Interactive Visual Analytics of Production Data - Predictive Manufacturing

Juhani Heilala, Paula Järvinen, Pekka Siltanen, Jari Montonen, Markku Hentula, Mikael Haag

VTT Technical Research Centre of Finland Ltd, Espoo, Finland, http://www.vttresearch.com/ {juhani.heilala, paula.jarvinen, pekka.siltanen, jari.montonen, markku.hentula, Mikael.haag}@vtt.fi

Abstract

Manufacturing creates a lot of data, and this is increasing due to digitalization of manufacturing, industrial Internet of Things (IIoT) and needs for product traceability as well as predictive maintenance. Typically data from production material flow is not analyzed and thus the improvement potential is not found. There is need for interactive analytics tools that can turn raw data from heterogeneous data sources e.g. starting from sensor data, manufacturing IT systems, (e.g. Enterprise Resource Planning, ERP, Manufacturing Execution System, MES and Supervisory Control And Data Acquisition, SCADA), into meaningful information and predictions—and presented on easy-to-use interfaces. This paper presents a feasibility study focusing on interactive visual analytics of manufacturing data set carried out at VTT Technical Research Centre of Finland Ltd.

Keywords: manufacturing industry, statistical analysis, machine learning, visual analytics, industrial internet of things

1 Introduction

The role of data in manufacturing has traditionally been understated. Manufacturing generates about a third of all data today (Simafore, 2013), and this is certainly going to increase significantly in the future. Data forms the backbone of all Digital Manufacturing technologies, which will be the centerpiece of the strategy for advancing Manufacturing in the 21st century (Simafore, 2013).

Manufacturing companies are facing global competition – they have to be better, cheaper and faster. In order to manage they need a productivity leap. Manufacturing companies collect huge amount of data of their manufacturing processes. Even though utilization of the data could potentially enable a big productivity leap, this data is poorly used. A recent survey published (MESA, 2016) find out that only 14% of respondents are using manufacturing data in analytics. This is largely because manufacturing companies lack tools and expertise needed to analyze the data. On the other hand, Industrial Internet of Things (IIoT), analytics and simulation methods,

collaboration and visualization tools are mature enough to be used and worldwide interest for applying them exists.

Many of current analytic tools require data analytics expertise and are mainly used as a desktop, "island of analysis". Current analytics tools are also static, pre-programmed, focusing mainly on business issues and targeted for upper management level. Typically those tools are expensive and aimed for large organizations. The current tools also have poor synchronous collaborative analysis and decision making features. There are several useful analytics methods that are not included in the current tools.

Advanced analytics refers to the application of statistics, machine learning, data mining and other mathematical methods to manufacturing and business data in order to assess and improve practices.

Predictive analytics is about extracting information from existing data, in order to determine patterns and predicting potential trends and outcomes. Predictive analytics forecasts what might happen in the future. The goal is to go beyond descriptive statistics and reporting on what has happened to providing a best assessment on what will happen in the future. The end result is to streamline decision making and produce new insights that lead to better actions.

In manufacturing, operations managers can use advanced predictive analytics to take a deep dive into historical process data, identify patterns and relationships among discrete process steps and inputs, and then optimize the factors that prove to have the greatest effect on yield.

For networking manufacturers, the IIoT becomes a full ecosystem when software, cloud computing (or in-house servers), and analytics tools are combined to turn raw data into meaningful information or predictions—and when it's presented on easy-to-use interfaces (such as dashboards or mobile Apps) enabling users to monitor, and in some cases, automate response actions or remotely control equipment or systems (PwC, 2015).

The research question is how to convert manufacturing big data to business and manufacturing advantage? In this article we present a feasibility study of applying visual analytics to manufacturing data. For

the purpose we use VTT OpenVA visual analytics platform for measurement data (Järvinen *et al*, 2013). The objectives are twofold: getting experience and guidelines for applying visual analytics in manufacturing and analyzing the feasibility of the VTT OpenVA platform with manufacturing data.

1.1 Impact estimations

Manufacturers taking advantage of advanced analytics can reduce process flaws, saving time and money. Gains will likely show up in both labor productivity and resource productivity: The impact estimations are (MGI, 2011).

- Sensor data-driven operations analytics: -10-20% operation costs, up to +7% revenue
- "Digital Factory" for lean manufacturing: -10-50% assembly cost, + 2% revenue

For extended enterprise real-time visibility between suppliers and the production line allows key value chain participants to optimize material flow and reduce process cycle time. Furthermore, the use of predictive and prescriptive analytics using real-time data allows the enterprise to rectify future bottlenecks and eliminate high costs associated with operational downtime.

General Electric estimates that full-scale exploitation of the industrial internet potential will bring an annual one percentage point increase in global production for the next 15–20 years. One percentage point might not sound much, but calculating one percent growth over fifteen consecutive years as compounding growth we end up with a global increase of ten to fifteen trillion dollars in national product – that is to say, an increase in products and services to the tune of 10,000 or 15,000 billion dollars each and every year. If even a part of this can be realized, we will have gone some way beyond mere hype! (Ailisto 2014).

An efficient workforce with strong data analytical skills and cross-domain expertise can facilitate transition into a smart factory.

2 Visual Analytics in Manufacturing Domain

One key change in the manufacturing process will come in the form of visual analytics (Riley 2015). For many industrial companies, the Industrial Internet of Things (IIoT) is the main source of Big Data. IIoT connects data produced by different objects – such as sensors, devices, machines, humans, other assets, and products – to different applications. IIoT provides access to data generated and manipulated by e.g. intelligent equipment with an IP address, machine-to-machine (M2M) communications, mobility, cloud computing, analytics, and visualization tools.

Inside the factory, having the ability to utilize data masses from orders and machine status allows production managers to optimize operations, factory scheduling, maintenance, and workforce deployment (Noor, 2014).

2.1 Visual Analytics

Visual analytics (Thomas and Cook, 2009; Järvinen *et al*, 2009; Keim *et al*, 2010; Järvinen, 2013,) provides visual and interactive tools to support analytical reasoning and finding insight from data. It combines the human capabilities to interpret visualizations with automatic data processing. A visual analytics tool shows the information in the form of interconnected and interactive visualizations, making the analysis easy for non-experts in data analysis. Behind the visualizations are statistical, data mining and machine learning methods. Users can look for patterns, trends, anomalies, similarities and other relevant features from the visualizations. Visual analytics is an iterative process, (see Figure 1), where users launch analysis, browse and navigate in visualizations, and highlight and select important areas for further study.

The use of visual analytics is still rare in manufacturing operation management. Examples on visual analytics at plant monitoring are shown by (Aehnelt *et al*, 2013; Tack *et al*, 2014).
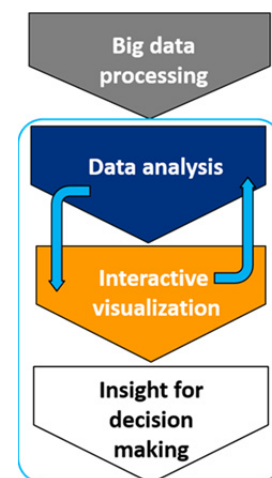

Figure 1. Steps in interactive visual analytics

2.2 Heterogeneous data sources in manufacturing domain

Heterogeneous data stems from various data sources and that comes in a multiplicity of data formats. In the domain of manufacturing, for example enterprise resource planning systems (ERP) are used to manage information about orders and personnel, while manufacturing execution systems (MES) are employed to collect and evaluate data about the production process. For an integrated analysis, different data access interfaces to the different data sources must be used (Aehnelt *et al*, 2013).

In addition to MES and ERP, there are various other systems at factory floor having useful information, such as different sensor systems, SCADA (supervisory control and data acquisition) and various automated systems log files. The most difficult data source is the human at the factory floor, how to get real-time information from him or her. ERP doesn't provide a real time feedback loop from production floor to the planning level. There is latency, delays on submitting the progress data to the ERP and potentially many human interactions are required, thus manual errors are possible. The time stamps from production phases can have inaccuracies. Typically the work phases status are entered at the end of the working shift; there could be the same time for starting and finishing the work phase or some of the work phase recordings are missing.

In case of manual reporting on paper, it requires another human action to type data to the information systems. Information on work progress status, exceptions, missing parts and quality etc. are needed for e.g. real-time control of manufacturing (Järvenpää et al, 2014a). Methods for analyzing this historic data are typically missing (Järvenpää et al, 2014b).

2.3 VTT OpenVA concept

VTT OpenVA is a visual analytics platform for measurement data by VTT Technical Research Centre Of Finland Ltd. It consists of a data base, a library of visualization and analysis methods, an interactive user interface and a visual analytics engine that delivers data and analysis requests between the different components.

The database stores the application data in a domain independent form. The data base contains data of background variables, measured variables and indicators, and it is populated with application specific metadata.

The analysis and visualization library contains a selection of analysis and visualization methods, and it is extendable. The visual analytics tool adapts itself to each application with the help of the metadata stored into the database. The data to be analyzed is loaded from external sources to the databases through a uniform data interface.

3 Feasibility Test Case – Automated Material Handling System Analysis

The feasibility of OpenVA platform for analyzing manufacturing data was tested with a small sample of industrially relevant data set. The motivation for the feasibility test was to see if the platform would help in the following goal: Advance from descriptive and diagnostic analytics towards predictive analytics – with prescriptive analytics as the next generation capabilities. This could be described as a move from traditional questions of "what happened?" and "why did it happen?" towards a questions "what will happen?" and "how can we make it happen?"

In the feasibility test case, the focus was on visual and interactive analytics, from status monitoring to predictive analytics. The problems of getting data were not studied. The data sets were structured data from a simulation study of an automated material handling system. The feasibility test data was similar to real industrial data that is typically automatically collected from automated equipment and robotics e.g. working time, disturbances, set-up and process times, utilization rate, Overall Equipment Efficiency (OEE) data etc.

In the feasibility test we were utilizing the following parameters from the automated material handling system: capacity (pieces/hour), various equipment (3 machines, 2 robots and lifter) utilization rate, operating time, and storage content (Figure 2). The monitoring of equipment data shows, that production (pieces/hour), utilization rates of robots are low while the utilization rates of the machines are high.
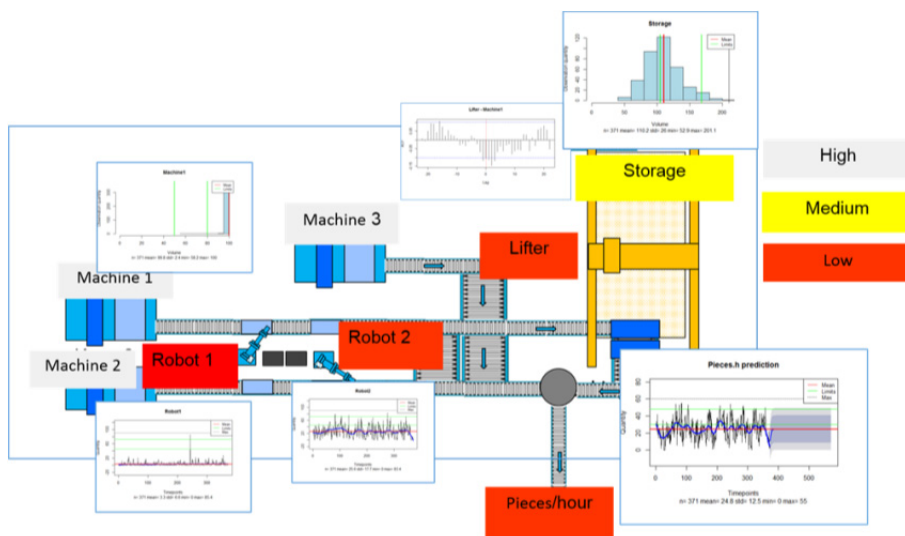

Figure 2. Automated material handling system and equipment's data visualization.
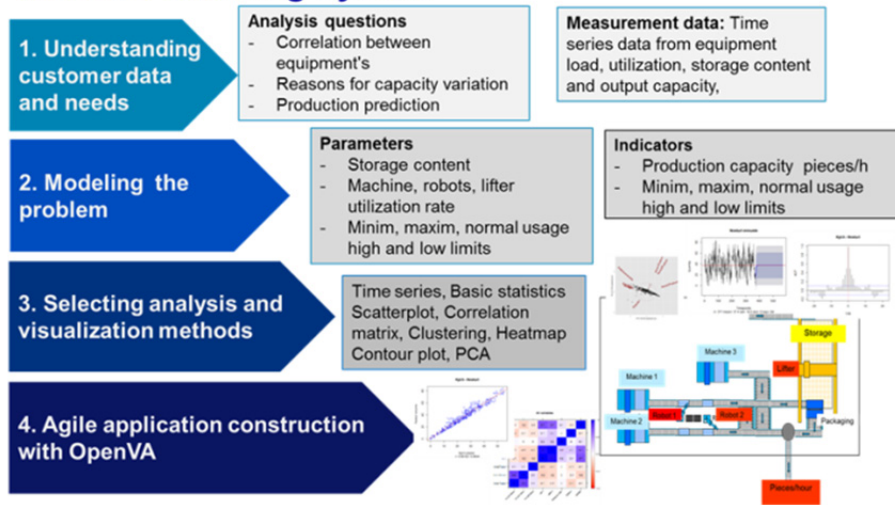
Figure 3. Setting up OpenVA test on automated material handling system

4   Methodology for Setting Up Interactive Visual Analytics

OpenVA is applied by a step-by-step configuration process (**Error! Reference source not found.**). The first step is to understand the customer business, the analysis needs and find out what data is available. Then the next step is to define the phenomena that are followed, to identify the variables that might explain the system behavior and form indicators from the variables. In the third step appropriate analysis and visualization methods are specified. A set of methods is already provided by OpenVA, but new methods can be added. In the final step the analytics application is constructed by configuring the OpenVA platform and loading the application onto the platform database.

In our feasibility study the analysis questions were to study the efficiency of the automated material handling system, to predict the production and find bottlenecks. The selected variables were the utilization rates of each production line component. The indicator chosen for the system production output was finished products/hour. The set of analysis and visualization methods included time series, histogram, contour plot, scatterplot, cross-correlation, correlation matrix, Principal component analysis (PCA), clustering and logistic regression.

The analysis with OpenVA is performed as an iterative reasoning process. First, the user is shown the current status of the performance indicators and the most important variables (Figure 2). The user can study indicators and the other variables in detail with the help of visualizations.

The user starts the analysis by formulating an analysis questions, e.g. "What explains the low production?" Next, the user selects the variables and indicators that might give answers to the questions. The tool suggests suitable analysis and visualization methods to the user based on the number and type of the selected variables. Then the user launches the analysis and gets the results in visual form.

In the analysis of the automated material handling system the most interesting result is shown by the correlation matrix (Figure 4). It shows that the lifter and the production indicator (pieces/hour) have a complete positive correlation, suggesting that the lifter might be the bottleneck of the system. The PCA and logistic regression formula confirm the result. Thereby answers to the analysis questions are: the efficiency of the production line is low, the lifter is the bottleneck of the system and the capacity of the production line is predicted by the lifter alone.

5   Discussion

Predictive analytics applies different analysis methods to predict the possible outcomes of the events that the data describes. It does not give exact measures of what will, could or did happen, just the possibilities of what may occur.

In small series production, different requirements (e.g. high mix, low volume, multiple jobs with different due dates, routing and process time requirements) need to be concurrently processed through the same production setup competing for shared resources of limited capacity. Extreme complexity is characteristic in the discrete made-to-order production, where availability of materials, machines and manpower creates dynamic or moving bottlenecks.

Data quality is a well-known problem in data mining research (Rahm and Do, 2000). The problem is partly caused by missing or erroneous measurements, as well
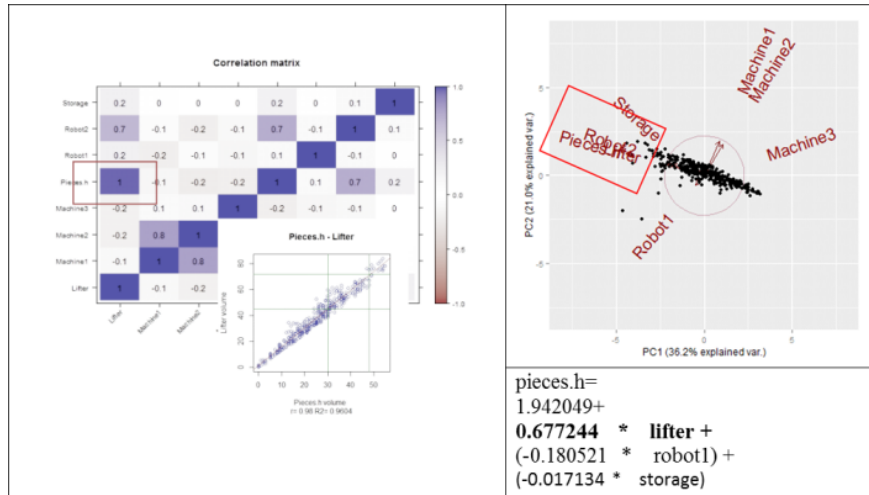
Figure 4. Correlation matrix, cross correlation, PCA and logistic regression

as disparate data formats when combing data from several sources In practical implementations, for example when comparing and estimating measurements from different data collecting applications, the data quality problem is faced immediately. The applications have their own databases, each storing the monitored data in different format. In our test case, the data resulted from a simulation study, and therefore data was cleaner than in a normal industrial case.

The production managers need easy to use tools for finding and eliminating those moving bottlenecks and doing manufacturing process improvements. The production manager needs to consider two different time frames: on the one hand development of equipment and control principles with longer planning and implementation time and on the other hand daily operative decision making. Even if amount of data is low (e.g. because of low production volumes), the predictions do give valuable insight to potential near future events.

As discussed in this paper, the visual and interactive tools can be used to support analytical reasoning and for finding insight from data. However, the visualizations should be chosen case by case so that they are focused on the task at hand, and support exactly those decisions that must be made.

6 Conclusions

In this study our focus was on interactive visual analytics. The key findings regarding the future use of analytics are:

- Instead of looking manually multiple time history plots, numerical tables and reports, one can use the power of interactive visual analytics.
- Getting information from heterogeneous sources, collected from factory floor, manufacturing ICT

systems or production engineering design systems for decision making and improvement planning.

- Comparing high and low productivity days, finding correlations, patterns, exceptions, digging deeply to the facts behind events, learning, reasoning and pin pointing improvement areas.
- Visual analytics can diminish the need of multiple customer interviews or simulation studies in a development project by getting findings from existing data.
- Visual analytics can be used with simulation studies, to enhance analysis of the simulation run results in order to get deeper understanding, as done in the feasibility study.

OpenVA platform concept proved out feasible for manufacturing data analysis.

6.1 Future Research

This was a small test with promising results and gives ideas for future research. One of them is synchronous collaborative visual analytics in manufacturing. Analysis and visualization of data are traditionally made asynchronously by individual users with their local tools. Research on synchronous collaborative visual analytics is still in its infancy. Supporting synchronous, multi-party collaboration over networks is becoming increasingly important in order to increase the efficiency of data analysis. An important objective is to combine the best ideas of collaborative work with those of Visual Analytics, i.e. to support interactive collaborative visualizations in multi-party settings. This enables the decision makers to get best possible support from e.g. data scientist without geographical limitations.

For collaborative visualization, integration of data access, communication and messaging functionalities in the Visual Analytics tools are needed. This will support the exchange of opinions and information over

the problem at hand. In addition it requires support solutions for viewing and processing the design collaboratively, linking comments to the specific parts of the object (documents, drawings, 3D designs, visualizations, data sets, etc.), as well as accessing, classifying and filtering those comments at any desired way.

The other development need is reliable and real-time access to data from heterogeneous data sources, from factory floor, sensors, devices, human operators and manufacturing information systems.

For building an IIoT system, all following topics are needed: connectivity, data management, analytics, and interoperability. Reference architectures as well standardization in this domain are evolving e.g. Reference Architecture Model for Industrie 4.0 (RAMI4.0) and the Industrial Internet Reference Architecture (IIRA).

The ability to analyze large amounts of complicated, heterogeneous data with custom-written visual analytics will be key component in the future business and industrial intelligence – analytics. Data-driven decision-making in manufacturing enables productivity leap.

Predictive manufacturing analytics enables users to:

- Progress from monitoring to predictive analytics, optimization and to "how can we make it happen?"
- Near real-time warnings of potential problems, embedded dashboard to factory floor, etc.
- Analyze production characteristics and business performance.

Acknowledgment

References

M.Aehnelt, H.Schulz and B.Urban. Towards a Contextualized Visual Analysis of Heterogeneous Manufacturing Data. In G. Bebis et al. (Eds.): *ISVC 2013, Part II, LNCS 8034,* pp. 76–85, Springer-Verlag Berlin Heidelberg, 2013.

H.Ailisto. Industrial Internet - hype or revolution? 2014. Available at https://vttblog.com/2014/05/07/industrial-internet-hype-or-revolution/

E.Järvenpää, H.Tokola, T.Salonen, M.Lanz, M.Koho and R. Tuokko. R. Requirements for Manufacturing Operations Management and Control Systems in a Dynamic Environment. *In Proceedings of the 24th International Conference on Flexible Automation & Intelligent Manufacturing* -FAIM 2014. San Antonio, Texas; University of Texas at San Antonio. Center for Advanced Manufacturing and Lean Systems. 2014a, http://dx.doi.org/10.14809/faim.2014.1135

E.Järvenpää, M.Lanz, H.Tokola, T.Salonen, M.Koho, J.Backman, K.Katajisto and H.Reinilä LeanMES: Tuotannonsuunnittelu ja -ohjaus suomalaisissa valmistavan teollisuuden yrityksissä. Nykytila, haasteet ja tarpeet, 2014. LeanMES project report 2014b

P.Järvinen. A Data Model Based Approach for Visual Analytics of Monitoring Data. Licentiate thesis. Aalto University School of Science. Department of Information and Computer Science. 2013. Available at http://lib.tkk.fi/Lic/2013/urn100763.pdf

P.Järvinen, K.Puolamäki, P.Siltanen and M.Ylikerälä. Visual Analytics. Final report. *VTT WORKING PAPERS* 117. 2009. Available at http://www.vtt.fi/inf/pdf/workingpapers/2009/W117.pdf

P.Järvinen, P.Siltanen, K.Rainio. Framework for Visual Analytics of Measurement Data. INFOCOMP 2013: *The Third International Conference on Advanced Communications and Computation*. ISBN: 978-1-61208-310-0, 2013.

D.Keim, J.Kohlhammer, G.Ellis and F.Mansmann, (eds), Mastering the information age: solving problems with visual analytics. First edition edn. Goslar, Germany: Eurographics association, 2010

MESA International and LNS Research. Survey on Metrics that Matter in the manufacturing world. MANUFACTURING METRICS IN AN IoT WORLD. Measuring the Progress of the Industrial Internet of Things, 2016.

MGI. McKinsey Global Institute. Big data: The next frontier for innovation, competition, and productivity 2011.

A.K.Noor. Putting Big Data to Work. American Society of Mechanical Engineers (ASME), 2014.

PwC. The Internet of Things: what it means for US manufacturing. 2015. Available at http://www.pwc.com/us/en/industrial-products/next-manufacturing/big-data-driven-manufacturing.html

E.Rahm and H.H.Do. Data cleaning: Problems and current approaches, *IEEE Data Engineering Bulletin 2000,* 23(4): 3-13, 2000.

S.Riley. Visual Analytics – The Future Of Manufacturing Processes. Manufacturing Business Technology. 2015. Available at http://www.mbtmag.com/article/2015/08/visual-analytics-%E2%80%93-future-manufacturing-processes ,

Simafore. How predictive analytics can shape manufacturing of the future. 2013. Available at http://www.simafore.com/blog/bid/118789/How-predictive-analytics-can-shape-manufacturing-of-the-future

T.Tack, A.Maier and O.Niggemann. On Visual Analytics in Plant Monitoring. In J.-L. Ferrier et al. (eds.), *Informatics in Control, Automation and Robotics, Lecture Notes in Electrical Engineering 283*, 2014. Doi: 10.1007/978-3-319-03500-0_2,

J.Thomas and K.Cook. Illuminating the path: The research and development agenda for visual analytics. 1st ed. Los Alamitos, 2009.