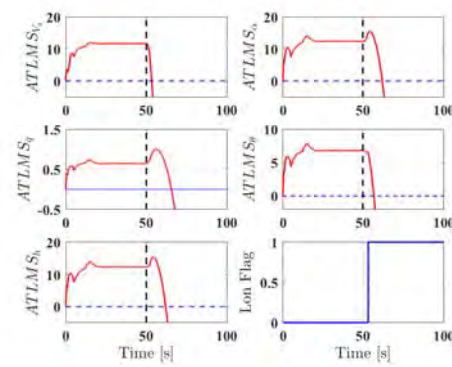

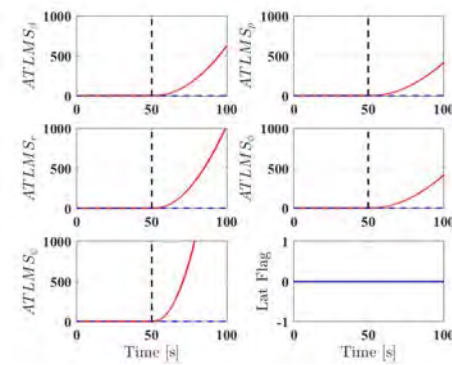Figure 5: *Dedicated adaptive thresholds and fault flag for longitudinal measurements with an incipient fault in pitch angle θ.*

Figure 6: *Dedicated adaptive thresholds and fault flag for lateral measurements with an incipient fault in pitch angle θ.*
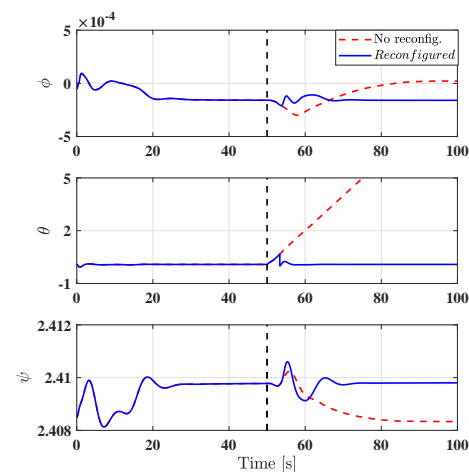

Figure 7: *Faulty and reconfigured aircraft roll, pitch and yaw angles (rad) for an incipient fault in pitch angle θ.*

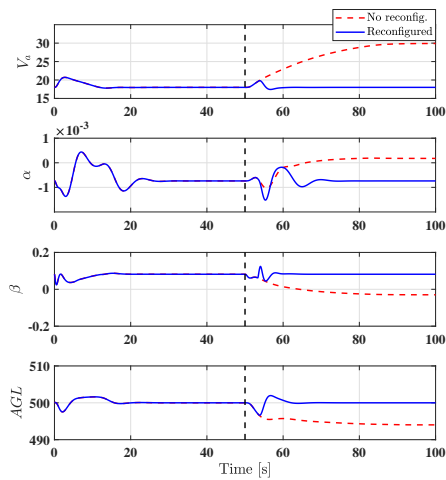reconfigured aircraft angular rates (rad/s) for the second simulation scenario.

Figure 8: Faulty and reconfigured aircraft airspeed (m/s), aerodynamic angles (rad), and altitude above ground (m) for an incipient fault in pitch angle θ.
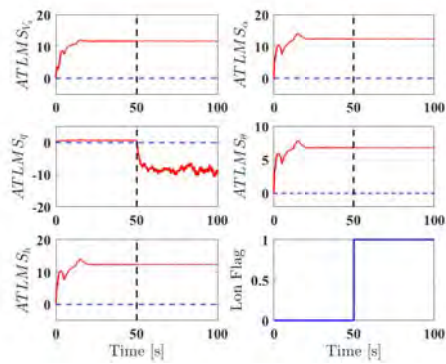

Figure 9: Dedicated adaptive thresholds and fault flag for longitudinal measurements with extra noise fault in angular rate q.
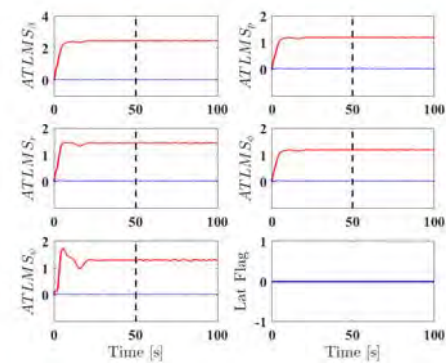

Figure 10: Dedicated adaptive thresholds and fault flag for lateral measurements with extra noise fault in angular rate q.
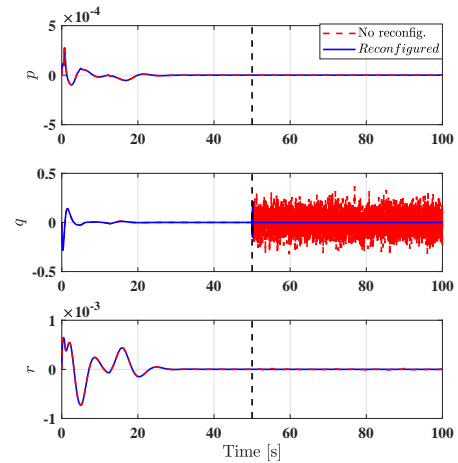

Figure 11: Faulty and reconfigured aircraft angular rates (rad/s) for extra noise fault in angular rate q.
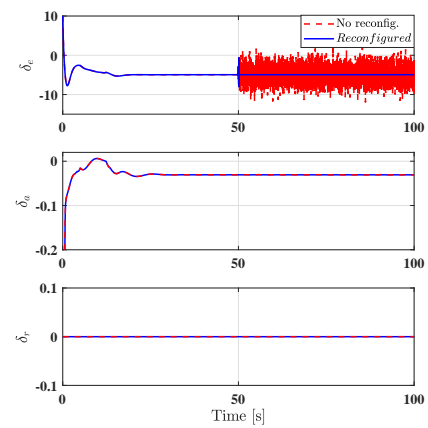
Figure 12: Faulty and reconfigured signals of aircraft control surfaces (degrees) for extra noise fault in angular rate q.

3.4 Abrupt fault in roll angle ϕ

The third simulation scenario illustrates an abrupt fault in roll angle ϕ at $50s$. As expected, a fault in roll angle affects the lateral alarm flag *Lat Flag* but not the longitudinal alarm flag *Lon Flag* (fig.14), since longitudinal and lateral linear models are decoupled as presented in equations (1) and (2). In the case of the proposed abrupt fault occurrence, the overall aircraft airspeed and altitude would remain in an acceptable flight condition even if the reconfiguration action does not take place (fig. 16). However, this is not the case for the aircraft attitude once roll and yaw angles would deviate from normal condition as illustrated in figure 15. Figures 15 and 16 present the faulty and reconfigured aircraft attitude angles (rad), airspeed (m/s), aerodynamic angles (rad), and altitude above ground (m).
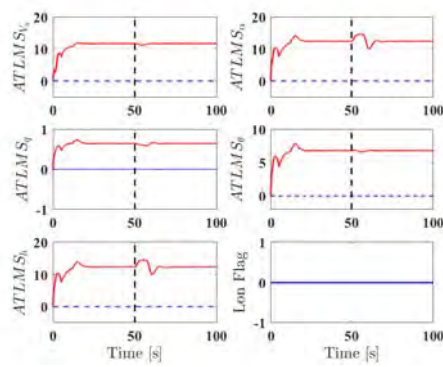
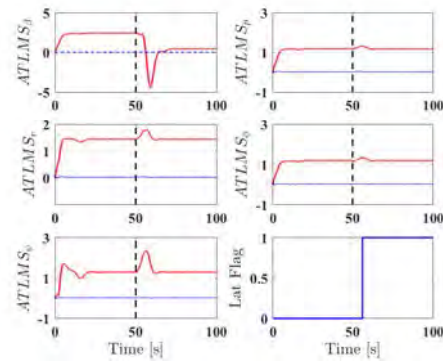*Figure 13: Dedicated adaptive thresholds and fault flag for longitudinal measurements with an abrupt fault in roll angle ϕ.*

Figure 14: Dedicated adaptive thresholds and fault flag for lateral measurements with an abrupt fault in roll angle ϕ.
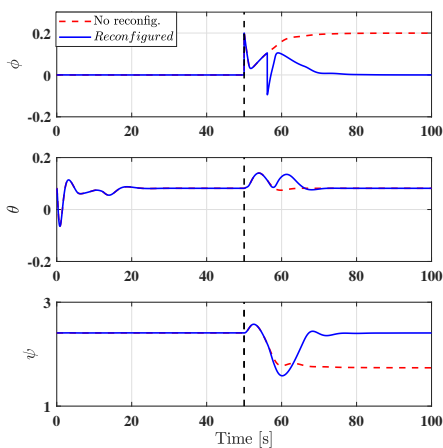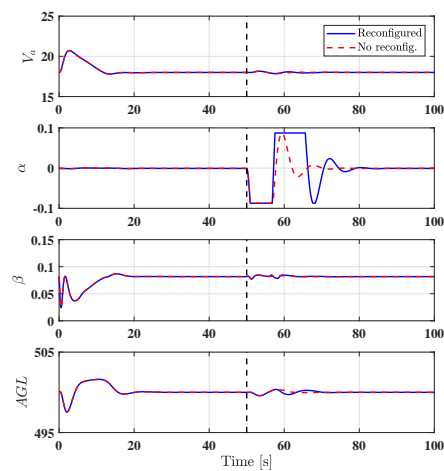

Figure 15: Faulty and reconfigured aircraft roll, pitch and yaw angles (rad) for an abrupt fault in roll angle ϕ.

Figure 16: Faulty and reconfigured aircraft airspeed (m/s), aerodynamic angles (rad), and altitude above ground (m) for an abrupt fault in roll angle ϕ.

be used in the Elektra 2 Solar aircraft. An adaptive threshold (ATLMS) based on the Least Mean Squares algorithm was used to detect different types of faults in IMU measurements. One of the major advantages of the ATLMS technique is the possibility - not fully explored in this work - of parameter tuning. This may result in a better fault detection performance. Besides that, the adaptive threshold is able to to handle dynamic effects on the insensitive residual. Such effects may arise due to plant perturbations caused by control actions or due to the dynamic nature of the residual generator. As a consequence, non-zero residual values can be generated in fault-free scenario, resulting in a set of undesired false alarms. This drawback could also be minimized with suitable tuning of ATLMS parameters. Due to time constraints, the proposed FDD approach have not been integrated and tested in real flights yet.

References

[1] G. HEREDIA and A. OLLERO. Sensor fault detection in small autonomous helicopters using observer/kalman filter identification. In *IEEE International Conference on Mechatronics*, 2009.

[2] M. BASSEVILLE. Detecting changes in signals and systems - a survey. *Annual Reviews in Control*, 24:309–326, 1988.

[3] R. ISERMANN. Supervision, fault-detection and fault-diagnosis methods - an introduction. *Control Engineering Practice*, 5:639–652, 1997.

[4] R. ISERMANN. Model-based fault-detection and diagnosis - status and applications. *Annual Reviews in Control*, 29:71–85, 2005.

4   Conclusion

The main contribution of this work was the development of a model-based sensor fault detection and diagnosis approach to

[5] P. FRANK. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy. *Automatica*, 26:459–474, 1990.

[6] R. PATTON and J CHEN. Observer-based fault detection and isolation: Robustness and applications. *Control Engineering Practice*, 5:671–682, 1997.

[7] A. CARVALHO. *Fault detection, diagnosis and reconfiguration for optimized design of safe space vehicles applied to the PMM and planetary rovers*. Ph.d. dissertation, Instituto Nacional de Pesquisas Espaciais (INPE), Sao Jose dos Campos, Brazil, 2012.

[8] R. ISERMANN. *Fault-diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer, 2006.

[9] S. SIMANI, C. FANTUZZI, and R. PATTON. *Model-based fault diagnosis in dynamic systems using identification techniques*. Springer, 2003.

VIII Major Cooperative Projects

Swedish-Brazilian Cooperation in Aeronautics in Santa Catarina State
André Luiz Meira de Oliveira, Edemar Morsch Filho, Amir A. M. Oliveira Jr.,
Talita Possamai and Victor Juliano De Negri

*Analysis of the Impact of the Conclusion of the Collaborative Professional
Master's Program on Embraer and ITA*
Paulo Lourencao, Orlando Souza Net, Flavio Bussamra and Roberto Gil Silva

Swedish-Brazilian Cooperation in Aeronautics in Santa Catarina State

André L. M. de Oliveira[1], Edemar Morsch Filho[2], Amir A. M. Oliveira Jr.[2], Talita S. Possamai[3], Victor J. De Negri[2]

[1]Metrology and Instrumentation Center – Centers of Reference in Innovative Technologies, Florianópolis, Brazil
[2]Department of Mechanical Engineering - Federal University of Santa Catarina, Florianópolis, Brazil
[3]Department of Mobility Engineering - Federal University of Santa Catarina, Joinville, Brazil
E-mail: aeo@certi.org.br, edemar@labcet.ufsc.br, amir.oliveira@gmail.com, talita.possamai@ufsc.br,
victor.de.negri@ufsc.br

Abstract

In this article, the strategic view leading to the creation of a convergence center for aerospace technologies in the State of Santa Catarina, Brazil, the business and operation plans, as well as the predicted impacts in the state´s economy are presented. Development and innovation are a result of a favorable ecosystem that nurtures the integration of technological creation, proof of concept and human resources. The transformation of new technologies in market products requires the close cooperation of R&D and academic institutions, tech companies and the market. This convergence of stakeholders is achieved by developing a common language and a sense of community, based on work and trust, built along the operation of a cycle, linking market demands to institutional offers, and multiplying the community efforts by cooperation. This is the ground, over which the SC2C.Aero, the Santa Catarina's Center of Convergence for Aerospace Technologies, has been rooted. The center has been configured as a network of tech groups aiming at supporting fast connections and R&D&I agreements. This article reviews the operational model that executes the center's key activities, such as the management of national and international agreements, customer and partner relationships, international matchmaking events, and a database mapping demands, offers and funding sources. The results achieved so far demonstrate the large acceptance this model received from both industry and R&D&I tech groups. These results have created visibility for the State of Santa Catarina and stimulated the expansion of the cooperation between Swedish and Brazilian institutions. Based on the center's strategic planning, five milestones are foreseen in the annual operation of SC2C.Aero that will support Santa Catarina as a top three destiny of investments from the aerospace sector in Brazil in the future.

Keywords: Technology Networks, Aeronautics, University-Industry Cooperation, Convergence Center

1 Introduction

The aerospace is a multifaceted sector, whose development requires diverse engineering specialties, a fast learning ecosystem with a capacity for fast growing, a network of technology and service providers, as well as a qualified work force. Santa Catarina fulfils these requirements. The state is home for three large public universities and more than a dozen community and private universities, more than thirty technology clusters and incubators, and public services and infrastructure that foster innovation and entrepreneurship. This scenario has resulted in estimated 15.5 billion Reals (3.7 billion Euros) of revenue from the technology sector in the state in 2017, about 5.6 % of the state gross product. The IT sector alone comprises more than 3,000 companies and 20,000 jobs in Santa Catarina [1].

The state has also been home of recent investments aiming at the aerospace sector, such as the recently inaugurated EMBRAER's office in Florianópolis; capital of Santa Catarina, the establishment of CERTI as an EMBRAPII Unit for Intelligent Systems; the cooperation of research groups from UFSC in joint projects with the University of Linkoping from Sweden and the company SAAB, as well as the cooperation with several Brazilian universities, such as ITA, UFMG and USP-São Carlos in several activities aiming at the technology transfer and spillover effects from the acquisition of the Gripen fighters by the Brazilian Defense Ministry [2-4]. The development of the FloripaSat, a CubeSat carrying payloads to measure effects of solar radiation on on-board electronics, in cooperation with the University of New South Wales (Australia) and the company NanoXplore, by a team of researchers from UFSC and private entrepreneurs; the inauguration in 2009 of the UFSC Joinville Mobility Center offering an undergraduate degree in Aerospace Engineering,

and the creation by SENAI of the Innovation Institute in Embedded Systems in 2018, completing its regular offering of the Airplane Maintenance technical degree since 2009, are also important events that have strengthened aerospace development in the state.

This several initiatives moved the departments of *Mechanical Engineering* and *Aerospace Engineering* from the *Federal University of Santa Catarina* (UFSC) and the *Centers of Reference in Innovative Technologies' Foundation* (CERTI), with the support of the *Research and Innovation Support Foundation of Santa Catarina State* (FAPESC), to start actions in 2016 for the constitution of a Brazil-Sweden Cooperation Center. The main objective was to improve the quality and intensity of the interactions between companies, government and R&D institutions. Secondary objectives were to propose topics for R&D and to provide support for the execution of research activities already in progress involving teams from both countries. It was a common understanding that cooperation among institutions, connecting skills and sharing existing infrastructures, is very effective way to achieve concrete results in shorter time and to optimize development costs [5].

In 2017, the center was created with its present name, the *Santa Catarina's Center of Convergence for Aerospace Technologies*, SC2C.Aero. The project focused on attracting public and private investments from the aerospace sector to Santa Catarina through collaborative work as a technology convergence center. It was configured as a network of technological R&D&I groups, with the mission of placing Santa Catarina as one of the top 3 states considering investments in the aerospace sector in Brazil. From 2018, this initiative is also supported by a project with the Ministry of Science, Technology, Innovation and Communications (MCTIC).

The SC2C.Aero is presently running its second annual cycle. In its present form, it aims to expand nationally and internationally through the resulting cooperation between Brazil and Sweden in the field of aerospace technologies, stronger since the Gripen purchase, as well as in other highly technological industrial sectors. It evolves in the perspective that a strong regional aerospace industry will result in extensive impacts in all areas of science and technology. The aerospace sector is technology-intensive and requires highly skilled labor. The investment in this sector is prone to create spillover opportunities to other sectors, such as IT, and result in a strong positive impact on the States' trade balance.

This article presents the key aspects of the design and operation of the SC2C.Aero, as well as, a report of its results up to now.

2 Methodology

In a worldwide context, other countries have already started initiatives to boost the aeronautical sector towards the next generations of aircrafts based on Technology Demonstrators, creating organizations to address the challenges of the sector. In US, the NextGen effort is responsible for modernization of air transportation system [6]. CleanSky is an European research and innovation program that aims to strength

European aero-industry collaboration, promote global leadership and competitiveness [7]. France has also worked to develop the technologies necessary to improve its environmental footprint and boost competitiveness through the CORAC committee by stimulating its members for future shape of aviation, consequently giving them a lead when a new program emerges [8]. Within the Strategic Research Agenda (SRA), from the Advisory Council for Aeronautics Research in Europe (ACARE), there is also important inspiration for the actuation of SC2C.Aero, which contemplates the integration of all stakeholders from this context to define the future shape of European aeronautics. Further, the society needs in aviation are translated and categorized in technological solutions to face the challenges of the field [9].

The work group from SC2C.Aero studied the main actions taken by those countries and identified the following trends: production of new aircraft with a noticeable reduction in the use of fuels, with drastic reduction of the noise level and emission of pollutants, coupled with strict safety and comfort standards.

2.1 The proposal of the convergence center

A preliminary model for the SC2C.Aero was advanced in 2017 from a grant of the *Research and Innovation Foundation of the State of Santa Catarina* (FAPESC) to build upon the initiatives between Brazil and Sweden in the aeronautical field. Encouraged by the *Brazilian Ministry of Science, Technology, Innovation and Communication* (MCTIC) in 2018, the focus of the center was amplified to cover R&D&I in aerospace in Brazil and international collaboration in general.

This initiative started by identifying the ongoing collaborations among academia, industry and govern in this field involving both countries. It was found that the aeronautical segment in Brazil and Sweden had been collaborating for at least 20 years in individual projects connected to different agents. A stronger connection of the whole chain (govern, industry and academia) was still missing despite the successful results of each individual cooperation. A new model was then identified with focus on the network of research groups combining efforts to a common goal - to increase the quality and impact of the interactions between companies, government and R&D institutions.

This model aimed at a center able to converge efforts of different stakeholders. Under the context of the triple helix model (see Fig. 1), the stakeholders in the aerospace sector were identified as:

- Project executors: technical groups, members of the network, that are interested in the development of the market to sell their projects of R&D&I;

- Partners and project funders: companies of the aeronautical sector that invest in R&D&I projects and are interested in their sustainability and in the development of the segment;

- Influencers and funders: government agents, other networks and associations in the aeronautical sector,

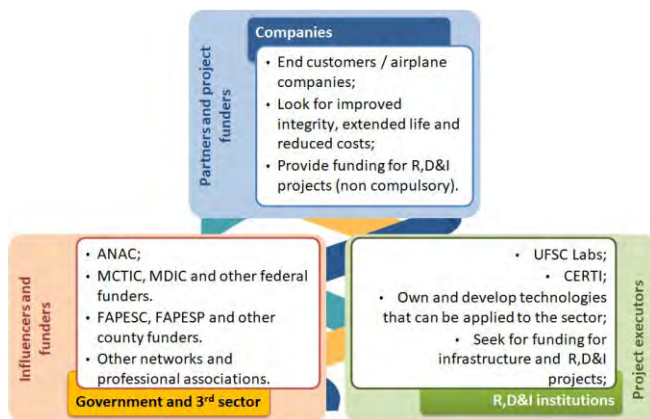interested in the development of the sector and society in all its dimensions.

Figure 1: The stakeholders of the convergence center in a triple helix.

To better understand the stage where these three main agents cooperate for the success of an attractive environment for the aeronautical field, critical scenarios involving this initiative were studied based on a SWOT analysis, where the best strengths and weakness in the sector were balanced to predict the competences required for a solid convergence center. As result of the analysis, the management of stakeholders surfaced as the priority zero action for a center in this context. The central node of the entity should also support marketing strategies and a communications plan, develop a list of competences directed to the sector, as well as, test sites and simulation facilities, keep national and international cooperation up to date, provide storage, communication and processing facilities, customer database management, tracking of opportunities, matchmakings' meeting organization, update funding sources and advocate with public agents.

Based on the long run perspective for the project, formal guidelines for the establishment of a convergence center were stated as:

- Mission:
  - To attract public and private investments from Aerospace Sector to the State of Santa Catarina through collaborative work as a technology convergence center.
- Vision:
  - In 2030, the State of Santa Catarina will be on the top three states in Brazil considering the investments in the Aerospace Sector - and the network and its members will be organized as an important driver to keep this position.
- Short term goals:
  - Inaugurate basic infrastructure facilities and hire management personnel by 2019;
  - Contract five new projects with its members in the aerospace sector until 2020;
  - To reach sustainability of its operations by 2021.

2.2 Operation of SC2C.Aero

The operation of the SC2C.Aero is based on attracting and connecting the stakeholders from the triple helix in annual cycles, where 5 milestones are foreseen (see Fig. 2): Demand research, Offer capacity, Matchmaking event, Contracting and Launch of projects. In this study, only the 3 initial steps of the annual cycle will be discussed in the following sections together with the main results of each one.
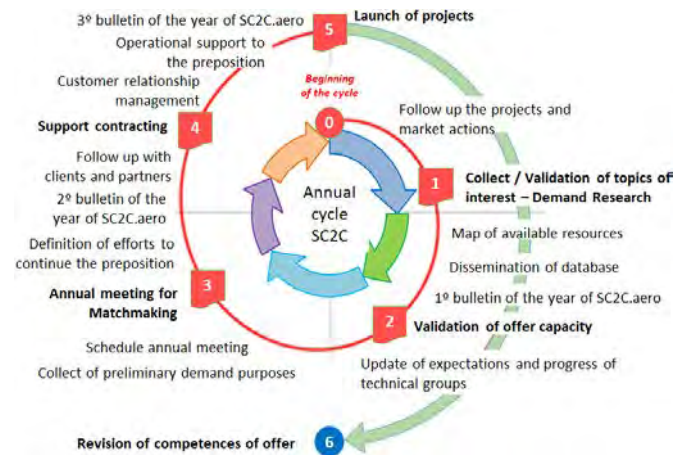

Figure 2: Annual cycle of SC2C.Aero.

2.2.1 Demand research

The demand research is the first milestone of the annual cycle of SC2C.Aero and has as main objectives to collect technological concerns from industry for a biennium (current and next year) and to create a networking with potential stakeholders. Secondary objectives are to disseminate the initiative and collect opinions to adjust the cooperation model of SC2C.Aero and to create motivation for academic research groups and stakeholders towards SC2C.Aero goals.

The group understood the need to use a common language among the players of the triple helix and the ASTERA taxonomy was chosen as a base system for communication [10]. This taxonomy divides the aeronautical areas as a first level and knowledge areas as second and third levels. In Fig. 3 the areas of first level are shown. Second and third levels become very specific and counts tens of items. It became the common language to organize the information received from the industry and academia by the Center of Convergence.
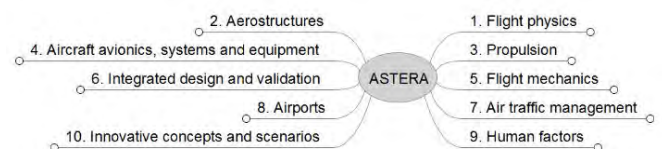

Figure 3: The first level of ASTERA taxonomy [10].

In order to collect the demand from companies, the SC2C.Aero work group focus answering the following question: **What are the main concerns, or "pains", in the aerospace industry that can be developed as projects by SC2C.Aero?**

The research to address this question follows a methodology characterized by its typology, population, sample and data. The research can be classified as descriptive conclusive [11].

The survey is qualitative because of its objectives and responses for the question, however some quantitative data may be collected to support the development of the project. In this research, the population are managers of R&D&I in companies of the aerospace segment stationed in Brazil and legal representative authorities.

For the first cycle, executed in 2017, the size of the sample was 14 companies and 2 federal ministries. Given the focus of that survey, time and cost limitations, companies located in the State of Santa Catarina, São José dos Campos in the State of São Paulo and representatives of the govern in Brasília were chosen. For that occasion, the sample size had 2 companies in Santa Catarina (WEGA Aircraft and Horus), 6 companies in São José dos Campos (Konatus, Embraer, Avibras, Akaer, Boeing, and Eleb), the Ministry of Defense and the Ministry of Science, Technology, Innovation and Communication.

The method and technique to collect data was the in situ research with application of face-to-face interview. The in-person meetings at companies and ministries were booked with the following agenda: Arrival and presentation/greetings (15 min), Presentation of the project and preliminary model of SC2C.Aero/gathering of opinions (15 min), Presentation of the objective and procedures (5 min), Collection of workable topics of interests (25 min).

The results of the interview for each company were later grouped and indexed following the ASTERA taxonomy, generating a highly valuable spreadsheet to be shared with the technical groups and partners members of SC2C.Aero. They included Company, Topic of interest or pain, Taxonomy and Theme (Education, Raw material, Process, Products and components, Technological services, Regulation). More than one company mentioned similar topics of interest, as well as technological issues and need of human resource programs.

2.2.2 Offer capacity

The next step and milestone of the annual cycle is the Offer capacity (see Fig. 2). This topic aims at collecting the offer capacity based on taxonomy and TRL. The secondary objectives include crossing of demand with offer through taxonomy, motivate the research groups to support the progress of the sector in Santa Catarina, develop marketing material and create a competence database model to be applied in cooperation with Sweden and other countries.

As mentioned above, ASTERA is the taxonomy for common language for topics of interest from industry. Moreover, the work group felt the need to measure the readiness level of each technical group in respect to the technologies they develop and adopted the Technology Readiness Levels (TRL) as a tool for the Offer Research. This scale has 9 levels to assess the maturity level of a given technology, where 1 is the lowest and 9 is the highest technology readiness level [12].

To collect the offer from Santa Catarina's groups, the research was structured to answer the following question: **What are the highlight competencies and skills of interest of the aerospace industry that can be provided by the technical groups on SC2C.Aero?**

Some of the R&D&I performance criteria analyzed are:

-Project history with the aerospace industry, which demonstrates the group's ability to provide strong solutions for the aerospace sector;

-Research infrastructure, including equipment and buildings;

-Existence of a specific research focus;

-Number of students in the group;

-Group leader entrepreneurship ability, related to publications, partners and customers.

The population of the research are the leaders of the groups that are members of SC2C.Aero. The size was 17 groups (14 from UFSC and 3 from CERTI). Leaders of the groups were invited to watch a presentation about the SC2C.Aero objectives and operational plan, their role in the triple helix, results from demand research and the Offer research. Later, a digital structured questionnaire was created and sent to the leaders of each group in the list.

In addition to the items mentioned by companies, the following questions were included in the form to favor the crossing of demand and offer capacity:

- Items of the ASTERA taxonomy that the lab can work nowadays, including the TRL;

- Example of projects (with companies), related or not with the aerospace sector, and the TRL reached;

- Facilities and competences to test and simulate aeronautical systems and products;

- The perspective for future by addressing the taxonomy and TRL that they would like to work with.

The results surveyed in this step were matched to the results of the demand survey/topics of interest.

2.2.3 Matchmaking event

After the demand and offer research, a matchmaking event took place to promote the matchmaking between demand and offer. The event was named Workshop of Santa Catarina's Convergence Center for Aerospace Technologies and is part of the annual cycle of operation of SC2C.Aero. Its objective is connecting companies, govern agencies and technology institutions in Brazil and Sweden in the aerospace sector. Bringing together official representatives from Sweden and Brazil, the connections between the two countries are further strengthened. The target audience are companies of the aerospace sector, technical groups of Santa Catarina, technology institutes focused in the aerospace segment in Brazil and Sweden, members of Santa Catarina state government and research foundation, as well as representatives of government and innovation supporting agencies in Brazil and Sweden.

The 1st Workshop SC2C.Aero was held in May 3rd and 4th, 2018. Topics presented and discussed were the SC2C.Aero results obtained from 2017 until the event, the perspectives and challenges in the sector, the established connections, the center members and next steps. Representatives from companies such as Embraer, Saab, ELEB, Glenair, Akaer,

and several others attended the event. Representatives from the Brazilian and Swedish government as well as the academic institutions UFSC, LiU, ITA, SENAI, IFSC, and UFABC also participated. The workshop included speeches from CISB - Swedish-Brazilian Research and Innovation Centre, AEB- Brazilian Aerospace Agency, and Innovair-Swedish Strategic Innovation Programme for Aeronautics.

3 Results

On the research demand, a total of 51 topics of interest were listed by companies, matching with 104 (including repeated) items of the taxonomy (see Fig. 4). Except for the item 8-Airports, all other primary level of the taxonomy had at least one element mentioned by companies. The areas most cited were 2-Aerostructures, 4-Aircraft avionics, systems and equipment and 6-Integrated design and validation, respectively.
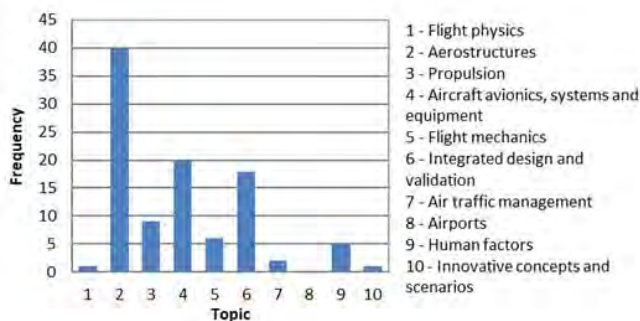
Figure 4: Histogram with topics of interest organized in the ASTERA taxonomy.

In the offer demand, a total of 11 groups answered positively and shared their information to form the offer capacity of SC2C.Aero. A total of 51 professionals were reached in the survey including faculty, post-docs, doctor and master's students. Regarding the topics of interest in the ASTERA taxonomy, a total of 75 items were listed by the groups. By observing the classification of demand or topics of interest (demand research), a good superposition of offer and demand was found, creating diverse opportunities for project (see Fig. 5).
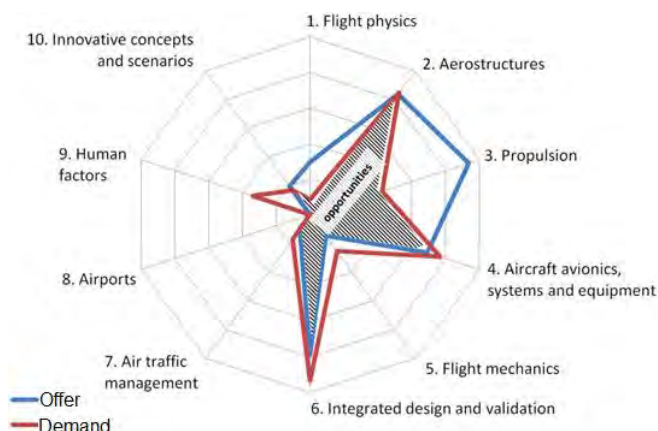

Figure 5: Crossover of topics of interest and offer with data collected in 2017.

Except in Flight Physics and Propulsion, where there is greater offer capacity than topics of interest, the other classifications are quite coherent and coincident. There are even opportunities of connection with other groups of UFSC in areas without mapped competencies (like Human Factors), while Aerostructures, Aircraft avionics and Integrated design and validation show capacities already mapped to meet the demands of the aeronautical sector. These databases are considered dynamic as new cycles begin, for both constant update of technical groups and introduction of new groups or international cooperation, creating a solid base to be explored in the solutions of demand of the sector.

4 Conclusions

After the consolidation period of SC2C.Aero, estimated as 3 years, the center should opt for professional network with permanent connection, and not through research fellowships, since the work of development ends but the operation has to be handled with continuity. At this moment, a new analysis and revision of this plan is under way.

The matchmaking event, part of the annual operational cycle, is a platform for marketing of competencies found in the State of Santa Catarina that are of value for competitive companies of the aerospace segment. Once this value is recognized, the companies will search for and find qualified human resources in the state, resulting in more jobs, income and wealth.

The deadlines for submission of funding sources must be monitored, but also directly handled with the agents, looking for the recognition of the importance of SC2C.Aero for the involved institutions, for the state and for the country.

The SC2C.Aero just completed its second year of existence. In its present form, it impacts not only the field of aerospace technologies, but also other highly technological industrial sectors. It will evolve in the perspective that a strong regional aerospace industry will result in extensive impacts in all areas of science and technology. The investment in the aerospace sector is prone to create spillover opportunities to other sectors, such as IT, and result in a strong positive impact on the state´s trade balance, economic and social indicators.

Acknowledgments

References

[1] Federação das Indústrias do Estado de Santa Catarina: Santa Catarina em Dados. Available at https://www.observatoriofiesc.com.br/sc-em-dados

[2] DE NEGRI, V. J.; OLIVEIRA, A. A. M. de; SCHNEIDER, C. A.; MANTOVANI, T.; PEREIRA, J.

C. C.; SALMORIA, G. V.; FREDEL, M. C.; DONATELLI, G. D.; SOETHE, V. L.; AL-QURESHI, H. A.; PEREIRA, R. S. F. Caracterização das competências no estado de Santa Catarina para o setor aeronáutico, 2013.

[3] Deiaco, E., Ek, I., Román, E. Nurturing spillover from the Industrial Partnership between Sweden and Brazil a case study of the Gripen project. Report 2016/106. Available at www.growthanalysis.se.

[4] Estudo Prospectivo Aeronáutico ABDI: 17/02/11 5. Industria Aeronáutica - ABDI: 12/11/10 6. Programa Demonstrativo para Inovacao em Cadeia Produtiva Selecionada Industria Aeronáutica Brasileira Serie Documentos Técnicos - CGEE set/2014.

[5] SC2C.Aero: SC2C.Aero. Available at https://sc2c.ufsc.br/

[6] Federal Aviation Administration: NextGen implementation plan. Washington,2012.

[7] Clean sky governing board: Clean sky annual implementation plan 2010. Bruxelas, 2010.

[8] Corac: Proposition sectorielle de Charte d'Organisation et de Gouvernance des plataformes de demonstration technologique proposées par le secteur aéronautique dans le cadre du pro-gramme d'investissements d'avenir. Available at http://www.aerorecherchecorac.com

[9] ACARE: Strategic Research Agenda, Volume 1. Available at: ww.acare4europe.org/sites/ acare4europe.org/files/document/volume1.pdf

[10] EASN: Aeronautical research & technology areas. Available at https://easn.net/research-technology-areas/

[11] Frauze N. Mattar: Pesquisa de marketing. 4th. Ed. São Paulo: Editora Atlas, 2007.

[12] NASA: Technology Readiness Level. Available at https://www.nasa.gov/directorates/heo/scan/engineering /technology/txt_accordion1.html

Analysis of the impact of the conclusion of the Collaborative Professional Master's Program on Embraer and ITA.

Paulo Lourencao, Orlando Souza Neto, Flavio Bussamra* and Roberto Gil Silva*
Engineering Strategy/Embraer, S J Campos, SP/Brazil
E-mail: paulo.lourencao@embraer.com.br, orlando.neto@ embraer.com.br, flaviobu@ita.br, gil@ita.br
*Division of Aeronautical Engineering/ITA, SJ Campos, SP/Brazil

Abstract

The creation of Embraer resulted from a long-term investment of the Brazilian Government in developing the necessary conditions for designing and manufacturing airplanes in the country. In each decade after its creation in 1969, Embraer launched a different type of aircraft. In 1999, Embraer launched the program for Regional Jets ranging from 70 -110 passengers, which implied a high demand of new engineers. One alternative to cope with that was the creation of the Professional Master's Program in Aeronautical Engineering in 2002, in partnership with ITA (Aeronautics Institute of Technology). The objective of the program has been to provide specialized knowledge to recently graduated engineers, in order to prepare them to work in multidisciplinary teams in Embraer engineering areas. The collaboration between Embraer and ITA is present from the selection process, through curriculum design up to Master's theses. The objective of this paper is to present results of an evaluation of the impact of the Master's Program on alumni's performance on their work, on the company (technology, process, products, services and patents) and on the Institute. Since 2002, the program has already prepared more than 1400 students. Out of them, more than 800 are still working for Embraer (Feb 2019) and more than 400 have their Master's theses approved. The methodology adopted in this paper consisted in interviews and surveys with former students of the program, their direct supervisors and both their industrial and academic advisors. This paper presents, as results, the most significant contributions both to the company and to ITA.

Keywords: Pre-projects, Conceptual Design, Engineering Education

1 Introduction

This session initially presents the context of creation of the Technical Center of Aeronautics (CTA), of Aeronautics Institute of Technology (ITA) and Embraer. This is important to understand the conditions that led the creation of the object of analysis of this paper, the Professional Master's Program of ITA in partnership with Embraer, to be conceptualized and implemented.

In continuation, the objective and the structure of the paper are presented.

1.1 Context

After the end of World War II, the Brazilian Government decided to create the necessary conditions to design and manufacture airplanes in the country. In addition to national security, the size of the nation and the need of connecting remote areas clearly justified that decision. CTA was conceptualized by prof. Richard H. Smith from the Department of Aeronautical Engineering of the MIT in 1945 [1].

The original plan proposed the creation of a "Technical Center that would consist of two technically autonomous Institutes working under the same coordination: the first for higher technical education and the second for research and cooperation with an aircraft industry that would have both military and commercial aviation branches. According to the original plan, when the Institutes laboratories could deliver a potential product then a plant would be established" [1].

In the 50's, ITA was established as a School of Aeronautical Engineering. In 1954, the Institute of Research and Development (IPD) was created, employing many engineers graduated from ITA. IPD conducted R&D in many areas such as Aeronautics, Electronics, Materials, Aircraft Systems, and others related to the field. IPD was responsible for the designing and building the "Bandeirante" aircraft that led to the creation of Embraer in 1969 [2].

Since its creation, decade after decade, Embraer has launched new aircraft. Considering commercial aviation only, in the 70's, the company developed the EMB 110 – Bandeirante – a turboprop commuter for 18 passengers. In the 80's, Embraer launched Brasilia - a turboprop pressurized aircraft for 30 passengers - its first FAR 25 certified plane. In the 90's, the ERJ 145 - its first Regional jet, and in 1999 Embraer launched the E-Jet 170-190 program, a clean sheet design jet for 70 - 118 passengers [3].

From 1990 to 1996, the company faced a big crisis that culminated with its privatization in 1994. Due to the crisis, big layoffs took place, reducing the number of employees from more than 11000 to 3200.

When the E-Jets 170 -190 program was launched, there was a need of hiring a large number of engineers to cope with the complexity of this new aircraft. The Professional Master's Program in Aeronautical Engineering (PEE) was launched to cope with this challenge [4].

The postsecondary degree system in Brazil is comprised of three levels: Bachelor's, Master's and Doctoral degrees. Typically, undergraduate engineering courses last 5 years, Master's Programs 2 years and Doctoral degree 4 years. The difference between Academic and Professional Master's Programs is that the thesis of the latter could be focused on technology problems instead of science challenges. Also, Professional Master's Programs allow students to be assisted by both academic and industrial advisors.

1.2 Objective

The objective of this paper is to present results of an evaluation of the perception of the most important stakeholders on the impact of the Master's Program as a whole on : i) the alumni's performance on their work, ii) the company (technology, process, products, services and patent) and iii) ITA (Research and Education).

The paper is structured as follows: in literature review, some basic concepts related to Case Study Research and Evaluation Methods for assessing the impact of educational programs are presented; in session 3, the current structure of PEE is presented, as well as some data about the program.

2 Literature Review

2.1 Case Study Research

The case study research method can be used in many circumstances in order to understand their social, political or organizational events in depth. Case studies tend to be more suitable when the understanding of "why" or "how" some phenomena work in certain circumstances, requiring some extensive descriptions. Case studies may also be a combination of both qualitative research and quantitative research [5].

2.2 Evaluation of Training

The importance of evaluation of training, or any development program, lies in determining its effectiveness, which can be meaningful for the ones responsible for the program, its clients and its stakeholders. One of the main reasons to evaluate a training program is to gather relevant information on how to improve it. Moreover, it is paramount to acquire some perspective on how it contributes to the company's goals and objectives [6].

Programs can be evaluated in four levels, according to Kirkpatrick [6], which are: 1) Reaction, 2) Learning, 3) Behavior and 4) Results. The first level can be summarized into "customer satisfaction". In other words, it is a measure of how the participant react to it, which may influence their learning throughout the training program in either a positive or a negative way. The second level, Learning, consists in the improvements and changes on the participant skills, attitudes and knowledge after taking the program. The third level, Behavior, is related to the consequence of one's change of behavior due to the conclusion of a training course, which requires some conditions in order for the change to occur. In this sense, it means to check if the participant behaves and is using the attitudes, knowledge and skills developed in the course, in his or her activities. At last, the fourth level, which is the main concern of the present work, is related to the results of one's attendance to a program. It can be noticed by an improvement on quality, decrease of rework, costs, and increase in production [6].

Another way to measure the contribution of training programs is the measurement of the Return on Investment (ROI), whose methodology has been adopted by many organizations in manufacturing, government, educational settings and others. It includes a fifth level of evaluation of training to the previous model developed by Donald Kirkpatrick: 1) *Reaction and Planned Action* – related to the satisfaction; 2) *Learning* – related to the learned skills; 3) *Application and Implementation* – focused on whether the learned content was either applied or not by the participants on their jobs; 4) Business Impact – related to the business impact of the training program; 5) *Return on Investment* – related to the monetary value of the program compared to its costs [7]. Fig. 1 summarizes the approach suggested by the ROI Institute [8].
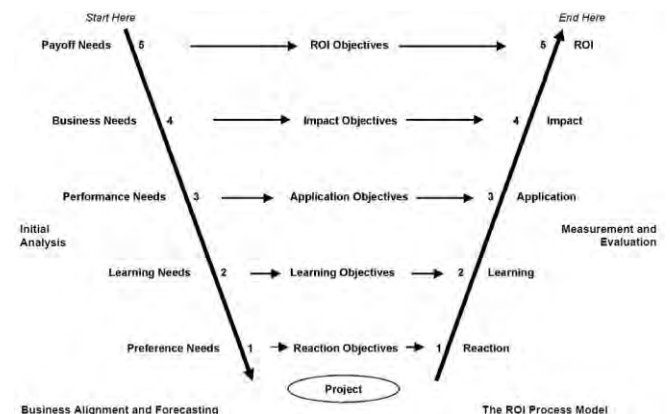

Figure 1 – The ROI Methodology Model. Photo: The ROI Methodology in 10 Easy Steps. (Adapted from [8]).

3 Master's Program Overview

The first class of PEE was launched in 2001. Until 2019, the program has already prepared more than 1500 engineers, out of whom more than 900 concluded their Master's theses. Based on an internal report issued in July 2019, more than 500 of the latter are still working in the company.

In fact, PEE is an Embraer's deliberate strategy of attracting, selecting, developing and retaining young engineers from all over Brazil, to get them involved in multidisciplinary teams to develop new aircraft. It is also an effective way of knowledge management, in the sense that experienced engineers, working in the program as instructors and mentors, share their knowledge with the future generations.

3.1 Pedagogical Model

One of the key factors of the success of PEE is the capacity of attracting and selecting talented engineers. The recruitment is done in both websites of ITA and Embraer.

The selection process of future students (adopted by Embraer and ITA) takes into consideration a set of competencies organized in the following categories: Engineering fundamental knowledge, personal skills, interpersonal skills and conceive-design-build-test engineering systems [9]. It comprises three parts.

In the **first part**, the candidates take an online test to assess their proficiency in English and their competence on handling questions of logic. Over the last five years, the PEE selection process has been able to attract more than 5000 recently graduated candidates from Engineering Schools all over Brazil. This first part brings this number down to 1500 candidates.

The **second part** in the selection process is a written exam held in more than 10 capitals and cities of several Brazilian states. The exam covers topics usually addressed in the first two years of any undergraduate Engineering course in Brazil. Questions are specially prepared by ITA Faculty, on Calculus, Differential Equations, Physics, Strength of Materials, Fluid Dynamics, Thermodynamics, and Heat Transfer, among other topics. The candidates are also tested in English, Technical Drawing and basic knowledge on Computer Science. From these results, the number of candidates is reduced to 200.

The **third part** is composed of interviews and group dynamics, which are held at Embraer facilities in São José dos Campos; ITA professors and Embraer employees are the examiners. Borrowing some ideas from the CDIO Approach [10], candidates are evaluated on their competencies on personal, professional and interpersonal skills. The final result of the selection process is a group of 30 to 40 engineers.

The academic part of the program is comprised of **four phases**, represented in fig. 2. During the **first phase**, all the students take the same classes. From ITA, courses are related to basic knowledge on Aeronautical Engineering (Aerodynamics, Performance, Flight Mechanics, Propulsion and Structures), Material and Manufacturing Processes, Control, Integrated Logistic Support and Integrated Product Development. In addition to ITA academic disciplines, the students take a series of Embraer courses: Introduction to Aeronautical Engineering, Introduction to Aircraft Systems, Aviation Business, System Engineering, etc. These Embraer courses are taught by very experienced engineers from the company, as well as recently retired specialists.
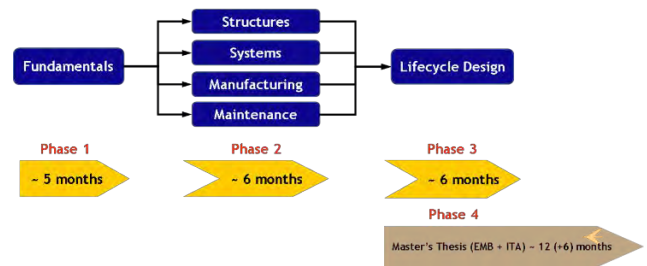

Figure 2 – General Structure of the Professional Master's Program.

In the **second phase**, students are separated in four tracks, according to their own preferences and the company's needs. On the **Structures** track, the students learn about Structural Analysis and Design, Fatigue and Fracture Mechanics, Structure Dynamics and Aeroelasticity, and Analysis and Design of Composite Structures. On the **Aircraft Systems** track, courses address different areas such as Electrical Systems, Communication Navigation and Surveillance, Hydraulic Systems, Modelling and Simulation, and Sensors and Measurement Systems. The **Manufacturing** track comprises courses on Production Management, Design and Manufacturing of Composite Structures, Modelling and Simulation of Discrete Systems, Manufacturing of Complex Surfaces and Topics in Advanced Manufacturing. Finally, the **Maintenance** track includes courses on Aircraft Maintenance, RAMS (Reliability, Availability, Maintainability and Safety), e-Maintenance, Reliability Centered Maintenance, Logistics in the Development, Acquisition and Operation of Complex Systems and Modelling and Simulation of Discrete Systems. Still during phase 2, students have to select the theme of their Master's theses, associated with a pair of advisors: academic (ITA) and industrial (Embraer).

The **third phase**, named by ITA "Industrial Internship", which lasts one semester, is characterized by a "Conceive – Design" exercise in which the students are divided into two teams to simulate the Conceptual and Preliminary Design of a new aircraft [11]. The deliverables of each team, during phase 3, are High Level Requirements, Value Proposition, Conceptual Design and Preliminary Design, including Business Case to demonstrate Technical Viability and Economical Feasibility. The real "magic" behind this exercise is that the two teams design and develop a completely new aircraft, with the support of more than 15 mentors from the company and from ITA, see fig. 3. Mentors represent different areas and technologies: Marketing, Lean Development, Project Management, Aerodynamics, Design, CATIA, Structures, Systems, Manufacturing, Maintenance, Interiors, Materials, Product Integrity, Business Plan, etc. Embraer considers this exercise pivotal for social-technical integration of the future engineers.

During the **fourth phase**, the students have to finalize their Master's theses. Since the company has already hired them at this moment, they have some hours per month set aside for this purpose, which are classified officially as training hours.

Figure 3 – Group of mentors of the Conceive-Design phase of the Professional Master's Program.

4 Methodology

This section presents the methodology applied in the current case study research, consisting of face-to-face interviews and online questionnaires involving **triads** composed of the PEE alumni, their direct supervisors and industrial advisors (see fig. 4). In order to evaluate the impact of the Master's Program conclusion, a case study research was conducted with the classes 22, 23 and 24 of PEE. The survey on the individual perceptions was composed of two parts. The first part consisted in interviews with a focus group of selected students from class 23, their correspondent supervisors and industrial advisors. Four triads were selected to participate in face-to-face, semi-structured interviews.

The objective of this pre-test was to identify the perception of the interviewees on the main contribution of the Master's conclusion on personal aspects, impacts on their work sector of the company and on Embraer as a whole (technology, processes, products, services and patents). During this part, the technology areas, in which the PEE alumni developed their Master's theses, were classified according to ASTERA taxonomy adopted by the European Aeronautics Science Network (EASN) [10].

Figure 4: The triad. Photo: the authors.

The pre-test for the interviews contained both open and closed questions. The closed questions followed the Likert scale [5], i.e. the respondents were asked to indicate their level of agreement or disagreement with each affirmative statement presented, ranging from "strongly agree" to "strongly disagree".

Each group of the triad had its specific outline for the interviews. Thus, supervisors could share their perception on the alumni's work at Embraer, while the industrial advisors would discuss on the Master's thesis impact on the company.

The alumni's outline covered the identification of the respondent, the technology areas (EASN) in which their Master's theses were developed and their contribution to the work on their sector of the company and possible impacts on Embraer (if the results of their theses were implemented by the company).

The supervisor's interview outline consisted in both open and closed questions, covering the identification of the respondent, the contribution of their respective alumni's thesis to their work and to Embraer. The closed questions addressed the main contribution to the alumni's work in terms of quality, cost, time, rework and process.

The industrial advisor's outline covered the identification of the respondent and the possible impact of the Master's thesis results on the company (if implemented by Embraer). The possibility of the industrial advisors not being able to evaluate the work of the alumni on the company was also considered during the elaboration of the outline.

The results of the pre-test allowed the elaboration of the questionnaires of the second part of the case study research, which was conducted inside the company through an online platform (SharePoint). The questionnaires of Part 2 were also specific for each group of the triad. The alumni's questionnaire contained the identification of the respondent, the relation between the theme of the thesis and the alumni's work sector in the company, the implementation of the results of their thesis and the technology area in which the thesis was developed. The supervisor's questionnaire covered the identification of the respondent and closed questions on the contribution of the thesis to the professional maturity of the alumni, to their work and possible impacts on the company. The industrial advisor's questionnaire was composed of the identification of the respondent and closed questions on the possible impacts of the results of the thesis on Embraer. Also, there were closed questions on the contribution of the results of the thesis on the alumni's work at their work sector, if the advisor felt comfortable to answer that.

At the end of every questionnaire, there were some open questions so the respondents could make observations on their answers and include other aspects that were not mentioned in the interview outline.

5 Results and Discussion

5.1 Embraer

The present case study covered the three last PEE classes to finish their Master's theses, considering only the alumni who were still working at Embraer by the time the questionnaires were answered and also by their respective supervisors and industrial advisors, as shown in tab.1. It is important to notice that some supervisors and industrial advisors could have more than one employee or graduate student from these classes.

For the pre-test, 4 triads were selected based on some aspects regarding the alumni's performance and leadership skills shown throughout the program, their supervisor's experience in the company and their advisor's mentoring experience.

Thus, the interviews covered the most qualified members of each group in order to obtain relevant data for designing the questionnaires of the second part of this case study.

Table 1 – Responses and total numbers of each group.

Group	Responses	Total
Alumni	62	81
Supervisors	27	58
Industrial Advisors	32	47

Out of the results obtained from the **alumni's interviews**, the main contributions from the Master's Program conclusion were related to aspects of personal and professional development, technical competence and networking. Possible impacts on the company would be the improvement of processes and products.

Based on the **supervisors' interviews**, the main contributions from the Master's Program conclusion to their employee's work were related to aspects of academic maturity, deeper knowledge related to the subjects of his/her work sector and technical competence. Some possible impacts on the company would be the development of new technologies and improvement of processes and products.

Finally, the results of the **industrial advisors' interviews** showed that the possible impacts on the company would also be related to the development of new technologies and to the improvement of processes and products.

The number of alumni who answered the second part of the questionnaire is presented in fig. 5.
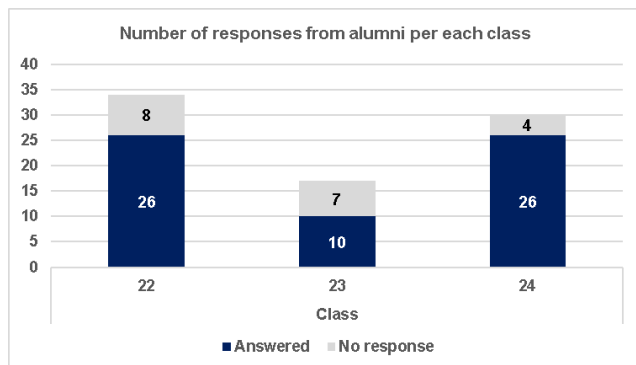
Figure 5 – Number of responses from alumni per each class.

5.1.1 Alumni

The results of the alumni's answers to the questionnaire of the second part of the present work are summarized in the following figures.

The technology areas [12] in which the Master's theses were developed are represented in fig. 6. These results are consistent with the fact that most of the engineers are enrolled in the tracks Systems and Structures. In addition to that, manufacturing processes are placed under the Aerostructures rubric.
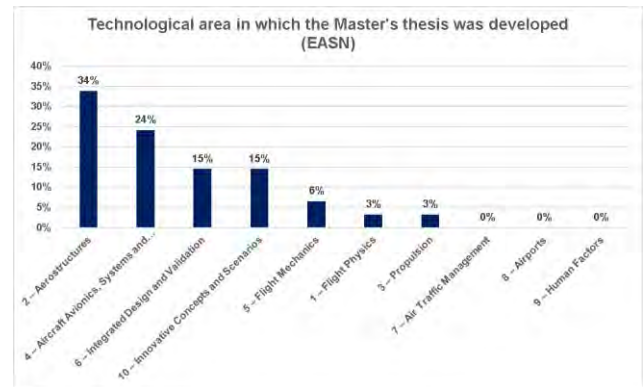

Figure 6 – Technology areas in which the Master's theses were developed.

The alumni answers indicate that the majority of them, 71% actually, planned to acquire a Master's degree before joining the program. This result reinforces the claim that PEE is a strategy of attracting talented people to Embraer.

Fig. 7 shows that, under alumni perspective, the majority of the Master's results were not implemented in the company. Based on other surveys done by the authors with former students from PEE, from previous classes, there are two possible explanations for those answers. One is that it usually takes time for the ideas of PEE's alumni work to be implemented. Another possibility is that the results in terms of new technologies, processes, products etc. are much more related to individual development of the engineers, and not so much to the themes of their theses.
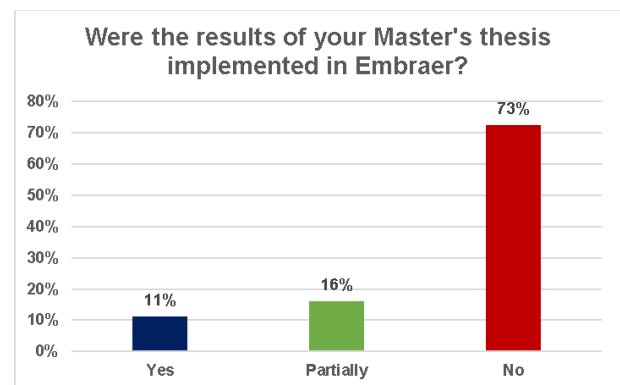

Figure 7 – Implementation of the results of the Master's theses in Embraer.

In terms of contributions of the Master's theses conclusion to the alumni's skills and competences, most of the respondents seem to agree on the improvement of their academic maturity, their knowledge of the company's work sector where they were allocated and their technical competence, as shown in fig. 8.
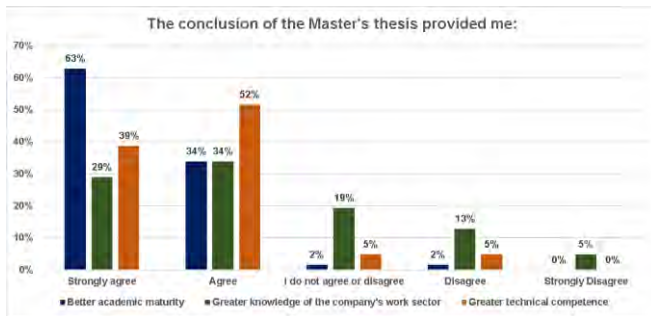
Figure 8 – Influence of the conclusion of the Master's theses on alumni academic maturity and technical competency.

They also seem to agree on the contribution to their personal development and their networking, according to fig. 9.
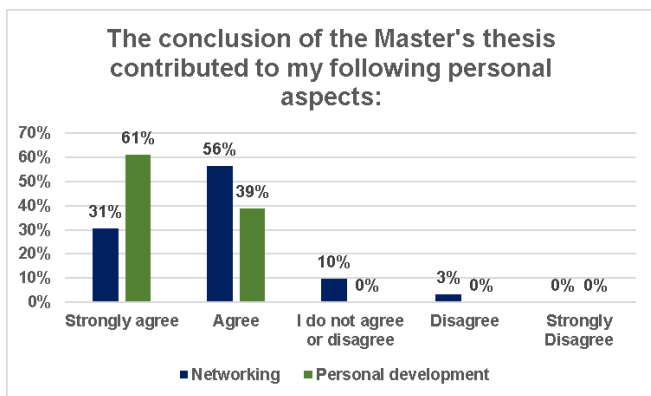

Figure 9 – Contribution of Master's thesis on networking and personal development.

Fig. 10 shows that majority of the alumni agree that there was some contribution to the quality of their work. However, it also implies that most of them do not agree or disagree with contributions to rework reduction, decreased in time demanded or with process improvement.
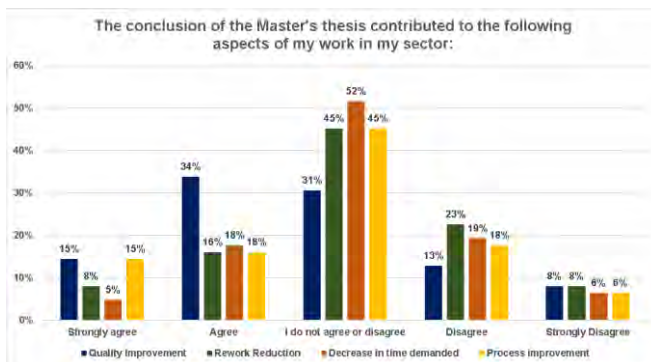
Figure 10 – Contribution of Master's thesis to some aspects of alumni work.

Fig. 11 shows that the majority of the alumni agree that the results of their Master's theses would affect the company on aspects of development and mastery of new technologies, processes and products, though not so much on the development of services and patents.
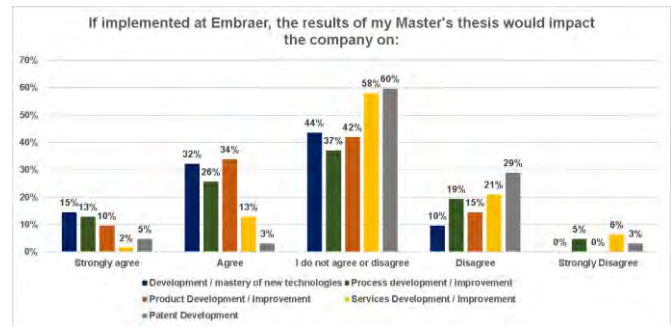

Figure 11 – Implementation of Master's theses on technology, process, product, service and patent.

Fig. 12 shows their perception on which Business Units of Embraer the results of their Master's theses would have impact.
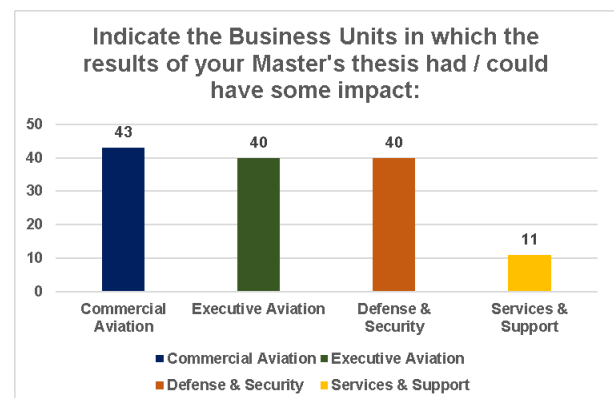

Figure 12 – Possible impact of the Master's theses on Business Units

It is interesting to observe that the perception of impact is evenly distributed, which is very much related to the fact that Embraer has a Chief Engineer organization to ensure harmonization of specialized personnel and knowledge among the different business areas.

Fig. 13 shows that the majority of the alumni developed articles in Congresses, Symposiums or similar and indexed journal articles related to their Master's theses.
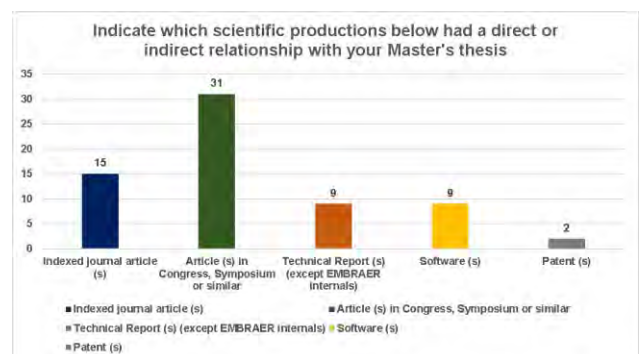

Figure 13 – Scientific productions as results of Master's theses

Results shown in fig. 13 are closely related to the metrics that Brazilian Ministry of Education uses to evaluate academic graduate programs in Brazil [13].

DOI
10.3384/ecp19162031   **Proceedings of the 10th Aerospace Technology Congress**
October 8-9, 2019, Stockholm, Sweden   **268**

5.1.2 Industrial advisors

Fig. 14 shows that the majority of the industrial advisors agree that the better academic maturity, the deeper knowledge of the company's sector and the greater technical competence of their assisted alumni are the main contributions of the Master's Program conclusion.
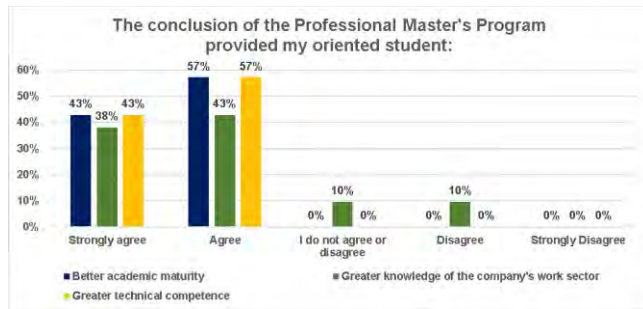

Figure 14 – Perception of industrial advisors on alumni competencies.

According to fig. 15, the majority of the industrial advisors agree on the contribution to the development and mastery of new technologies, processes, and products. However, it seems that most of them are not able to recognize the impact on patent development and do not agree on the services development / improvement.
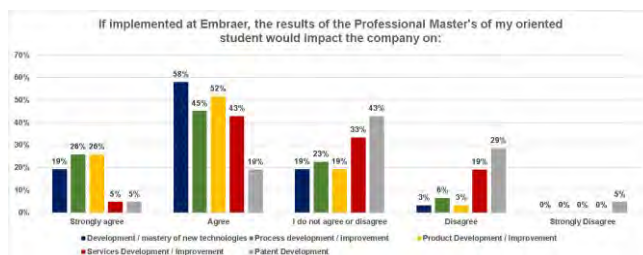

Figure 15 – Perception of industrial advisors on the impact of their Master's students work on the company.

5.1.3 Supervisors

In the context of this paper, supervisors are considered the students' direct managers after the company hires them. One of the most important roles of supervisors in leading their teams is to be aware of the level of competencies of his/her employees.

Fig. 16 shows that the majority of the supervisors agree that there were contributions of the Master's Program conclusion on their employee's academic maturity, knowledge of the company's sector and technical competence. Since the supervisors involved in this survey have other people under their command, this perception is even more important and precise than the alumni's opinion on the same topic.

Figure 16 – Direct supervisors' opinions on the results of the conclusion of the Master's Program on employee's competencies

Fig. 17 shows that the majority of supervisors agree with the contributions of the Master's Program to their employee's work in their sector, in terms of quality improvement, rework reduction and process improvement; however, there is a considerable number of respondents who did not agree or disagree with these aspects.
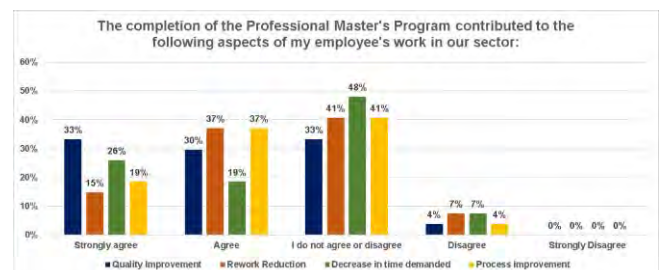

Figure 17 – Perception of supervisors on the contribution on the conclusion of his alumni Master's Program on their work.

Fig. 18 shows that the majority of the supervisors also agree that there would be some impact on Embraer in terms of technologies, processes and products. However, a considerable number of respondents do not agree or disagree with these aspects nor with the impact on services and patents. These results could be explained by the fact that Embraer is still a very "product oriented" company, and the culture that patents is new are much more related to the Technology Development sector than to Engineering areas.
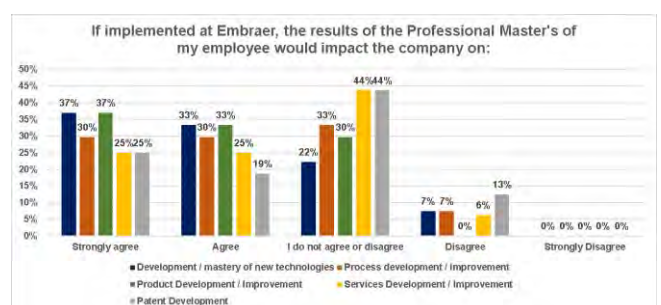

Figure 18 – Perception of Supervisors on the impact of their employee's Master's theses on the company as a whole.

In general, the triads (student, industrial advisor and direct supervisor) agree that the conclusion of the Master's theses contribute to the engineers in terms of both technical skills as well as networking. It could also be observed that industrial advisors and supervisors are more positive with respect of impact on the company than the students are. This could be a consequence of their larger experience.

5.2 ITA

In order to evaluate possible impacts on ITA caused by the conclusion of the PEE Master's theses, a questionnaire was developed for ITA Faculty, who have been the most frequent academic advisors of the alumni of the program.

A pre-test was also carried out in this case study with the academic advisors from ITA. Professors responsible for the Management of the Master's Program were also interviewed so that the main aspects of impacts and contributions from the Master's Program conclusion on the Institute could be identified. Based on those results, a questionnaire was created with closed questions structured with Likert scale, covering impacts on the undergraduate engineering courses of the Institute, the research lines of the graduate programs and cooperative projects. Nine Professors have responded the survey.

Fig. 19 summarizes ITA Professors' perception on the influence of their participation in the Professional Master's program on the Institute. It is noticeable that none of the respondents disagree with the contributions to ITA. In addition, the majority of the Professors agree with the influences of the Master's Program on the conduction of their undergraduate courses, on their research lines in graduate programs and in cooperative projects involving FAPESP, FINEP, VINNOVA and others.

Some of the respondents claimed that their interaction with Embraer, during the development of the alumni's theses, allowed them to adapt the syllabus of their undergraduate courses according to the topics that were relevant for the company, in order to prepare the Engineering students to work in the Aeronautical Industry as a whole.
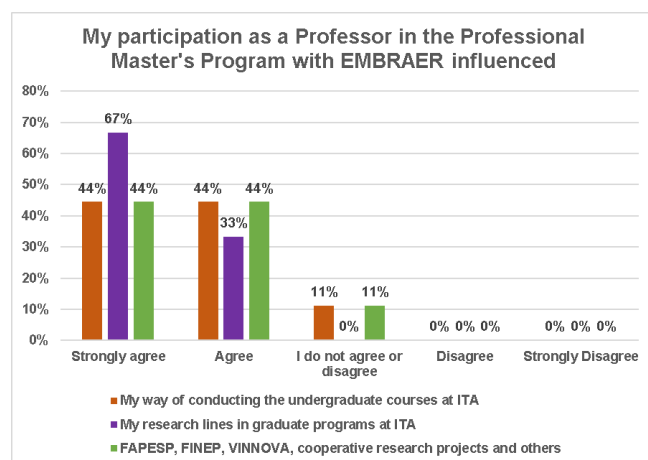

Figure 19 – Perceptions of the influence of the Professional Master's Program on ITA.

Others have stated that their interaction with Embraer specialists highlighted the main technological challenges that are considered strategic for the Aeronautical field in the near future, hence influencing their research lines at ITA. In addition, some respondents have commented that their contact with the company through the PEE Master's Program brought them a broader view of the technological challenges of the Aeronautical sector and some maturity for proposing applied projects that were relevant.

At ITA, the majority of the respondents agree that since its creation in 2002, the professional Master's Program has had a strong influence on updating courses contents in the undergraduate level, to be more adequate to current challenges in Aeronautical areas. In the graduate level, working together with experienced people from the company (industrial advisors) provided them not only some interesting challenges, but also new ways of raising funds.

6 Comments and Conclusions

This paper was prepared to show some results about the perceptions of the most important stakeholders involved with the Master's Program in Aeronautical Engineering, on the students' development, on Embraer in terms of new technologies, processes, products, services and patents, and on ITA regarding education and research.

The presented results indicate that Embraer employees, as well as ITA Faculty, recognize that the Professional Master's Program (PEE) generated positive outcomes for both organizations.

Since the total retention rate (from all graduated classes) of PEE students as Embraer employees is around 50%, it is reasonable to suppose that they have strong influence on the quality of products launched in the last decade, such as KC-390, Praetors and E-2 Family [14]. A recent internal survey indicated that among the engineers who developed those planes, around 30% of them are PEE alumni.

ITA and Embraer have recently embarked on a new education initiative, an Academic Professional Doctorate, similar to PEE, in which the students also have two advisors (academic and industrial) whereas the dissertation should address a challenge proposed by the company and accepted by ITA. Funding support is shared between ITA and Embraer, and the company contributes to the student's scholarship.

References

[1] History of CTA, 2019. Available: http://www.iae.cta.br/index.php/history. [Accessed Jul 19].

[2] EMB 100 Bandeirante, 2019. Available: https://historicalcenter.embraer.com/global/en/emb-100-bandeirante [Accessed Jul 19].

[3] Historical Aircraft of EMBRAER, 2019. Available: https://historicalcenter.embraer.com/global/en/historical-aircraft [Accessed in Jul 19].

[4] D. Andrade, S.F. Almeida, L.C.S. Goes, "Experience in Strategic Partnership: Professional Master's in Aeronautical Engineering, ITA-Embraer". COBENGE, 2003.

[5] R. K. Yin, *Case Study Research and Applications: Design and Methods*. Thousand Oaks: Sage Publications, 2017.

[6] D. Kirkpatrick and J. Kirkpatrick, *Evaluating Training Programs - The Four Levels*. San Francisco: Berrett-Koehler Publishers, 2006.

[7] P. Phillips, J. J. Phillips, R. D. Stone, and H. Burkett. *The ROI Fieldbook – Strategies for Implementing ROI in HR and Training*. Burlington: Butterwoth- Heinemann, 2007.

[8] The ROI Methodology in 10 Easy Steps – Application Guide. The ROI Insititute, 2013.

[9] CDIO Initiative, 2019. [Online]. Available: http://cdio.org/ [Accessed Jul19].

[10] CDIO Syllabus 2.0, 2019. [Online]. Available: http://cdio.org/benefits-cdio/cdio-syllabus/cdio-syllabus-topical-form [Accessed Jul 19].

[11] P. Lourencao, F. Rosa and O. Resende. "Use of Conceive-Design Learning Environments to Prepare Engineers for the Development of Complex and Highly Integrated Aeronautical Systems". Proceedings of the 12th International CDIO Conference, Turku: Finland, 2016.

[12] Aeronautical Research and Technology Areas. Available: https://easn.net/research-technology-areas/ [Accessed Jul 19].

[13] Ministry of Education – Evaluation Criteria: http://www.capes.gov.br/avaliacao [Accessed Jul 19].

[14] Embraer´s webpage - https://embraer.com/global/en [Accessed Aug 19]