

## ACT LAUNCH Project No 299662



The LAUNCH project is funded through the ACT programme (Accelerating CCS Technologies, Horizon2020 Project No 294766). Financial contributions are made from: Netherlands Enterprise Agency (RVO), Netherlands; Bundesministerium für Wirtschaft und Energie (BMWi), Germany; Gassnova SF (GN), Norway; Department for Business, Energy & Industrial Strategy (BEIS) together with extra funding from NERC and EPSRC research councils, United Kingdom; US-Department of Energy (US-DOE), USA.  
All funders are gratefully acknowledged.



**Lowering Absorption process **UN**certainty, risks and **C**osts by predicting and controlling amine degradation**

### **Deliverable Nr. D1.2.1**

**Big Data tools for identifying key degradation predictors**

Dissemination level	Public	
Written By	Juliana Monteiro (TNO),	Date: 19.03.2023
Checked by WP1 Leader	Andreas Grimstedt (SINTEF)	Date: 22.03.2023
Approved by the coordinator	Peter van Os (TNO)	Date: 13.04.2023
Issue date	13.04.2023	





## Executive summary

In this deliverable, we have explored the use of Big Data tools for improving the knowledge on solvent degradation, and particularly for identifying key degradation predictors. Initially, process and solvent analysis data from the ALIGN-CCUS RWE campaign with MEA were used. However, and despite working with data from the longest ever open MEA campaign, not enough *degradation data* was available. The attempts to analyse the process data did not lead to any new insights. There is a clear need to generate larger data sets on solvent degradation. However, doing that by solvent sampling and laboratory analysis would become prohibitively expensive.

A large data set on solvent composition can be generated by using a technique for online solvent monitoring. In this work, we describe the development of the hardware and methodology to use a mini-ATR-FTIR for this purpose. The final system configuration includes not only the analytical equipment itself (ATR-FTIR), but also a methodology for switching between solvent and water circulation loops. The innovative feature of the methodology developed is in the use of de-ionized water (DI) as the medium for taking backgrounds (instead of air). Equipping the system with automated pumps and valves for switching between solvent and DI water flow has the added benefit of ensuring a clean crystal, thus lowering the probability of measurement errors due to solvent accumulating on the crystal surface (amine solutions are “sticky”).

Despite all the progress on developing the methodology, the online test realized in a pilot plant using MDEA/PZ solvent gave erroneous measurements of solvent composition. The errors arise from software issues. In order to correct that, a more robust methodology for spectra analysis is required. In this work, we have tested different machine learning (ML) models for that end. The most successful models – convolutional neural networks – allow for the application of explainability methods which led to knowledge on key degradation predictors for MEA (wavenumbers 1350-1650  $\text{cm}^{-1}$  explain most of the degradation).

In conclusion, a robust methodology for generating online “Big Degradation Data” is needed, and the use of the mini-ATR seems like a promising route to achieve this goal. Once data is available, ML models can be used to interpret the spectra to derive information on solvent composition and degradation state. While the hardware side of the solution described seems sufficiently developed, the software side needs further attention. This report outlined next steps in the development so that a robust tool is available.



## Table of Contents

<b>1</b>	<b>INTRODUCTION .....</b>	<b>5</b>
<b>2</b>	<b>BIG DATA TOOLS FOR EVALUATING PLANT DATA.....</b>	<b>6</b>
2.1	APPLYING CHEMOMETRICS TO SOLVENT DEGRADATION DATA.....	8
2.2	HOW TO GENERATE BIG DEGRADATION DATA?.....	8
2.2.1	<i>MINI-ATR, offline operation .....</i>	<i>10</i>
2.2.2	<i>MINI-ATR, preparations for inline operation .....</i>	<i>12</i>
2.2.3	<i>Mini-ATR, operation at a pilot plant .....</i>	<i>19</i>
<b>3</b>	<b>BIG DATA TOOLS APPLIED TO SOLVENT MANAGEMENT .....</b>	<b>22</b>
<b>4</b>	<b>CONCLUSIONS AND FUTURE WORK .....</b>	<b>26</b>
<b>5</b>	<b>REFERENCES .....</b>	<b>27</b>



## 1 Introduction

Amine-based post combustion CO<sub>2</sub> capture (PCCC) is a mature technology, demonstrated at full scale, and currently entering implementation phase at various industries, including power plants, waste-to-energy facilities, chemical industries and ships. This means that a large data set will be available, from each of these plants individually, containing information on process data as well as solvent degradation data. Process data is generally available at a short frequency, with the various measurements in the plant (temperatures, pressures, flow, etc.) typically being logged many times per hour. Direct information on solvent degradation, on the other hand, derives from solvent sampling and offline analysis. This information is not available at real time, is much less frequent (typical intervals can be weekly, bi-weekly or monthly), and is expensive to generate.

Aqueous amines suffer from degradation – oxidative and thermal – under different operating conditions in the capture plant [1]. In this report, we evaluate the possibility of using big data tools for identifying key degradation predictors. The degradation mechanism of interest here is oxidative degradation, which is the main cause of amine solvent losses due to degradation. In literature, simple molecules that are easy to analyse for, such as formate, have been proposed as a key component for tracking solvent oxidative degradation [2]. Formate is a common degradation product of all amines, which allows for developing a generalized approach regardless of the solvent of choice. Formate is typically analysed for by using ion chromatography, and within the same analysis, other heat stable salt anions can be identified, like acetate and oxalate.

In the current report, data from a MEA campaign from RWE is used, consisting on process and solvent analysis data.

## 2 Big Data tools for evaluating plant data

The target of this activity was to identify, amongst a set of process and solvent composition data, key parameters that would enable to predict the degradation behaviour of amines. In a first exercise, the MEA campaign at the RWE pilot plant [3] was selected. This campaign represented the largest and most complete dataset for a MEA campaign available within LAUNCH. Moreover, after a period of ca. 200 days of operation, the MEA degradation spiralled out of control, and an acceleration in solvent loss and the accumulation of degradation products could be observed. This is illustrated in Figure 1.

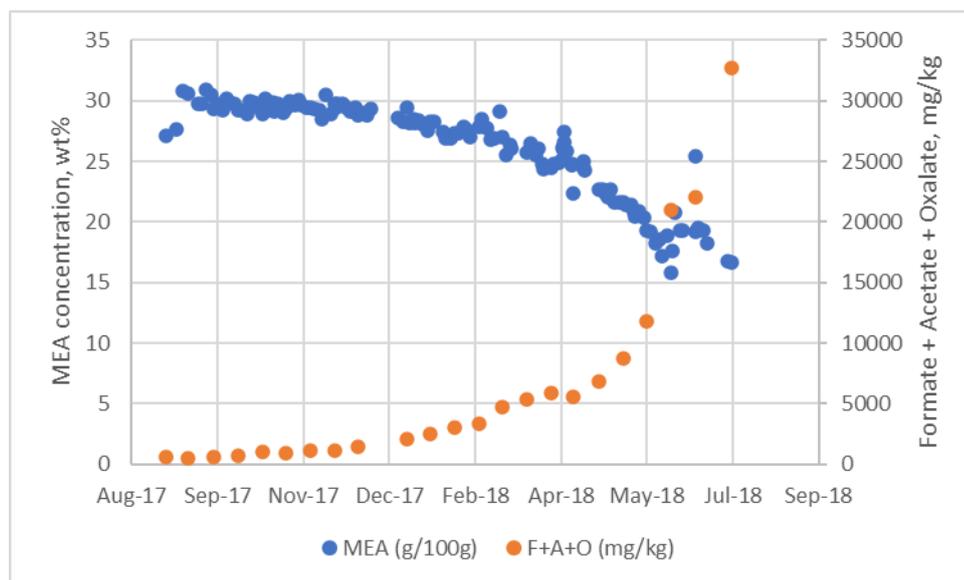


Figure 1. MEA concentration and sum of concentrations of formate, acetate and oxalate as measured during RWE's MEA campaign of 2017-2019.

To perform the analysis in question, we have received process data from the RWE pilot, consisting on hourly averages of 64 process signals (temperatures, flows, pressures, etc.). The data represented the period between the 24<sup>th</sup> of August 2017 and the 1<sup>st</sup> of March 2019, with a total of 13298 entries for each signal. This leads to 851072 data points.

For period between 24<sup>th</sup> of August 2017 and the 18<sup>th</sup> of July 2019, data on solvent composition is shared. This data was measured offline, and is available at a lower frequency. The water and CO<sub>2</sub> contents in the solvent was measured on a daily basis. The MEA content was initially measured weekly, but from September 2017 onwards, the frequency increased to twice a week. Iron, total sulphur, chloride, formate, nitrate, sulphate, acetate and oxalate were measured every 2 weeks. Finally, total Nitrosamines, No-HEGly, NDELA, OZD, HEIA and HEEDA concentrations were measured every 6 weeks (for clarity, these components names and CAS numbers are given in Table 1). This added to a total of 922 data points.

Table 1 – List of components names and CAS numbers

Abbreviation	Name	CAS number
MEA	monoethanolamine	141-43-5
No-HeGly	N-Nitroso(2-hydroxyethyl)glycine	80556-89-4
NDELA	N-nitrosodiethanolamine	1116-54-7
OZD	2-Oxazolidone	497-25-6
HEIA	N-(2-hydroxyethyl)imidazolidinone	3699-54-5
HEEDA	N-(2-hydroxyethyl)ethylenediamine	111-41-1

The nature of the process and the control loops of the plant create dependencies between the process entries. Moreover, operators' actions greatly influence the behaviour of the system. And finally, process equipment limitations may mask correlations. In Figure 2, it can be seen that the solvent circulation rate is inversely related to the MEA concentration. This is logical, and a result of the operators actions – as the solvent degrades and loses cyclic capacity, the solvent circulation rate is increased by the operators in order to keep the CO<sub>2</sub> capture rate constant. This is done consistently between February and May 2018, but at that point the pump approached its maximum capacity, and the solvent circulation rate remained high but somewhat constant until July that year, when a major intervention (bleed and feed) took place to try to control the accelerated degradation rate of MEA.

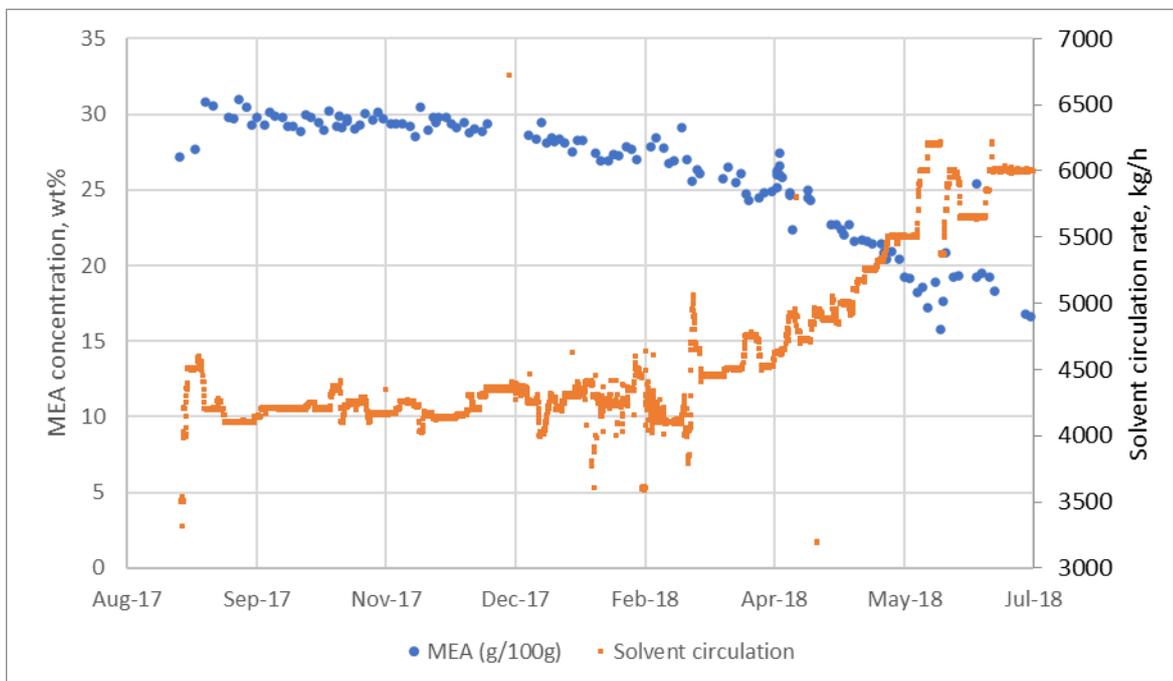


Figure 2. MEA concentration and solvent circulation rate.

## 2.1 Applying chemometrics to solvent degradation data

While the dataset generated at RWE during the 18 month MEA campaign (the longest ever open campaign with this solvent) is impressive, it is not a “Big Data” set. Particularly not when looking into degradation data, i.e., the accumulation of solvent degradation products over time. There are many data gaps, mostly deriving from the fact that performing solvent analysis is not only laborious but also very expensive. For total Nitrosamines, No-HEGly, NDELA, OZD, HEIA and HEEDA there are only 8 individual measurements for each of these components. And for formate, acetate and oxalate there are 23 data points for each. Moreover, for some major degradation compounds such as HeGly (N-(2-hydroxyethyl)glycine, CAS No: 5835-28-9), HEPO (4-(2-hydroxyethyl)piperazin-2-one, CAS No: 23936-04-1), and MEA-urea (N,N'-bis(2-hydroxyethyl)-urea, Cas No: 15438-70-7), no data was available.

This discouraged the use of Big Data analytics and pointed at the direction of more traditional chemometric tools. For that, only the solvent composition data was used. The different degradation products and impurities accumulated in the system are used as predictors, with MEA concentration being the independent variable.

The first step was data pre-treatment. Missing values were filled either with the value of the lower detection limit, the last value measured or by using linear interpolation. The choice had little influence on the results. Further data transformation included reporting the compositions as normalized weight percentages, and excluding water and CO<sub>2</sub> concentrations from the dataset.

After the data was scaled, a principal components analysis (PCA) was carried out. PCA is a common statistical technique for reducing the dimensionality of a dataset by linearly transforming the data such that the data variation can be described with fewer dimensions in the new coordinate system than the initial data. The PCA indicated that 3 principal components were sufficient to explain most of the variation in the data. Several degradation components seem to be equally important in predicting MEA degradation. A partial least square (PLS) analysis reveals that iron, formate, oxalate and acetate are the variables with the highest (absolute) regression coefficients. As mentioned, due to relative ease in analysing, the rate of accumulation of formate is commonly used as a proxy of degradation [2]. Our analysis indicates that the procedure is valid for MEA. Moreover, it has been pointed that, for various MEA campaigns, the rate of accumulation of iron in the solvent is well correlated to the rate of accumulation of formate [4]. This is also confirmed in our analysis.

In general, no new insights were obtained by this chemometrics analysis.

## 2.2 How to generate Big Degradation Data?

While manipulating the RWE dataset, it quickly became clear that we had taken the wrong approach in trying to apply Big Data analytics to degradation. Instead of gathering the somewhat scarce dataset available and trying to squeeze useful and non-obvious information out of it, it would be more useful to turn the problem around. We then asked ourselves how to generate a Big Dataset which contains useful information on degradation – that could then be processed to predict degradation.

An online measurement tool based on chemometrics was developed by TNO during the ALIGN-CCUS project, Figure 3. The setup consisted of different sensors, generating measurements that were then used to calculate the solvent composition and the loading of the samples. The method made use of a model relating physical properties of the solvent with its composition, and was built based on a chemometric approach [5], [6]. The sensors used were: NIR spectrometer, sound velocity meter, pressure sensors, temperature sensors, conductivity sensor, pH sensor, mass flow controller, UV vis spectrometer and refractive index . The system

was also equipped with a software with an automatic logging/operation system which allowed the operator to easily monitor the solvent composition in real time during operation, shown in Figure 4.



Figure 3 - TNO's online monitoring tool (Chemcube)

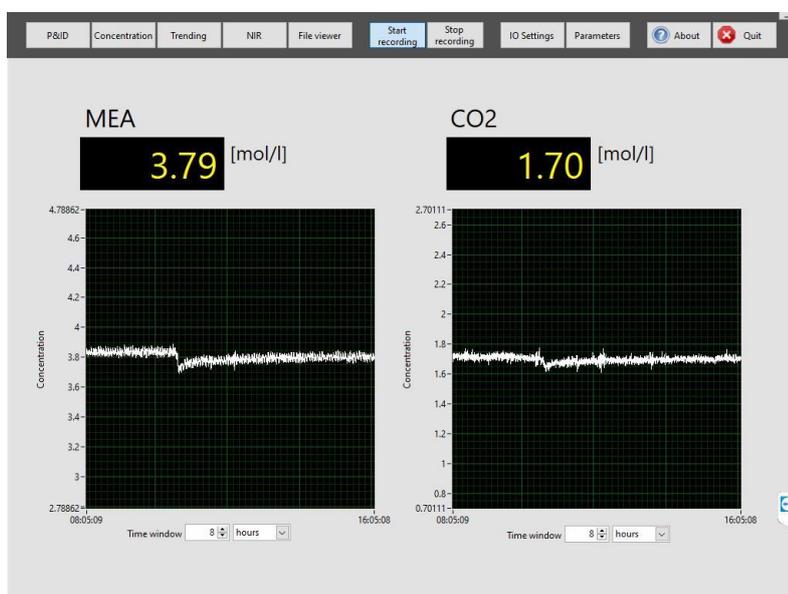


Figure 4: Real-time solvent monitoring

In operation, the Chemcube used a slipstream on the order of 20 kg/h. The Chemcube operation started in a lab-scale and developed to TRL7 during the ALIGN-CCUS project, being operated at RWE in Niederaussem both with MEA and CESAR1. During those campaigns, however, issues were encountered with the data acquisition which made the setup more complex than originally intended. Moreover, as the model relied on

the readings of many instruments, robustness became an issue. For instance, we have noticed that with drifting of some of the readings over time, particularly pH, the wrong concentrations were calculated.

At the end of ALIGN-CCUS, tests were done using a mini-ATR, the *Agilent Cary 630 FTIR* system, as shown in *Figure 5*, in order to evaluate if this tool would enable a more robust online solvent monitoring system.



Figure 5a and 7b: the mini ATR seen from above and in perspective

### 2.2.1 MINI-ATR, offline operation

FTIR analysis is extensively used at TNO for offline sample analysis and has consistently produced satisfying and reproducible results using the Nicolet iS50 FTIR Spectrometer, a stationary bench-top instrument confined to the laboratory. However, the instrument is relatively big, expensive and sensitive, making it less easy to transport from place to place.

Handheld ATR instruments exist and within ALIGN-CCUS initial steps have been taken to explore the possibility of employing such an instrument on-site in continuous monitoring of solvent quality. The mini-ATR holds several advantages over the Chemcube, the most important being its compactness and ease of handling, ease of operation and calibration and simple yet efficient construction.

The mini-ATR was calibrated for the same solvents as the Chemcube: fresh 30 wt% MEA and CESAR-1. Both calibrations were highly linear. The MEA solvent has an accuracy of  $\pm 0.4$  wt% and  $\pm 0.17$  wt% for MEA and  $\text{CO}_2$  respectively, while the CESAR-1 solvent has an accuracy of  $\pm 0.2$  wt%,  $\pm 0.1$  wt%, and  $\pm 0.1$  wt% for AMP, PZ, and  $\text{CO}_2$  respectively. Figure 6 a, b, and c show the good fit between the real composition and the composition predicted by the mini-ATR, for AMP, PZ and  $\text{CO}_2$ , respectively.

With the promising results obtained in ALIGN-CCUS, it was decided to work with the mini-ATR in LAUNCH-CCUS, in detriment of the Chemcube. Since the interest was in generating “real-time” degradation data, the system had to be adapted so that it could be installed in-line, processing a slip-stream of the solvent.

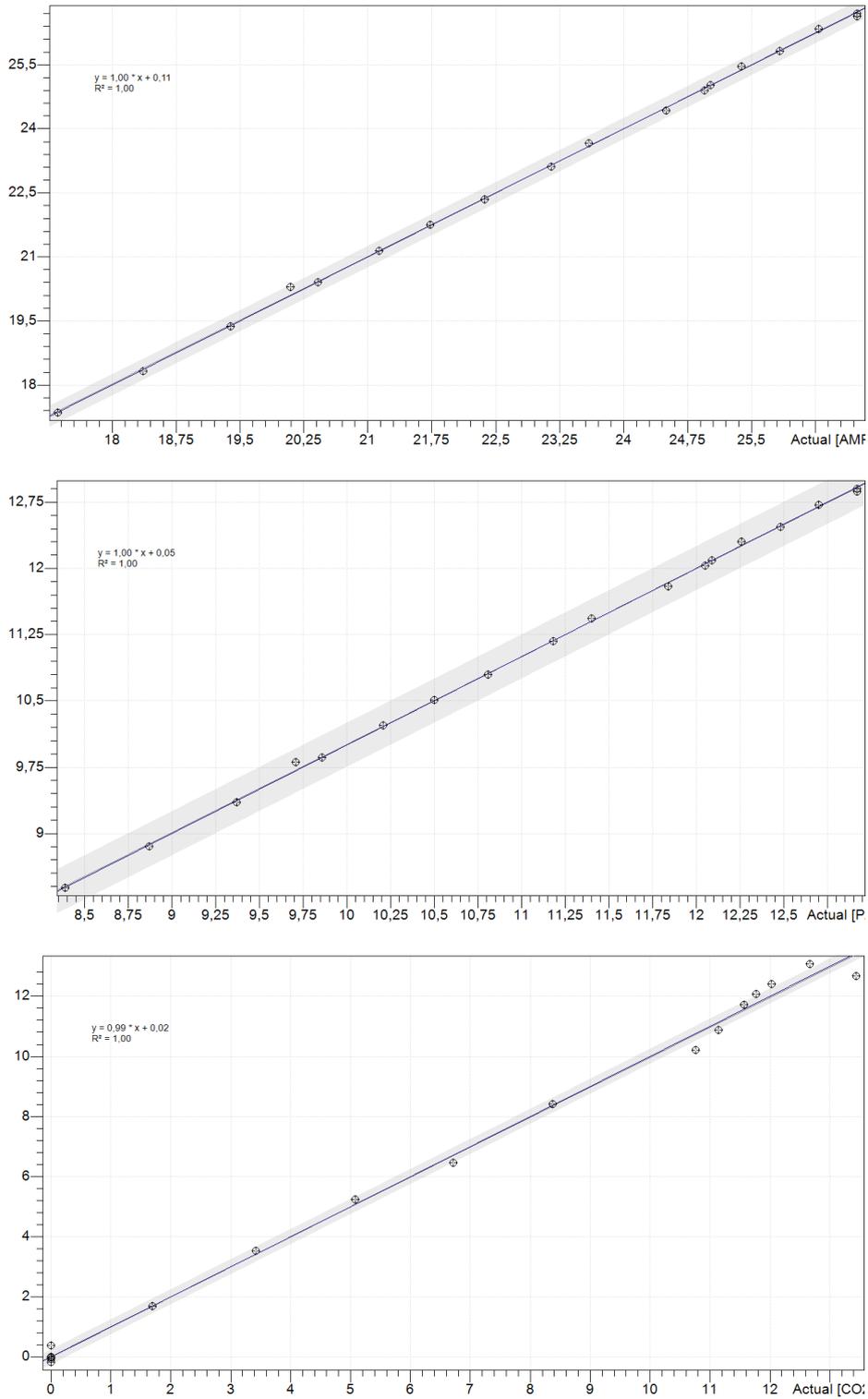


Figure 6 a, b, c: Predicted vs actual composition for AMP, PZ and CO<sub>2</sub> (from top), for the calibration set data

### 2.2.2 MINI-ATR, preparations for inline operation

Agilent, the mini-ATR provider, also provides a flow cell that allows for a solvent stream to be connected to the mini-ATR. This hardware was purchased by TNO, and a new operational methodology was developed.

In operation, the mini-ATR continuously pumps a slip stream to a flow-through cell placed on the ATR's measuring location. Measurements are taken at a fixed interval that can range from every five seconds to hourly and the concentration of solvent and dissolved CO<sub>2</sub> is reported instantaneously. Test measurements with the ATR have been conducted at TNO, shown in Figure 7, with promising results.

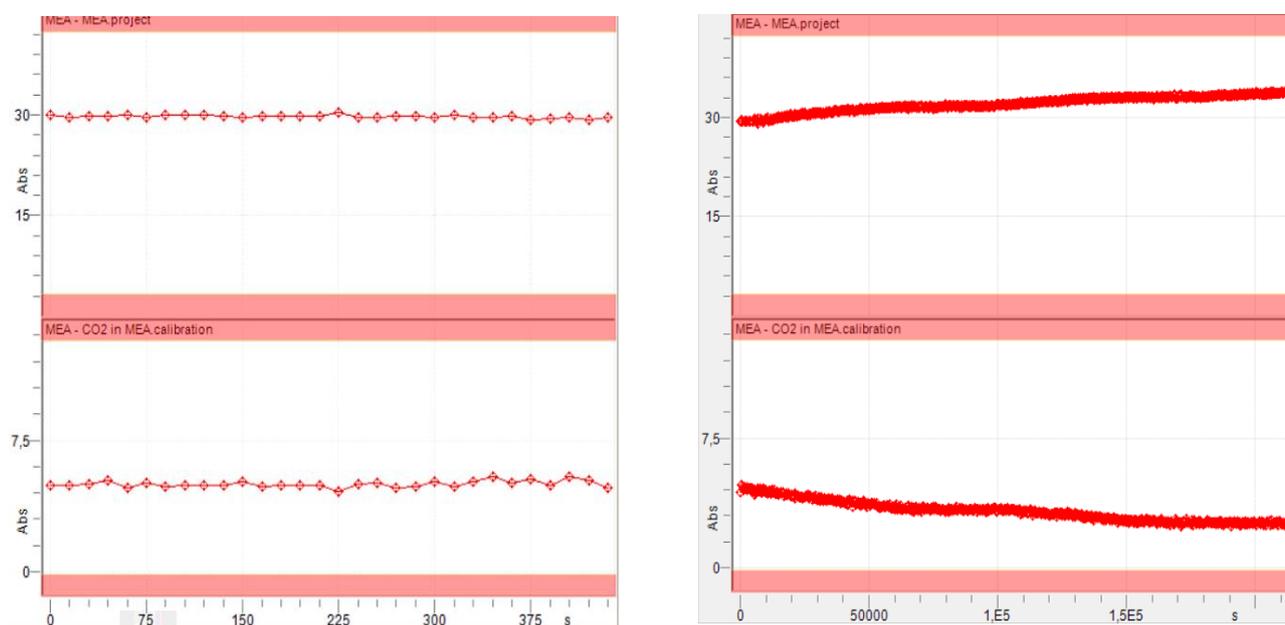


Figure 7 a) and b): Test runs with mini-ATR. MEA concentration on top and CO<sub>2</sub> concentration below. Run over 6 minutes to the left, and run over 2 days to the right

In Figure 7 a), the measurement is accurate and stable over a period of about 6 minutes. However, for more extensive trials, as shown in Figure 7 b) lasting for ca. 6 days, some problems were encountered including drifting of the reported concentration and the software crashing. Figure 7 b) shows that the MEA concentration reading starts at 30 wt% with CO<sub>2</sub> loading of around 5 wt% (as expected). However, the concentrations drifted towards 40 wt% and 3 wt%, respectively over the course of 48 hours.

The most likely culprit of this drifting is a lack of reference spectra: In manual operation, a reference spectrum is manually taken before analysing any sample, to remove the spectrum of the background medium, which normally is air. In this manner results are reproducible in different background conditions. When operating continuously with a solvent flowing through the flow cell, there are no opportunities to take new backgrounds. Therefore, the methodology was modified so that water is used as the background.

To automate the background procedure, a three-way valve was installed prior to the flow cell, allowing us to periodically pump deionized water through the system. The DI water is also used for cleaning the glass, avoiding deposition of impurities overtime. Once the glass is considered clean (i.e., after running with DI water

for a period understood as sufficient, as per our tests around 5 minutes), a new background can be taken. This is a relatively simple solution yet potentially very efficient. The P&ID of the mini-ATR system for inline measurements is shown in Figure 8, with the three-way valve allowing for switching between water and solvent. It also includes another, optional three-way valve, that would allow for the water waste to be disposed of. In operations with large scale plants, this extra step would not be necessary, as the waste water flow consists of a 20 L/h flow, over 5 minutes run, thus adding 1,7 L of water to the plant inventory. This would be a negligible addition, and the extra water could easily be evaporated via the water wash. However, in small scale plants, particularly in LAUNCH rigs, the flow through the mini-ATR is of the same magnitude of the solvent circulation flow, and the added water would disturb the solvent composition, generating a significant dilution. Therefore, in small scale systems, the disposal of the waste water is necessary.

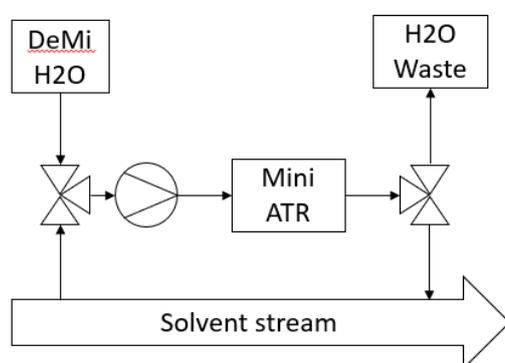


Figure 8. Schematic of continuous solvent monitoring with mini-ATR

To facilitate online solvent monitoring, the full process of measurement of single beam spectrums, transformation into absorbance, quantification of the spectrum and visualization have been automated. The spectra measured by the ATR are automatically uploaded to a TNO server at fixed intervals. This is done by means of using a scheduled task on a Windows server that starts a Python script. Copying of the raw source files (raw measurements and valve position data) to the processing server ensures traceability and enables backup of the data, before processing steps begin. This ensures that a reliable Big Data set is collected over time, that can then be analysed for.

The processing server sorts the raw measurements into background files and sample files based on the position data of the inlet valve. Next, the script transforms the single beam measurements into absorbance spectrums. The absorbance spectrums are quantified using a partial least squares model present in the Microlab Expert software accompanying the Cary 630 FTIR system.

The quantification software requires a Windows environment, so it runs on a separate server (the quantification server). The absorbance spectra generated on the processing server are sent to the quantification server automatically every 30 minutes. A scheduled task script uses the quantification software to predict the amine content and CO<sub>2</sub> loading in the solvent being monitored. Results are pulled by the processing server that stores the predicted solvent composition in a time-series Database (InfluxDB).

The database's accompanying visualization platform (Grafana) provides visualization capabilities of all the logged data. Figure 9 shows a screenshot of an interactive dashboard used for visualizing solvent

composition. Operators, analysts and others can easily make and store their own dashboards, suited to the tasks that they need to perform. This is a significant improvement in time and tediousness over the earlier process where a lot of manual work was required, such as manual copying and (Excel) spreadsheet manipulation of data.



Figure 9. Example of solvent composition visualized in Grafana

The dynamic measurements, calculation and visualization by means of the hardware modification and TNO's software allow for the use of the ATR-FTIR as an online tool for solvent monitoring. The setup has been successfully tested in a laboratory environment with MEA. Following that, and in combination with the Dutch national project DECIPHER, the mini-ATR system was tested at a large scale pilot plant at the HVC waste-to-energy facility in Alkmaar. The plant operates with a blend of MDEA and PZ, therefore a new calibration was required.

In the schematic representation of the mini-ATR system shown in Figure 8, only one pump is used. In practice, it was noted that this configuration led to occasional air flow into the system, causing measurement errors. Therefore, the flowsheet was modified according to Figure 10, to include a solvent sump (a small vessel with ca. 500 mL), and a second pump. A picture of the final configuration of the system is given in Figure 11.

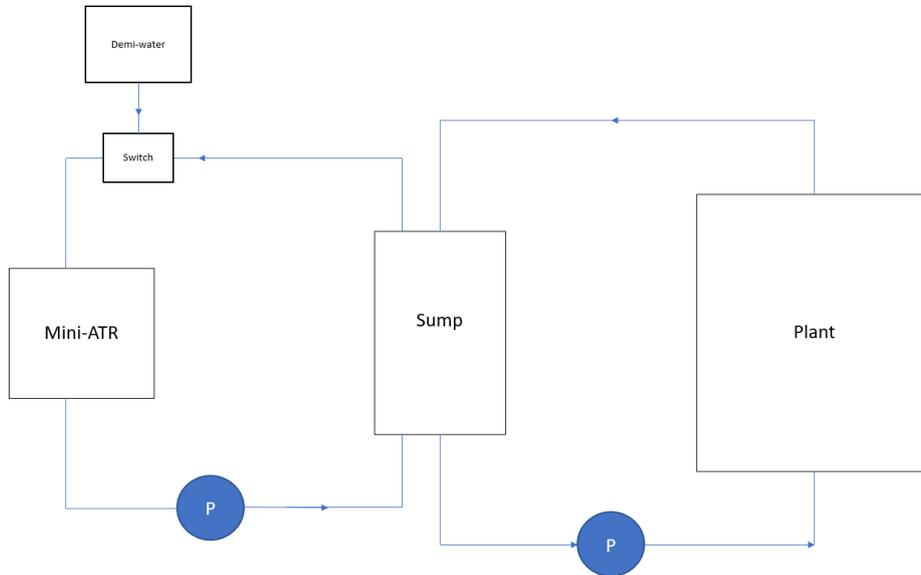


Figure 10. Scheme of the Mini-ATR configuration. Pumps represented by the blue circles with a “P”; 3-way valve represented by the “Switch” box.



*Figure 11, Mini-ATR setup*

To determine the concentration of amine (MDEA/Piperazine in this case), water and CO<sub>2</sub>, a new calibration was made. The calibration lines, for CO<sub>2</sub> (Figure 12), water (Figure 13), MDEA (Figure 14) and Piperazine (Figure 15) are given next. The x-axis shows the wt% of the components in each of the samples prepared to perform the calibration, whereas the y-axis show the model predictions.

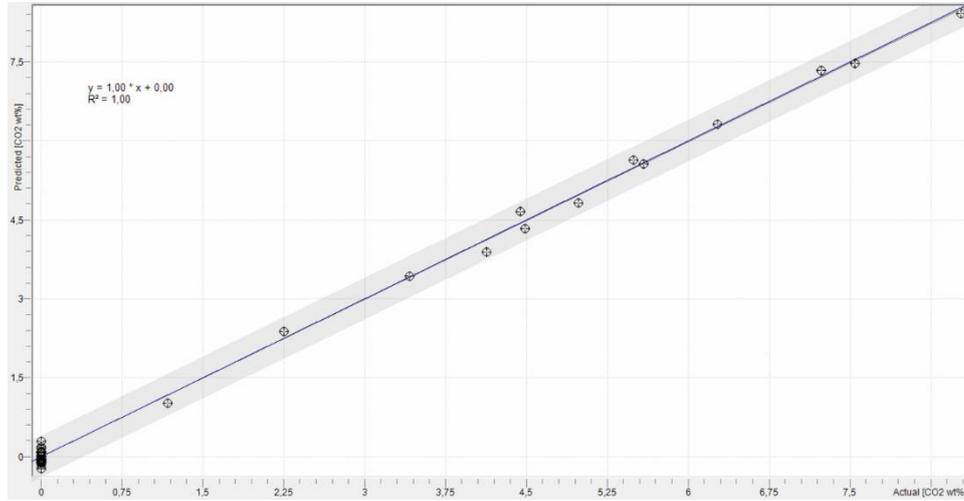


Figure 12, Calibration line for CO<sub>2</sub> in the mini-ATR

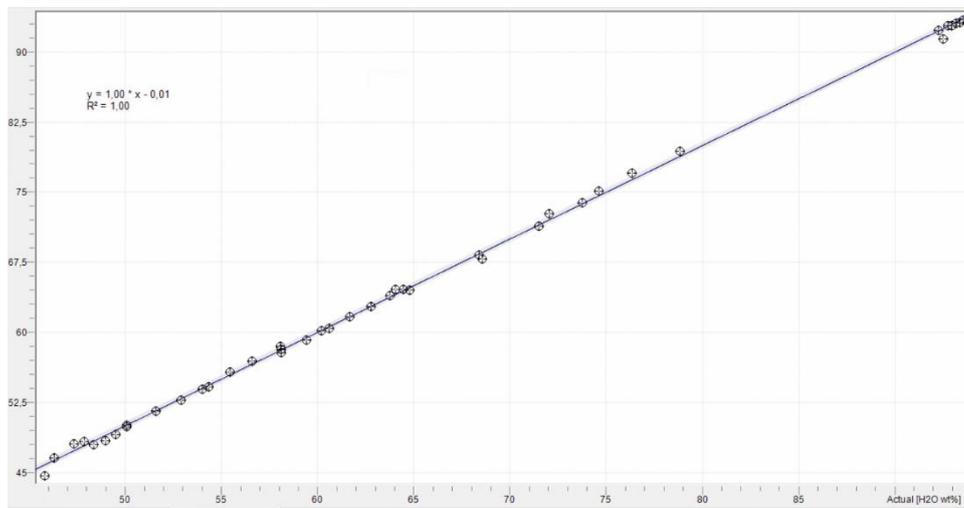


Figure 13, Calibration line for H<sub>2</sub>O in the mini-ATR

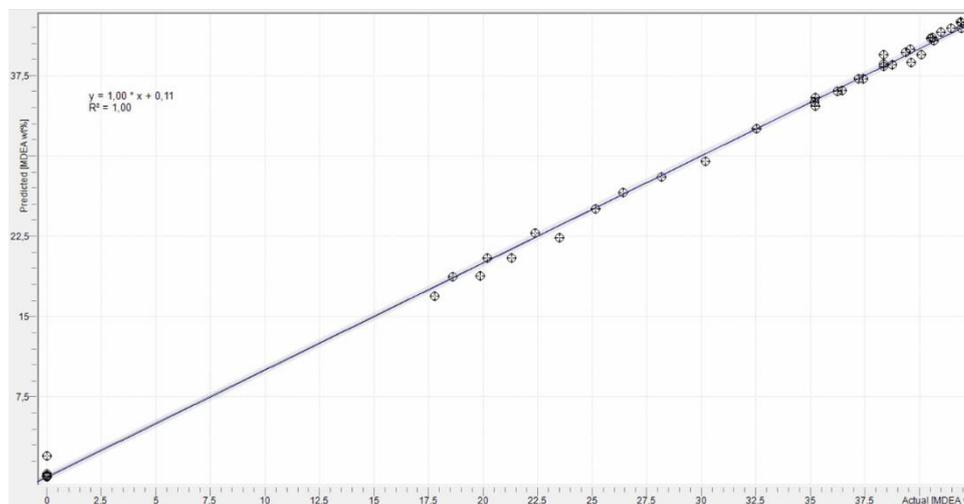


Figure 14, Calibration line for MDEA in the mini-ATR

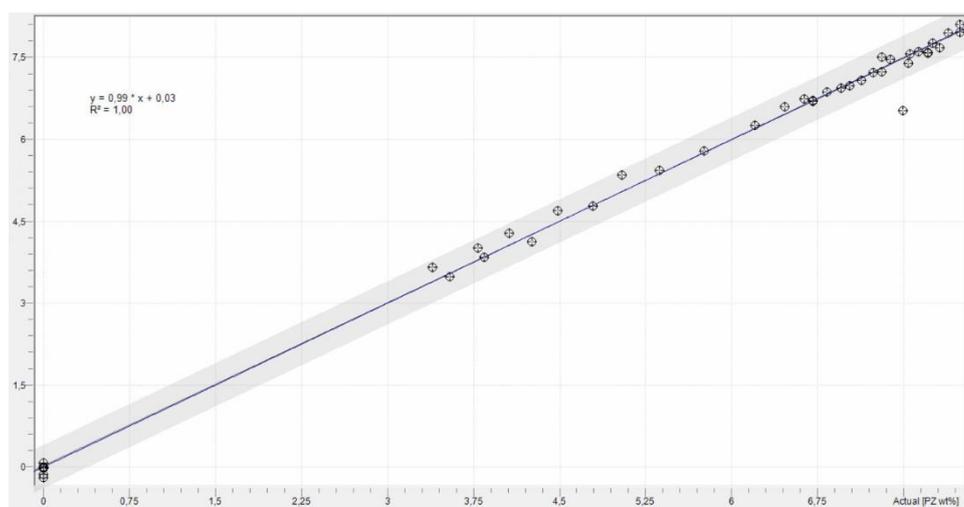


Figure 15, Calibration line for Piperazine in the mini-ATR

To validate the mini-ATR calibration, a first test was to put a droplet on the mini-ATR cell (i.e., to operate without the flow cell). The result of this is in Table 2, and show good agreement between the mini-ATR measurements and the prepared solvent.

Table 2, Mini-ATR validation tests

Compound:	Prepared solvent (wt%):	Measured 1 (wt%):	Measured 2 (wt%):
MDEA	33.64	32.70	32.76
PZ	4.87	4.76	4.94
H <sub>2</sub> O	59.01	59.66	59.39
CO <sub>2</sub>	2.49	2.88	2.91

### 2.2.3 Mini-ATR, operation at a pilot plant

After the equipment was validated in the lab, the Mini-ATR was taken to HVC and connected to the CO<sub>2</sub> capture pilot plant, taking a slip stream of the rich solvent via a sampling port. The equipment was at HVC for 10 days, and the collected data is shown in Figure 16. Unfortunately, stable operation was only achieved during a couple of hours in the beginning of the campaign, Figure 17, and for two consecutive days at the end of the campaign, Figure 18.

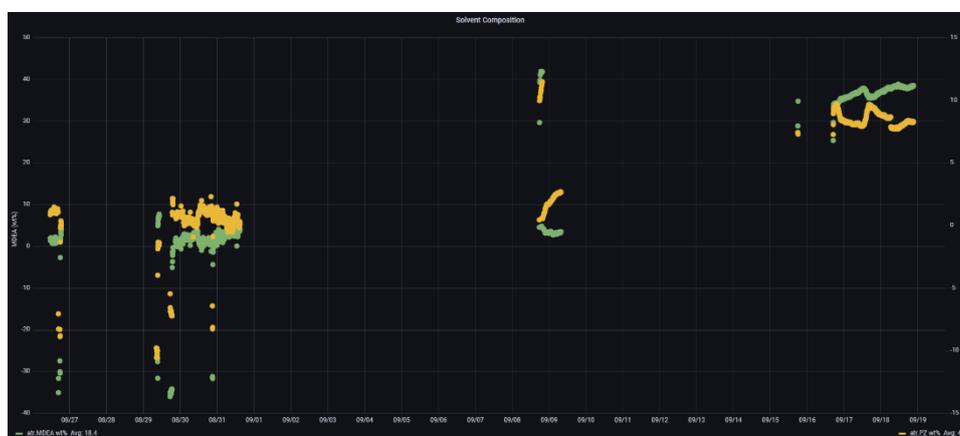


Figure 16, Overview of Mini-ATR campaign at HVC. Green: MDEA wt% composition (left y-axis) Yellow: PZ wt% composition (right y-axis)

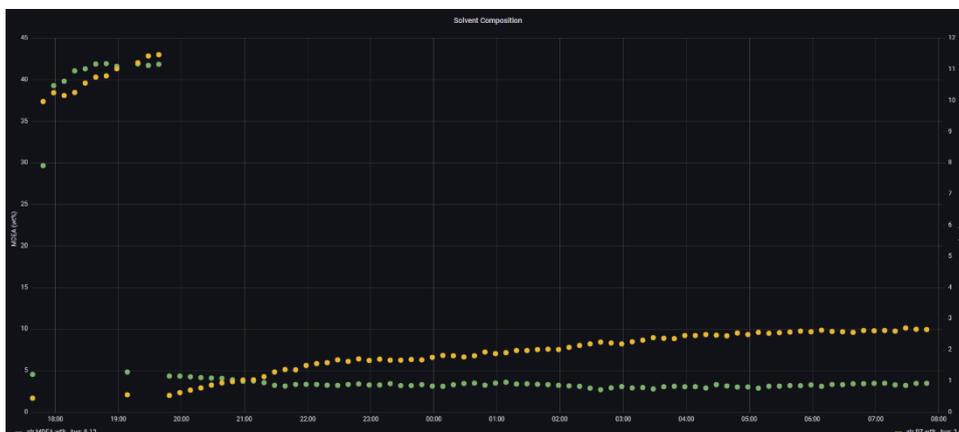


Figure 17, Mini-ATR operation during times with device issues. Green: MDEA wt% composition (left y-axis) Yellow: PZ wt% composition (right y-axis)

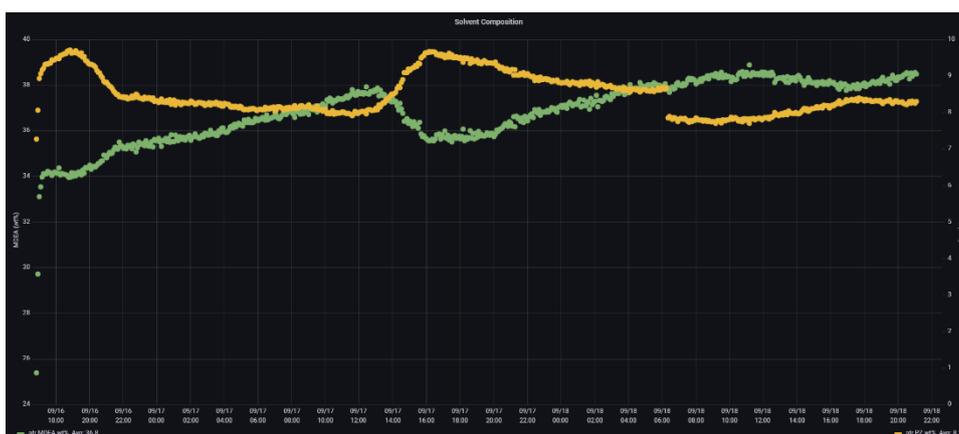


Figure 18, Example of stable Mini-ATR operation during the campaign. Green: MDEA wt% composition (left y-axis) Yellow: PZ wt% composition (right y-axis)

In order to validate the results of the Mini-ATR, the data was compared with the offline sample analysis performed at TNO using Ion Chromatography (IC), and specifically cation chromatography. The comparison was done for 3 different periods and it is shown on Table 3. It is possible to observe that the results did not agree with the IC results. Some of the reasons identified include:

1. The technology used is based on the spectra of the components. In the case of MDEA and Piperazine, some of the peaks of the spectra overlap which means that the concentration could be read as both MDEA or Piperazine. This introduces errors to the measurement.
2. During the period of the measurements, the solvent was already degraded which has an impact on the accuracy of the measurements due to the fact that the calibration lines are made using fresh solvent. Improving the calibration line by adding degraded samples to it can increase the accuracy of the results for future experiments.

3. The Agilent software introduces a time delay that interferes in the measurements. Over time, due to the delay, the background is mistakenly measured (instead of data) and the following measurements will show a significant deviation. This can be observed in Figure 17 in which the concentrations for both MDEA and Piperazine were higher in the beginning of the measurement period, but decreased abruptly and with no apparent cause. This software delay error was detected, but no solution was yet found. One possibility could be not to use the Agilent software, as described in Chapter 3.

Table 3. Comparison of the Mini-ATR and IC results for MDEA, Piperazine, CO<sub>2</sub> and water

	Component	Mini-ATR		IC		Error
		wt%	mol/L	wt%	mol/L	
8-sep	MDEA	41.8		30.6		37%
	PZ	11.4		5.2		119%
	CO <sub>2</sub>	1.7	0.9		0.6	50%
	H <sub>2</sub> O	45.1		48.0		-6%
16-sep	MDEA	31.8		30.5		4%
	PZ	7.4		5.1		45%
	CO <sub>2</sub>	3.9	0.9		0.7	29%
	H <sub>2</sub> O	56.9		48.0		19%
19-sep	MDEA	37.7		28.9		30%
	PZ	8.4		4.8		74%
	CO <sub>2</sub>	2.3	0.5		0.7	-19%
	H <sub>2</sub> O	51.6		49.0		5%

Online operation of the Mini-ATR was accomplished enabling the concentration to be monitored in real-time during operation of the pilot plant. Therefore, the proof of concept was achieved. Longer operation of the Mini-ATR needs to be realized once the issues discussed above regarding the measurement accuracy are solved. In order to overcome such issues, TNO is developing a new data analysis technique using AI (machine learning), as described in Chapter 3.



### 3 Big Data tools applied to solvent management

The spectra generated by the ATR-FTIR while monitoring the solvent in real time are typically used to calculate the solvent composition based on calibrations, as discussed in this report, item 2.2. Specific peaks in the spectra correspond to the components of interest, and these are normally analysed. However, the entire spectrum holds information about the concentrations of all degradation products that are formed in the solvent over time. Using artificial intelligence (AI) and machine learning (ML) methodologies, the entire spectra can be studied. Based on this information, an indication of solvent quality can be made.

One key objective of LAUNCH was to evaluate the use of AI/ML (Big Data tools) to improve the understanding of solvent degradation. If AI/ML methods can extract useful information from ATR FTIR spectra on the degradation of solvent over time, then the online spectra information can be interpreted by the plant operators to assess the state of the solvent, and support decision making regarding the use of solvent management techniques (e.g., when to start reclaiming? When to stop?). The current definition of a “dirty” solvent is arbitrary, and within LAUNCH we have shown that techniques to remove impurities from degraded solvent can actually lead to accelerated degradation. The mechanisms behind that are not yet fully understood, but there is evidence that certain (unknown) impurities act to retard degradation. Some impurities may act as oxygen and/or radical scavengers, some impurities may chelate and deactivate metals. The fact that the complex chemistry behind the growing body of evidence against “over-cleaning” the solvent is unknown gives further support to the use of AI/ML to perform data analysis.

As discussed, the traditional method for analysing ATR-FTIR data uses partial least squares (PLS) regression to interpret the spectra data and derive composition information. The software for making that analysis is supplied together with the equipment (Agilent is the supplier of the TNO equipment, but other suppliers are available such as Metler Toledo, Thermo Fisher, etc.). The user needs to select which ranges of the spectra would correspond to different components in the system, and then process a number of calibration samples (at least 20) from which data is derived and used in the regression. Normally, the calibration procedure for a new solvent takes a couple of days, so it is not very work-intensive. However, when a CO<sub>2</sub> capture plant is in operation, the solvent composition changes over time due to the accumulation of flue gas impurities, metals from the materials of construction and degradation products in the solvent. With that, more “noise” is present in the ATF-FTIR spectra and the predictive capabilities of the original PLS model is lost (see poor predictions shown in Table 3). Another important point to notice is that all the information in that “noise” is lost with the current approach.

Therefore, we looked into the possibility of using a ML methodology to evaluate the composition and degradation level of amine samples. For that, ATR-FTIR measurement data for amine solutions of different degradation levels (scores) was generated. The data was divided into training and test datasets. A machine learning model (artificial neural network) was trained to predict the degradation level based on the FTIR spectrum, and the model accuracy was evaluated with the test dataset. As LAUNCH ultimate goal is to generate more knowledge to improve understanding of degradation, explainability techniques were used to show which parts of the spectrum the model considers most informative when making predictions. With this, the ML model findings can be combined with experts inputs to generate new knowledge.

A total of 64 ATR-FTIR scans (21 samples, measurements in triplicates) were performed, with sample degradation level ranging from 0 (fresh) to 10 (fully degraded). The absorbance data as a function of the wavenumber for each sample is given in Figure 19. All tests were performed with 30wt% MEA.

The fully degraded solvent was obtained from the RWE pilot plant, after the solvent was operational for around 300 days. The fresh solvent was prepared by TNO using Sigma-Aldrich 99% purity MEA and DI water. The intermediate degradation levels were achieved by diluting fully degraded solvent with fresh solvent. The starting composition of the degraded solvent, measured by the ATR-FTIR was 25 wt% MEA, with a loading

of 0.22 mol/mol. To bring the composition to 30 wt%, MEA was added to the solvent. After that, the composition was determined again by FTIR to be 30 wt% MEA, and the resulting CO<sub>2</sub> loading was 0.17 mol/mol. The fresh MEA solvent was then prepared (gravimetrically) with this same composition and loading. It is therefore expected that all the solutions tested had the same MEA and CO<sub>2</sub> content. However, as previously discussed in this report, the presence of degradation products in the solution can cause the FTIR measurements to deviate. Hence, we are not able to discard the possibility that there would have been differences in the MEA content and CO<sub>2</sub> loading between the degraded sample (after composition correction) and the fresh solvent separated. In future studies, the composition should also be verified by ion chromatography.

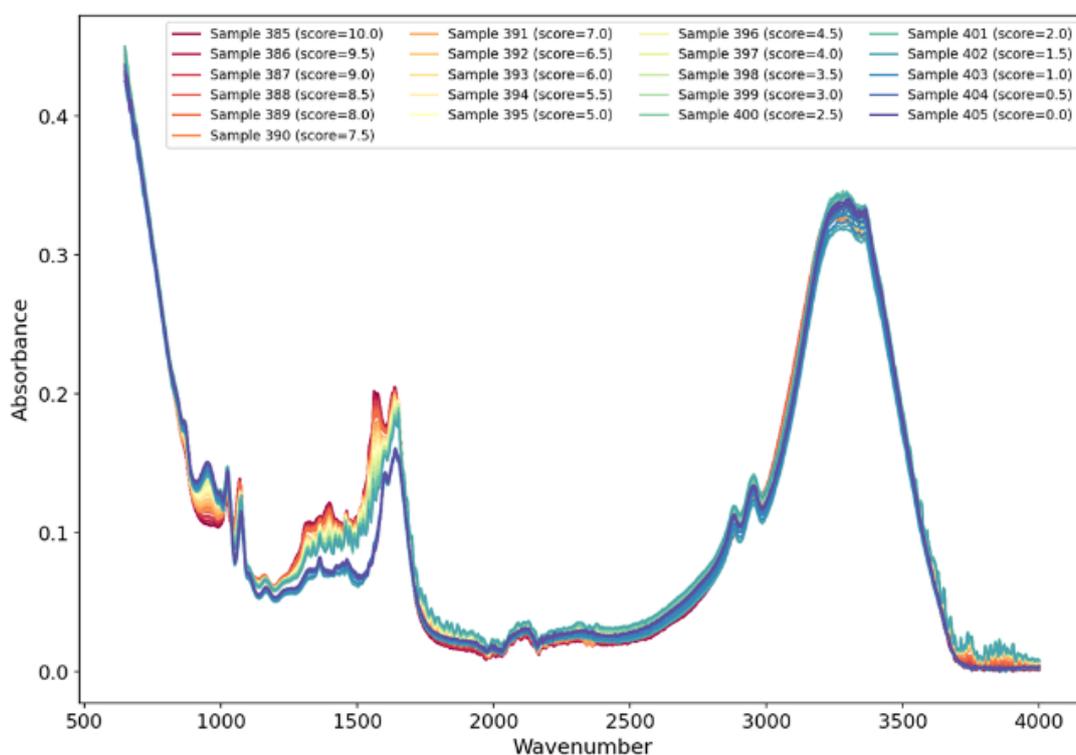


Figure 19. ATR-FTIR spectra of all 64 scans performed

Using the full FTIR spectra as input, the model target is to predict the degradation level. This can either be set as a regression or a classification problem. In regression, the scores are treated as continuous, and the model predicts a numeric value corresponding to the score. In a classification problem, the scores are treated as discrete classes and the model predicts a degradation group the spectrum belongs to.

It was decided to frame the problem as a regression problem for two main reasons: i) it allows for prediction of spectra in between current evenly spaced scores, which is likely to occur in practice; ii) classification of 21 classes with 3 measurements per class is a big challenge as there are very little examples per class to learn from, and the distinction between neighbouring classes is fuzzy.

Next, 3 different types of models were applied, namely: i) a “regular” feedforward artificial neural network (ANN); ii) a fully convolutional network (FCN); iii) a residual Network (ResNet). The latter two networks are so-called convolutional neural networks (CNN). CNN are specifically designed to process image data (for instance for image recognition) and use convolutional filters to extract feature maps from the images. These feature maps typically correspond to elementary parts of the image such as edges, lines, basic shapes, etc. The neural network then learns to connect these elementary parts to things/objects it can learn to recognize. By treating the FTIR spectra as a one dimensional image, these networks can be used to predict amine degradation levels.

All three models were trained on the same 45 spectra (70%) and tested on the same 13 (20%) spectra. The plots in Figure 20 show the predicted value versus the true value for each of the three networks.

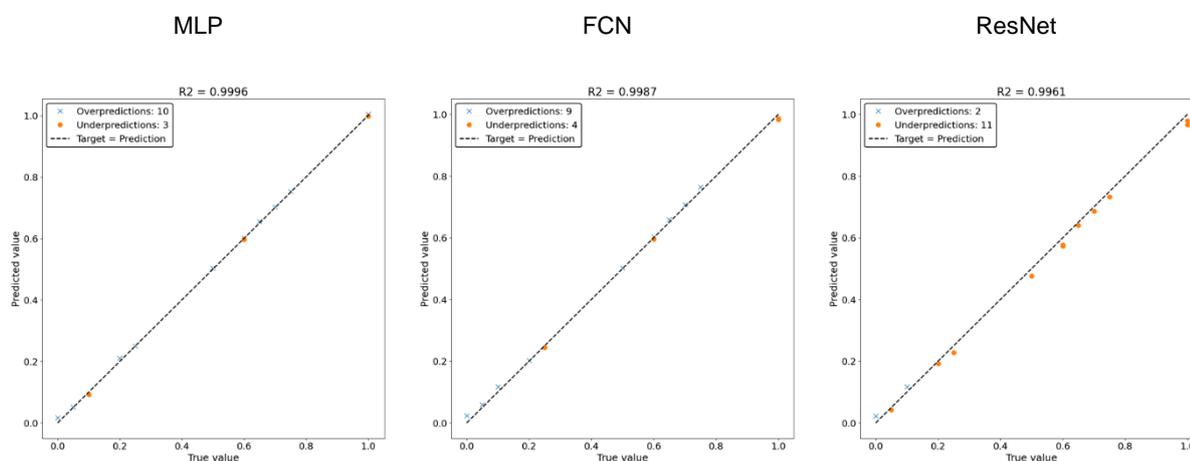


Figure 20. Predicted value versus the true value for each of the three networks

As seen in Figure 20, all network types can predict the degradation levels with very good fit ( $R^2 > 0.99$ ), but there were some differences in the ease of training and setting up the models. The “regular” ANN was the easiest to set up and quickest to train and didn’t require any sort of “tweaking” of the hyperparameters. In addition, it was also the most accurate (although all models were very close in accuracy). The FCN took longer to set-up and train and was quite sensitive to hyperparameters values. The ResNet took longest to train and was most sensitive to the specific hyperparameter values. There were many times that the ResNet never converged (training is a stochastic process so doesn’t always lead to the same outcomes), so in practice it was the least useable.

While the MLP is simplest and most accurate, the reason the FCN/ResNet were used was that they allow the use of explainability methods to see which parts of the spectrum is mostly used when making a prediction. With such methods, we aim to explain how the model comes to a certain prediction, as the internal parameters are learned from data alone and don’t always have a particular (physical) meaning, through the use of a so-called “activation map”. At the end of the convolutional network, a global average pooling layer (GAP) is added, which takes the average of all the feature maps at the last convolutional layer, and the weight between this layer and the output to indicate which part of the spectrum has the highest activation when a prediction is made. We can directly map this activation back onto the original input. This allows us to see which parts of

the input spectra contribute most to making a prediction and whether the model learn from the “correct” parts of the spectrum, i.e., the parts that the experts judge as important.

Figure 21 left shows the scaled activation map for all spectra using the FCN. Parts coloured red are the ones the model uses the most when making a prediction. As can be seen, there are a number of areas that FCN focusses on. These correspond to the following wavenumbers: 900-1100; 1350-1650; 2850-3050; 3500-3550  $\text{cm}^{-1}$ . However, as can be seen, there is quite some variability from one spectrum to the next; especially wavenumbers 1350-1650  $\text{cm}^{-1}$  show large variation, all the way from dark red to dark blue. These parts where the variation is highest are actually the most indicative wavenumbers, as a change in these parts leads to the highest change in model activation (and thus output prediction). Figure 21 right shows the average activation and 5<sup>th</sup>-95<sup>th</sup> percentile for all spectra combined. The wider the blue band, the most important the wavenumbers for the prediction. This means that the wavenumbers 1350-1650  $\text{cm}^{-1}$  explain most of the degradation information.

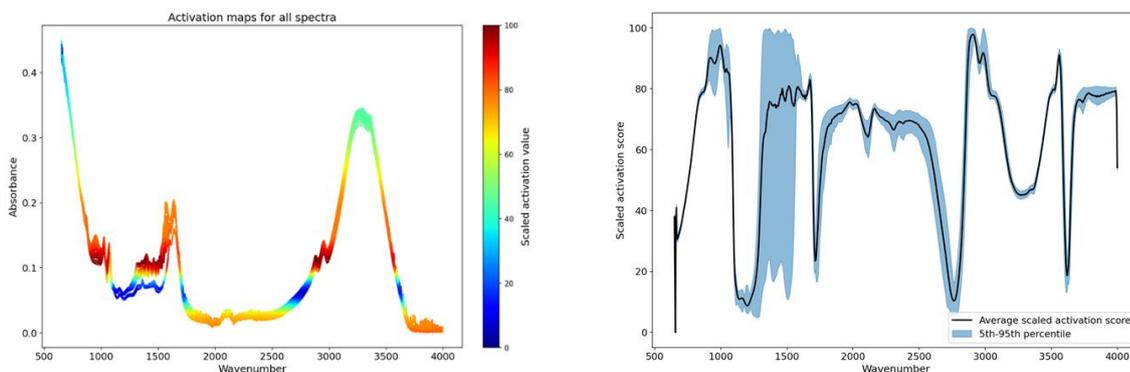


Figure 21. Left: activation maps; Right: average activation and 5<sup>th</sup>-95<sup>th</sup> percentile for all spectra. Wavenumbers given in  $\text{cm}^{-1}$

The results clearly indicate that ML methods are capable of extracting useful information from FTIR spectra of amine samples to indicate the level of degradation of the solvent. This is extremely useful as it will allow us to incorporate a more robust data analysis methodology to the online solvent monitoring tool under development, based on the use of the ATR-FTIR. This not only allows for analysing the composition of the solvent, but also its degradation level. In this sense, it can be used as a signal to trigger operator action for solvent management. Moreover, the use of “raw” absorbance data by the ML avoids the need to use the Agilent software. This circumvents the issues with data logging delays caused by that software.



## 4 Conclusions and future work

In this deliverable, we have explored the use of Big Data tools for improving the knowledge on solvent degradation, and particularly for identifying key degradation predictors. Initially, process and solvent analysis data from the ALIGN-CCUS RWE campaign with MEA were used. However, and despite working with data from the longest ever open MEA campaign, not enough *degradation data* was available. The attempts to analyse the process data did not lead to any new insights. There is a clear need to generate larger data sets on solvent degradation. However, doing that by solvent sampling and laboratory analysis would become prohibitively expensive.

A large data set on solvent composition can be generated by using a technique for online solvent monitoring. In this work, we describe the development of the hardware and methodology to use a mini-ATR-FTIR for this purpose. The final system configuration includes not only the analytical equipment itself (ATR-FTIR), but also a methodology for switching between solvent and water circulation loops. The innovative feature of the methodology developed is in the use of de-ionized water (DI) as the medium for taking backgrounds (instead of air). Equipping the system with automated pumps and valves for switching between solvent and DI water flow has the added benefit of ensuring a clean crystal, thus lowering the probability of measurement errors due to solvent accumulating on the crystal surface (amine solutions are “sticky”).

Despite all the progress on developing the methodology, the online test realized in a pilot plant using MDEA/PZ solvent gave erroneous measurements of solvent composition. The errors arise from software issues. In order to correct that, a more robust methodology for spectra analysis is required. In this work, we have tested different machine learning (ML) models for that end. The most successful models – convolutional neural networks – allow for the application of explainability methods which led to knowledge on key degradation predictors for MEA (wavenumbers 1350-1650  $\text{cm}^{-1}$  explain most of the degradation).

In conclusion, a robust methodology for generating online “Big Degradation Data” is needed, and the use of the mini-ATR seems like a promising route to achieve this goal. Once data is available, ML models can be used to interpret the spectra to derive information on solvent composition and degradation state. While the hardware side of the solution described seems sufficiently developed, the software side needs further attention.

Within LAUNCH, the ML models were only tested on MEA solutions derived from the degraded solvent from the RWE pilot. The different degradation levels were artificially generated by diluting a degraded solvent with fresh solvent. Ideally, the model should be used with spectra generated over time at different operational plants. For that, long-term Big Data sets from the ATR-FTIR is needed. Moreover, work with other solvents (CESAR1, MDEA/PZ blends) is also needed. This will prove the robustness of the methodology developed. The ML model predictions should be validated against offline solvent analysis, to validate the methodology for the different solvent tested.

Future development steps should also include investigating whether the FTIR data can be integrated with other real-time plant data (e.g. operating conditions) to predict degradation over time, and assess the impact of solvent management strategies.

## 5 References

- [1] S. A. Bedell, C. M. Worley, K. Darst, and K. Simmons, "Thermal and oxidative disproportionation in amine degradation-O<sub>2</sub> stoichiometry and mechanistic implications," *Int. J. Greenh. Gas Control*, vol. 5, no. 3, pp. 401–404, 2011, doi: 10.1016/j.ijggc.2010.03.005.
- [2] A. Z. Plantz and G. T. Rochelle, "Potassium Iodide as an Amine Oxidation Inhibitor in CO<sub>2</sub> Capture," *Ind. Eng. Chem. Res.*, vol. 61, no. 40, pp. 14910–14926, Oct. 2022, doi: 10.1021/acs.iecr.2c01571.
- [3] P. Moser *et al.*, "Results of the 18-month test with MEA at the post-combustion capture pilot plant at Niederaussem – new impetus to solvent management, emissions and dynamic behaviour," *Int. J. Greenh. Gas Control*, vol. 95, p. 102945, 2020, doi: <https://doi.org/10.1016/j.ijggc.2019.102945>.
- [4] S. Dhingra *et al.*, "Understanding and Modelling the Effect of Dissolved Metals on Solvent Degradation in Post Combustion CO<sub>2</sub> Capture Based on Pilot Plant Experience," *Energies*, vol. 10, no. 5, p. 629, May 2017, doi: 10.3390/en10050629.
- [5] A. Kachko *et al.*, "Real-Time Process Monitoring of CO<sub>2</sub> Capture by Aqueous AMP-PZ Using Chemometrics: Pilot Plant Demonstration," *Ind. Eng. Chem. Res.*, vol. 54, no. 21, pp. 5769–5776, Jun. 2015, doi: 10.1021/acs.iecr.5b00691.
- [6] A. C. van Eckveld, L. V van der Ham, L. F. G. Geers, L. J. P. van den Broeke, B. J. Boersma, and E. L. V Goetheer, "Online Monitoring of the Solvent and Absorbed Acid Gas Concentration in a CO<sub>2</sub> Capture Process Using Monoethanolamine," *Ind. Eng. Chem. Res.*, vol. 53, no. 13, pp. 5515–5523, Apr. 2014, doi: 10.1021/ie402310n.