

Application of multivariate data analysis of Raman Spectroscopy spectra of 2-oxazolidinone

Federico Mereu^{1,2} Jayangi D. Wagaarachchige³ Klaus-J. Jens² Zulkifli Idris^{2,*}

¹Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

²Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Kjølnes Ring 56, 3918 Porsgrunn, Norway

³Department of Electrical, IT and Cybernetics, University of South-Eastern Norway, Kjølnes Ring 56, 3918 Porsgrunn, Norway

*corresponding author: zulkifli.idris@usn.no

Abstract

Chemical absorption of carbon dioxide (CO₂) using amine solution is considered as the readiest technology available for capturing CO₂ gas from industrial processes. The well-known amine for this process is 2-aminoethanol (MEA) which is normally mixed with water to a typical concentration of 30 wt%. MEA degrades over time producing non-reactive chemicals such as 2-oxazolidinone (OZD) due to exposure to impurities and high process temperature. It is thus important to find a suitable method for OZD qualification and quantification. In this work, we approach this challenge by means of Raman spectroscopy and multivariate data analysis. We started by collecting Raman spectra of 40 OZD samples and applying Principal Component Analysis to study these samples.

Keywords: multivariate data analysis, MEA, Raman spectroscopy, CO₂ capture, degradation

1 Introduction

Due to the economic development and the subsequent increase in world population, the global demand for energy will continue to rise in the following decades. The dependence on fossil fuels, the primary source of energy, emitting copious amount of CO₂, is the main cause of global warming. Even if large investments are underway to decarbonise the world energy production, renewable electricity may not be suitable for certain applications, such as the cement, iron and steel, and chemical sectors.

Carbon capture and storage (CCS) and its ability to avoid CO₂ emissions at their source, represents a solution in the fight against climate change. Among all the different alternatives, post-combustion capture by using amine-based solvent is considered to be the most advanced technology (Sexton and Rochelle, 2011). This process relies on the ability of the amine solution to chemically react with CO₂ in the flue gas. The best absorbents are the ones with high net cyclic capacity, fast reaction with CO₂, low heat of reaction, high chemical stability, low vapor pressure and minimally corrosive (Hartono et al., 2017). Of

the many solvents tested, 2-aminoethanol (MEA) is the most used due to its good operational properties and relatively low price. The solvent used in operating plants simply consists of water and amines, whose concentration is usually made based on operating experience (typical concentration range values goes from 12% wt to a maximum of 32% wt (Kohl and Nielsen, 1997)).

A typical chemical absorption process for CO₂ capture plant is shown in Figure 1.

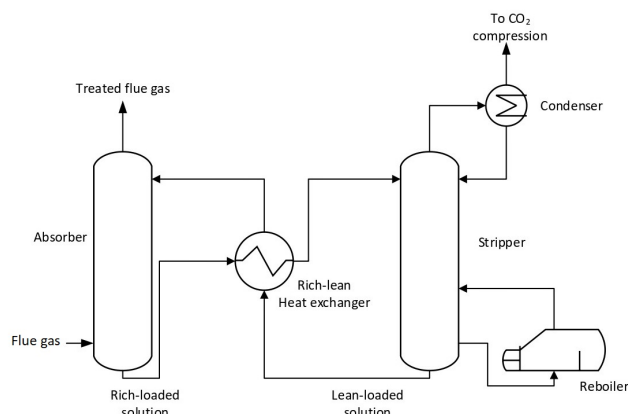
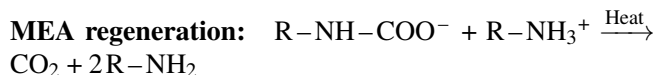
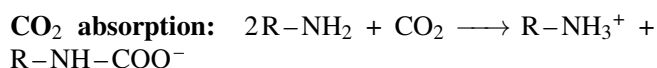


Figure 1. Schematic of a chemical absorption process for CO₂ capture.

After a preliminary purification from NO_x, SO_x, and particulate matter, the flue gas enters the absorber. Through contact with MEA solvent, part of CO₂ in the flue gas is absorbed into amine solution, forming a weakly bonded and quite stable compound, carbamate. The scrubbed gas is then washed with water to remove the solvent and discharged into the atmosphere. Then, the rich-loading solvent (with absorbed CO₂) passed through a cross-heat exchanger and pumped up to the head of the stripper. In the stripper, the high temperature and pressure generated by a reboiler cause the carbamate to dissociate back to MEA and CO₂. The obtained product stream with high CO₂ purity is conveyed to compression for trans-

portation to storage sites. At the bottom of the stripper, the high temperature lean-loading is conveyed to a heat exchanger to decrease the temperature of the lean-loading solution before entering the absorber again.

The entire process chemistry is complex, and the two main reactions taking place in the absorber and the stripper are:



For simplicity, MEA is expressed by R-NH₂, where R stands for OH-CH₂-CH₂. The first reaction shows that only half a mole of CO₂ is absorbed per mole of MEA, leading to the formation of carbamate. In the second equation, under the application of heat, the carbamate dissociates to give back CO₂ and amine sorbent.

However, there is a main problem associated with this process, which is degradation of the solvent caused by heat exposure and impurities in the exhaust gas. This leads to foaming, fouling, increased viscosity, corrosion and formation of different degradation compounds that are unreactive towards CO₂. In the case of MEA, one of the main degradation products is 2-oxazolidinone (OZD), a heterocyclic five-membered ring organic compound, which formation pathway is shown in Figure 2.

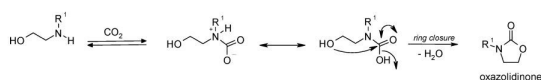


Figure 2. Oxazolidinones formation (R¹, R²: H, alkyle) (Lepaumier et al., 2009).

The formation of OZD starts with a reaction between MEA and CO₂, which leads to the formation of carbamate complex, as shown in the first equation above. Elimination of a water molecule from the carbamate complex during a ring closure reaction yields an OZD molecule. The formation of OZD is a problem because it is unstable and will react giving other degradation products (namely HEEDA, HEIA, AEHEIA, BHEI (Gouedard et al., 2012)) that must be purged from the system to prevent their build-up.

For this purpose, it is essential to find a procedure for the conversion of the molecule to its precursor amine. This requires a preliminary identification and quantification step.

Raman spectroscopy is a valuable technique for qualitative and quantitative analyses, since there is a relationship between intensity of the Raman band, chemical information and the concentration of a sample being analyzed (Larkin, 2011). Raman spectrums are generally plotted as intensity against Raman Shift (or wavenumber). Vibrations of functional groups of a molecule appear in a Raman spectrum at characteristic Raman shift, which is similar for all molecules containing the same functional group.

Chemometric multivariate analysis is an advanced statistical method that can be used to extract this huge information by building specific model for specific chemical species.

The approach in this paper started with the analysis of OZD samples at different concentrations using Raman spectrometer. Principal component analysis (PCA) was then performed on these samples to check for any outliers, relevant peaks for OZD, and monitor changes in the OZD at different concentrations.

2 Materials & Methods

2.1 Sample preparation and Raman analysis

The first big part of this work consisted of sample preparation. Stock solution of OZD was prepared dissolving 2-Oxazolidinone (Sigma-Aldrich, purity 98%) in Milli-Q® water (18.2 MΩ ·cm at 25°C). Samples of increasing concentration from 5 to 815 mM were then prepared by diluting the stock solution in distilled water.

The amount of OZD and water needed were weighted using a Mettler-Toledo MS 105 balance.

The Raman scans were taken using a Kaiser Raman Rxn2 analyzer of 785 nm laser wavelength, 400 mW laser power and 150-3425 cm⁻¹ spectral range. In a typical experiment, a vial containing OZD solution was placed inside a black sample holder to avoid light disturbance and the top part of the sample holder was also covered with aluminum foil to further reduce any possible disturbance from fluorescence of external light sources. A fiber-optic immersion probe (optic of ¼ inch) from Kaiser Optical Systems Inc. was used for the measurement. To avoid contamination, the probe was first washed with deionized water followed by acetone before each measurement to remove any possible impurities/leftovers on the probe tip. The Raman probe was kept at the same depth and same temperature (20 °C) for all the measurements to ensure consistency and to avoid changes of acquisition background. In order to improve sample sensitivity for off-line analysis of each measurement, maximum laser power (400 W) was used with exposure time of 30 seconds and an average of six scans. iC Raman software from Kaiser Optical Systems Inc. was used for the acquisition of the spectra.

2.2 Principal Component Analysis (PCA)

PCA is a data simplification technique used in multivariate statistics. The aim of the technique is to reduce the high number of variables describing a set of data to a smaller number of compressed variables, called Principal Components, PCs, which describe the variation and structure of the data. The PCs can then be plotted to visualize the relationship between samples and variables through the use of scores (which describe the relationship between observations) and loading plots (which show the relationship of the variables) (Wold et al., 1987).

The data is seen as a matrix, called data matrix or X

matrix, composed by n objects (samples) and p variables (the measurement for each object) (Esbensen, 2012). This data matrix can be represented in a Cartesian co-ordinate system of dimension p . Considering the first variable, X_1 , its entries can be plotted along a 1-dimensional axis. This approach can be extended considering the next variable, X_2 , resulting in a 2-dimensional plot and so on, until all p variables are covered. This p -dimensional co-ordinate system is the variable space.

To better understand, it is assumed an X matrix with n objects and 3 variables. Its variable space will be composed by 3 axes: one for each variable. And for each object in the variable space, its x -value will be plotted, meaning that all the objects can be as a point in the variable space. When all the points are plotted, the result is a swarm of points. It is then possible to recognize a linear behaviour, which can be described by a line that lies along the direction of maximum variance in the data set, called the first Principal Component, PC1. Further PCs can be plotted; the second principal component will lie along the direction of the second largest variance, and it will be orthogonal to the first PC. The third PC will be orthogonal to both PC1 and PC2, lying along the direction of the third largest variance and so on for the subsequent PCs. This PCs system will constitute a new coordinate system, where each PC will represent successively smaller and smaller variances. The PCs are uncorrelated with each other since they are mutually orthogonal.

There are two main parameters used in PCA: loadings and scores. The loadings are coefficient of linear combination for each PC, namely p_{ka} , where k is the index for p variables and a is the index for principal component direction coefficients. All the loadings constitute a matrix, P , which expresses the transformation between the initial variable space and the new space formed by the PCs. These loading vectors, namely the columns in P , are orthogonal. In synthesis, loadings describe the relationship between the initial p variables and the PCs.

The score is the distance between the object and its projection into the PC, and it is called score for object i , t_{i1} , if it refers to PC1. The projection of object i onto PC2 will give the score t_{i2} , and so on. The projected object i will correspond to a point in the new co-ordinate system, an A -dimensional surface. Each object will thus have its own set of scores in this dimensionality-reduced subspace. The NIPALS (Nonlinear Iterative Partial Least Squares) algorithm (Wold, 1966) is one of the several methods used to find the score and loading vectors. In this study, NIPALS algorithm was applied when using PLS toolbox with MATLAB® software.

3 Results & Discussion

3.1 Pre-processing of raw spectra

Raw spectra from 40 different OZD samples in water at different concentration are shown in Figure 3.

The raw spectra contain important information on

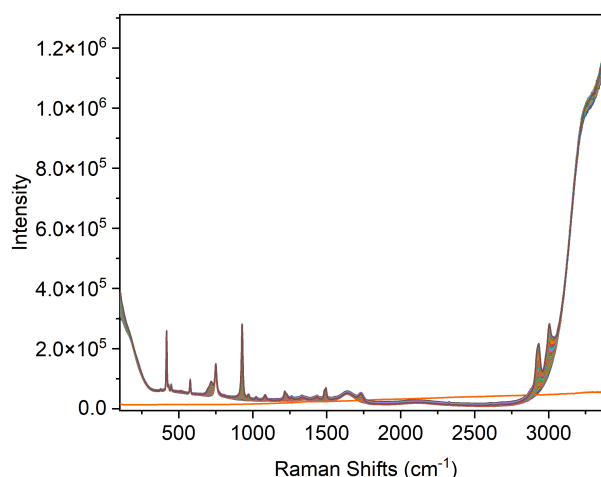


Figure 3. Raman raw data of the 40 analysed samples.

chemical fingerprints of the samples but also noise from background and instrument. Pre-processing of the raw spectra can be applied to extract useful information and to remove offset and irrelevant signals.

The raw spectra were subjected to a baseline correction technique by applying Automatic Whittaker filter with lambda equals to 100 and P equals to 0.001. The Whittaker filter used is an extended version of Eilers, 2003, available in the PLS toolbox in MATLAB, where a weighted least square method was applied to remove baseline variations and background noise. The factor lambda controls the curvature allowed for the baseline, while the P factor governs the extent of asymmetry required of the fit (Eilers, 2003).

Baseline corrected spectra of OZD samples are shown in Figure 4. As can be seen, as the concentration of OZD increases, the intensity values of some peaks also increase, suggesting that OZD concentration is proportional to the peak intensity, according to the Beer-Lambert law.

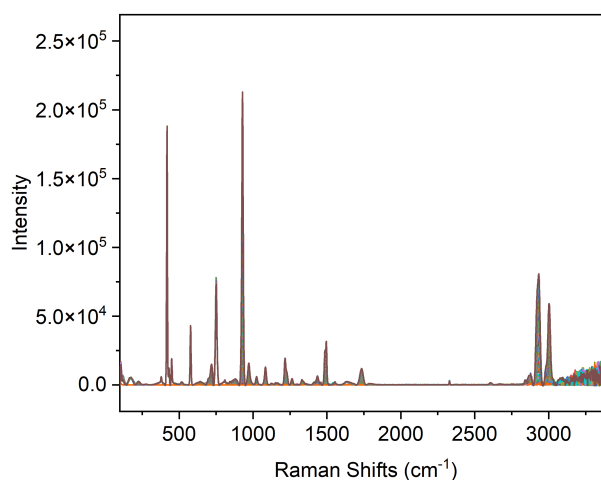


Figure 4. Baseline corrected Raman spectra of the 40 analysed samples.

The peaks that change according to changes in the OZD

concentration are now easily identified and their signals and band assignments are listed in Table 1.

Observed frequency [cm^{-1}]	Vibrational mode
719	C-C stretch
928	C-C stretch
1083	C-H ₂ rock
1216	C-H ₂ twist
1333, 1261	C-H ₂ wag
1436	C-H ₂ scissor
1495	C'-N stretch; C=O stretch
1733	C=O stretch
2932	C-H ₂ symmetric stretch
3003	C-H ₂ asymmetric stretch

Table 1. Vibrational assignments of OZD (McDermott, 1986) (C' = carbonyl carbon).

All the band assignments were referenced to earlier work from McDermott (1986) from the spectra of γ -butyrolactone and 2-pyrrolidinones, which are cyclic esters, like 2-oxazolidinone.

There are also strong peaks at wavelengths 418, 577, and 750 cm^{-1} that do not change according to the changes in OZD concentration and these peaks can be assigned to the noise from the Raman instrument. These peaks were also seen previously in earlier publication from Jinadasa (2019).

Concerning water, its characteristic peaks are cut off from the range of interest, since it usually shows bands below 300 cm^{-1} corresponding to the hydrogen bond bending and stretching motions and strong bands above 3000 cm^{-1} typical of the O-H stretching region; the low intensity peak at 1650 cm^{-1} arises from the intramolecular bending motion (Franks, 1972).

3.2 Initial PCA Analysis

Using the whole spectra as a starting point, the pre-processed OZD spectra were then subjected to initial PCA analysis. Figure 5 illustrates the cumulative variance of the PCA model. PC1 is defined as the first principal component which relates to the maximum variance of the data, while PC2 is the second principal component which corresponds to the second largest variance. The number of PCs corresponds to the number of orthogonal variables in the spectral data set. As can be seen, PC1 explains 92.58% of the total variance, while PC2 describes an additional 6.88%. These two PCs make up 99.46% of the variation in the model, suggesting that they are probably sufficient to determine the most important variables for the description of OZD samples.

In Figure 6, a score plot of PC1 versus PC2 for the pre-processed OZD spectra is shown. The dotted circle represents a 95% confidence level. As can be seen, one of the samples is outside the area of interest meaning that this sample is most likely an outlier. By checking the raw spectra of OZD samples, this sample is confirmed to be an

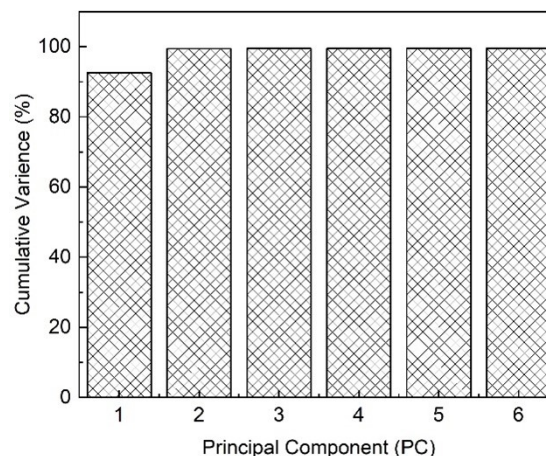


Figure 5. Cumulated percentage variation explained.

outlier and it is probably coming from an error when using the Raman instrument. The outlier was thus removed.

The pre-processed OZD spectra as shown in Figure 4 also show some noise in the region of $>3000 \text{ cm}^{-1}$ Raman shift and this region was also removed in the next PCA analysis.

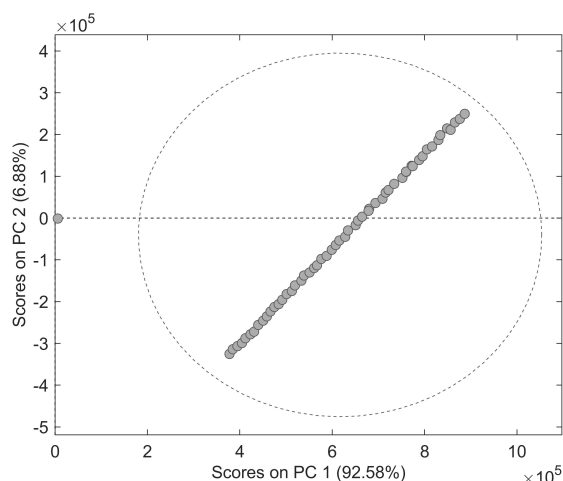


Figure 6. PCA analysis for preprocessed Raman data, first score plot of PC1 vs PC2.

The loading plot for PC 1 for the PCA model is shown in Figure 7. As mentioned by Wold et al. (1987), loading plots define what a principal component represents. The higher the loading value, the higher the contribution of the variable to the PC. In the case of this work, these plots will represent OZD concentrations in the samples. Figure 7 indicates that significant contribution comes from peaks at 418, 577, and 750 cm^{-1} . These peaks however do not correspond to OZD or water, and thus most likely coming from the instrument. The fact that these peaks have higher loading values even though they do not really represent the actual components in the samples necessitates further correction to the PCA model. These peaks were therefore excluded from the model.

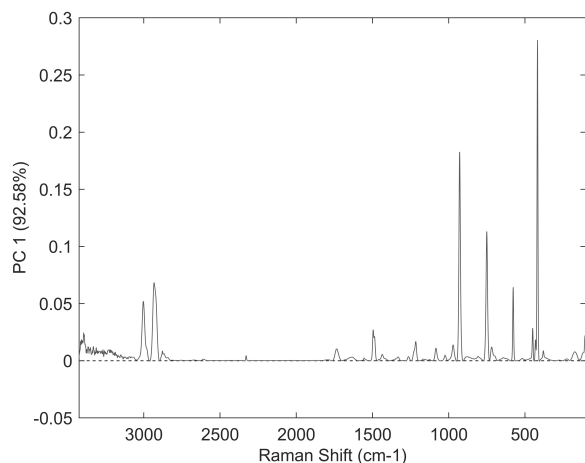


Figure 7. Loading plot for PC1.

3.3 Optimized PCA with Variable Selection

Based on earlier considerations, the PCA model was recalibrated, and Figure 8 displays the new cumulative variance of the model. PCA model was recalibrated by selecting the variable range of OZD to optimize the PC1, which mainly describes the OZD concentration variation.

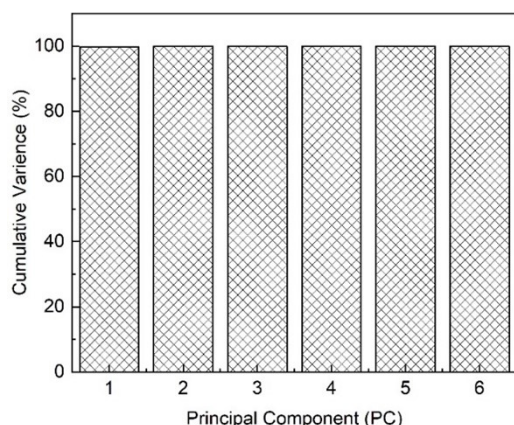


Figure 8. Cumulated percentage variation for different principal components.

The new score plot is shown in Figure 9. Based on the figure, PC1 and PC2 account for 99.78 and 0.17% of the model variance, respectively. These two principal components already make up 99.95% of the cumulative variance for the model suggesting that it is very likely that the OZD changes are sufficiently described by PC1.

With the elimination of outliers, all samples are now at 95% confidence level. Values of PC1 are always positive, whilst values of PC2 change from positive to negative for all the samples. The samples also show a linear trend suggesting that there is a linear trend between Raman intensity and OZD concentration and that the PCA model can be used to classify OZD.

The loading plot for PC1 is illustrated in Figure 10. According to the plot, the sharp OZD peak at 928 cm^{-1}

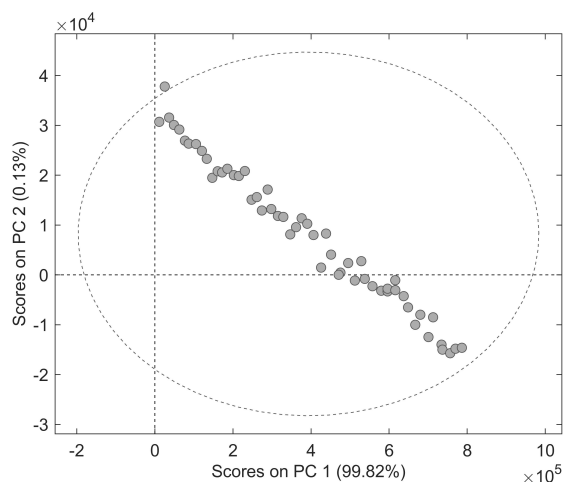


Figure 9. PCA analysis for preprocessed Raman data, score plot of PC1 vs PC2 after removal of variables below 650 cm^{-1} .

gives the highest contribution to PC1. This indicates that this peak can be used as an indicator for the presence or changes in OZD concentration in a sample. Other peaks that positively contribute to PC1 loading plot include 3003, 2933, 1736, 1496, and 720 cm^{-1} and these peaks are observed as relevant functional group peaks for OZD, as shown in Table 1.

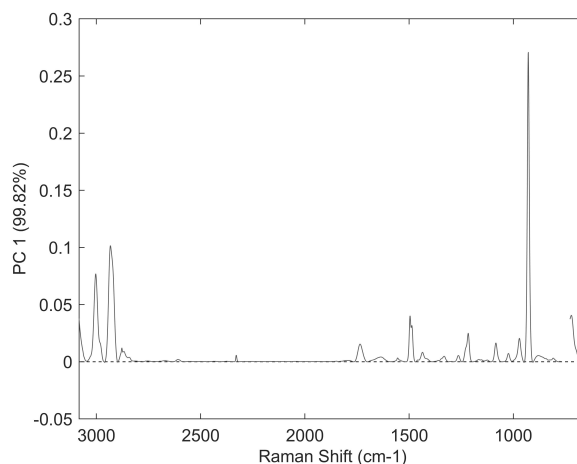


Figure 10. Loading plot for PC1 with variable selection.

4 Conclusion

This paper aims to analyze Raman spectra of 2-oxazolidinone samples by using Principal Component Analysis to detect relevant peaks, monitor changes in the samples at different concentrations and remove outliers.

After spectra acquisition and a preliminary baseline correction, the data were subjected to PCA analysis. The first two PCs, which made up 99.46% of the variation in the model, were considered for the analysis. After that, outlier removal was performed and the PCA model was recalibrated by selecting relevant variable range of OZD

to optimize PC1, which describes the OZD concentration variation. With these considerations, the two PCs made up 99.95% of the cumulative variance, an increase of 0.49 percentage point.

Finally, according to the loading plot for PC1, it was found out that the sharp OZD peak at 928 cm^{-1} gave the highest contribution to PC1, indicating that this peak can be used as an indicator for the presence or changes in OZD concentration in a sample.

By using PCA, we have shown in this work that we can systematically identify with precision any outliers, relevant peaks for OZD, and monitor changes in the OZD at different concentrations.

References

- Paul H. C. Eilers. A perfect smoother. *Analytical Chemistry*, 75(14):3631–3636, 2003. doi:10.1021/ac034173t. PMID: 14570219.
- K. H. Esbensen. *Multivariate data analysis in practice: an introduction to multivariate data analysis and experimental design*. 5th ed. edition, 2012. ISBN 978-8299333030.
- Felix Franks. *The Physics and Physical Chemistry of Water*. Springer New York, 1972. ISBN 978-1-4684-8336-9 978-1-4684-8334-5. doi:10.1007/978-1-4684-8334-5.
- C. Gouedard, D. Picq, F. Launay, and P.L. Carrette. Amine degradation in CO₂ capture. i. a review. *International Journal of Greenhouse Gas Control*, 10:244–270, 2012. doi:https://doi.org/10.1016/j.ijggc.2012.06.015.
- Ardi Hartono, Solrun Vevelstad, Arlinda Ciftja, and Hanna Knuutila. Screening of strong bicarbonate forming solvents for CO₂ capture. *International Journal of Greenhouse Gas Control*, 58:201–211, 2017. doi:https://doi.org/10.1016/j.ijggc.2016.12.018.
- M.H. Wathsala N. Jinadasa. *Process analytical technology for real-time quantitative speciation of aqueous phase CO₂ capture solvents*. PhD thesis, University of South-Eastern Norway, 2019.
- Arthur L. Kohl and Richard B. Nielsen. *Gas Purification*. Gulf Professional Publishing, fifth edition, 1997. ISBN 978-0-88415-220-0. doi:https://doi.org/10.1016/B978-088415220-0/50002-1.
- Peter Larkin. *Infrared and Raman Spectroscopy*. Elsevier, 2011. ISBN 978-0-12-386984-5. doi:10.1016/B978-0-12-386984-5.10001-1.
- Helene Lepaumier, Dominique Picq, and Pierre-Louis Carrette. New amines for CO₂ capture. i. mechanisms of amine degradation in the presence of CO₂. *Industrial & Engineering Chemistry Research*, 48(20):9061–9067, 2009. doi:10.1021/ie900472x.
- Dana P. McDermott. Vibrational assignments and normal-coordinate analyses of γ -butyrolactone and 2-pyrrolidinones. *The Journal of Physical Chemistry*, 90(12):2569–2574, 1986. doi:10.1021/j100403a006.
- Andrew J. Sexton and Gary T. Rochelle. Reaction products from the oxidative degradation of monoethanolamine. *Industrial & Engineering Chemistry Research*, 50(2):667–673, 2011. doi:10.1021/ie901053s.
- H. Wold. Estimation of principal components and related models by iterative least squares. *Multivariate analysis*, pages 391–420, 1966.
- Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2:37–52, 1987. doi:https://doi.org/10.1016/0169-7439(87)80084-9.