Principal Component Analysis Applied to CO_2 Absorption by Propylene Oxide and Amines

M. H. Wathsala N. Jinadasa[1], Klaus J-Jens[1], Carlos F. Pfeiffer[1], Sara Ronasi[2], Carlos Barreto Soler[2], Maths Halstensen[1]

[1]Faculty of Technology, Natural Sciences and Maritime Sciences – University College of Southeast Norway, Post box 235, N-3603 Kongsberg, Norway, {Wathsala.jinadasa, Klaus.J.Jens, carlos.pfeiffer, Maths.Halstensen}@usn.no

[2]Norner Research, Asdalstrand 291, N-3962, Stathelle, Norway, {sara.ronasi, carlos.barreto}@norner.no

Abstract

Carbon dioxide absorption by mixtures of propylene oxide / polypropylene carbonate at 60^0C was monitored by Raman spectroscopy at 20, 40 and 60 bar in a 2 L autoclave reactor. Multivariate preprocessing techniques were used to process raw Raman spectra and Principal Component Analysis was performed. Simulation data from the Peng- Robinson equation of state were used to model the absorbed CO_2 amount and spectroscopic signals. Results showed that Principal Component Analysis can be used to explore the dynamics of the system at different pressure levels and to track the CO_2 absorption. A similar analysis was carried out to monitor CO_2 absorption by four different amines at room temperature and pressure in a batch reactors. The CO_2 content was determined from titration and was used to model the spectroscopic data. Principal Component Analysis proved to be able to identify CO_2 absorption capacity in the amines. This feasibility study confirms that Raman spectroscopy together with multivariate analysis can effectively report chemical information and dynamics in these CO_2 absorption systems and hence can be used for developing regression models for online monitoring and control.

Keywords: principal component analysis, CO_2 absorption, propylene oxide, amines

1 Introduction

Carbon dioxide (CO_2) is known to be the primary greenhouse gas contributing more than 60% of global warming. Capturing CO_2 from power plants and industrial sources and utilization them to produce usable products is of paramount importance from a standpoint of "waste to money". Absorption of CO_2 by amines is one of the most popular technologies for CO_2 capture. Amines are categorized as primary, secondary or tertiary amine based on their chemical structure. The reaction between amines and CO_2 is complex (McCann *et al*, 2009). However, when considering the CO_2 mass balance, it can be seen that once absorbed by a primary amine, CO_2 will remain in the form of carbonate, bicarbonate, carbamate or molecular CO_2 as given in (1). When it is a tertiary amine, there is no carbamate formation (2).

Synthesis of polypropylene carbonate (PPC) by reaction of CO_2 and propylene oxide (PO) in the

presence of a catalyst has become a fascinating research area as a CO_2 utilization technique to produce a polymer out of a waste greenhouse gas (Jiang *et al*, 2014). In the presence of a catalyst, the chemical reaction of PPC synthesis takes place as given in (3).

CO_2 absorption capacity by an amine or by in the liquid phase PO is a key performance criteria in industrial scale CO_2 capture and polymerization processes. However, the measurement of CO_2 absorption in these mixtures are challenging and require proper understanding of the chemistry behind reaction (1), (2) and (3). Several offline analytical instruments and chemical methods are available such as titration, Nuclear Magnetic Resonance spectroscopy and gas chromatography to determine the CO_2 absorption in both applications above. Most of these methods are time consuming. A fast, online method to detect CO_2 absorption is important in process monitoring and control. Considering the in-situ performance, Raman spectroscopy can be suggested as a competitive approach for this purpose. It gives chemical information of a sample as a function of Raman wavenumber and scattered light intensity. When converting the information given by a Raman spectroscopy, multivariate calibration is required to transform the spectroscopic measurement into informative output. Raman spectra contain several wavenumbers or group of wavenumbers which are chemically important and needed to be included in the multivariate regression models. However, it is often misleading to use

traditional multilinear methods such as ordinary least square for calibration, when a single wavenumber (X variable) is not sufficient to predict the useful quantity (y variable); when X variables are highly correlated or when there is no adequate information to understand which X variables are correlated to the y variable. In such instances, multivariate analysis gives the advantage of overcoming the collinearity problems while preserving useful information hidden in collinear data. In this study, Principal Component Analysis (PCA) which is a fundamental multivariate analysis tool, has been used as a data compression and exploratory method to investigate the feasibility of Raman spectroscopy as a viable analytical technology to quantify CO_2 absorption by amines and propylene oxide. Eight experimental cases have been used in this analysis. Four of them are related to CO_2 absorption by PO and a mixture of PO and PPC. These experiments were meant to compare CO_2 absorption in the CO_2-PO system with respect to the CO_2-PO-PPC system at some selected process conditions. The other 4 experiments were used to identify CO_2 uptake by four liquid amine solvents. These solvents are currently in research interest to capture CO_2 from flue gas in power plants and industries (Leung *et al*, 2014).

2 Methods

Experimental description of 8 test cases are presented in Table 1. Six organic chemicals were used in the experiments and they are given in Table 2. Case 1-4 were carried out in a closed 2L steam jacketed autoclave reactor equipped with a stirrer while the pressure was increased gradually by adding CO_2 to the reactor. Case

Table 1: Description of test cases.

Case Number	CO_2 loaded solution	Description
1	PO in non-stirred condition	Each case has one sample in a 2L reactor at 60^0C. Tested pressure levels :20, 40 and 60 bar Stirrer speed = 400 rpm
2	PO in stirred condition	
3	PO+PPC in non-stirred condition	
4	PO+ PPC in stirred condition	
5	MEA 37 samples	Each sample in 10 mL glass reactor. Reaction between CO_2 and amine took place at room temperature and pressure
6	3AP 42 samples	
7	3DMA1P 41 samples	
8	MDEA 41 samples	

1 and 2 were PO-CO_2 binary mixtures while Case 3 and 4 were PO-PPC-CO_2 ternary mixtures. A Raman immersion probe was connected through the bottom of the reactor and signals were acquired continuously with time. In case 5-8, CO_2 absorption on liquid amines was observed under equilibrium condition at room temperature and pressure. Raman signals were recorded by immersing the Raman probe into sample reactors after allowing each sample to reach equilibrium.
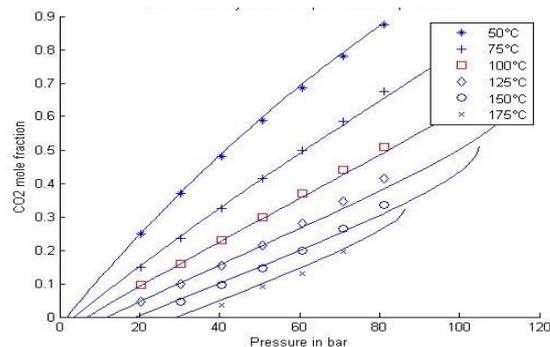

Figure 1. CO_2 mole fraction of PO-CO_2 system at different pressures and temperatures (Peng-Robinson model with binary interaction parameter equal to 0.281).

Table 2. Description of materials.

Name	Abbreviation	Chemical structure	Chemical category
Propylene oxide	PO		epoxide
Polypropylene carbonate	PPC		a copolymer of CO2 and PO
2- Aminoethanol	MEA		Primary amine
3-Amino-1-propanol	3-AP		Primary amine
3-dimethylamino-1-propanol	3DMA1P		Tertiary amine
Methyl diethanolamine	MDEA		tertiary amine

2.1 CO_2 in polymer solutions – from thermodynamic models

In this study, Raman signals (X variables), were calibrated with the absorbed CO_2 content (y variable). Reliable measurement of y variable in Case 1-4 using an analytical method is challenging as CO_2 quickly desorbs if a sample is taken out from the reactor for analysis. Therefore, the CO_2 content data at required pressure and temperature were calculated from the vapour-liquid equilibrium (VLE) data of CO_2-PO system generated using the Peng-Robinson equation of state. The Peng-Robinson model was fitted using experimental data reported in (Chen *et al*, 1994; Shakhova *et al*, 1973). Figure 1 shows predictions of the CO_2 mole fraction in PO-CO_2 system using Peng-Robinson model simulated in Aspen Plus V7.2 software which shows that the absorption of CO_2 at a constant temperature gives a linear behavior with pressure. This linear relationship was taken to model the CO_2 mole fraction at 60^0C at which the experimental cases of 1-4 were carried out.

2.2 CO_2 in amine solutions – from titrations

In experiments from case 5-8, each sample contained 30 % of solvent (solvent weight/total weight of water and solvent) but different amounts of CO_2 added. They were prepared in 10 mL glass reactors and after reaching equilibrium a titration method was carried out to measure its true CO_2 content in units of moles CO_2 per mole solvent.
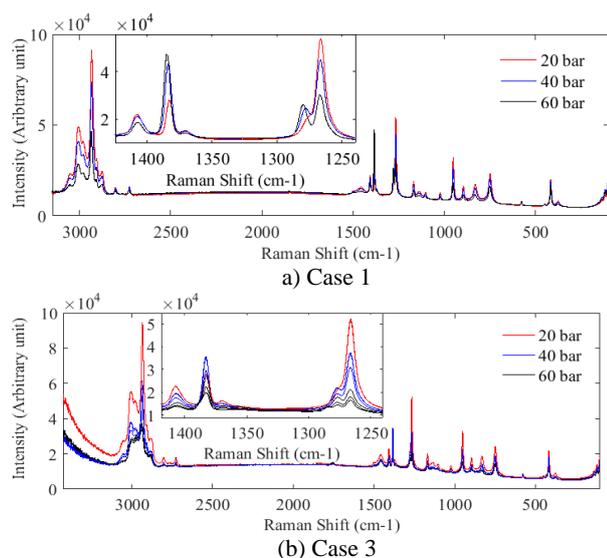
a) Case 1

(b) Case 3

Figure 2. Raman signals of CO_2 loaded polymer samples.

(a) 100-3325 cm^{-1} (b) 950-1550 cm^{-1}

Figure 3. Raw spectra of CO_2 loaded samples (Case 5-8).

Figure 4. Raw spectra of CO2 loaded MEA samples (Case 5).

2.3 Raman Spectroscopy

Raman spectroscopy used in this study was Kaiser RXN2 Analyser of 785 nm laser wavelength, 400 mW laser power and 100-3425 cm^{-1} spectral range. An immersion optic probe which is connected to the RXN2 Analyser via a fibre optic cable, carries the laser light to the sample and in-elastically scattered Raman light is conveyed back to the instrument. The instrument output is a plot of intensity of scattered light versus energy difference (given by wavenumber in cm-1) which is called a Raman spectrum. Peaks and their intensity in a Raman spectrum carry information about the chemicals and their composition respectively.

2.4 Data Processing

For a set of n objects (eg: different samples or signals with time), a Raman spectroscopy measurement generates a data matrix of n x p where p is 3326 Raman wavenumbers. This data matrix contains useful information about the chemical fingerprint of objects as well as noise. They are also called residuals which can be due to the interference of other chemical components, laser input variations or instrument noise. Unless any data conditioning method is applied to remove this unwanted structure from the data matrix, calibration of spectroscopic signals will not be reliable and do not really generate a model which really represent the variation of analyte of interest.

Three data pre-processing techniques were applied for raw Raman data. These were baseline-whittaker filter, standard normal variate (SNV) and mean centering. The baseline-whittaker filter available in PLS toolbox in Matlab is an extended version of (Eilers, 2003) where a weighted least square method is applied to remove background noise and baseline variations. A detailed description of the algorithm can be found in the original work (Eilers, 2003) and (Atzberger et al, 2010). Some
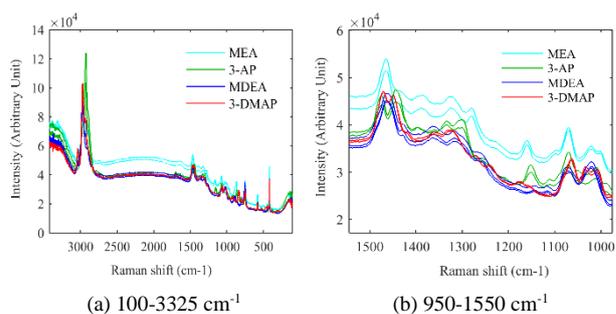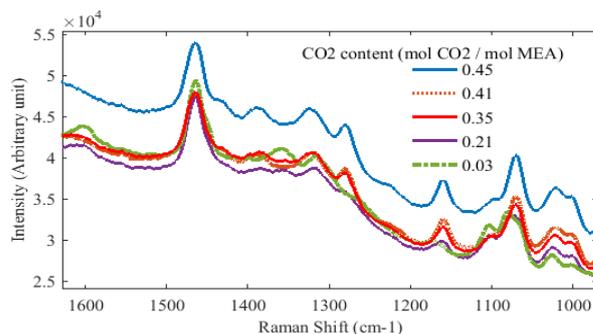
spectra which should be otherwise identical, become different due to baseline and pathlength changes. SNV was applied to remove these scatter effects in the spectra which were specially observed in case 1-4. The algorithm is similar to autoscaling row wise and hence corrects each spectrum individually (Barnes et al, 1989). By mean centering of data, each column in the data matrix is centered by subtracting the mean. It is reported that by mean centering, rank of the model is reduced, data fitting accuracy is increased and offset is removed (Bro et al, 2013).

2.5 Principal Component Analysis (PCA)

Principal component analysis is one of the most important data analysis methods providing a platform for advanced chemometrics methods. As stated in (S. Wold et al, 1987) PCA can have many goals; simplification, data reduction, modelling, outlier detection, variable selection, classification, prediction and unmixing. It can be used to understand general characteristics of data set and guide further investigation through more refined techniques (Wentzell et al, 2012). PCA reduces the dimension of data by calculating principal components (PCs) which reflect the structure of data corresponding to maximum variance. These PCs can then be plotted to visualize the relationship between samples and variables through the use of scores and loading plots. A tutorial review on PCA can be found in (Bro et al, 2014). Decomposing a data matrix X into a structure part which consists of a score matrix (T) and a

loading matrix (P) and noise part or residual matrix (E), is shown in (4) and (5).

$$X = TP^T + E \qquad (4)$$
$$X = t_1 P_1{}^T + t_2 P_2{}^T + t_3 P_3{}^T + \cdots + t_A P_A{}^T + E \quad (5)$$
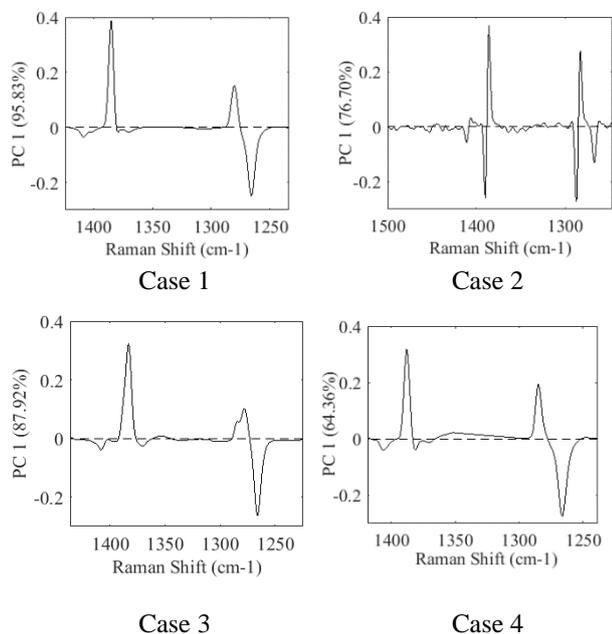

Figure 5. Loading plots of the first principal component for case 1-4 (Region : 1225 – 1450 cm-1).

t_A and P_A are score vector and loading vector for PCA respectively. PC1 is the first principal component which relates to the maximum variance of the data, and PC2 is the second principal component which corresponds to the second largest variance etc. Score values provide information about sample variations while loading value explains the relationship between variables. Residuals provides information as to what spectral variations have not been explained. There are different ways to decompose a matrix to score and loading vectors. NIPALS (*Non-linear Iterative Partial Least Squares*) algorithm (H. Wold, 1966) uses iterative sequence of ordinary least square regression to calculate PCs and was used in this study.

2.6 Important variables related to CO₂ absorption

CO₂ absorbed PO, PPC and amine mixtures exhibit several sharp overlapping peaks in the region of 300 to 1500 cm-1 and 2700 to 3250 cm-1. The focal point in this study is to investigate CO_2 absorption and hence only the peaks related to absorbed CO_2 are considered in the model development. In case 3-4, the monomer PO and the polymer PPC were added into the autoclave reactor and the CO_2 was absorbed into this mixture. Therefore, CO2 bands related to dissolved CO_2 in the PO or PO/PPC mixture were followed in this study.
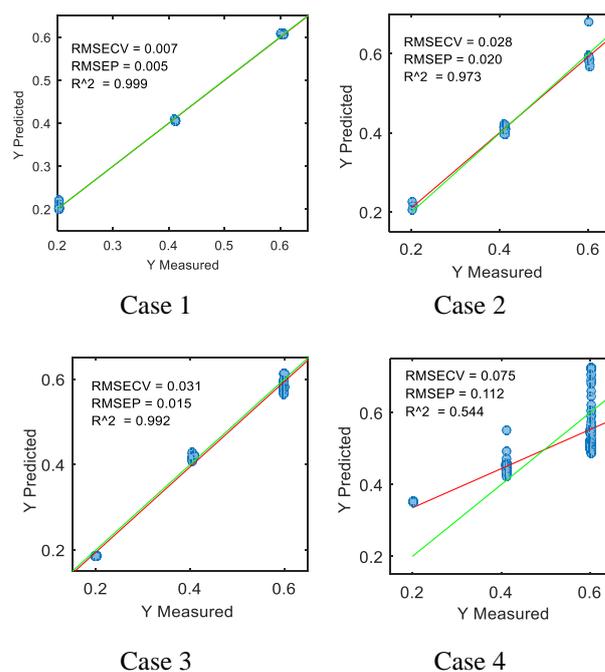

Figure 6. Development of linear regression model using PC1 score values and thermodynamic model data.

(Y measured = CO₂ mole fraction predicted by VLE data; Y predicted = CO₂ mole fraction predicted by PC1 scores; red line= best fitted line based on calibration points; green line=1:1 target line; RMSE (CV/P)= root mean square error of (cross validation/prediction)
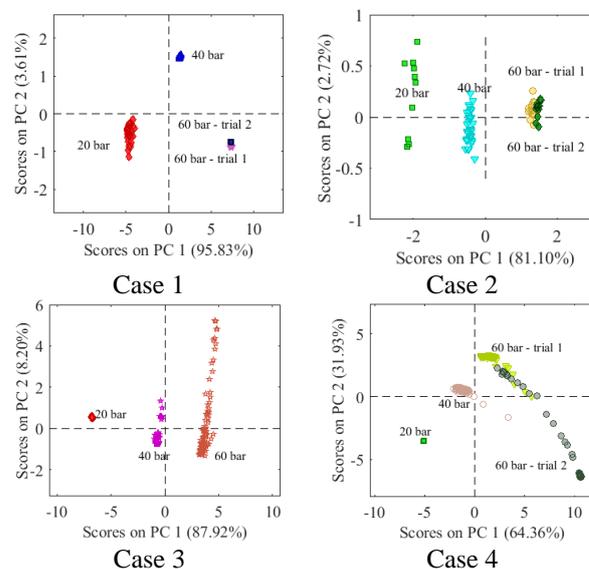

Figure 7. Score plots – PC1 vs PC2 for case 1-4.

Literature reports such Raman wavenumbers of 1264, 1284, 1369, 1387, 1408 cm-1 (Hanf *et al*, 2014). In case 5-8, peaks related to carbonate, bicarbonate, carbamate and dissolved CO_2 fall in the region of 1000-1500 cm-1 ((Vogt *et al*, 2011), (Wong *et al*, 2015)). Therefore, for development of PCA models, the region between 1000-1500 cm-1 and 1225-1450 cm⁻¹ were selected for case 1-4 and case 5-8 respectively.
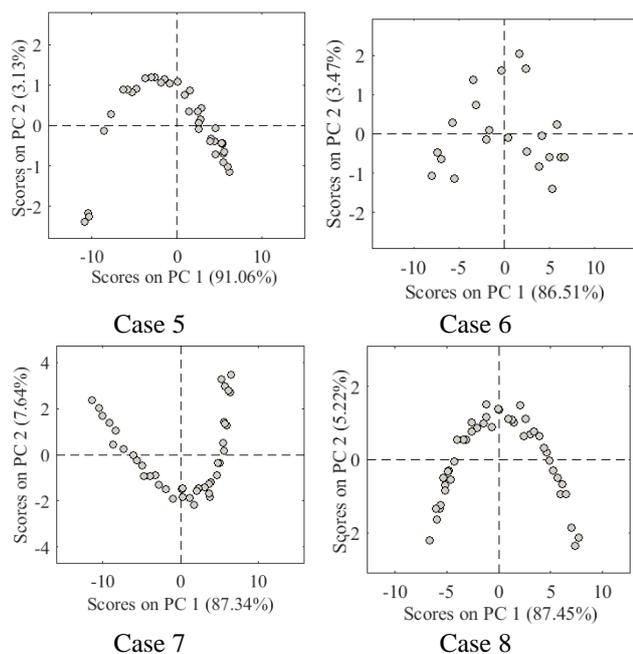
Figure 8. Score plots for case 5-8.

3   Results and Discussion

Figure 2 (a) and (b) show raw spectra for case 1 and 3 respectively highlighting spectral variation with
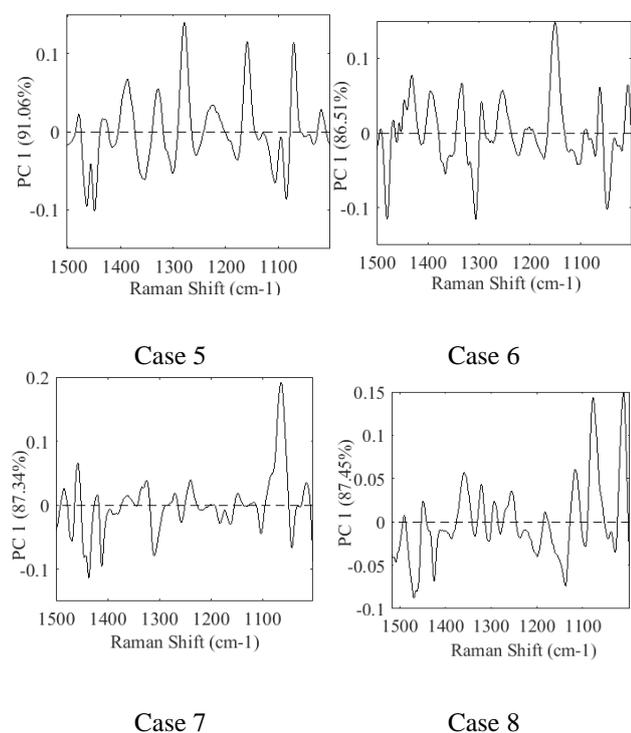

Figure 9. Loading plots of the first principal component for case 5-8.

increasing pressure in the region of 1225 to 1450 cm^{-1}. CO_2 peaks at 1264, 1284, 1369, 1387, 1408 cm-1 can be identified in this figure. A similar spectral behavior was observed for case 2 and 4 in the same region. Figure 3
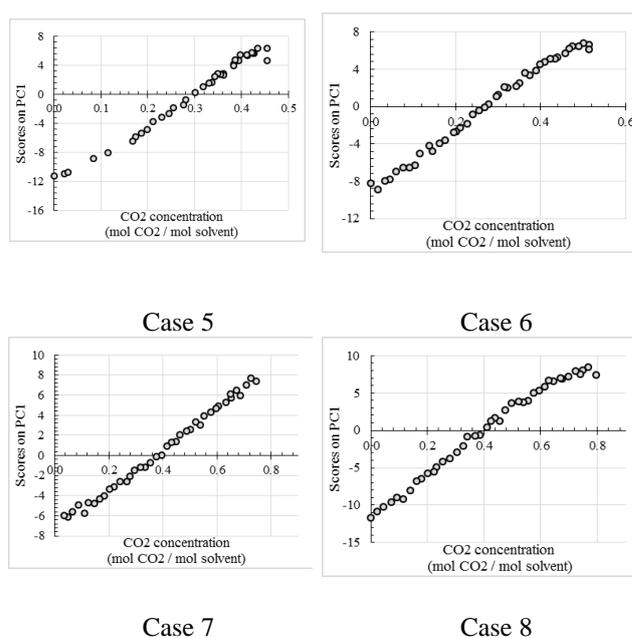

Figure 10. Development of linear regression model using PC1 score values and CO_2 content given by titration data.

gives raw Raman signals observed for CO_2 loaded 4 different amine solvents. Only two spectra from each solvent are shown. Figure 4 shows how the intensity of Raman bands varies with the CO_2 content for MEA samples (Case 5). Both Figure 3and Figure 4 claim that spectral evolution in the region between 1000 to 1500 cm-1, for case 5-8 with respect to case 1-4 is complex due to curved baseline, baseline offsets and overlapping bands. The reason is that the chemical products when CO_2 is reacted with the solvent appear with overlapping peaks in this region. Therefore, when quantifying the total amount of CO_2 absorbed in solvent, all these peaks are needed.

All the Raman signals under each case were first smoothed using baseline-whittaker smoother, then SNV and finally mean centered. PCA was performed for processed data. First principal component was identified as the dimension explaining the largest variance of data in each case. Finally, score values of PC1, were compared with the mole fraction of CO_2 predicted by thermodynamic models for case 1-4 and CO_2 amount determined from titration for case 5-8. Loading plot, score plot and comparison of PC1 score value with CO_2 content under each case were used to explain characteristics in each system.

3.1   Case 1-4

With reference to Figure 5 loading plots of case 1, 3 and 4 almost give similar information about important variables (Raman shifts) while case 2 is different. This is caused by exposing the Raman sensor to both gas and liquid phases as a result of high stirrer speed and

development of vortex in case 2. There is also low viscosity in the medium at low pressures, which creates high turbulence. Score plots of PC1 vs. PC2 as given in Figure 7, show clear distinguish of recorded signals between the three pressure values of 20, 40 and 60 bar. PC2 direction explains only a small variation of data for case 1-3. Experiments for 60 bar, were conducted in replicates and their overlap in score values could be observed in case 1 and 2.

Figure 6 shows how closely PC1 score values are related to VLE data. Plots in this figure were derived by linear regression between PC1 score values as X variables and predicted CO_2 content from VLE data as y variables. From VLE data, CO_2 mole fractions at 20, 40 and 40 bar are 0.202, 0.411 and 0.601 respectively. These values are represented as 'Y measured' in Figure 6. PC1 score values at 3 pressure conditions follow the linear trend given by the mole fraction of CO_2 predicted by thermodynamic models at case 1 and 2. In the presence of PPC (case 3-4), even though pressure and temperature were maintained constant, a significant time was needed to achieve equilibrium condition of CO_2 absorption by the solvent especially at higher pressure region. For example, at 40 bar and 60 bar, PC1 score value of the initial spectra is less than the final recorded spectra at that condition. Therefore, although the reactor is maintained at the required pressure, the score plot gives the hint whether the equilibrium condition has been achieved or not. The significance of the above fact can be clearly understood when examining the score plot for case 4 (Figure 7). In this trial, we see that only 20 bar condition shows a compressed data swarm while at 40 bar, PC1 score values increases with time and this variation is more significant for 60 bar. This is further assured by Figure 6 (case 4) where the thermodynamic model satisfies the trend of final recorded data for 60 bar condition, but highly deviate from the initial recorded data at this condition. PC1 score values positively correlate with the amount of absorbed CO_2 by PO-CO_2 and PO-PPC-CO_2 systems.

3.2 Case 5-8

Absorption of CO_2 by amines (case 5-8), features several important variables in the region 1000-1500 cm-1 as given by loading plots in Figure 9 and this is the result of several parallel equilibrium reactions happening in each system. Each sample carries different information which mean different amount of CO_2 absorption and hence the concentration of chemical species produced during these reactions are different. Therefore, a data spread in score plot of PC1 vs PC2 can be observed in the score plots in each case as presented in Figure 8. However, similar to polymer-CO2 system, PC1 explains the largest variation of data structure and therefore PC1 score values were compared with total CO_2 absorbed by the system. Results are shown in

Figure 10. With the increasing amount of CO_2, there is a gradual increase of PC1 score value highlighting that PC1 score value is an indication of the level of CO_2 absorbed by the sample.

4 Conclusions

Monitoring CO_2 in liquid phase of PO-CO_2 system or PO-PPC-CO_2 system by analytical techniques is challenging as the CO_2 quickly desorbs if the pressure is lowered in sample taking. Therefore, online analysis such as spectroscopy is more favorable For CO_2-amine systems, an in situ characterization of CO_2 absorption gives credits to process monitoring and control ability. Based on this study, combination of Raman spectroscopy and PCA claims that PC1 score value explains variation of data structure corresponded to absorbed CO_2 amount. PCA plots give an indication of CO_2 composition, process dynamics and equilibrium conditions in these two chemical systems and hence can be used as an efficient tool to analyse collinear process data. Further investigation of PCA model development under different process parameters is recommended to validate the findings from this feasibility study. Experiments to develop advanced chemometrics tools such as partial least square regression can now be recommended for both polymer-CO_2 system and amine-CO_2 system.

References

C. Atzberger and P. H. C. Eilers. Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements *International Journal of Remote Sensing,* 32(13): 3689–3709, 2010. doi:10.1080/01431161003762405

R. J. Barnes, M. S. Dhanoa, and S. J. Lister. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy,* 43(5): 772-777, 1989. doi:10.1366/0003702894202201

R. Bro and A. K. Smilde. Centering and scaling in component analysis. *Journal of Chemometrics,* 17(1): 16–33, 2013. doi:10.1002/cem.773

R. Bro and A. K. Smilde. Principal component analysis. *Anal. Methods,* 6: 2812-2831, 2014. doi:10.1039/C3AY41907J

L. B. Chen and X. G. Fang. Phase equilibria and CO2 distribution in the CO2-epoxide-toluene systems. *Natural Gas Chemical Industry,* 3: 255-266, 1994.

P. H. C. Eilers. A Perfect smoother. *Analytical Chemistry,* 75(14): 3631-3636, 2003. doi:10.1021/ac034173t

S. Hanf, R. Keiner, D. Yan, J. Popp, and T. Frosch. Fiber-enhanced Raman multigas spectroscopy: A versatile tool for environmental gas sensing and breath analysis. *Analytical Chemistry,* 86(11): 5278-5285, 2014. doi:10.1021/ac404162w

X. Jiang, F. Gou, and H. Jing. Alternating copolymerization of CO2 and propylene oxide catalyzed by C2v-porphyrin cobalt: Selectivity control and a kinetic study. *Journal of Catalysis,* 313: 159-167, 2014. doi:10.1016/j.jcat.2014.03.008

D. Y. C. Leung, G. Caramanna, and M. M. Maroto-Valer. An overview of current status of carbon dioxide capture

and storage technologies. *Renewable and Sustainable Energy Reviews,* 39: 426-443, 2014. doi:10.1016/j.rser.2014.07.093

N. McCann, D. Phan, X. Wang, W. Conway, R. Burns, M. Attalla, G. Puxty, and M. Maeder. Kinetics and mechanism of carbamate formation from CO2(aq), carbonate species, and monoethanolamine in aqueous solution. *The Journal of Physical Chemistry A,* 113(17): 5022–5029, 2009. doi:10.1021/jp810564z

S. F. Shakhova, O. L. Rutenberg, and M. N. Targanskaya. Liquid-gas equilibrium in the epoxypropane–carbon dioxide system. *Russ. J. Phys. Chem,* 47(6), 1973.

M. Vogt, C. Pasel, and D. Bathen. Characterisation of CO2 absorption in various solvents for PCC applications by Raman spectroscopy. *Energy Procedia,* 4: 1520-1525, 2011. doi:10.1016/j.egypro.2011.02.020

P. D. Wentzell and S. Hou. Exploratory data analysis with noisy measurements. *Journal of Chemometrics,* 26(6): 264–281, 2012. doi:10.1002/cem.2428

H. Wold. Estimation of principal components and related models by iterative least squares. *Journal of Multivariate Analysis* 1: 391-420, 1966.

S. Wold, K. H. Esbensen, and P. Geladi. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems,* 2(1): 37-52, 1987. doi:10.1016/0169-7439(87)80084-9

M. K. Wong, M. A. Bustam, and A. M. Shariff. Chemical speciation of CO2 absorption in aqueous monoethanolamine investigated by in situ Raman spectroscopy. *International Journal of Greenhouse Gas Control,* 39: 139–147, 2015. doi:10.1016/j.ijggc.2015.05.016