UC Berkeley
UC Berkeley Previously Published Works

Title

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

Permalink

https://escholarship.org/uc/item/3cd037qh

Journal

Science Advances, 9(1)

ISSN

2375-2548

Authors

Jablonka, Kevin Maik
Charalambous, Charithea
Sanchez Fernandez, Eva
et al.

Publication Date

2023-01-06

DOI

10.1126/sciadv.adc9576

Peer reviewed

APPLIED SCIENCES AND ENGINEERING

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

Kevin Maik Jablonka[1]†, Charithea Charalambous[2]†, Eva Sanchez Fernandez[3], Georg Wiechers[4], Juliana Monteiro[5], Peter Moser[4], Berend Smit[1]*, Susana Garcia[2]*

One of the main environmental impacts of amine-based carbon capture processes is the emission of the solvent into the atmosphere. To understand how these emissions are affected by the intermittent operation of a power plant, we performed stress tests on a plant operating with a mixture of two amines, 2-amino-2-methyl-1-propanol and piperazine (CESAR1). To forecast the emissions and model the impact of interventions, we developed a machine learning model. Our model showed that some interventions have opposite effects on the emissions of the components of the solvent. Thus, mitigation strategies required for capture plants operating on a single component solvent (e.g., monoethanolamine) need to be reconsidered if operated using a mixture of amines. Amine emissions from a solvent-based carbon capture plant are an example of a process that is too complex to be described by conventional process models. We, therefore, expect that our approach can be more generally applied.

INTRODUCTION

The most well-known and broadly used benchmark solvent to capture CO_2 is monoethanolamine (MEA) (*1*). 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 (*2*, *3*). At present, we do not have a clear understanding of these amine emissions from a capture plant operating with these new solvent mixtures such as CESAR1 (*4*, *5*).

Amine emissions from carbon capture plants are one example of an industrial process for which the plant's design, control, and optimization require detailed knowledge of how the process parameters interact and affect the operation of the plant and what the (chemical) mechanisms and rate constants are. Because of the complexity of such plants, process models typically focus on capturing the steady-state operation (*6*). 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 (*7*, *8*, *9*, *10*). This requires tools that fully capture the dynamic and multivariate behavior of the plant away from its steady-state operation. The classical analysis techniques, such as response function fits (*11*, *12*), or chemometrics approaches (*13*) give some insights into the typical response to the different perturbations. However, these techniques cannot take the full multivariate, nonlinear nature of the time-dependent behavior of a complex plant into account. In addition, conventional causal analysis techniques cannot be used without an understanding of the mechanisms (i.e., the causal graph) (*14*) or additional experiments,

which interpretation, however, is not trivial because one cannot easily compare to a baseline (i.e., the behavior of the plant under the same environmental and solvent conditions without a particular change).

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 (in real time) given its current and historic behavior, even if it is operated far from its steady-state conditions, without a detailed understanding of the underlying process. These forecasts can subsequently be used to model potential emission mitigation scenarios and to understand experimental observations.

Experimental campaign

To mimic the intermittency expected for the operation of future power plants, we carried out an experimental campaign that involved a series of stress tests on the pilot capture plant at Niederaußem. Because of its size and the fact that it has been operating on a slipstream of flue gas from a raw lignite-fired power plant (*15*, *16*) with CESAR1 solvent for more than 12 months (see Fig. 1 for a schematic flow diagram), it provides an ideal real-life example of the difficulties of understanding amine emissions (*12*). These stress tests were based on eight different scenarios of how intermittency can affect the operation and hence the amine emission of the capture plant (see note S1 for more details on the rationale of these scenarios).

Figure 2 shows that the sequence of stress tests causes emissions substantially higher than those under normal operating conditions. The other interesting observation is that 2-amino-2-methyl-1-propanol (AMP) and piperazine (Pz) have different emission profiles. Such a campaign gives us a wealth of experimental data on the behavior of a capture plant. These data would even be more valuable if we were able to use them for quantitative predictions on future emissions. However, we cannot even make qualitative predictions. For example, during most of the stress tests, interventions of the operators were required to ensure the safe operation of the plant. These interventions make even a qualitative interpretation of the data a

[1]Laboratory of Molecular Simulation (LSMO), École Polytechnique Fédérale de Lausanne (EPFL), Sion, Switzerland. [2]The Research Centre for Carbon Solutions (RCCS), School of Engineering and Physical Sciences, Heriot-Watt University, EH14 4AS Edinburgh, UK. [3]Solverlo Ltd, EH42 1TL Dunbar, UK. [4]RWE Power AG, Ernestinenstraße 60, 45141 Essen, Germany. [5]TNO, Leeghwaterstraat 44, 2628 CA Delft, Netherlands.
*Corresponding author. Email: berend.smit@epfl.ch (B.S.); s.garcia@hw.ac.uk (S.G.)
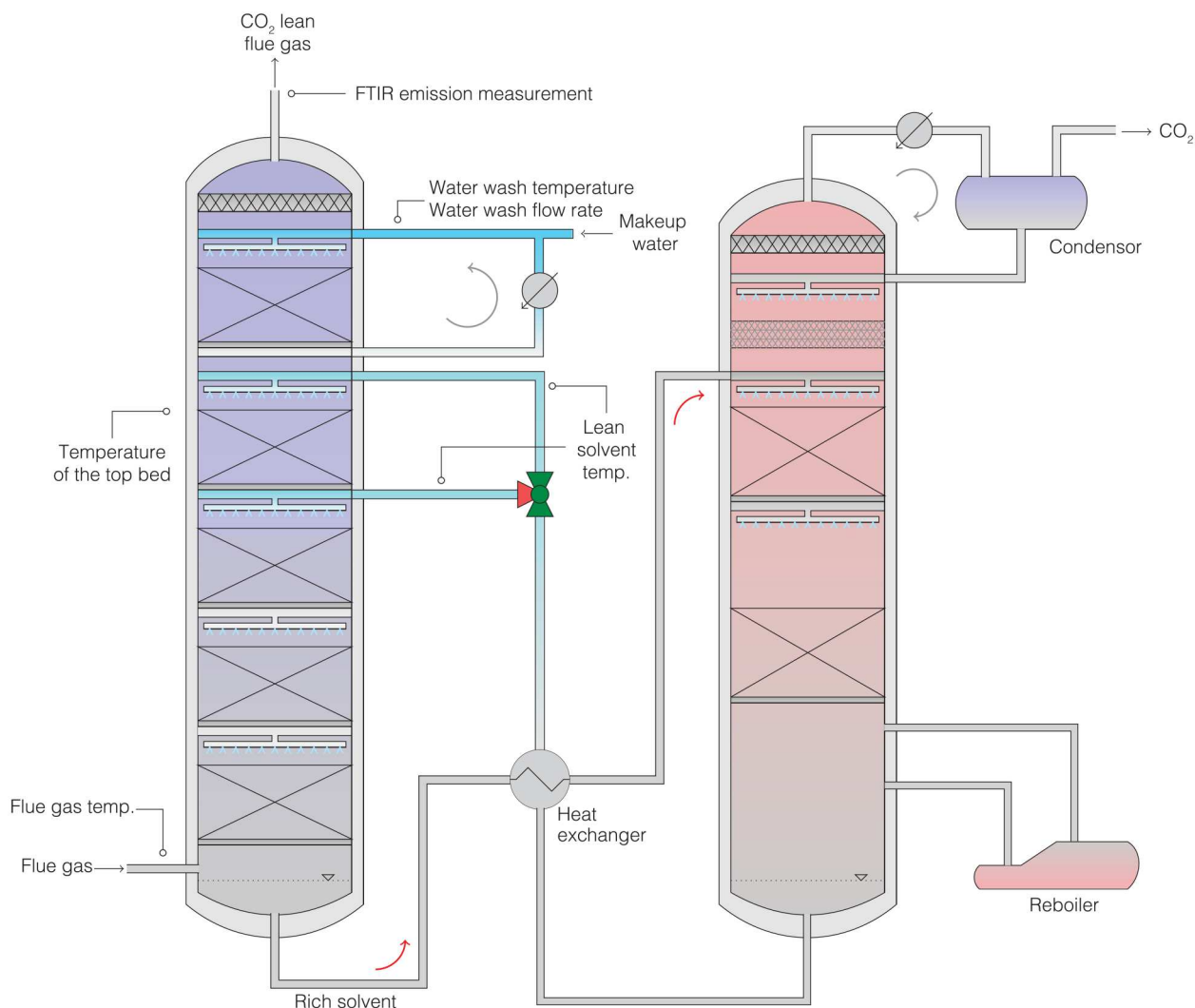†These authors contributed equally to this work.

Fig. 1. Simplified process flow diagram of the postcombustion carbon capture pilot plant at Niederaußem. The plant uses a slip stream from the coal-fired power plant. The positions of the process parameters discussed in the main text are indicated in the figure. A complete piping and instrumentation diagram can be found in fig. S2. FTIR, Fourier transform infrared.

challenge, as we cannot disentangle the effects of these interventions from the operational changes induced by the stress test.

Therefore, we have a case in which we have a large amount of valuable experimental data, but where the complexity of the operation of the pilot plant does not allow any other conclusion that these emissions are problematic. In particular, we cannot draw any statistically relevant conclusions on why our stress tests caused such a marked increase in emissions and which countermeasures we could take to reduce emissions.

Machine learning model

During the experimental campaign, data were taken every minute. This provides us with a large dataset. Such a dataset allows us to use data science methods and develop a machine learning model to analyze the data. In this section, we summarize the main features of our approach; for details, we refer to Methods and the Supplementary Materials. Our machine learning approach is based on the observation that we can build a forecasting model by thinking

of the time-dependent process and emission data as an image (i.e., matrix of data; see Fig. 3). 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 $x(t)$ with p elements representing the process variables (e.g., flue gas temperature and water wash temperature). If we take the state vectors of the plant for t timestamps, we have a matrix of $t \times p$ entries, which can be seen as an "image" that is connected to a future emission profile, $y(t)$.

The next step is to link the pattern in the image of the history of the plant to a particular future emission. For this, we have adopted a gradient-boosted decision tree (17, 18) model that is trained on a feature vector of concatenated historic data of process parameters and emissions (i.e., we combine the rows, characterizing the different parameters and emissions, into a long vector). We train these models using quantile loss (19, 20) to obtain uncertainty estimates. We have also adopted a temporal convolutional neural network with Monte Carlo dropout for uncertainty estimation and show
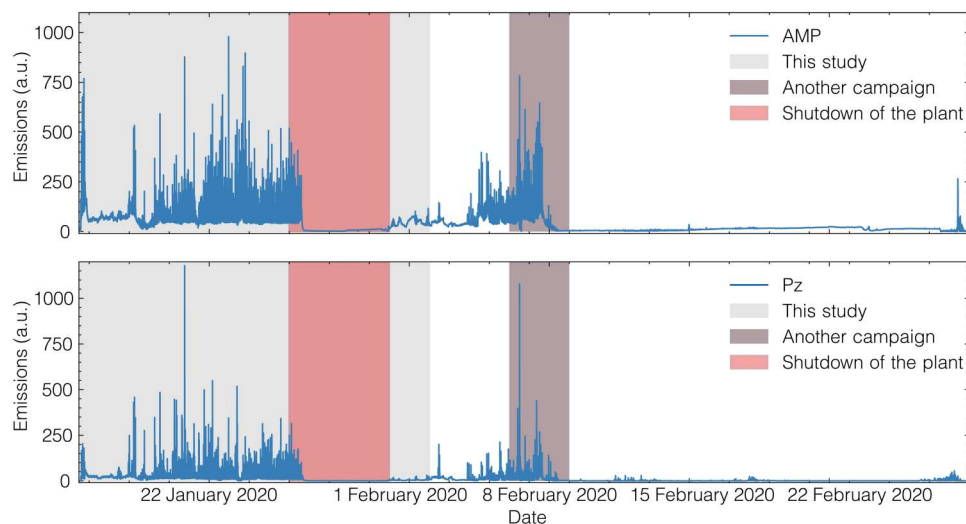
Fig. 2. Amine emissions during and after the experimental campaign. Time frame of the stress tests is highlighted in gray. The power plant was shut down from 25 to 30 January (red region), which explains the very low emissions around that time. In this work, we only used the data generated before the shutdown of the plant. In the period of 6 to 8 February (gray region) other experiments were carried out at the pilot plant, but these were not part of our campaign. This figure shows that applying the different scenarios cause the plant to emit much more compared to its steady-state operation. A preliminary analysis of the data has been reported by Charalambous *et al.* (*12*). a.u., arbitrary units.

results (equivalent to those obtained with the gradient-boosted decision tree) obtained with this model in note S8.

Insights into amine emissions from machine learning

We apply our machine learning model for different purposes, and each of them requires us to forecast the emissions but each with a different aim and time horizon:

1) (Real-time) prediction of future emissions: The aim here is to predict what the emissions x hours in the future will be given the historic and current operation and emissions.

2) Causal impact analysis of the data: To measure the impact of a particular stress test on the amine emissions, we need a reference, i.e., a baseline that gives us the emissions that would occur without the changes directly induced by the stress test. Without this baseline, it is impossible to correctly quantify the effect of the different stress tests on the observed emissions.

3) Emissions mitigation: To understand and identify how we can mitigate emissions, we use our model to predict emissions in "what-if" scenarios. For example, we predict how the overall emissions would change if we ran the entire experimental campaign with a lower temperature of the water wash section.

In the next sections, we show how we use our machine learning model to forecast amine emissions for the three different purposes mentioned above. The basic model architecture we use is the same; however, the way we apply and train the model for the different ways of forecasting is different.

Prediction of future emissions

The machine learning model that we introduced in the previous section takes some historic data to predict future emissions. For example, we use a sequence of input data (e.g., 2 hours) and predict the emissions, say, 10 min, 1 hour, or 2 hours in the future. For doing so, we use a sliding window; for the next

prediction, we update the input sequence with the observed emissions (see Fig. 4).

The model can be used for making predictions for any time horizon; however, one can expect the accuracy to decrease for longer time horizons compared to shorter ones. To quantify the accuracy of our prediction, we use the data that we have not used in our training (and validation) set. One has to be careful in making this comparison. Our machine learning model makes predictions on the likely emissions, given the plant data preceding these predictions, and in the testing step, we use the measured data in the test set. However, this validation is overly pessimistic with respect to potential real-world application because the validation and test set contain, by construction, step changes that have not been seen in the training set. In addition, these stress tests are designed to take the plant outside normal operations. Also, the moment such a stress test will be applied has no logical relation with the historic data of the plant and hence cannot be learned. This makes our validation overly pessimistic, as we are rather testing how well our learning extends to very extreme conditions in the stress test.

Causal impact analysis

The key motivation for performing our experimental campaign is to understand what changes to the plant have a significant impact on the amine emissions. This understanding is essential to identify those parameters that need to be tightly monitored and controlled to mitigate emissions. In statistics, the gold standard for answering such a question requires control experiments (*21*) to establish a baseline. At present, such a baseline is impossible to obtain. As the pilot plant receives the flue gas from a commercially operated coal-fired power plant, it is impossible to precisely reproduce the varying conditions of the plant. For this, one would need to run two identical pilot plants at the same time.

Similar problems exist, for example, in finance where one might want to measure the impact of some political intervention and
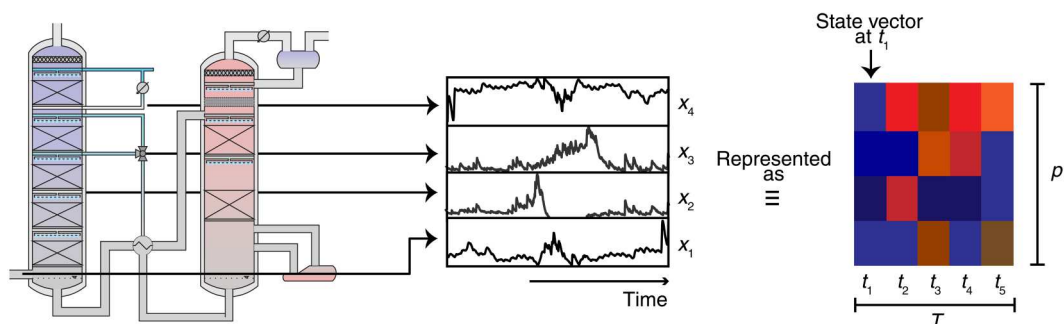
Fig. 3. Schematic illustration of the data representation. The dataset can be thought of an "image" with "width" equal to length of the input sequence (T) and "height" equal to the number of parameters, p. We represent with colors the value of parameter x_j at a time t_i. We then use a machine learning model to learn how this image characterizing the history, and current state of the plant is connected to its future emissions.

where it is equally impossible to duplicate society for a control experiment. For these problems, causal impact analysis (*22*) can be used to construct a so-called counterfactual baseline of the behavior of the system without the intervention. For this, we use our machine learning model to "rerun" the campaign but without the stress tests. The fact that we now can obtain a reasonable performance baseline is one of the major technical insights of our approach.

Let us assume that we have a perturbation on variable $x_2(t)$, e.g., we apply a step change for variable $x_2(t)$ at times $t \in [t_{\text{startstep}}; t_{\text{endstep}}]$. To obtain a prediction of the baseline, we then train our model on the training data but without any input from $x_2(t)$ [i.e., we also remove (Granger) causally related features]. We then have a model that predicts the emissions worse than if $x_2(t)$ was included in the training (as fewer features are used as input for the model), but it does give us our best prediction for the normal operation of the plant irrespective of the actual value of $x_2(t)$. This is the best approximation of the baseline operation we are interested in. Similarly, we train our model for all other variables that are changed during

the different stress tests conducted in the experimental campaign. We then use each of these models to predict the baseline (see Fig. 5).

Emission mitigation
To shed light on how we can reduce the overall amine emissions during, for instance, a given experimental campaign, we have used our model to run what-if scenarios. These scenarios were inspired by the outcome of the causal impact analysis, which highlighted some of the variables that affected the emissions the most. An example of such a scenario could be: "What-if we run the entire stress test with an increase in variable $x_2(t)$ of 10%." For this scenario, we replace in the input of our model $x_2(t) \rightarrow 1.1 x_2(t)$ (see Fig. 6) to compute the predicted emissions $y'(t)$. We can then compute the change in total emissions from

$$\text{rel. emission change} = 100 \times \frac{\int dt \, [y'(t) - y(t)]}{\int dt \, y(t)}$$
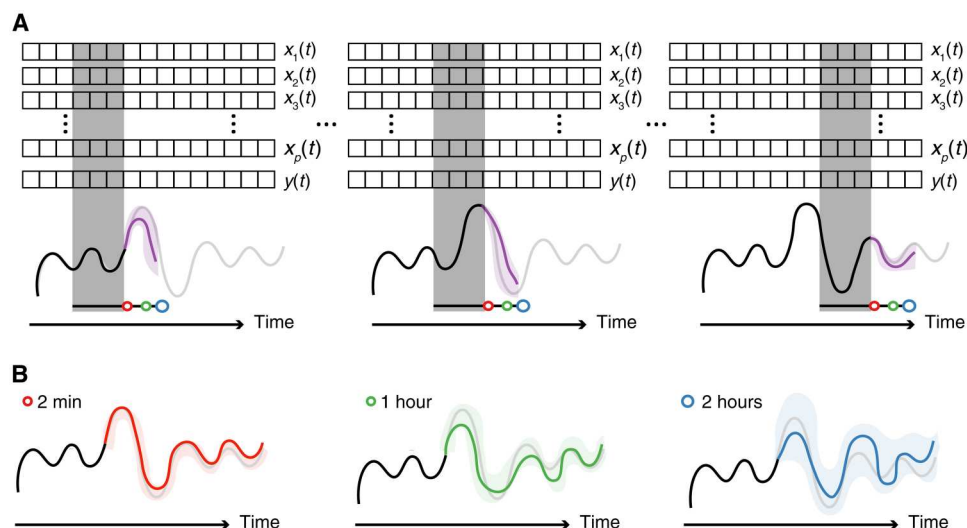

Fig. 4. Predicting future emissions. In this figure, $x_i(t)$ represents the input data of the plant (e.g., temperatures, pressures, etc. in different parts of the plant) and $y(t)$ the emissions. The gray box represents the data used by the model to predict the future emissions. The black curve represents the measured past emissions and the gray curve the measured future emissions. The purple curve represents the "real-time" predicted future emissions and the shaded purple area the uncertainty of the predictions. We mimic these real-time predictions by sliding our gray box over the data, i.e., the measurements of the current time are added, and the oldest data are no longer seen by the window, and we make a new prediction. In the bottom figure, we collect the predictions for the different time horizons (2 min, red; 1 hour, green; and 2 hours, blue). For the training of the model, we use the first half of the dataset.
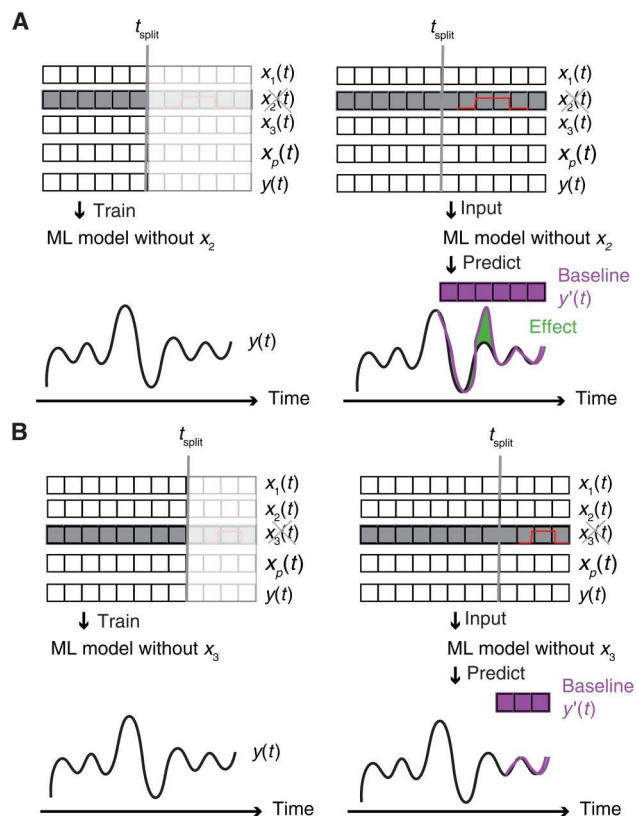
Fig. 5. Causal impact analysis. The left column shows the training of the model, in which we use all the data preceding a particular step change to train our model on predicting the capture plant's performance. In this example, we make a step change of variable $x_2(t)$ for example (**A**), and we train a model without variable $x_2(t)$. For example (**B**), we make a step change in variable $x_3(t)$ and hence train another model without variable $x_3(t)$. In this calculation, we have assumed that the other variables are not causally related with $x_2(t)$ and $x_3(t)$, respectively; if there is a causal relation, then these variables also need to be removed. The right column shows how we compute the baseline, the step changes are indicated by the vertical lines. The black curve gives the actual plant data, $y(t)$, obtained from the experimental campaign. The violet curves, $y'(t)$ give the machine learning predictions of normal operation without the stress test, i.e., the baseline. The predictions show that the x_2 step change test caused a real reduction in emissions, whereas the change in x_3 showed no effect.

To compute these scenarios, we need to predict the emissions $y(t)$ given an input $\mathbf{x}(t)$. For doing so, we retrain the model using all available data from the experimental campaign to ensure the highest possible accuracy from our model. For a more detailed discussion, see also note S11.

RESULTS AND DISCUSSION
Prediction of future emissions
In Fig. 7, we compare the measured AMP and Pz emissions with the predicted emissions for different forecasting horizons. For the short-horizon predictions (top row), we observe that the measured emissions are typically within our prediction interval (shaded area) and that our model even correctly captures the spikes in the emission profile [AMP mean absolute percentage error (MAPE) of 2.4% and overall percentage error (OPE) of 0.38%; Pz MAPE of 4.3% and
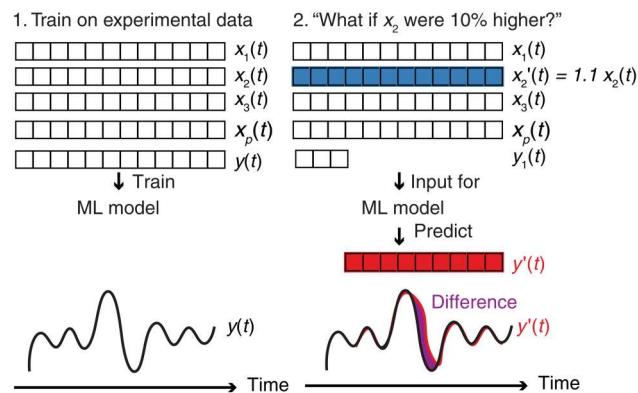

Fig. 6. Emission mitigation. To predict the effects of a given variable on the total emissions, we train our machine learning model on the entire dataset (left). We then use this model to run a what-if scenario, i.e., to use the model trained in the first step to predict the emissions $y'(t)$ (red) for this changed input. For example, what are the emissions if we replace input $x_2(t)$ (blue array) by, say, $x'_2(t) = \alpha_2(t)$ (here, $\alpha_2 = 1.1$). We can then calculate the difference between the actual measured emissions $y(t)$ and the predictions $y'(t)$ (green). If we perform this for different α, we can estimate and plot the change in emissions as a function of α.

OPE of 2.0%; see note S7]. We can also make predictions for a longer time horizon. For 1- to 2-hour windows, we can correctly forecast the trends, but as expected, we lose accuracy on the events (such as spikes) that happen on a short time scale (AMP MAPE of 9.5% and OPE of 3.8%; Pz MAPE of 21% and OPE of 10%). It is interesting to zoom in on some of the areas where our predictions deviate significantly from the actual measurements. These deviations are associated with a stress test that was not seen in the training, yet it is encouraging to see that our model did learn something as we do predict the trends. In addition, our model indicates at those conditions a very large uncertainty, which is exactly how the stress test is designed, to take the plan far outside normal operations. It is therefore very encouraging that our model recognizes this and that this is correctly reflected in the uncertainties.

Of course, a stress test is far from ideal to test our model to make (real-time) predictions, but these results do indicate that our model, if applied under normal operating conditions, can be used to make predictions about the emissions on a 2-hour window, which does give the operators a window to take actions if emissions are predicted to exceed specification limits.

Causal impact analysis
The first step in our analysis of the experimental data is computing the baseline for all stress tests. In Fig. 8, we compare the measured emissions for three of the stress tests with the predicted baseline (which we predicted with the model architecture we validated in Fig. 7). We can see the importance of these baselines in Fig. 8 (C and D). At the first black vertical line, the lean solvent temperature was increased from 43° to 52°C and put back to normal at the second vertical line. The measurements (black lines) suggest an increase in emissions during and after the intervention. However, we find that this behavior is notably similar to the prediction without the intervention (within the prediction interval) for AMP. Applying the same analysis to those stress tests that involved changes in the water wash flow rate or the solvent and water wash temperature
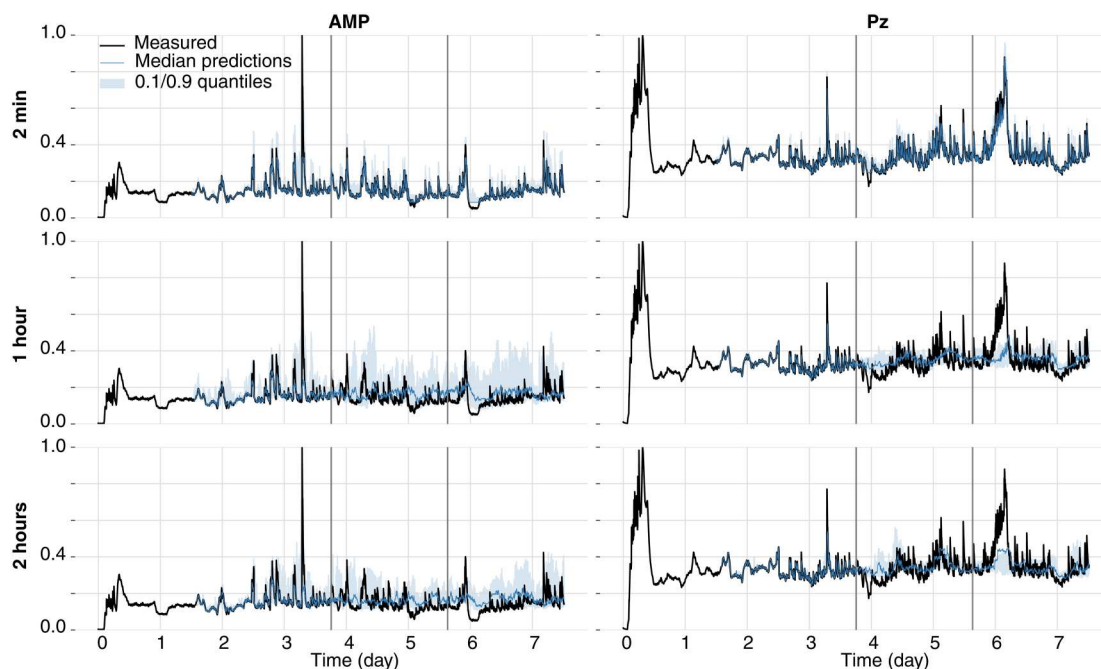
Fig. 7. Amine emissions as predicted by the machine learning model. To test the performance of the model for the amine emissions of AMP (left) and Pz (right), we trained the model on the first part of the data, used a subsequent part for hyperparameter search, and tested the performance on the final part. The splits are indicated with gray vertical lines. The gap without predictions is due to the fact that the model needs to be initialized with a part 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 we simulate what the predictions would have been if one used the model with the forecasting horizon with an updated dataset (i.e., the model sees the actual emissions for making forecasts and does not have to use its predictions, but we do not retrain the model). In the rows, we show the predictions for different forecasting horizons and we can observe, as one would expect, that the predictions for shorter forecasting horizons are better than for longer ones. The shaded areas fill the range between the 10 and 90% quantiles.

(Fig. 8, A, B, E, and F), we observe a significant effect. In note S10, we show the analysis for all interventions investigated in our campaign.

It is interesting that the causal impact analysis reduces this extremely complex emission behavior (see Fig. 2) into an unexpectedly simple conclusion that controlling the water wash and solvent temperature as well as the water wash flow rate are the most promising handles for emission mitigation. However, without the counterfactual baseline, we would have concluded that many other interventions that show a change in emissions during the intervention are also good handles for emission control. This shows how machine learning techniques can be used to extract insights from complex experimental datasets that remained opaque to conventional approaches.

Emission mitigation

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

Figure 9 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 let our model predict the emissions. The heatmaps

then show the difference with the measured emissions for which reason the center (0,0) of the heatmaps is gray.

These figures point to the most important conclusion from our experimental campaign. Figure 9A suggests that lower AMP emissions are obtained when operating at a lower solvent temperature. However, under these conditions, we do not have the minimum Pz emissions. On the other hand, minimum Pz emissions are predicted for increased lean solvent temperature and increased temperature at the top bed, under which conditions AMP emissions are predicted to increase. Similar conclusions can be drawn from the other scenarios (see note S11). These results suggest that Pz and AMP have different emission mechanisms. If volatility were the only mechanism, one would expect the amine emissions to increase with increasing solvent temperatures. This is what we observe for AMP. Because AMP is more volatile, the AMP partial pressure throughout the column is expected to be around two orders of magnitude higher than that of Pz (23). One would not expect significant emissions of Pz if volatility were the only mechanism. However, one can also have emissions through aerosols (24). These aerosol emissions are thought to be related to supersaturation in the column, which can be caused by a temperature bulge in the column profile that can be influenced by a change in lean solvent temperature (5, 25). Absorbed in these aerosol droplets, Pz and AMP are forming nonvolatile carbamates, and (pure component) studies have shown that the kinetics of this reaction is much faster for Pz (23). Moreover, because of steric hindrance, the AMP carbamates are short-lived and AMP is present as a protonated species in equilibrium with the free amine (26). This leads to a situation in which there is a
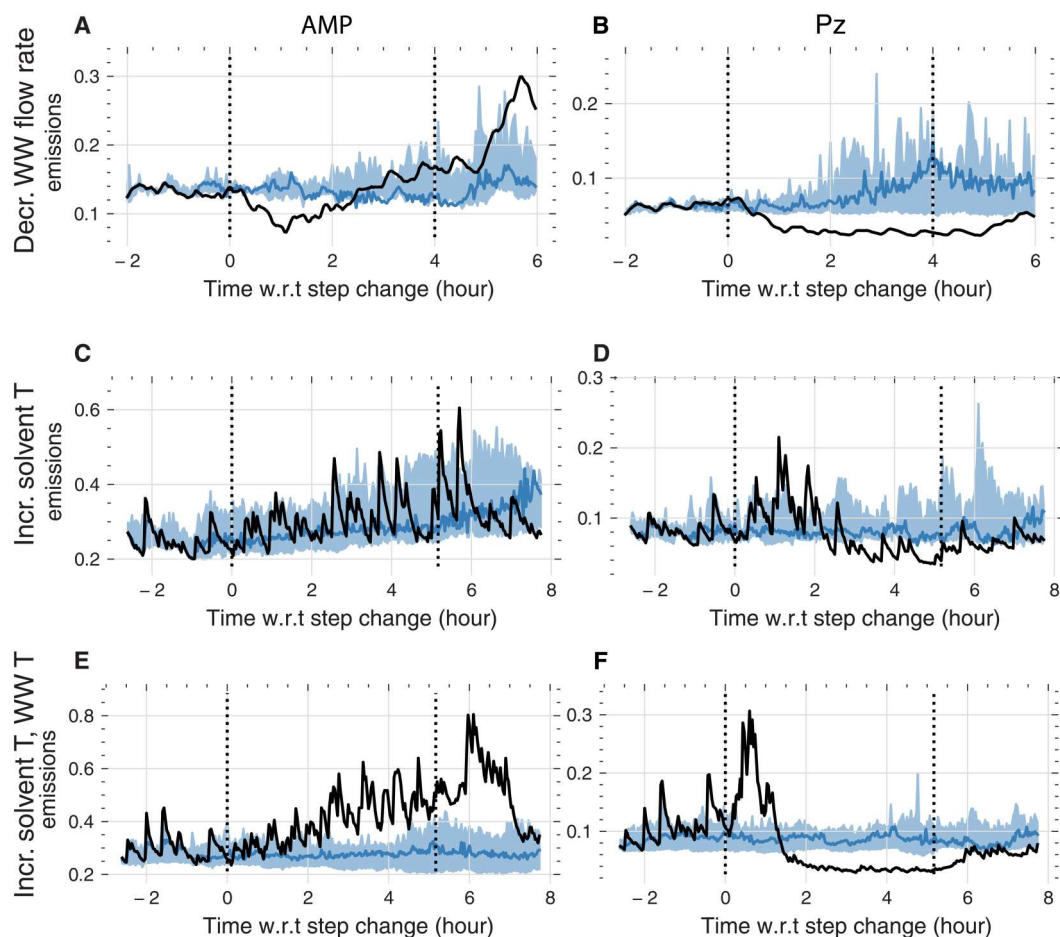
Fig. 8. Causal impact analysis for three of our dynamic experiments. In causal impact analysis, we use the machine learning model to predict what the emissions [(**A**), (**C**), and (**E**) AMP emissions; and (**B**), (**D**), and (**F**) Pz] were without intervention (blue). The difference between the prediction and the actual measurement (black) is the effect size. That is, if we observe no difference between the measurement (black) and prediction (blue), then there is no effect. (A) and (B) show the measurement and predictions for the step decrease (decr.) in water wash (WW) flow rate. One can observe that also the counterfactual model forecasts an increase (incr.) in amine emissions compared to the actual observations. (C) and (D) show the effect of the increase (incr.) in lean solvent temperature. One can observe that Pz (D), in contrast to AMP (C), shows a significant reduction in emission with respect to the baseline. (E) and (F) show 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 cover the area between the 0.05th and 0.95th percentiles. Dotted vertical lines indicate the start and end of the step change.

back-pressure buildup that hinders further AMP absorption in aerosol particles, which is not the case for Pz. Hence, one would expect the aerosol mechanism to be more relevant for Pz emissions, and we conclude that in our stress test, the aerosol mechanism seems 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, such as CESAR1, than single amine solvents. It is important to include these additional costs in the current discussion to replace conventional MEA-based capture plants with those based on more advanced solvent systems such as CESAR1. Even though we could not have derived this insight without our machine learning–based analysis, further experiments will be needed for a more detailed understanding of the causal mechanisms as our current model can only highlight predictive correlations.

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 on the amines and an understanding of the emission mechanisms. To make things worse, over the experimental campaign, 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 machine learning model allows us to not only forecast (in real time) the emissions of the plant but also gain insights into 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 from a carbon capture plant are just one example of an industrial process for which a better understanding
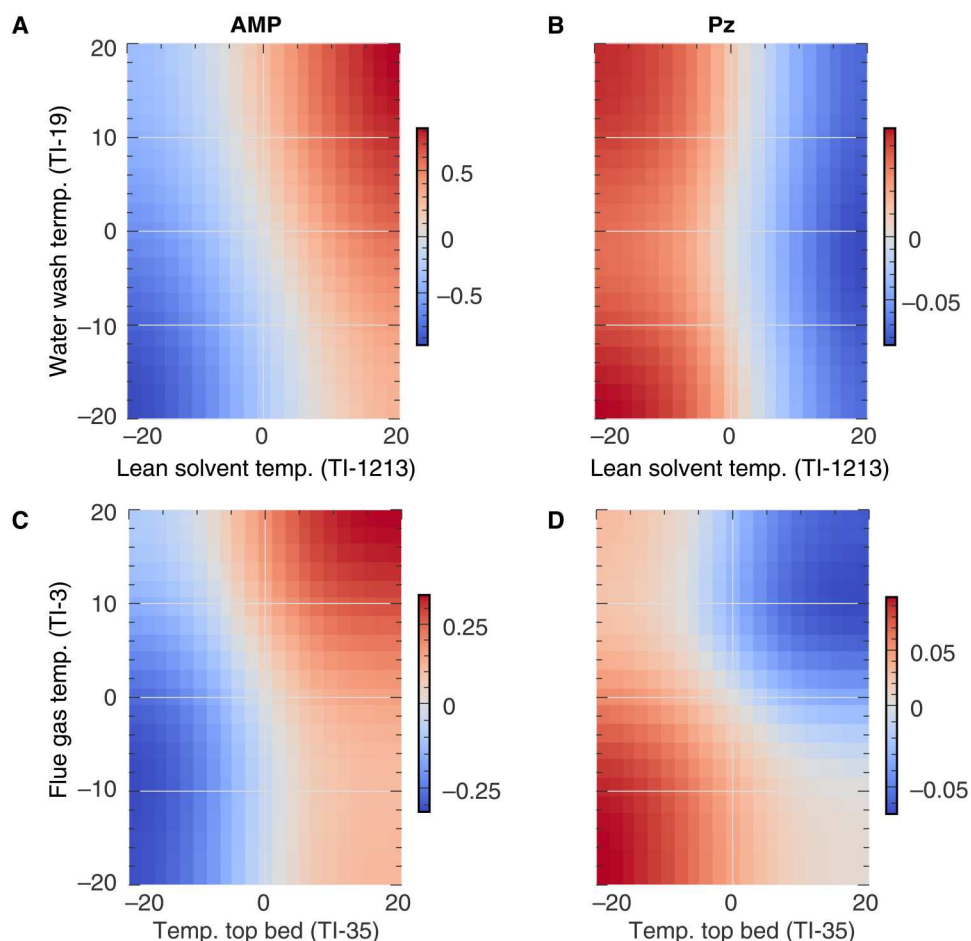
Fig. 9. Predicted changes in emissions. Ordinate and abscissa show the relative change in the process variable (in percent). The color indicates the cumulative change in normalized emissions over the full observation time compared to the actual emissions (i.e., not absolute emissions). To increase the reliability of the forecasts, we trained the model for this analysis on the complete dataset and use a short output sequence length. Left column (**A** and **C**) shows predicted changes in emissions for 2-amino-2-methyl-1-propanol (AMP). Right column (**B** and **D**) shows the predicted changes in emissions for Pz. (A) and (B) plot the changes in emissions for changes in water wash temperature (temp.) and lean solvent temperature. For AMP, the highest predicted reduction in emissions is for decreased water wash temperature and decreased temperature of the lean solvent, whereas for Pz, the highest predicted reductions in emissions are possible for increased solvent temperature. (C) and (D) show the predicted change in emissions for a change in the temperature of the flue gas upstream of the adsorber column and the temperature of the top bed. Here, decreasing the temperature at the top bed and the temperature of the flue gas yields the lowest AMP emissions but high predicted Pz emissions.

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 safe operational limits. These tests can take many months before a plant can be put into 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 analyze 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 experiment (*27, 28*), which we expect to greatly reduce the time to operability and, in contrast to conventional approaches, can easily (via retraining) adjust to changes in the plant (e.g., solvent degradation). This power of machine learning in chemical engineering also highlights the need to share data in a machine-actionable form (*29, 30, 31*).

Machine 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 perceive images with our brain. However, in an industrial plant, we often lack understanding of the underlying mechanisms, but with machine learning, we can find the underlying rules of the mapping from the parameters to observables and make predictions for phenomena we could not predict thus far.

METHODS
Pilot plant
Figure 1 shows a schematic 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

desulfurization plant, and a direct contact cooler 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 packings, 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 (32, 33). The amine degradation, due to the presence of oxygen and other impurities such as nitrogen oxides, as well as elevated temperatures during solvent regeneration, can result in other gaseous emissions of degradation compounds such as ammonia (34).

The flue gas upstream of the absorber was analyzed using a BA5000 Bühler infrared spectroscope. The CO_2-lean flue gas downstream of the water wash outlet was analyzed using a GasMET CX/DX 4000 analyzer (i.e., CO_2, CO, O_2, AMP, Pz, NH_3, and H_2O).

Experimental campaign: 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 that drive our stress tests, we focus on those (combinations of) changes, of which our previous study on MEA (15) has shown that they can affect the emissions. The time scale and the magnitude of the changes are based on the expected intermittency (7) and typical requirements of the grid services (10, 35, 36), respectively. A more detailed description is given in note S1.

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 note S6). For all our modeling, we removed deterministic trend components from the data using linear regression, which is motivated by the fact that the characteristic time scale of these components is beyond the one captured by our dataset (and analysis). In addition, removed outliers using a z-score filter ($z = 3$), performed exponential window smoothing (window size of 16 min) and downsampled the data to a frequency of 2 min. The impact of the preprocessing is shown in fig. S5. For use in the models, we additionally standardized the data using min-max scaling. We did not retrain models for historical forecasts.

Quantile regression using gradient boosted decision tree models

To forecast the emissions, we used gradient-boosted decision tree models in which the feature vector is constructed by concatenating lagged time series for process parameters and emission. In this approach, we train a new gradient-boosted decision tree [as implemented in the LightGBM library (37)] for every forecasting horizon using the darts package (38) To obtain uncertainty estimates, we use quantile regression. We tune the hyperparameters of the gradient-boosted decision tree and the number of lags using hyperparameter optimization on a validation set using Bayesian optimization (see note S7.2.). For all models, we scaled the data

(emissions and process variables) based on statistics computed on the training dataset.

Causal impact analysis

For the causal impact analysis, we remove causally related covariates and trained models on the data of the days preceding the step change and following the step changes. For every model, we performed a new hyperparameter optimization (using the shorter sequence preceding or following the step change as a validation set). We also attempted to use Bayesian-structured time series models as in the original implementation of the causal impact analysis technique (22) and found qualitative agreement.

We made use of the following Python (39) libraries: pandas (40), sklearn (41), scipy (42), statsmodels (43), matplotlib (44), jupyter (45), numpy (46), pytorch (47), darts (38), lightgbm (37), and shap (48).

REFERENCES AND NOTES

1. M. Bui, C. S. Adjiman, A. Bardow, E. J. Anthony, A. Boston, S. Brown, P. S. Fennell, S. Fuss, A. Galindo, L. A. Hackett, J. P. Hallett, H. J. Herzog, G. Jackson, J. Kemper, S. Krevor, G. C. Maitland, M. Matuszewski, I. S. Metcalfe, C. Petit, G. Puxty, J. Reimer, D. M. Reiner, E. S. Rubin, S. A. Scott, N. Shah, B. Smit, J. P. M. Trusler, P. Webley, J. Wilcox, N. M. Dowell, Carbon capture and storage (CCS): The way forward. *Energ. Environ. Sci.* **11**, 1062–1176 (2018).

2. A. J. Reynolds, T. V. Verheyen, S. B. Adeloju, E. Meuleman, P. Feron, Towards commercial scale postcombustion capture of CO_2 with monoethanolamine solvent: Key considerations for solvent management and environmental impacts. *Environ. Sci. Technol.* **46**, 3643–3654 (2012).

3. K. Veltman, B. Singh, E. G. Hertwich, Human and environmental impact assessment of postcombustion CO_2 capture focusing on emissions from amine-based scrubbing solvents to air. *Environ. Sci. Technol.* **44**, 1496–1502 (2010).

4. IEA Greenhouse Gas R&D Programme, Environmental Impacts of Amine Emission During Post Combustion Capture (2010); www.globalccsinstitute.com/archive/hub/publications/106171/environmental-impacts-amine-emissions-post-combustion-capture.pdf.

5. P. Khakharia, J. Mertens, M. Abu-Zahra, T. Vlugt, E. Goetheer, in *Absorption-Based Postcombustion Capture of Carbon Dioxide*, P. H. M. Feron, Ed. (Elsevier, 2016), pp. 465–485.

6. L. Biegler, *Systematic Methods of Chemical Process Design* (Prentice Hall PTR, 1997).

7. IEAGHG, *Valuing Flexibility in CCS Power Plants* (Tech. Rep. 2017-09, 2017); https://ieaghg.org/exco_docs/2017-09.pdf.

8. N. E. Flø, H. M. Kvamsdal, M. Hillestad, Dynamic simulation of post-combustion CO2 capture for flexible operation of the Brindisi pilot plant. *Int. J. Greenh. Gas Control.* **48**, 204–215 (2016).

9. J. Gaspar, J. B. Jorgensen, P. L. Fosbol, Control of a post-combustion CO2 capture plant during process start-up and load variations. *IFAC-PapersOnLine* **48**, 580–585 (2015).

10. H. Chalmers, M. Leach, M. Lucquiaud, J. Gibbins, Valuing flexible operation of power plants with CO2 capture. *Energy Procedia* **1**, 4289–4296 (2009).

11. N. E. Flø, H. M. Kvamsdal, M. Hillestad, T. Mejdell, Dominating dynamics of the post-combustion CO2 absorption process. *Comput. Chem. Eng.* **86**, 171–183 (2016).

12. C. Charalambous, A. Saleh, M. van der Spek, G. Wiechers, P. Moser, A. Huizinga, P. Gravesteijn, J. Ros, J. G. M.-S. Monteiro, E. Goetheer, S. Garcia, Analysis of flexible operation of CO2 capture plants: Predicting solvent emissions from conventional and advanced amine systems. *SSRN Electron. J.* , (2021).

13. A. Kachko, L. V. van der Ham, L. F. G. Geers, A. Huizinga, A. Rieder, M. R. M. Abu-Zahra, T. J. H. Vlugt, E. L. V. Goetheer, Real-time process monitoring of CO_2 capture by aqueous AMP-PZ using chemometrics: Pilot plant demonstration. *Ind. Eng. Chem. Res.* **54**, 5769–5776 (2015).

14. J. Pearl, *Causality* (Cambridge Univ. Press, 2009).

15. P. Moser, S. Schmidt, G. Sieder, H. Garcia, T. Stoffregen, Performance of MEA in a long-term test at the post-combustion capture pilot plant in Niederaussem. *Int. J. Greenh. Gas Control* **5**, 620–627 (2011).

16. P. Moser, S. Schmidt, K. Stahl, 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* **4**, 473–479 (2011).

17. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS' 17 (Curran Associates Inc., 2017), pp. 3149–3157.

18. J. H. Friedman, Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **29**, 1189–1232 (2001).

19. K. Das, M. Krzywinski, N. Altman, Quantile regression. *Nat. Methods* **16**, 451–452 (2019).

20. R. Koenker, G. Bassett Jr., *Econometrica* **46**, 33–50 (1978).

21. M. A. Hernán, J. M. Robins, *Causal Inference* (CRC, 2010).

22. K. H. Brodersen, F. Gallusser, J. Koehler, N. Remy, S. L. Scott, Inferring causal impact using Bayesian structural time-series models. *Ann. Appl. Stat.* **9**, 247–274 (2015).

23. A. Hartono, H. F. Svendsen, H. K. Knuutila, Impact of absorption kinetics of individual amine components in CESAR1 solvent on aerosol model performance in the absorber, in *Proceedings of the 15th Greenhouse Gas Control Technologies Conference* (SSRN, 2021).

24. J. Mertens, H. Lepaumier, D. Desagher, M.-L. Thielens, Understanding ethanolamine (MEA) and ammonia emissions from amine based post combustion carbon capture: Lessons learned from field tests. *Int. J. Greenh. Gas Control.* **13**, 72–77 (2013).

25. P. Khakharia, L. Brachert, J. Mertens, C. Anderlohr, A. Huizinga, E. S. Fernandez, B. Schallert, K. Schaber, T. J. Vlugt, E. Goetheer, Understanding aerosol based emissions in a post combustion CO_2 capture process: Parameter testing and mechanisms. *Int. J. Greenh. Gas Control.* **34**, 63 (2015).

26. A. F. Ciftja, A. Hartono, E. F. da Silva, H. F. Svendsen, Study on carbamate stability in the Amp/CO_2/H_2O system from ^{13}C-NMR spectroscopy. *Energy Procedia* **4**, 614–620 (2011).

27. K. M. Jablonka, G. M. Jothiappan, S. Wang, B. Smit, B. Yoo, Bias free multiobjective active learning for materials design and discovery. *Nat. Commun.* **12**, 2312 (2021).

28. T. Lookman, P. V. Balachandran, D. Xue, R. Yuan, Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput. Mater.* **5**, (2019).

29. K. M. Jablonka, L. Patiny, B. Smit, Making the collective knowledge of chemistry open and machine actionable. *Nat. Chem.* **14**, 365–376 (2022).

30. A. M. Schweidtmann, E. Esche, A. Fischer, M. Kloft, J.-U. Repke, S. Sager, A. Mitsos, Machine learning in chemical engineering: A perspective. *Chem. Ing. Tech.* **93**, 2029–2039 (2021).

31. J. M. Weber, Z. Guo, C. Zhang, A. M. Schweidtmann, A. A. Lapkin, Chemical data intelligence for sustainable chemistry. *Chem. Soc. Rev.* **50**, 12013–12036 (2021).

32. P. Moser, S. Schmidt, K. Stahl, G. Vorberg, G. A. Lozano, T. Stoffregen, F. Rösler, Demonstrating emission reduction – Results from the post-combustion capture pilot plant at Niederaussem. *Energy Procedia* **63**, 902–910 (2014).

33. A. Rieder, S. Dhingra, P. Khakharia, L. Zangrilli, B. Schallert, R. Irons, S. Unterberger, P. Van Os, E. Goetheer, Understanding solvent degradation: A study from three different pilot plants within the OCTAVIUS Project. *Energy Procedia* **114**, 1195–1209 (2017).

34. E. F. da Silva, K. A. Hoff, A. Booth, Emissions from CO_2 capture plants; An overview. *Energy Procedia* **37**, 784–790 (2013).

35. M. Bui, I. Gunawan, V. Verheyen, P. Feron, E. Meuleman, S. Adeloju, Dynamic modelling and optimisation of flexible operation in post-combustion CO_2 capture plants—A review. *Comput. Chem. Eng.* **61**, 245–265 (2014).

36. M. Bui, N. E. Flø, T. de Cazenove, N. M. Dowell, Demonstrating flexible operation of the Technology Centre Mongstad (TCM) CO_2 capture plant. *Int. J. Greenh. Gas Control.* **93**, 102879 (2020).

37. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, LightGBM: A highly efficient gradient boosting decision tree. *Adv. Neural. Inf. Process Syst.* **30**, 3146 (2017).

38. J. Herzen, F. Lässig, S. G. Piazzetta, T. Neuer, L. Tafti, G. Raille, T. V. Pottelbergh, M. Pasieka, A. Skrodzki, N. Huguenin, M. Dumonal, J. Kościsz, D. Bader, F. Gusset, M. Benheddi, C. Williamson, M. Kosinski, M. Petrik, G. Grosch, Darts: User-friendly modern machine learning for time series. *J. Mach. Learn. Res.* **23**, 1–6 (2022).

39. G. Van Rossum, F. L. Drake, *Python 3 Reference Manual* (CreateSpace, 2009).

40. W. McKinney, Data structures for statistical computing in python, in *Proceedings of the 9th Python in Science Conference* (SciPy 2010), Austin, Texas, 28 June to 3 July 2010, vol. 445, pp. 51–56.

41. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

42. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt; SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

43. S. Seabold, J. Perktold, Statsmodels: Econometric and statistical modeling with Python, in *9th Python in Science Conference* (SciPy 2010), Austin, Texas, 28 June to 3 July 2010.

44. J. D. Hunter, Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

45. T. Kluyver, B. Ragan-Kelley, F. Pérez, B. E. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. B. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing; Jupyter Development Team, Jupyter Notebooks-a publishing format for reproducible computational workflows, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides, B. Schmidt, Eds. (IOS Press, 2016), vol. 2016, pp.87–90.

46. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, Array programming with NumPy. *Nature* **585**, 357–362 (2020).

47. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett, Eds. (Curran Associates Inc., 2019), pp. 8024–8035.

48. S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2**, 56–67 (2020).

49. J. Monteiro, J. Ros, E. Skylogiani, A. Hartono, H. Svendsen, H. Knuutila, P. Moser, G. Wiechers, C. Charalambous, S. Garcia, Accelerating Low carboN Industrial Growth through CCUS Deliverable Nr. D1.1.7 Guidelines for Emissions Control (Tech. Rep. D1.1.7, ALIGN-CCUS, 2021).

50. E. Mechleri, A. Lawal, A. Ramos, J. Davison, N. M. Dowell, Process control strategies for flexible operation of post-combustion CO_2 capture plants. *Int. J. Greenh. Gas Control.* **57**, 14–25 (2017).

51. M. Bui, I. Gunawan, V. Verheyen, P. Feron, E. Meuleman, Flexible operation of CSIRO's post-combustion CO_2 capture pilot plant at the AGL Loy Yang power station. *Int. J. Greenh. Gas Control.* **48**, 188–203 (2016).

52. S. Bai, J. Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv:1803.01271 [cs.LG] (4 March 2018).

53. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. arXiv:1512.03385 [cs.CV] (10 December 2015).

54. L. Zhu, N. Laptev, Deep and confident prediction for time series at Uber. *IEEE Int. Conf. Data Mining Workshops* , 103–110 (2017).

55. J. Hron, A. G. de G. Matthews, Z. Ghahramani, Variational Gaussian Dropout is not Bayesian. arXiv:1711.02989 [stat.ML] (8 November 2017).

56. J. Hron, A. G. de G. Matthews, Z. Ghahramani, Variational Bayesian dropout: Pitfalls and fixes. arXiv 1807.01969 [stat.ML] (5 July 2018).

57. B. Lim, S. O. Arik, N. Loeff, T. Pfister, Temporal fusion transformers for interpretable multi-horizon time series forecasting. 37, 1748–1764 (2021).

58. C. Molnar, *Interpretable Machine Learning* (2019);https://christophm.github.io/interpretable-ml-book/.

59. S. M. Lundberg, S.-I. Lee, in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett, Eds. (Curran Associates Inc., 2017), pp. 4765–4774.

60. M. Eichler, Granger causality and path diagrams for multivariate time series. *J. Econom.* **137**, 334–353 (2007).

61. N. Artrith, K. T. Butler, F.-X. Coudert, S. Han, O. Isayev, A. Jain, A. Walsh, Best practices in machine learning for chemistry. *Nat. Chem.* **13**, 505–508 (2021).