Principal Component Analysis Applied to CO₂ Absorption by Propylene Oxide and Amines

M. H. Wathsala N. Jinadasa¹, Klaus J-Jens¹, Carlos F. Pfeiffer¹, Sara Ronasi², Carlos Barreto Soler², Maths Halstensen¹

¹Faculty 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
²Norner 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°C 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 CO₂ 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 CO₂ absorption. A similar analysis was carried out to monitor CO₂ absorption by four different amines at room temperature and pressure in a batch reactors. The CO₂ content was determined from titration and was used to model the spectroscopic data. Principal Component Analysis proved to be able to identify CO₂ 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 CO₂ absorption systems and hence can be used for developing regression models for online monitoring and control.

Keywords: principal component analysis, CO₂ absorption, propylene oxide, amines

1 Introduction
Carbon dioxide (CO₂) is known to be the primary greenhouse gas contributing more than 60% of global warming. Capturing CO₂ 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 CO₂ by amines is one of the most popular technologies for CO₂ capture. Amines are categorized as primary, secondary or tertiary amine based on their chemical structure. The reaction between amines and CO₂ is complex (McCann et al., 2009). However, when considering the CO₂ mass balance, it can be seen that once absorbed by a primary amine, CO₂ will remain in the form of carbonate, bicarbonate, carbamate or molecular CO₂ as given in (1). When it is a tertiary amine, there is no carbamate formation (2).

\[
\text{PPC} + \text{CO}_2 \xrightarrow{\text{catalyst}} \text{cyclic carbonate}
\]

Synthesis of polypropylene carbonate (PPC) by reaction of CO₂ and propylene oxide (PO) in the presence of a catalyst has become a fascinating research area as a CO₂ 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).

\[
\text{PPC} + \text{CO}_2 \xrightarrow{\text{catalyst}} \text{cyclic carbonate}
\]

CO₂ absorption capacity by an amine or by in the liquid phase PO is a key performance criteria in industrial scale CO₂ capture and polymerization processes. However, the measurement of CO₂ 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 CO₂ absorption in both applications above. Most of these methods are time consuming. A fast, online method to detect CO₂ 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
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 CO₂ to the reactor. Case 1 and 2 were PO-CO₂ binary mixtures while Case 3 and 4 were PO-PPC-CO₂ 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, CO₂ 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.

Table 1: Description of test cases.

<table>
<thead>
<tr>
<th>Case Number</th>
<th>CO₂ loaded solution</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PO in non-stirred condition</td>
<td>Each case has one sample in a 2L reactor at 60°C. Tested pressure levels: 20, 40 and 60 bar</td>
</tr>
<tr>
<td>2</td>
<td>PO in stirred condition</td>
<td>Stirrer speed = 400 rpm</td>
</tr>
<tr>
<td>3</td>
<td>PO+PPC in non-stirred condition</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>PO+PPC in stirred condition</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>MEA 37 samples</td>
<td>Each sample in 10 mL glass reactor. Reaction between CO₂ and amine took place at room temperature and pressure</td>
</tr>
<tr>
<td>6</td>
<td>3AP 42 samples</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3DMA1P 41 samples</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>MDEA-41 samples</td>
<td></td>
</tr>
</tbody>
</table>

1 and 2 were PO-CO₂ binary mixtures while Case 3 and 4 were PO-PPC-CO₂ 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, CO₂ 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.

2.1 CO₂ in polymer solutions – from thermodynamic models

In this study, Raman signals (X variables), were calibrated with the absorbed CO₂ content (y variable). Reliable measurement of y variable in Case 1-4 using an analytical method is challenging as CO₂ quickly desorbs if a sample is taken out from the reactor for analysis. Therefore, the CO₂ content data at required pressure and temperature were calculated from the vapour-liquid equilibrium (VLE) data of CO₂-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 CO₂ mole fraction in PO-CO₂ system using Peng-Robinson model simulated in Aspen Plus V7.2 software which shows that the absorption of CO₂ at a constant temperature gives a linear behavior with pressure. This linear relationship was taken to model the CO₂ mole fraction at 60°C at which the experimental cases of 1-4 were carried out.

2.2 CO₂ 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 CO₂. They were PO-CO₂ binary mixtures while Case 3 and 4 were PO-PPC-CO₂ 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, CO₂ 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.

Table 2: Description of materials.

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
<th>Chemical structure</th>
<th>Chemical category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propylene oxide</td>
<td>PO</td>
<td>HO–(CH₂)₂–OH</td>
<td>epoxide</td>
</tr>
<tr>
<td>Polypropylene carbonate</td>
<td>PPC</td>
<td>HO–(CH₂)₃–CH₃</td>
<td>a copolymer of CO₂ and PO</td>
</tr>
<tr>
<td>2- Aminoethanol</td>
<td>MEA</td>
<td>HO–NH₂</td>
<td>Primary amine</td>
</tr>
<tr>
<td>3-Amino-1-propanol</td>
<td>3-AP</td>
<td>H₃N—CH(CH₃)₂</td>
<td>Primary amine</td>
</tr>
<tr>
<td>3-dimethylamino-1-propanol</td>
<td>3DMA1P</td>
<td>H₃N—CH₂—CH₂—NH₂</td>
<td>Tertiary amine</td>
</tr>
<tr>
<td>Methyl diethanolamine</td>
<td>MDEA</td>
<td>HO–(CH₂)₂–NH–CH₃</td>
<td>Tertiary amine</td>
</tr>
</tbody>
</table>

Figure 1. CO₂ mole fraction of PO-CO₂ system at different pressures and temperatures (Peng-Robinson model with binary interaction parameter equal to 0.281).
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).

\[ X = TP^T + E \]  \hspace{1cm} (4)

\[ X = t_1P_1^T + t_2P_2^T + t_3P_3^T + \cdots + t_A P_A^T + E \]  \hspace{1cm} (5)

**Case 1**

**Case 2**

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

**Case 3**

**Case 4**

\( t_A \) 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 provide 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 CO\(_2\) absorption**

CO\(_2\) 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, CO\(_2\) 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 = CO\(_2\) mole fraction predicted by VLE data; Y predicted = CO\(_2\) 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.
3 Results and Discussion

Figure 2 (a) and (b) show raw spectra for case 1 and 3 respectively highlighting spectral variation with 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 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 3 and 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-CO$_2$ 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


