

Deep learning for industrial processes: Forecasting amine emissions from a carbon capture plant

Kevin Maik Jablonka,^{†,⊥} Charithea Charalambous,^{‡,⊥} Eva Sanchez Fernandez,[¶]
Georg Wiechers,[§] Peter Moser,[§] Juliana Monteiro,^{||} Berend Smit,^{*,†} and
Susana Garcia^{*,‡}

[†]*Laboratory of Molecular Simulation, École Polytechnique Fédérale de Lausanne (EPFL), Rue de
l'Industrie 17, CH-1951 Sion, Switzerland*

[‡]*The Research Centre for Carbon Solutions (RCCS), School of Engineering and Physical Sciences,
Heriot-Watt University, EH14 4AS Edinburgh, United Kingdom*

[¶]*Solverlo Ltd, EH42 1TL Dunbar, United Kingdom*

[§]*RWE Power AG, Ernestinenstraße 60, 45141 Essen, Germany*

^{||}*TNO, Leeghwaterstraat 44, 2628 CA Delft, The Netherlands*

[⊥]*These authors contributed equally.*

E-mail: berend.smit@epfl.ch; s.garcia@hw.ac.uk

Abstract

One of the main environmental impacts of amine-based carbon capture processes is the emission of the solvent and degradation products into the atmosphere. To mimic the mounting importance of intermittent operations of power plants we performed a stress test in which we measured the amine emissions from a pilot plant that has been in operation using a mixture of amines (CESAR1) in a slipstream from a coal-fired

power plant. Understanding how changes in the operation far from the steady-state of the plant affect the emissions is key to designing emission mitigation strategies. However, conventional process modelling techniques struggle to capture the full dynamic, multivariate, and non-linear nature of this data. In this work, we report how a data-intensive approach can be used to learn the mapping between process and emissions from data. The resulting model can forecast the emissions, can be used to analyse the data and also perform *in silico* stress tests. By doing so, we reveal that emission mitigation strategies that work well for single component solvents (e.g. monoethanolamine) need to be revised for a mixture of solvents such as CESAR1. We expect that the combination of large amounts of data with flexible learning algorithms will impact the way we design and operate industrial processes, as we can now harvest information at conditions where conventional approaches fail.

Main

The design, control, and optimisation of industrial processes requires detailed knowledge of how process parameters interact and impact the operation of a plant. Due to the complexity of such plants, process models typically focus on capturing the steady-state operation.¹ However, there are many cases in which operation beyond the steady-state is required. For instance, the design and operation of current and future power plants will need to constantly adapt to the increased share of intermittent renewable energy generation.²⁻⁵ This requires tools that fully capture the dynamic and multivariate behaviour of the plant away from its steady-state operation. The classical analysis techniques, such as response function fits,^{6,7} or chemometrics approaches⁸ give some insights into the typical response to the different perturbations. However, these techniques cannot take the full multivariate, non-linear nature of the time-dependent behaviour of a complex plant into account. In this work, we show that data science methods that are typically used for dynamic pattern recognition and predictions of financial data can successfully be adapted

to forecast the performance of a plant given its current and historic behaviour, even if it is operated far from its steady-state conditions. These forecasts can subsequently be used to model scenarios, understand experimental observations, and perform control operations.

Here, we use such data science techniques to understand and forecast amine emissions from a solvent-based post-combustion carbon capture pilot plant under dynamic operation. The most well-known and broadly used benchmark solvent is monoethanolamine (MEA).⁹ Its main disadvantages are that a significant amount of energy is required for its regeneration¹⁰ and the comparatively high degradation rate. Research has focused on amine mixtures that have better energy performance, e.g., the CESAR1 solvent, which is a blend of 2-amino-2-methyl-1-propanol (AMP) and piperazine (Pz).^{11–14} Energy efficiency, however, is not the only criterion that is important in selecting a solvent for a carbon capture process. Amine emissions are equally important, as these may require cost-incurring gas treatment strategies to meet the operational permits and address environmental concerns.^{15,16} At present, we do not have a clear understanding of such amine emissions from a capture plant operating with these new solvent mixtures such as CESAR1.^{17,18}

Stress-test campaign

To mimic the intermittency expected for future power plant operation, we carried out a stress test on the pilot capture plant at Niederaussem. This plant has been operating on a slipstream of flue gas from a raw lignite-fired power plant^{19,20} with the CESAR1 solvent for over 12 months (see Extended Data Fig. 1 for a schematic flow diagram) and therefore provides an ideal real-life example of the difficulties of understanding amine emissions.⁷ This test involved a sequence of eight different scenarios (see Supplementary Note 1 for more details on the rationale of these scenarios). Extended Data Fig. 2 shows that these daily scenarios cause significantly higher emissions than with normal operating conditions. During the stress test, several interventions of the operators were required to ensure the safe operation of the plant. Such interventions make the subsequent data analysis very

challenging as these interventions need to be disentangled from the operational changes induced by the scenarios. Hence, we have a case in which we have a large amount of valuable experimental data, but where the complexity of the pilot plant operation does not afford any other conclusion that these emissions are problematic. In particular, we cannot draw any statistically relevant conclusions on which countermeasures could be most effective in reducing emissions. Of course, one can—to analyse the data with conventional techniques—always reduce the extent of the perturbations and allow for enough time for the plant to return to steady-state. This would make the campaign significantly more expensive and is of limited, if any, use to deal with the complexity of intermittency scenarios. Therefore, from an operational point of view, the lack of reliable forecasts implies that the de-risking and the permitting of the plant for this novel solvent would require large “engineering margins”, which would drastically increase the cost of the plant.

Modelling approach

We can build a forecasting model by representing the time-dependent process and emission data as an image. This representation allows us to use the most powerful machine-learning techniques for pattern recognition. In this representation, the state of the plant at a given time t defines a “state” feature vector with m elements representing the process variables (e.g., flue gas temperature, water wash temperature). If we take the state vectors of the plant for t timestamps we have a matrix of $t \times m$ entries which can be seen as an “image” that is connected to a future emission profile. We now have to link the pattern in the image of the history of the plant that leads to a particular future emission. For this, we have adapted a convolutional neural network, which is often used for computer vision. The idea of a convolution neural network is to learn predictive representations from an image in steps. In the first step we look at correlations between neighbouring parts of the image, and we then increase the length scale in each step by adding “holes” into the convolutional

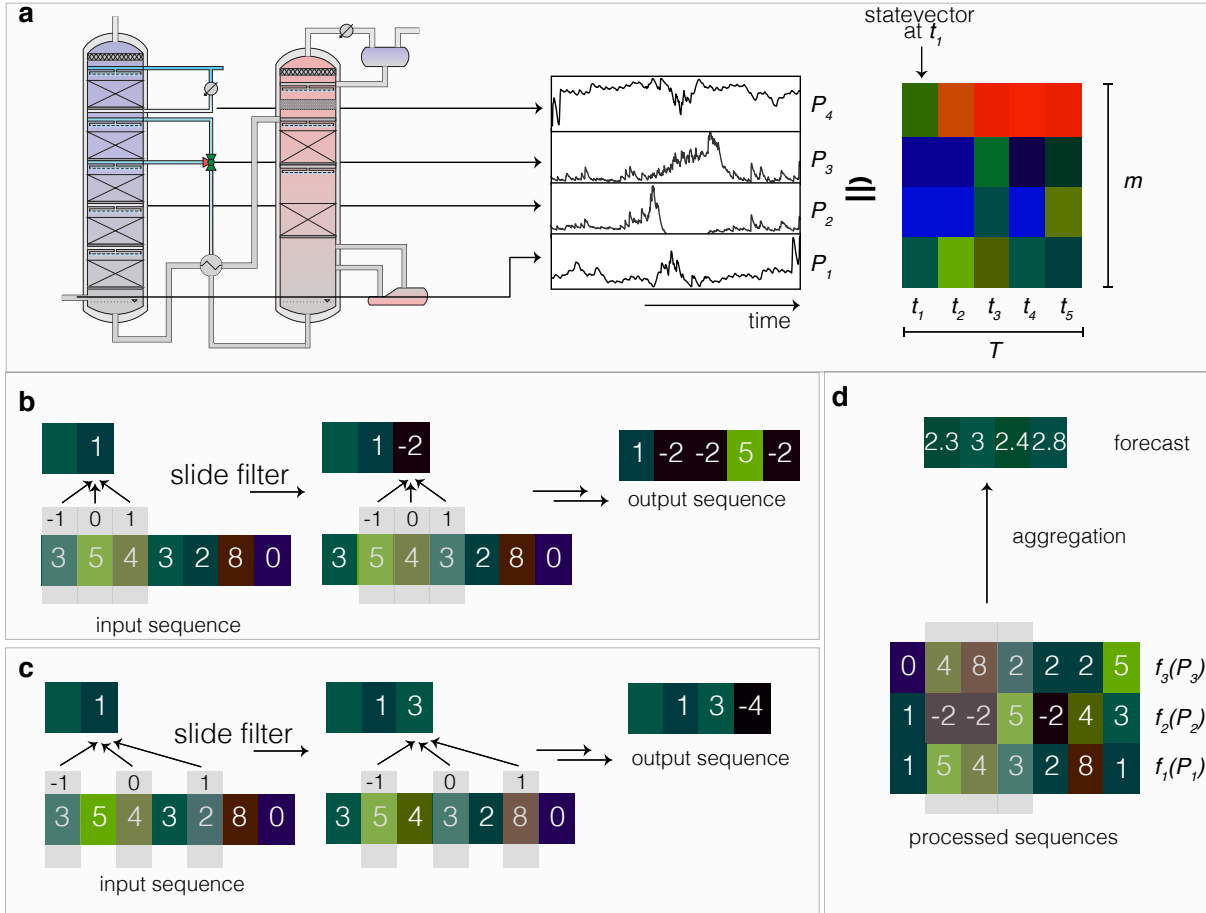


Fig. 1 | Schematic illustration of the modelling approach. **a**, Mapping of the data of the plant onto an image; the data set can be thought of an “image” with “width” equal to length of the input sequence (T) and “height” equal to the number of parameters, m . We represent with colours the value of parameter P_j at a time t_i . As the predictions should be invariant to the order of the rows, we only apply the pattern learning via convolutional filters (light grey) in the time direction, wherefore the image of the plant should be seen as m one-dimensional images. **b**, Convolutional kernels are slid over the m images as part of the pattern recognition algorithm. The weights of the kernels (of the f th filter), $W_f = [w_{-1}^f, w_0^f, w_1^f]$, are initially set randomly and learned in the training procedure. In the first layer, the kernel operates on directly neighbouring values. In order to allow for the model to learn different representations (patterns) we use 128 learnable filters per layer of the neural network,^{21,22} i.e., the layer outputs 128 one-dimensional “images”. In total, we apply eight layers and the output of one layer is fed into the next layer as an input. **c**, To allow learning of correlations across large time scales, as they are expected to be relevant in industrial processes, we add “holes” to the kernels (dilated convolutions) that operate on the output of preceding convolutional layers. **d**, The results of all the kernel operations (after applying operations of the forms of **a** and **b** eight times) are all collected via a “2D” convolution into a predicted emission. From this schematic, it follows that our output sequence cannot be longer than T , the length of the input sequence. To deal with the “edges”, we apply (causal) padding with zeros at the front of the input sequence (not shown in the figure).

kernels. As the emissions are independent of the way we order the m process variables, we apply this method only in the time direction of the image which gives us new time series in which relevant patterns are highlighted. In Fig. 1 we show how the training of such a sequence of convolution operations results in a prediction of the emissions.

In Fig. 2 we compare the measured AMP (a) and Pz (b) emissions with the forecasted emissions. For this figure, we trained the model on the data of the first four days (before red vertical line) and let the model predict the emissions for the following days together with the estimate of the uncertainty. We observe that the measured emissions are typically within our tolerance interval (shaded area) and that our model even correctly captures the spikes in the emission profile. We provide the predictions for CO₂ and NH₃ as well as numerical metrics and learning curves in Supplementary Note 8.

Causal impact analysis

The key motivation for performing our stress test campaign is to understand what changes to the plant have an impact on the emissions. This understanding is essential to identify those parameters that need to be tightly monitored and controlled to mitigate emissions. One of the difficulties of analysing the data is that conventional techniques cannot disentangle the effects of the scenarios from other responses of the plant that one would expect to happen even *without* any intervention. In statistics, the gold standard for answering such a question are control experiments for which we would need to operate a copy of the power plant. Similar problems exist, for example, in finance where one might want to measure the impact of some political intervention and where it is equally impossible to duplicate society for a control experiment. Interestingly, for these problems causal impact analysis²⁴ can be used to construct a so-called counterfactual baseline of the behaviour of the system *without* the intervention. For this, we trained our model on the “baseline” operation immediately preceding the dynamic experiments to forecast what the behaviour of the plant would have been without the intervention.

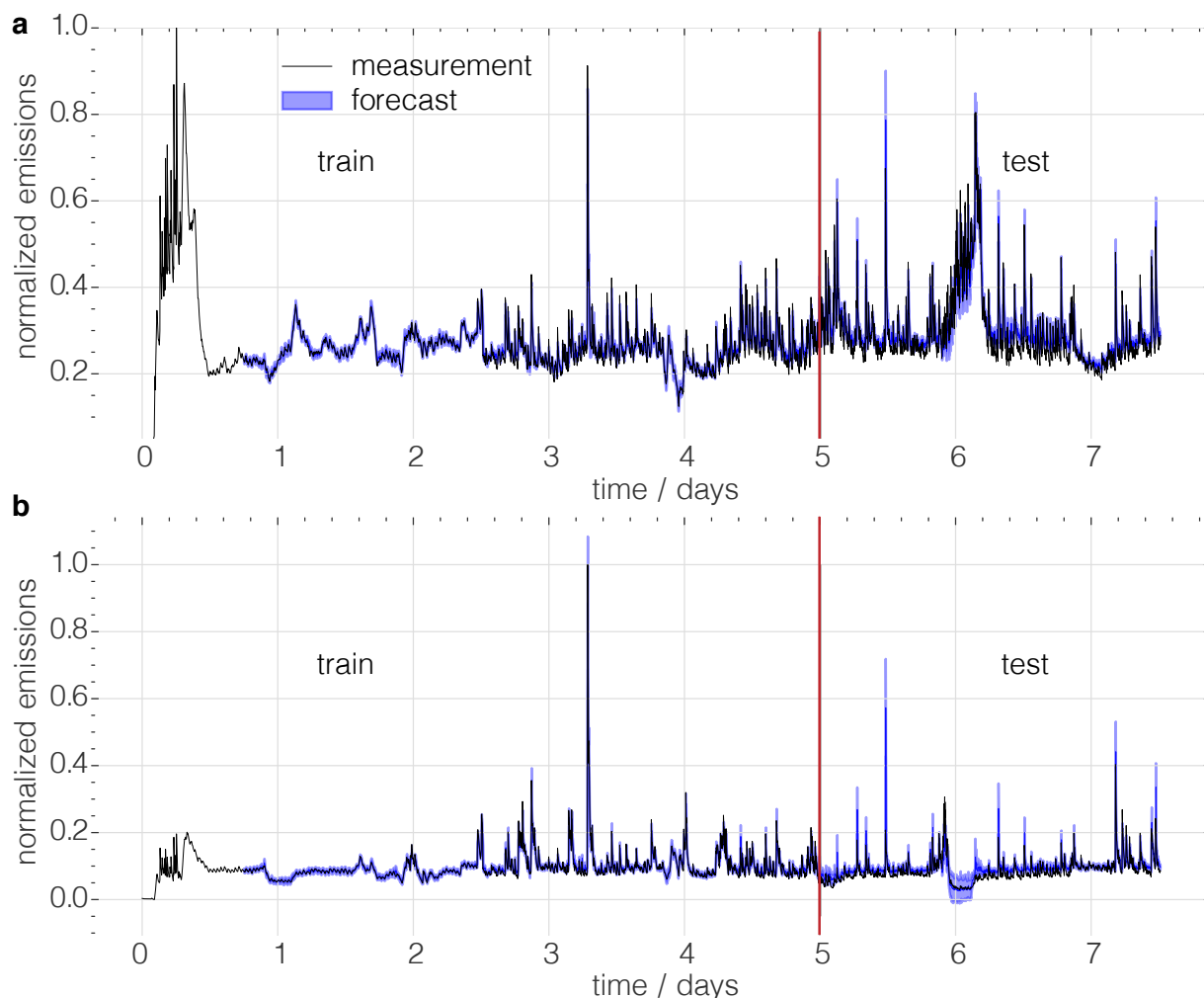


Fig. 2 | Amine emissions as predicted by the temporal convolutional neural network. To test the performance of the model for the amine emissions of AMP (a) and Pz (b), we trained the model on the first part of the data and tested the performance on the subsequent part. The split is indicated with the red vertical line. The gap without predictions is due to the fact that the model needs to be initialised with a part (in this case 10 %) of the sequence. The blue lines show so-called historical forecasts, which can be produced by an expanding window approach where the model is moved over the time series and the last forecasts (over 4 min) are used as the inputs of the subsequent ones. Shaded, blue areas indicate the 95 % tolerance intervals estimated via the standard deviation of 100 Monte-Carlo dropout²³ draws. The solid lines indicate the mean of the Monte-Carlo dropout draws.

In Fig. 3 we compare the measured emissions for three of the scenarios with the predicted baseline. We can see the importance of these baselines in Fig. 3 **a** and **b**. At the first black vertical line, the flue gas temperature was increased from 45 °C to 55 °C and put back to normal at the second vertical line. The measurements (black lines) suggest an increase in the emissions during and after the intervention. However, we find that this behaviour is strikingly similar to the prediction without the intervention. Interestingly, applying the same analysis for those scenarios that involved changes in the water wash flow rate or the ratio of solvent and water wash temperature (**c** and **d**) we observe a significant effect. In Supplementary Note 9 we show the analysis for all scenarios.

It is interesting that the causal impact analysis reduces this extremely complex emission behaviour (see Extended Data Fig. 2) into a surprisingly simple conclusion that controlling the water wash and lean solvent temperature as well as the water wash flow rate are the most promising handles for emission mitigation. It is important to note that without the counterfactual baseline, however, we would have concluded that many other interventions that show a change in emissions during the intervention are also good handles for emission control.

Emission mitigation

The causal impact analysis can give us insights into the significance and magnitude of effects for changes we actually performed on the plant. However, many other parameters were implicitly changed during the stress test. Using our model, we can use this data to investigate which changes to the operation of the plant would result in lower overall emissions during the stress test.

Fig. 4 shows the predicted cumulative change in amine emissions over the full campaign for the two sets of variables that caused some of the largest changes in our *in silico* experiments. In these *in silico* experiments we change the value of two parameters by a fixed percentage over the entire stress test, keeping also the dynamics unchanged, and

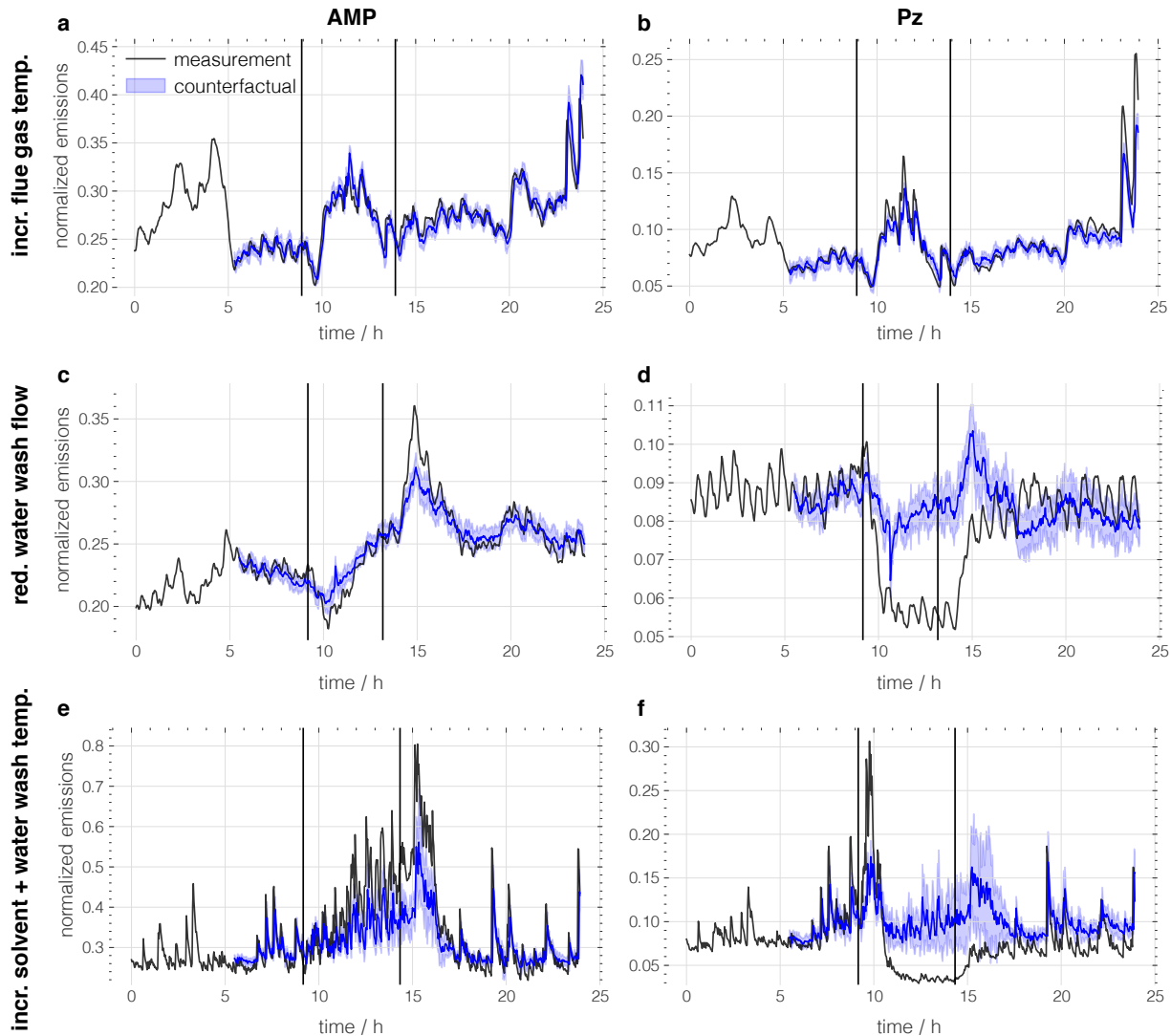


Fig. 3 | Causal impact analysis for three of our dynamic experiments. In causal impact analysis we use the machine learning model to predict what the emissions (**a, c, e** AMP emissions, and **b, d, f** Pz) were without intervention (blue). The difference between the prediction and the actual measurement (black) is the effect size. To avoid that the model is biased due to signals in covariates that are causally related ($p < 0.05$) to the intervention parameter, we automatically excluded all variables to which the intervention parameter is Granger causally related (see Supplementary Table 4). **a, b** Shows the measurement and predictions for the step increase in flue gas temperature. One can observe that also the counterfactual model forecasts an increase (incr.) in amine emissions. **c, d** Shows the effect of the reduction (red.) in water wash flow rate. One can observe that Pz (**d**), in contrast to AMP (**c**), shows a reduction in emission w.r.t. the baseline. **e, f** Shows the effect of an increase in water wash and lean solvent temperature. One can observe that Pz (**f**), in contrast to AMP (**e**), shows a reduction in emission w.r.t. the baseline. Shaded areas indicate the 95% tolerance intervals estimated via the standard deviation of 100 Monte-Carlo dropout draws. Black vertical lines indicate the start and end of the step change. Model trained in the period before the step change.

let our model predict the emissions. The heatmaps then show the difference with the measured emissions for which reason the centre (0,0) of the heatmaps is grey.

These figures point to the most important conclusion from our experimental campaign. Fig. 4a suggests that lower AMP emissions are obtained when operating at a lower water wash temperature and higher flue gas temperature. However, at these conditions, we do not have the minimum Pz emissions. On the other hand, minimum Pz emissions occurred at decreased temperature difference between the flue gas exiting the water wash section and the top of the absorber and increased low lean solvent temperature — at which conditions AMP emissions have increased. Similar conclusions can be drawn from the other scenarios (see Supplementary Note 10). These results suggest that Pz and AMP have different emission mechanisms. Indeed, from the literature²⁵ it is known that amine emissions can be volatile or in aerosol form. If one increases the temperature in the water wash section one would expect an increase in volatile emissions while a decrease in the lean solvent inlet temperature can create conditions that favour aerosol emissions.^{18,26} Therefore, we can conclude that in our stress test the aerosol mechanism is more relevant for Pz than for AMP. Because the two components in the CESAR1 mixture have different governing emission mechanisms, different mitigation strategies have opposite effects on the emission of the two components. Therefore, one needs to design the capture plant to be able to deal with both mechanisms. This is a more challenging task when considering blended solvents, like CESAR1, than single amine solvents. It is important to include these additional costs in the current discussion to replace MEA with CESAR1.

Conclusion

Even at steady-state, we would not have been able to develop a conventional process model to predict amine emissions from the carbon capture plant. For instance, we would need additional experiments as we lack relevant thermodynamic data of the amines and

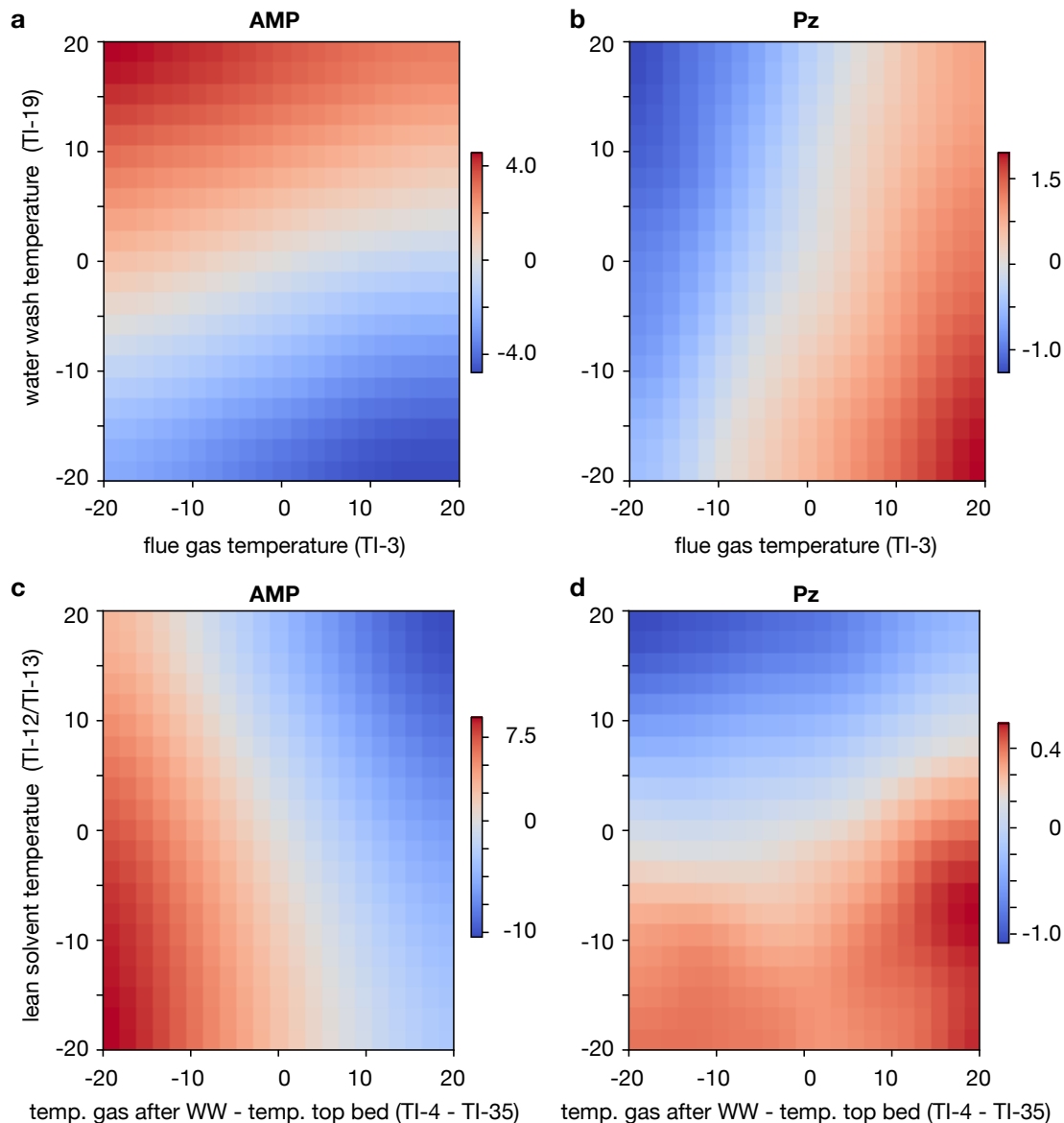


Fig. 4 | Predicted changes in emissions. Ordinate and abscissa show the relative change in the process variable. The colour indicates the cumulative change in normalised emissions over the full observation time compared to the actual emissions. To increase the reliability of the forecasts, we trained the model for this analysis on the complete dataset. Left column (a, c) shows predicted changes in emissions for 2-Aminopropanol (AMP). Right column (b, d) shows the predicted changes in emissions for piperazine (Pz). a, b plot the changes in emissions for changes in water wash (WW) temperature (temp.) and flue gas temperature. For AMP, the highest predicted reduction in emissions is for decreased water wash temperature and increased flue gas temperature whereas for Pz the highest predicted reductions in emissions are possible for increased water wash temperature and decreased flue gas temperature. c, d shows the predicted change in emissions for a change in of the difference in temperature of the flue gas and the top of the absorber bed and the gas exiting the water wash section and the temperature of the lean solvent. Here, decreasing the temperature difference and the increasing the lean solvent temperature leads to the lowest predicted emissions for Pz, but those conditions are not optimal for AMP.

understanding of the emission mechanisms. To make things worse, over the full stress test, the plant was far from steady-state. The current process models are too simple to deal with this complexity. In this work, we developed an alternative approach in which we start with the data and learn the mapping between the process and the emissions directly from the data. The resulting deep learning model allows us to not only forecast the emissions of the plant but also to gain insights in which parameters are key for emission mitigation. A similar approach can be used to forecast and understand other key performance parameters such as those related to the plant energy requirements.

Amine emissions in a stress test in a carbon capture plant is just one example of an industrial process for which a better understanding of its operation beyond its steady state is needed. Another example is the start-up of a plant during which one has to carry out many tests to identify the safe operational limits. These tests can take many months before a plant can be put in operation. Typically, during such a start-up phase or any other change to a new operating regime, there is a lot of data created and collected, but this data collection has outpaced our ability to sensibly analyse the data, let alone understand it. Our work shows that we could feed the data into an active learning model to harvest all the knowledge that has been collected during these experiments. Interrogation of this model can help us define the next most informative experiments^{27,28} which we expect to greatly reduce the time to operability.

Deep learning has the potential to make an even bigger impact in chemical and process engineering than it did in computer vision. In the case of computer vision, the basic features of an image that are learned by a model are often closely related to how we process an image with our eyes and brains. However, in an industrial plant, we often lack understanding of the relevant features, but with deep learning, we can discover the underlying rules of the mapping from the parameters to observables.

Methods

Pilot plant

Extended Data Fig. 1 gives a schematic process flow diagram of the capture plant at Niederaussem (Germany). The flue gas is supplied by a 965 MW_{el} raw lignite-fired power plant subjected to a state-of-the-art multistage electrostatic precipitator, a conventional wet limestone flue gas desulphurisation plant, and a direct contact cooler (DCC) located upstream of the absorber. The capture plant follows a conventional amine scrubbing process. The absorber column consists of four beds and is integrated with a flexible intercooling system and a water wash section. The flexible intercooler, which can be located either between the bottom and the second packing or between the second and third packing, controls the temperature rise in the absorber. A water wash section has been added to the pilot plant to reduce amine emissions to the atmosphere.^{29,30} The amine degradation, due to the presence of oxygen and other impurities like nitrogen oxides, as well as elevated temperatures during solvent regeneration, can result in other gaseous emissions of degradation compounds such as ammonia.³¹

The flue gas upstream of the absorber was analysed using a BA5000 Bühler infrared spectroscope. The CO₂-lean flue gas downstream of the water wash outlet was analysed using a GasMET CX/DX 4000 analyser (i.e., CO₂, CO, O₂, AMP, Pz, NH₃, and H₂O).

Stress test: Intermittency scenarios

As the baseline, we assume that the capture plant operates with the power plant at full-load, but that the intermittency associated with a future increase of renewables will cause regular changes of the load of the power plant. Variations in this load not only change the amount of flue gas that the capture plant has to process but can also change the amount of steam that is available for the capture plant. In the scenarios, we focus on those (combinations of) changes, of which our previous study on MEA¹⁹ has shown that they can impact the

emissions significantly. The time scale and the magnitude of the changes are based on the expected intermittency² and typical requirements of the grid services,^{5,32,33} respectively. A more detailed description is given in Supplementary Note 1.

Machine learning

To avoid overfitting and the exploitation of spurious correlations, the models were trained on a small feature set that was created using manual feature selection and engineering (see Supplementary Note 6). For all our modelling, we removed deterministic trend components from the data using linear regression, removed outliers using a z -score filter ($z = 3$), performed exponential window smoothing (window size 16 min) and downsampled the data to a frequency of 2 min. The impact of the preprocessing is shown in Supplementary Fig. 5. For use in the models, we also additionally standardised the data using min-max scaling.

Temporal convolutional neural networks

The most important advantage of using a neural network is that we do not limit our analysis to some functional form, which may bias the outcomes or might not be flexible enough. However, this full flexibility has the danger of overfitting the data. This overfitting is mitigated by using regularization techniques (such as dropout and multioutput training). To train the model on our dataset, we use a simple sliding window approach, where the model uses 80 min of data as input and then forecasts the next 4 min. Note that only the forecasting procedure, not the model architecture, needs to be changed to make forecasts over longer horizons (or to take correlations across longer times into account). A detailed description of the model architecture, the feature set, and the forecasting method can be found in Supplementary Note 8. Note that the forecasting performance decreases for longer horizons. Also, note that the predictive performance for CO₂ and NH₃ is worse than

for AMP and Pz. Some caveats of the modelling approach are discussed in Supplementary Note 11.

We used the darts library²¹ as a high-level wrapper for Pytorch,³⁴ to train temporal convolutional neural networks in which the convolutional layers are added with residual connections.^{21,22,35} We used 8 layers with 128 filters and a kernel size of 3. Between the layers, we dropout layers with a dropout probability of 0.3.³⁶ Additionally, we employed weight normalisation.³⁷ We trained the models with a batch size of 32 and a learning rate of 5×10^{-6} for 400 epochs using the Adam optimiser.³⁸

For estimation of epistemic uncertainty, we used the Monte-Carlo dropout scheme proposed by Gal and Ghahramani, i.e., activated the dropout layers also for inference.²³ Our models were trained to predict the emissions two time steps ahead and not re-trained on the inference dataset. Note that we also attempted the use of recurrent neural networks but could not achieve convincing forecasting performance. For this work, we trained two double-headed models: one for the amine emissions and another one for the ammonia and carbon dioxide emissions. This choice is motivated by the fact that the forecasting performance degraded when we trained the model to forecast all targets at the same time.

Causal impact analysis

For the causal impact analysis, we removed the features that Granger causal for the intervention variable ($p < 0.05$, determined using the statsmodels package,³⁹ considering a maximum lag of 10). This is of paramount importance such that the effect is not underestimated by some change in a covariate that is in causal relation to the intervention variable. The covariates we considered for the separate analyses as listed in Supplementary Table 4. Note that this step automatically reduces the learning capacity of the models, i.e., we expect the forecasts to be less accurate. We trained the TCN models on the baseline operation before the step change and also use the baseline operation immediately preceding the step change to start the model. Due to the smaller training set size, we reduced the model size

for this analysis (input length of 60 min, output length of 4 min), used a dropout rate of 0.3, four layers, 128 filters, kernel size of three (and trained for 200 epochs with a learning rate of 1×10^{-3}). We also attempted to use Bayesian structured timeseries models as in the original implementation of the causal impact analysis technique²⁴ but found much better predictive performance with TCN models. We also performed the analysis with the R package and found qualitative agreement, e.g., finding no significant effect for the change in flue gas temperature but significant effects for the changes in water wash flows and temperatures.

We made use of the following Python⁴⁰ libraries: pandas,⁴¹ sklearn,⁴² scipy,⁴³ statsmodels,³⁹ matplotlib,⁴⁴ jupyter,⁴⁵ numpy,⁴⁶ pytorch,³⁴ darts,²¹ lightgbm,⁴⁷ shap.⁴⁸

Conflicts of interest

There are no conflicts to declare.

Author contributions

K.M.J. built the machine learning framework with input from all authors under supervision of B.S. C.C., E.S.F., J.M., G.W., P.M., and S.G. contributed to feature engineering and interpretation of the results.

Data availability

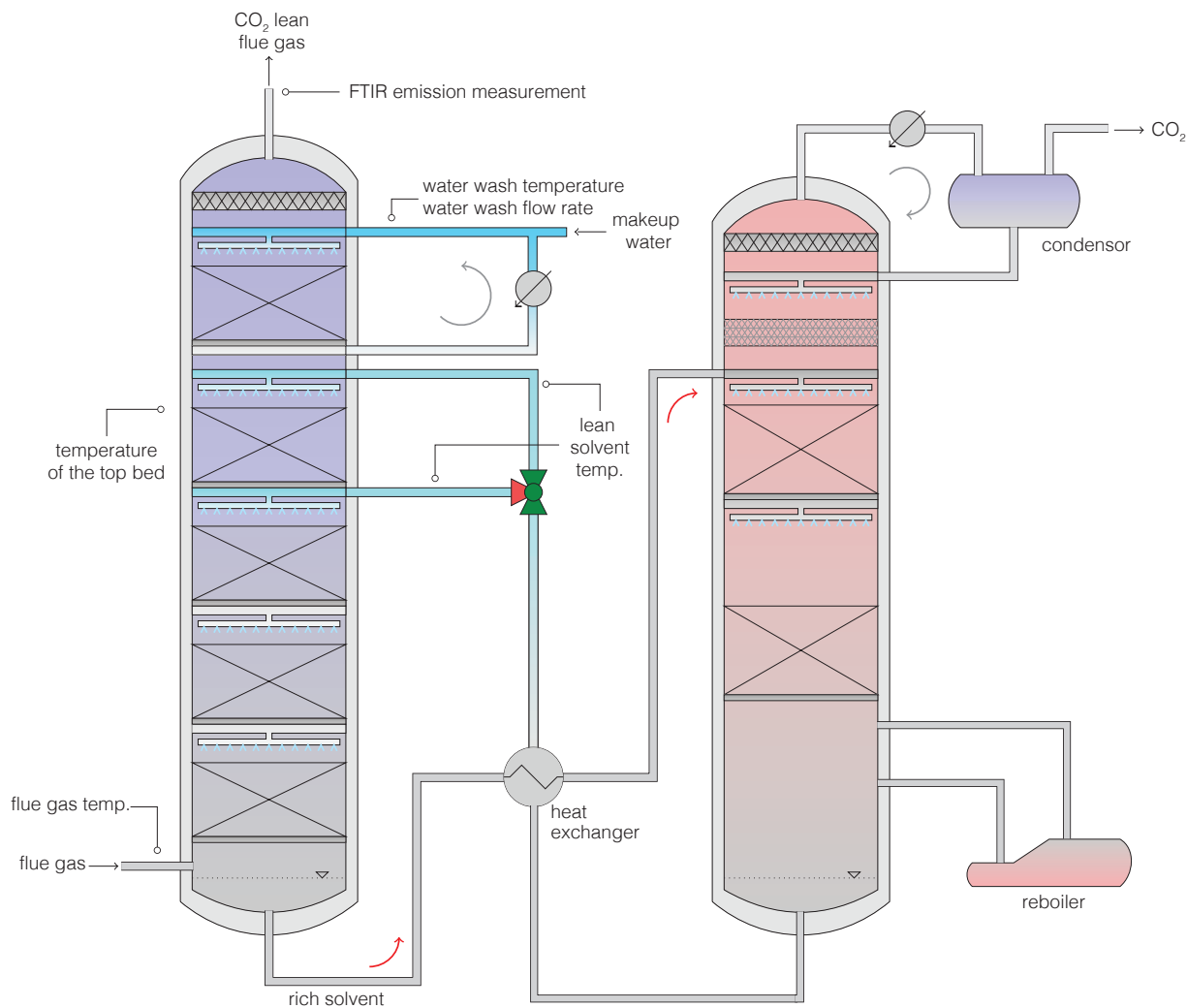
The raw data (emissions and process parameters) and model checkpoints are archived on Zenodo (DOI: 10.5281/zenodo.5153417).

Code availability

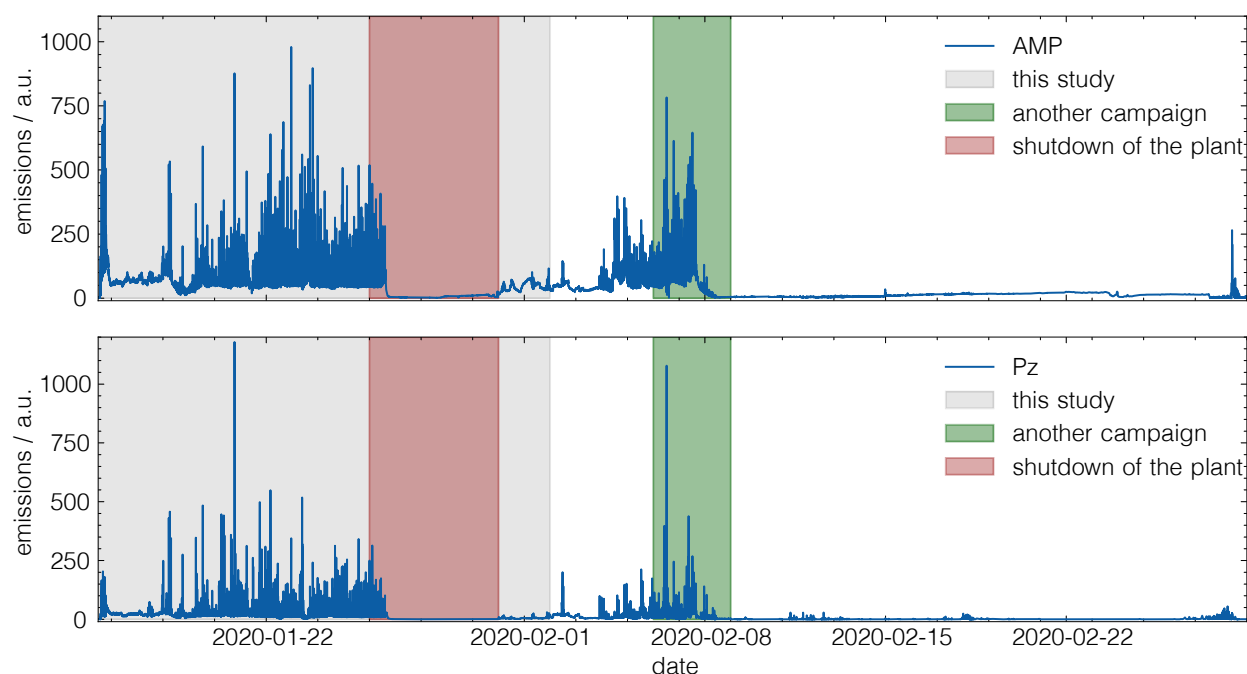
The code for our analysis is available on GitHub (github.com/kjappelbaum/pyprocessta) and archived on Zenodo (DOI: 10.5281/zenodo.5222270).

Acknowledgements

The authors would like to acknowledge the ACT ALIGN-CCUS Project (No 271501) and the ACT PrISMa Project (No 299659). The ALIGN-CCUS project has received funding from RVO (NL), FZJ/PtJ (DE), Gassnova (NOR), UEFISCDI (RO), BEIS (UK) and is cofunded by the European Commission under the Horizon 2020 programme ACT, Grant Agreement No 691712; www.alignccus.eu. The PrISMa Project is funded through the ACT programme (Accelerating CCS Technologies, Horizon2020 Project No 294766). Financial contributions made from: BEIS together with extra funding from NERC and EPSRC, UK; RCN, Norway; SFOE, Switzerland and US-DOE, USA, are gratefully acknowledged. Additional financial support from TOTAL and Equinor, is also gratefully acknowledged. The responsibility for the contents of this publication rests with the authors. Parts of the calculations were enabled by the Swiss National Supercomputing Centre (CSCS), under project ID s1019. K.M.J. was supported by the Swiss National Science Foundation (SNSF) under Grant 200021_172759 and thanks Brian Yoo for feedback on a draft of this manuscript.



Extended Data Fig. 1 | Simplified process flow diagram of the PCC pilot plant at Niederaussem. The position of the process parameters discussed in the main text are indicated in the figure. A complete P&ID diagram can be found in Supplementary Fig. 2.



Extended Data Fig. 2 | Amine emissions during, and after the stress test. Time frame of the stress test highlighted in grey. The power plant was shut down from 25–30 January (red region), which explains the very low emissions around that time. In this work, we only used the data generated before the shutdown of the plant. In the period of 6–8 February (grey region) other experiments were carried out at the pilot plant, but these were not part of our test. This figure shows that applying the different scenarios cause the plant to emit much more compared to its steady state operations. A preliminary analysis of the data has been reported in Charalambous et al.⁷

References

- (1) Biegler, L. *Systematic methods of chemical process design*; Prentice Hall PTR: Upper Saddle River, N.J, 1997.
- (2) IEAGHG, *Valuing Flexibility in CCS Power Plants*; 2017; available at https://ieaghg.org/exco_docs/2017-09.pdf.
- (3) Flø, N. E.; Kvamsdal, H. M.; Hillestad, M. Dynamic simulation of post-combustion CO₂ capture for flexible operation of the Brindisi pilot plant. *International Journal of Greenhouse Gas Control* **2016**, *48*, 204–215.
- (4) Gaspar, J.; Jorgensen, J. B.; Fosbol, P. L. Control of a post-combustion CO₂ capture plant during process start-up and load variations. *IFAC-PapersOnLine* **2015**, *48*, 580–585.
- (5) Chalmers, H.; Leach, M.; Lucquiaud, M.; Gibbins, J. Valuing flexible operation of power plants with CO₂ capture. *Energy Procedia* **2009**, *1*, 4289–4296.
- (6) Flø, N. E.; Kvamsdal, H. M.; Hillestad, M.; Mejdell, T. Dominating dynamics of the post-combustion CO₂ absorption process. *Comput. Chem. Eng.* **2016**, *86*, 171–183.
- (7) Charalambous, C.; Saleh, A.; van der Spek, M.; Wiechers, G.; Moser, P.; Huizinga, A.; Gravesteijn, P.; Ros, J.; Monteiro, J. G. M.-S.; Goetheer, E.; Garcia, S. Analysis of flexible operation of CO₂ capture plants: Predicting solvent emissions from conventional and advanced amine systems. *SSRN Electronic Journal* **2021**,
- (8) Kachko, A.; van der Ham, L. V.; Geers, L. F. G.; Huizinga, A.; Rieder, A.; Abu-Zahra, M. R. M.; Vlugt, T. J. H.; Goetheer, E. L. V. Real-Time Process Monitoring of CO₂ Capture by Aqueous AMP-PZ Using Chemometrics: Pilot Plant Demonstration. *Industrial & Engineering Chemistry Research* **2015**, *54*, 5769–5776.

- (9) Bui, M. et al. Carbon capture and storage (CCS): the way forward. *Energy & Environmental Science* **2018**, *11*, 1062–1176.
- (10) Moser, P.; Wiechers, G.; Schmidt, S.; Garcia Moretz-Sohn Monteiro, J.; Charalambous, C.; Garcia, S.; Sanchez Fernandez, E. 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. *International Journal of Greenhouse Gas Control* **2020**, *95*, 102945.
- (11) Rochelle, G.; Chen, E.; Freeman, S.; Wagener, D. V.; Xu, Q.; Voice, A. Aqueous piperazine as the new standard for CO₂ capture technology. *Chemical Engineering Journal* **2011**, *171*, 725–733.
- (12) Mangalapally, H. P.; Hasse, H. Pilot plant study of two new solvents for post combustion carbon dioxide capture by reactive absorption and comparison to monoethanolamine. *Chemical Engineering Science* **2011**, *66*, 5512–5522.
- (13) Feron, P. H.; Cousins, A.; Jiang, K.; Zhai, R.; Garcia, M. An update of the benchmark post-combustion CO₂-capture technology. *Fuel* **2020**, *273*, 117776.
- (14) Garcia, S.; Weir, H.; der Spek, M. V.; Saleh, A.; Charalambous, C.; Sanchez-Fernandez, E.; Ros, J.; Skylogianni, E.; Monteiro, J. G. M.-S.; Wiechers, G.; Moser, P.; Kvamsdal, H. M.; Haugen, G.; Eldrup, N. H.; Skagestad, R. Process Integration of Advanced Amine-based Solvents in Power and Industrial Plants: A New Benchmark for Post-combustion Carbon Capture? *SSRN Electronic Journal* **2021**,
- (15) Reynolds, A. J.; Verheyen, T. V.; Adeloju, S. B.; Meuleman, E.; Feron, P. Towards Commercial Scale Postcombustion Capture of CO₂ with Monoethanolamine Solvent: Key Considerations for Solvent Management and Environmental Impacts. *Environmental Science & Technology* **2012**, *46*, 3643–3654.

- (16) Veltman, K.; Singh, B.; Hertwich, E. G. Human and Environmental Impact Assessment of Postcombustion CO₂ Capture Focusing on Emissions from Amine-Based Scrubbing Solvents to Air. *Environmental Science & Technology* **2010**, *44*, 1496–1502.
- (17) IEA Greenhouse Gas R&D Programme, ENVIRONMENTAL IMPACTS OF AMINE EMISSION DURING POST COMBUSTION CAPTURE. 2010; available at <https://www.globalccsinstitute.com/archive/hub/publications/106171/environmental-impacts-amine-emissions-post-combustion-capture.pdf>.
- (18) Khakharia, P.; Mertens, J.; Abu-Zahra, M.; Vlugt, T.; Goetheer, E. *Absorption-Based Post-combustion Capture of Carbon Dioxide*; Elsevier, 2016; pp 465–485.
- (19) Moser, P.; Schmidt, S.; Sieder, G.; Garcia, H.; Stoffregen, T. Performance of MEA in a long-term test at the post-combustion capture pilot plant in Niederaussem. *International Journal of Greenhouse Gas Control* **2011**, *5*, 620–627.
- (20) Moser, P.; Schmidt, S.; Stahl, K. Investigation of trace elements in the inlet and outlet streams of a MEA-based post-combustion capture process Results from the test programme at the Niederaussem pilot plant. *Energy Procedia*. 2011; pp 473–479.
- (21) Unit8 SA, darts. 2021; <https://github.com/unit8co/darts/>.
- (22) Bai, S.; Kolter, J. Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv 1803.01271* **2018**,
- (23) Gal, Y.; Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *Proceedings of The 33rd International Conference on Machine Learning*. New York, New York, USA, 2016; pp 1050–1059.
- (24) Brodersen, K. H.; Gallusser, F.; Koehler, J.; Remy, N.; Scott, S. L. Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics* **2015**, *9*, 247–274.

- (25) Mertens, J.; Lepaumier, H.; Desagher, D.; Thielens, M.-L. Understanding ethanolamine (MEA) and ammonia emissions from amine based post combustion carbon capture: Lessons learned from field tests. *International Journal of Greenhouse Gas Control* **2013**, *13*, 72–77.
- (26) Khakharia, P.; Brachert, L.; Mertens, J.; Anderlohr, C.; Huizinga, A.; Fernandez, E. S.; Schallert, B.; Schaber, K.; Vlugt, T. J.; Goetheer, E. Understanding aerosol based emissions in a Post Combustion CO₂ Capture process: Parameter testing and mechanisms. *International Journal of Greenhouse Gas Control* **2015**, *34*, 63–74.
- (27) Jablonka, K. M.; Jothiappan, G. M.; Wang, S.; Smit, B.; Yoo, B. Bias free multiobjective active learning for materials design and discovery. *Nature Communications* **2021**, *12*.
- (28) Lookman, T.; Balachandran, P. V.; Xue, D.; Yuan, R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials* **2019**, *5*.
- (29) Moser, P.; Schmidt, S.; Stahl, K.; Vorberg, G.; Lozano, G. A.; Stoffregen, T.; Rösler, F. Demonstrating emission reduction - Results from the post-combustion capture pilot plant at Niederaussem. *Energy Procedia* **2014**, *63*, 902–910.
- (30) Rieder, A.; Dhingra, S.; Khakharia, P.; Zangrilli, L.; Schallert, B.; Irons, R.; Unterberger, S.; Van Os, P.; Goetheer, E. Understanding Solvent Degradation: A Study from Three Different Pilot Plants within the OCTAVIUS Project. *Energy Procedia*. 2017; pp 1195–1209.
- (31) da Silva, E. F.; Hoff, K. A.; Booth, A. Emissions from CO₂ capture plants; an overview. *Energy Procedia* **2013**, *37*, 784–790.
- (32) Bui, M.; Gunawan, I.; Verheyen, V.; Feron, P.; Meuleman, E.; Adeloju, S. Dynamic modelling and optimisation of flexible operation in post-combustion CO₂ capture plants - A review. *Computers & Chemical Engineering* **2014**, *61*, 245–265.

- (33) Bui, M.; Flø, N. E.; de Cazenove, T.; Dowell, N. M. Demonstrating flexible operation of the Technology Centre Mongstad (TCM) CO₂ capture plant. *International Journal of Greenhouse Gas Control* **2020**, *93*, 102879.
- (34) Paszke, A. et al. In *Advances in Neural Information Processing Systems 32*; Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R., Eds.; Curran Associates, Inc., 2019; pp 8024–8035.
- (35) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv 1512.03385* **2015**,
- (36) Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* **2014**, *15*, 1929–1958.
- (37) Salimans, T.; Kingma, D. P. Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks. *arXiv 1602.07868* **2016**,
- (38) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv 1412.6980* **2017**,
- (39) Seabold, S.; Perktold, J. statsmodels: Econometric and statistical modeling with python. 9th Python in Science Conference. 2010.
- (40) Van Rossum, G.; Drake, F. L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, 2009.
- (41) McKinney, W., et al. Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference. 2010; pp 51–56.
- (42) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V., et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research* **2011**, *12*, 2825–2830.

- (43) Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272.
- (44) Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in science & engineering* **2007**, *9*, 90–95.
- (45) Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B. E.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J. B.; Grout, J.; Corlay, S., et al. *Jupyter Notebooks-a publishing format for reproducible computational workflows.*; 2016; Vol. 2016.
- (46) Harris, C. R. et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362.
- (47) Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems* **2017**, *30*, 3146–3154.
- (48) Lundberg, S. M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J. M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* **2020**, *2*, 2522–5839.