Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Digital Discovery*, 2026, **5**, 1191

Machine learning-based time-series forecasting prevents electrode corrosion in organic electrochemistry

Josef Tausendschön, *[a] Michael Poelzl, [ab] Nikola Petrovic, [a] Jason D. Williams [a] and Elisabeth Fink [a]

A real-time monitoring and predictive modeling framework was developed to address electrode corrosion and process deviations in the single-step electrochemical synthesis of cortisone to adrenosterone. The approach combined inline Fourier-transform infrared (FT-IR) spectroscopy with partial least squares (PLS) modeling for rapid quantification of chemical species, alongside optical coherence tomography (OCT) for electrode surface monitoring. A novel OCT image analysis was introduced enabling efficient dynamic pixel-based quantification of electrode corrosion in real time. To complement this, machine learning-based time-series forecasting models, including convolutional neural networks (CNNs) and long short-term memory (LSTM) architectures, were optimized for long-term concentration forecast from real-time sensor data. The incorporation of domain-inspired additional features like the applied current value and the relative corrosion measure resulted in enhanced accuracy. To ensure chemical plausibility a concentration conservation constraint was employed. Despite challenges arising from fluctuations in FT-IR/PLS-derived concentration data and experimental variability, the combined monitoring–forecasting framework demonstrated robustness and predictive value. This work highlights the potential of integrating advanced sensing with machine learning-based predictive modeling for future predictive process control in electrochemical reactors.

Received 10th October 2025
Accepted 31st January 2026

DOI: 10.1039/d5dd00458f

rsc.li/digitaldiscovery

1. Introduction

Electrochemistry is emerging in many pharmaceutical and industrial applications.[1–4] Electro-organic synthesis is environmentally friendly and cost-effective and therefore offers a great opportunity for a transformation into green chemistry and sustainability.[2,5,6] The comparably easy application, precise control and safe conversion is very attractive in pharmaceutical production for enhanced safety and efficiency.[2,6]

However, applying electrochemical methods is not inherently sustainable, as environmentally harmful solvents or waste through sacrificial electrodes are problematic.[6] Furthermore, scaling up the electrochemical processes developed at the lab scale to industrial settings is not straightforward,[7] with much of the challenge stemming from reactor design, which must be informed by suitable reaction data to facilitate transfer between different scales. This work focuses on developing process data collection and predictive monitoring strategies at the laboratory scale, with specific emphasis on integrating corrosion data to address these challenges. This includes systems capable of monitoring reaction progress and identifying potential issues such as electrode corrosion, connectivity faults, or deviations from the expected process trajectory. Establishing such reliable monitoring systems, is a critical step toward enabling scalable and robust electrochemical synthesis.[8]

In the realm of pharmaceuticals, steroids and their derivatives play a crucial role as drug for anti-inflammatory, contraceptive, antibiotic or antihistamine treatments.[9–16] Sommer *et al.*[10] pioneered the electro-organic synthesis of C19 android steroids. The investigated reaction proceeds in a single step, is safe to perform, and is relatively inexpensive—making it a strong candidate for a model process for development and provides an opportunity to explore how corrosion-informed forecasting could improve process control.

In order to perform electrochemical process control, in-line or *operando* monitoring was established.[17–22] Chemical species concentrations must be recorded rapidly to ensure fast detection and response to disturbances, therefore spectroscopic methods are preferred over chromatographic ones.[23] For reactions involving complex molecules with multiple functional groups, such as adrenosterone and cortisone, spectroscopic monitoring is generally not straightforward, due to a large number of overlapping peaks. To facilitate rapid quantification, chemometric methods, such as partial least squares (PLS) modeling can be used to link spectroscopic data and chemical concentration. By calibrating a PLS model, the entire

[a]*Research Center Pharmaceutical Engineering GmbH, Inffeldgasse 13, Graz, 8010, Austria. E-mail: josef.tausendschoen@rcpe.at*
[b]*Graz Center for Machine Learning, Graz University of Technology, 8010, Graz, Austria*

"fingerprint" of a molecule is considered, rather than signals arising from individual functional groups. Fourier transform infrared (FT-IR) spectroscopy is especially suitable for such an approach, due to the characteristic fingerprint for a particular molecule, albeit often with many overlapping signals.[24]

Degradation of the electrodes due to fouling and corrosion can hinder process efficiency. The study of fouling and corrosion processes is crucial for understanding their causes and dynamics, which can aid in mitigating these issues. This work hypothesizes that real-time incorporation of corrosion data into predictive models will improve their ability to capture process dynamics. Real-time monitoring of the electrode surface during electrolysis provides valuable insights by simultaneously acquiring electrolysis data and quantitative information about surface changes, allowing for the establishment of cause-and-effect relationships. Recently, optical coherence tomography (OCT) has emerged as an *in situ* monitoring technique for electrochemical processes.[25] This technique enables the detection and quantification of critical challenges such as corrosion and fouling, which significantly impact the efficiency and commercial viability of electrochemical processes. These findings underscore the potential of OCT as a valuable tool for real-time analysis, offering insights into the conditions that lead to electrode degradation and enabling improved management of electrode performance in electrochemical applications.

In electrochemistry, thermodynamics and kinetics have a quantitative relationship,[26] allowing mathematical models to predict the effects of change in a system, *e.g.*, concentration or rate constant. By integrating mass transfer with the thermodynamics and kinetics of the reaction, it is possible to describe electrochemical systems in a rigorous manner. However, this generally leads to coupled systems of partial differential equations (PDEs), whose formulation, parametrization, and solution in a reactor context can be complex and computationally demanding, particularly when time-varying operating conditions are considered.[26–29]

For real-time applications, mathematically simpler alternatives such as ordinary differential equation (ODE) models or reduced-order models (ROMs) are often employed in electrochemical environments.[30–32] Traditional ODE-based models are usually not applied in multivariate time-series forecasting across extended horizons when the system is subject to non-linear parameter drift (*e.g.*, electrode corrosion) as this behavior is difficult to capture with fixed coefficients. ROMs, while computationally efficient, are commonly derived for specific system configurations and require significant effort to construct and re-validate when system inputs or operating regimes change. This limitation highlights the potential of flexible, data-driven approaches that can explicitly account for corrosion dynamics.

In this context, data-driven time-series models based on neural networks offer a flexible alternative. Such models can naturally incorporate multiple dynamic inputs, including applied current and corrosion-related variables, and can capture nonlinear, time-dependent behavior, aligning with the goal of real-time data processing and forecasting.

We propose that models incorporating corrosion data will better capture the time-dependent degradation of electrode performance. Time-series forecasting is a critical and rapidly evolving research area with significant implications for both academia and industry. Chandra *et al.*[33] thoroughly analyzed univariate time-series for multi-step-ahead prediction using simple recurrent neural networks (RNNs), long short-term memory (LSTM) networks, bidirectional LSTM networks, encoder-decoder LSTM networks, and convolutional neural networks (CNNs). Similarly, Lim *et al.*[34] reviewed attention-based[35] models for univariate time-series. These findings for univariate time-series can in general be extended to multivariate time-series.[36,37] However, optimal architecture varies depending on the specific time-series data. It is important to note that time-series data, such as weather patterns or consumer prices, which often underpin these studies, typically exhibit underlying seasonality and trends. In contrast, electrochemical reaction data, particularly when corrosion is a factor, requires models capable of handling non-stationary, nonlinear drift.

The application of neural networks to electrochemical reaction data began several decades ago,[38] laying the foundation for more recent developments in time-series forecasting within this domain. Luo *et al.*[8] developed an online predictive control system for an electrochemical reactor using an LSTM network to forecast the production rates of CO, C_2H_4, and H_2 one step ahead. Similarly, Zhang *et al.*[39] created a battery forecasting system employing Gaussian process machine learning models based on electrochemical impedance spectroscopy data. Davies *et al.*[40] leveraged machine learning to predict future product formation from time-series data collected *in situ* from twelve sensors, employing models such as random forest, gradient boosting regression, and LSTM neural networks. Ji *et al.*[41] treated steel corrosion, a process inherently involving electrochemical reactions, as a time-series prediction problem and proposed RNN models for time-variant corrosion rate forecasting. Furthermore, Chen *et al.*[42] introduced a hybrid prediction model and benchmarked its performance against classical machine learning models for forecasting iron and silicon content from time-series data.

Although models for time-series forecasting in some specific systems have been developed, they cannot be directly applied to the present system. Transfer learning is also not suitable, as no large body of pre-existing data from similar electrochemical processes exists to support such an approach. In addition, flexible and adaptive process modeling require forecasts with variable input and output lengths, which exceeds the capabilities of some models. Therefore, this study aims to develop a process model for an electrochemical reactor. Specifically, we design a machine learning-based time-series forecasting model in which the applied current, the main process parameter, and corrosion measurements are directly incorporated.

2. Experimental

2.1. Process overview

The electrochemical conversion of cortisone to adrenosterone is used here as a case example. This specific reaction has been explored by Sommer *et al.*[10] as part of a multistep synthesis of

C19 androgen steroids. In a typical experiment, cortisone and a supporting electrolyte (*e.g.*, sodium tetrafluoroborate) are suspended in a mixture of acetonitrile and water. A cortisone concentration of 0.05 M was selected, as this represents the maximum concentration at which a homogeneous solution is maintained.[43] Although exceeding the solubility limit could improve the yield-to-solvent ratio, it would generate a suspension that risks clogging the system or causing undesired solid deposition on the electrode surface.

The experiments utilize two different solvent ratios of acetonitrile to water, specifically 40 : 1 and 80 : 1 ($MeCN : H_2O$). The mixture is then electrolyzed using a graphite anode and a stainless-steel cathode with defined current density at room temperature. To ensure that the process runs optimally, three different data sources are monitored (see Fig. 1a): (1) the power supply defining the voltage and current settings, (2) the online FT-IR measurement providing information about the chemical composition of the solution and (3) the OCT measurement, providing images of the electrode's surface.

However, optimal control of such processes typically results in a tradeoff between productivity, product yield and electrode condition: high applied current generally provides a fast reaction rate (high productivity) but can detrimentally affect reaction selectivity (lower product yield). Fouling or corrosion of the electrode can occur with high current,[44] which reduces the process performance and can even cause irreversible damage, forcing a process-ending replacement. Furthermore, unnecessary waste is produced, and the production of new electrodes is energy intensive.

Real-time FT-IR analytics has been utilized by Sagmeister *et al.*[24] for multistep synthesis among other applications. In this study we want to add the OCT measurements[25] and aim to develop a neural network-based time-series forecasting model for the concentration of the reactant and the products. During the process, the model forecasts the reactant and product concentrations (and thus the product yield), based on real-time sensor data from an online FT-IR spectroscopy measurement (and a subsequent PSL regression model), the power supply and the corrosion measurement.

2.1.1. Reaction and reactor setup. The reaction vessel with the magnetic stirrer can be seen in Fig. 1b. The full experimental setup can be seen in Part A of the SI. In the reaction,[10] an oxidative cleavage of the C17 sidechain of cortisone is performed by applying galvanostatic current to the reaction mixture *via* a power supply at the electrodes. The performed transformation yields adrenosterone, an important precursor in synthesis of C19 steroids. The system consists of three quantifiable species: cortisone, adrenosterone and combined "side products". The main reaction can be seen in Fig. 3.

The reaction by-product, glycolic acid, is then proposed to decompose to CO_2 and formaldehyde, which are lost from the system and not observed. Water is present in the system in excess, so can be ignored in rate equations. Furthermore, the cathodic reduction ($2H^+ \rightarrow H_2$) is considered to occur freely, due to the availability of H^+ in the reaction. The reactor is filled with 10 mL of the mixture, which contains 50 mM cortisone. Side products are formed by reaction of either cortisone or adrenosterone, thereby the mole balance is conserved. This means that one can assume that the sum of concentration of cortisone, adrenosterone and the side products remains at around 50 mM.

2.1.2. FT-IR measurement and chemometric modeling. During the process, FT-IR online measurements are performed *via* a bypass at a recirculation rate of 1 mL min^{-1}. This setup allows for the acquisition of a full spectrum every 30 seconds. Fig. 2 illustrates such spectra captured at several time points during an experiment. The chemical composition can be inferred from the peaks in these spectra, which correspond to specific functional groups. However, due to the complexity of the molecules involved in this process, it is not possible to accurately quantify the reaction species based on their individual functional group signals alone. To enable this quantitation, a PLS model is calibrated.

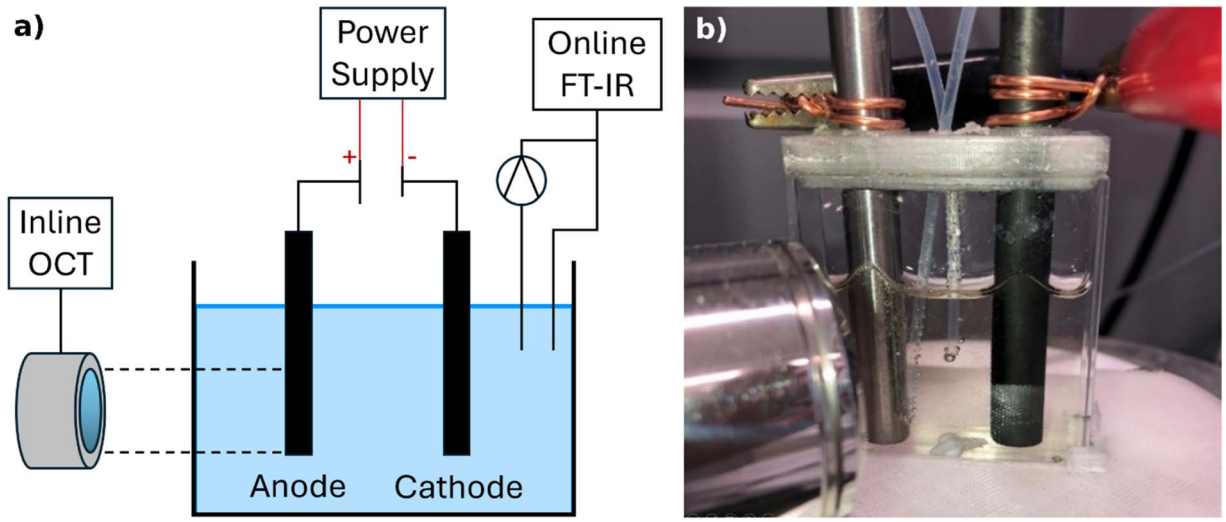

Fig. 1 (a) Process overview, including the different data sources and (b) the actual glass quartz cell used as electrochemical reactor. Further, the right side shows the anode and cathode, the OCT lens as well as the tubing used to circulate through the online FT-IR instrument.

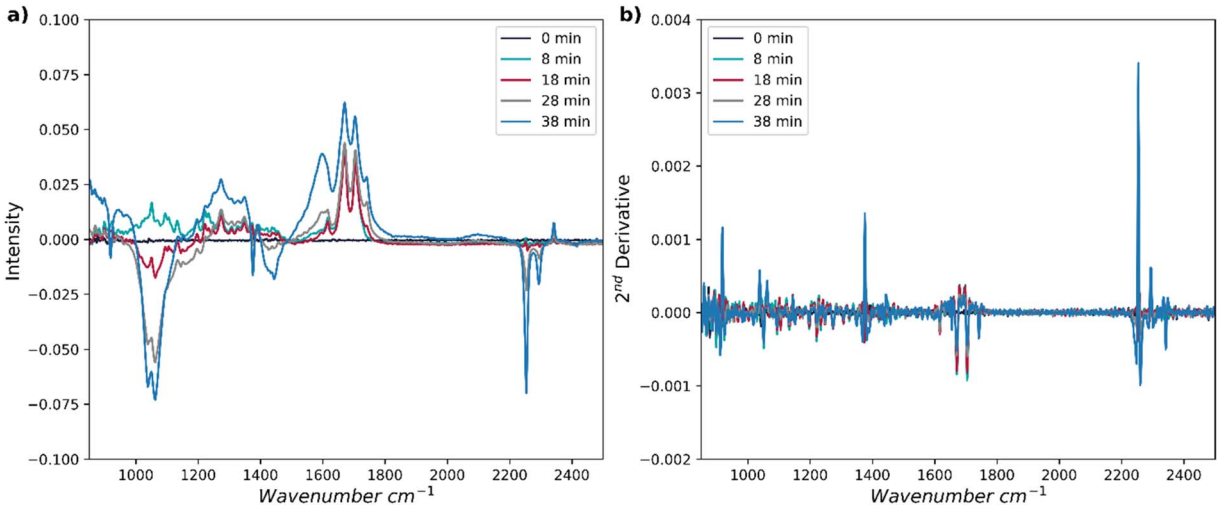

Fig. 2 (a) Representative FT-IR spectra over the process time in the desired range and (b) the second derivative of these spectra.

Offline measurements of the chemical composition are determined using high-performance liquid chromatography (HPLC). This method identifies the amounts of cortisone, adrenosterone, and any side products. These HPLC-derived concentrations are then tagged to the recorded FT-IR spectrum at that time point. A PLS model is then calibrated to correlate these HPLC measurements with the FT-IR spectra. The range considered in the FT-IR spectra was reduced to 850–2500 cm^{-1} (see Fig. 2), no baseline correction was applied and a second derivative (see Fig. 2b) was taken (to enhance the influence of small differences in the spectra and help distinguish shoulders of overlapping peaks). A rank of 9, 9 and 8 was taken for cortisone, adrenosterone and side products, respectively.

These model parameters are chosen for the ability of the resulting model to handle multicollinearity and to predict the concentrations of substances from new spectra accurately. As measures of accuracy and overfitting, the root mean squared error (RMSE) of calibration ($RMSE_C$) and cross validation ($RMSE_{CV}$) were used. For the chosen model, $RMSE_C = 2.73$ mM, 1.55 mM, 1.94 mM for cortisone, adrenosterone and side products, respectively. To demonstrate that the model was not overfitted, $RMSE_{CV} = 4.11$ mM, 3.13 mM, 3.05 mM for cortisone, adrenosterone and side products, respectively.

2.1.3. OCT measurement. Optical Coherence Tomography (OCT) is a non-invasive imaging technique based on low-coherence interferometry. It operates by measuring the optical path length differences between a reference beam and light backscattered from the sample. This measurement produces a depth profile (A-scan) at a single point. By performing lateral scans, these profiles are combined to form a 2D cross-sectional image; stacking these 2D images allows for the reconstruction of a 3D volumetric representation of surface and subsurface structures. This enables depth-resolved imaging with micrometer-scale resolution *in situ*. Further details on the principles of OCT can be found in Schippling[45] and applications regarding corrosion investigation are detailed in Fink *et al.*[25]

In this study a commercial 1D OCT system (OSeeT, Phyllon GmbH, Graz, Austria) was adapted to a 3D system. The 1D system employs a low-coherence light of wavelength of a range of 742–900 nm. The emitted light beam is further split into a reference and a sample measurement arm and the reflected light of the two compared to obtain a spectrum from which an image is reconstructed. The 1D system is a spectral-domain OCT system where the sample is moved in front of the sensor. By utilizing two galvo mirrors, the 1D OCT light beam can scan a stationary surface. Each horizontal movement of the mirrors results in a 2D scan across the surface of the electrode. Depending on the desired resolution the 2D images are then combined into a 3D depiction of the surface. The system can process one 3D scan of 2 GB, which resembles 512 2D scans, with each allocating 4 MB, per minute.

2.2. Experiments

All chemicals were obtained from standard commercial vendors and were used without any further purification. The electrolysis cell was powered with an adjustable BK Precision BK 1735A power supply (max. 30 V, 1 A). The Vaportec SF-10 peristaltic pump was used for transferring the solution between the electrolysis cell and the in-line FT-IR. The in-line FT-IR spectra were recorded using a Mettler Toledo ReactIR 15 FTIR, equipped with a flow cell (Mettler Toledo, Micro Flow Cell DS DiComp). The acquisition time was 30 s per data point, and the spectra were recorded between 600 cm^{-1} and 4000 cm^{-1} using a resolution of 4 cm^{-1}. Before use, the MCT detector was allowed to warm to

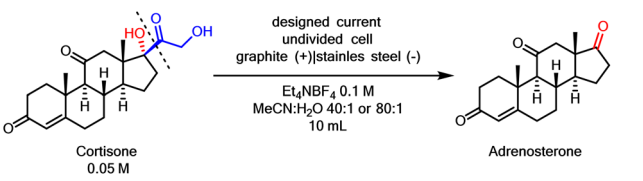

Fig. 3 The chemical structure of the reactant and the product as well as the reaction conditions (adjusted from ref. 10).

room temperature, then cooled with liquid N_2. It was ensured that the peak height was between 18 000 and 24 000 and that the signal to noise ratio was above 5000.

Two types of experiments were performed, constant current experiments and designed current path experiments. The selected values for the constant current experiments were 12.5, 25, 50, 75, 100 and 125 mA. For the designed paths see Part A in the SI. In the constant currents experiments HPLC samples were taken after each equivalent of electrons passed through the mixture, while in designed path experiments a HPLC sample taken after each change in the applied current.

The experiments were performed under constant mixing. The solutions were prepared by dissolving 180 mg of cortisone and 217 mg of supporting electrolyte in 10 mL of the solvent mixture in a 10 mL volumetric flask and then transferred into the electrolysis cell. For each performed experiment a new set of electrodes were used to avoid cross-contamination issues in the dataset produced. In total 40 experiments have been performed, for a comprehensive list and overview please see Part A of the SI.

3. Corrosion detection

The resulting images from the OCT measurement are then further pre-processed, as they contain speckle noise, which is common in OCT images.[46,47] The processing pipeline is illustrated in Fig. 4. Unprocessed images from the early stage of an experiment and from the later stage can be observed in Fig. 5a and b. The first step in the pre-processing pipeline is non-local

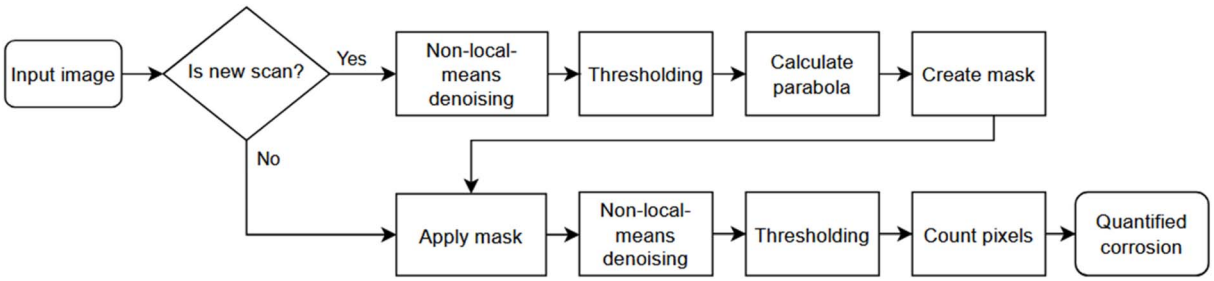

Fig. 4 OCT image processing pipeline.

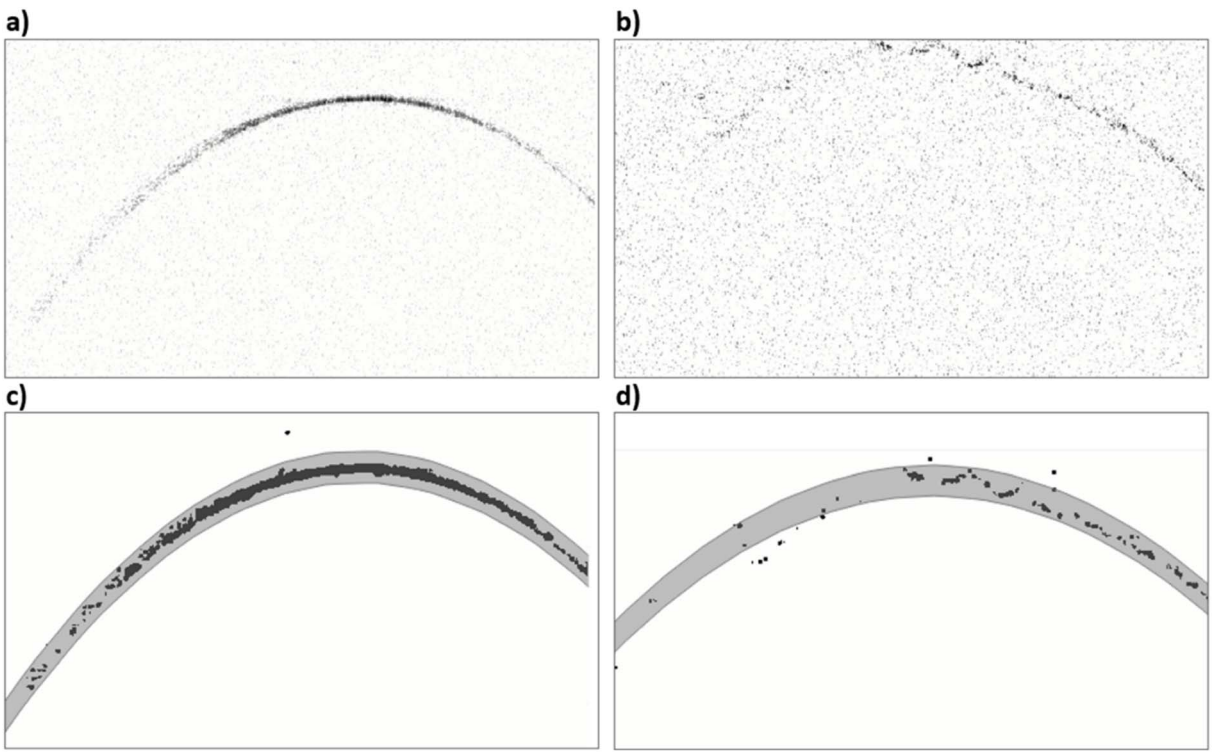

Fig. 5 Evolution of corrosion features in OCT over time. (a) Displays the OCT image from the first scan, whereas (b) from the fifth scan, approximately 30 minutes after the first scan. (c) Shows the image from the first scan after the processing pipeline is applied. The remaining pixels included in the calculated mask are relevant for corrosion quantification. (d) Also shows the same processing applied to the fifth scan. Clearly, fewer pixels fall within the mask—indicating that more abnormalities are present, which are not considered in the corrosion quantification. Only the pixels within the mask are used for this purpose. The surface becomes nearly invisible.

means denoising using a weighted average, which aims to reduce the speckle noise while preserving image details.[48]

In the next step, an adaptive thresholding operation is applied to account for variations in background illumination. The threshold value T is calculated for each individual image based on the mean pixel intensity (μ) of that image, defined as $T = \mu - 5$. Pixels below this calculated value are set to 1 (white), and those above are set to 0 (black). The constant offset of 5 was selected based on pilot data to maximize the contrast between the electrode and the smoothed background noise.

Following this step, a parabola is fitted to the highlighted region using a least-squares approach, minimizing the sum of squared deviations between the measured OCT signal and the fitted curve. This fitted parabola serves as the reference for generating a dynamic mask that restricts subsequent processing to the region of interest. Specifically, the mask is defined by vertically offsetting the fitted parabola by ± 20 pixels along the y-axis, thereby establishing upper and lower bounds around the estimated electrode surface. Pixels within this band are retained for further analysis, while pixels outside the mask are excluded.

The choice of the ± 20-pixel vertical extent is based on a statistical analysis of parabola height variability across experiments and is summarized in the SI (see Table S5). For the initial, minimally corroded stages of each experiment, the observed parabola heights and their standard deviations remain well within this range, ensuring that the mask reliably captures all relevant structural information without truncation. Restricting the analysis to this region substantially reduces the computational load while preserving the features of interest.

A new mask is generated for each scan to compensate for potential lateral or axial shifts caused by sensor movement or changes in the observation angle, which could otherwise result in partial loss of the electrode region when using a static mask. Representative examples of the dynamically generated masks are shown in Fig. 5c and d. Recomputing the mask for every scan enhances robustness and stability of feature detection across the dataset.

After these pre-processing steps, only the surface of the electrode remains clearly identifiable in the images. As corrosion progresses, the captured surface of the electrode begins to become less pronounced in the images – possibly due to increased light scattering caused by changes in surface roughness in the corroded areas. This leads to a reduction in the intensity of light transmitted back to the OCT sensor, resulting in fewer pixels surpassing the empirically determined threshold mentioned previously. To quantify the level of corrosion, the number of pixels of high intensity per image is calculated. This value serves as a feature for assessing the degree of corrosion.

Over the course of the experiment, the image of the electrode captured by the OCT device gradually degrades or appears to vanish, as shown previously. Consequently, the number of remaining pixels per image after the pre-processing steps may serve as an indicator of corrosion. This information can be incorporated as an additional feature to improve the prediction of concentration values during training.

4. Concentration forecast modeling

As outlined so far, we aim to develop a model to forecast the concentration of the reactant and the products during the process in real-time. As link to process control, we incorporate the applied current into the model and further we investigate whether the consideration of the corrosion measurement is beneficial or not.

To enhance accuracy and incorporate chemical reasoning, we employ domain-inspired neural networks.[49] A prevalent approach involves integrating a constraint into the cost function, derived from a domain-specific quantity. Hence, we add the relative conservation of concentration to the loss function. This constraint should ensure that the relative sum of the concentration remains at unity across all time steps (for details of the chemical reasoning see Section 2.1.1).

4.1. Data collection and preparation

A dataset was created from the data recorded during the experiments performed. This so-called "historical raw data" is reviewed and cleaned before further processing. To develop meaningful training, validation and test data, the experiments are randomly assigned to the according split. A relative amount of 60% is assigned to the training data set and 20% each in the validation and test data set. The data is then sequenced into a multivariate time-series, containing the concentrations of the reactant and the products at each timestep (the created dataset is publicly available[50]).

The optimal number of input and output steps in a sequence are searched within the hyperparameter optimization. Preprocessing the model input data involves normalization of the concentration of the reactant and the products based on the initial cortisone concentration in the solution. The concentration of cortisone at the beginning is equal to 50 mM in the solution. The applied current is also normalized by the maximum setting applied during any experiment (which is 125 mA).

The corrosion value extracted from each image, as presented in chapter 3, enhances the dataset by serving as a promising feature for improving concentration forecasting. To ensure comparability across experiments, the number of corrosion-related pixels per image is normalized using min–max normalization. The min value is set to 0, while the max value is searched in the training data. Fig. 6 illustrates the change in pixel count over six full scans of an experiment conducted at 100 mA. A decreasing trend in the number of pixels is observed across the scans, suggesting progressive corrosion or surface degradation.

4.2. Model structure and hyperparameters

RNNs are designed to handle sequential data, so are well-suited for handling time-series data. However, CNNs can also be adopted for sequential data. So, two types of neural networks were investigated: CNNs and LSTM networks as RNN-subtype. For all models the structure is adjusted to include the applied current (A/A_{max}) and the relative corrosion value (S_{rel}) as

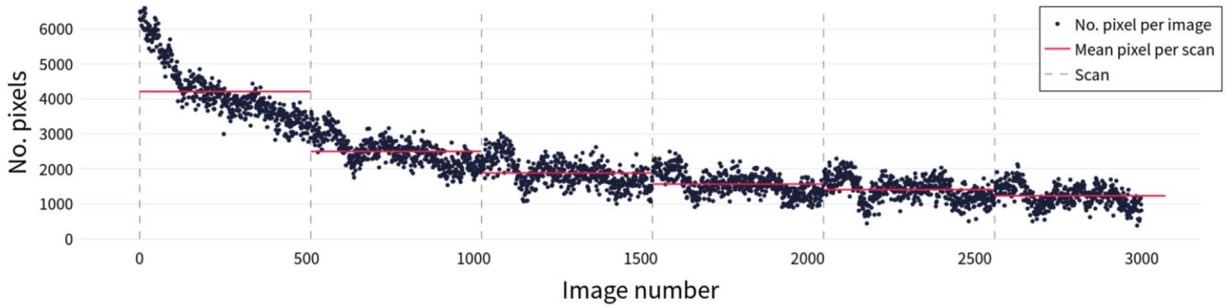

Fig. 6 This figure shows the number of pixels per image remaining after the preprocessing steps (vertical axis) as a function of the image number (horizontal axis). The image number corresponds to a time-series, where image 0 was acquired before image 1, and so on. The red line indicates the mean number of pixels per scan. Each scan consists of 512 images and is separated by a dashed grey line.

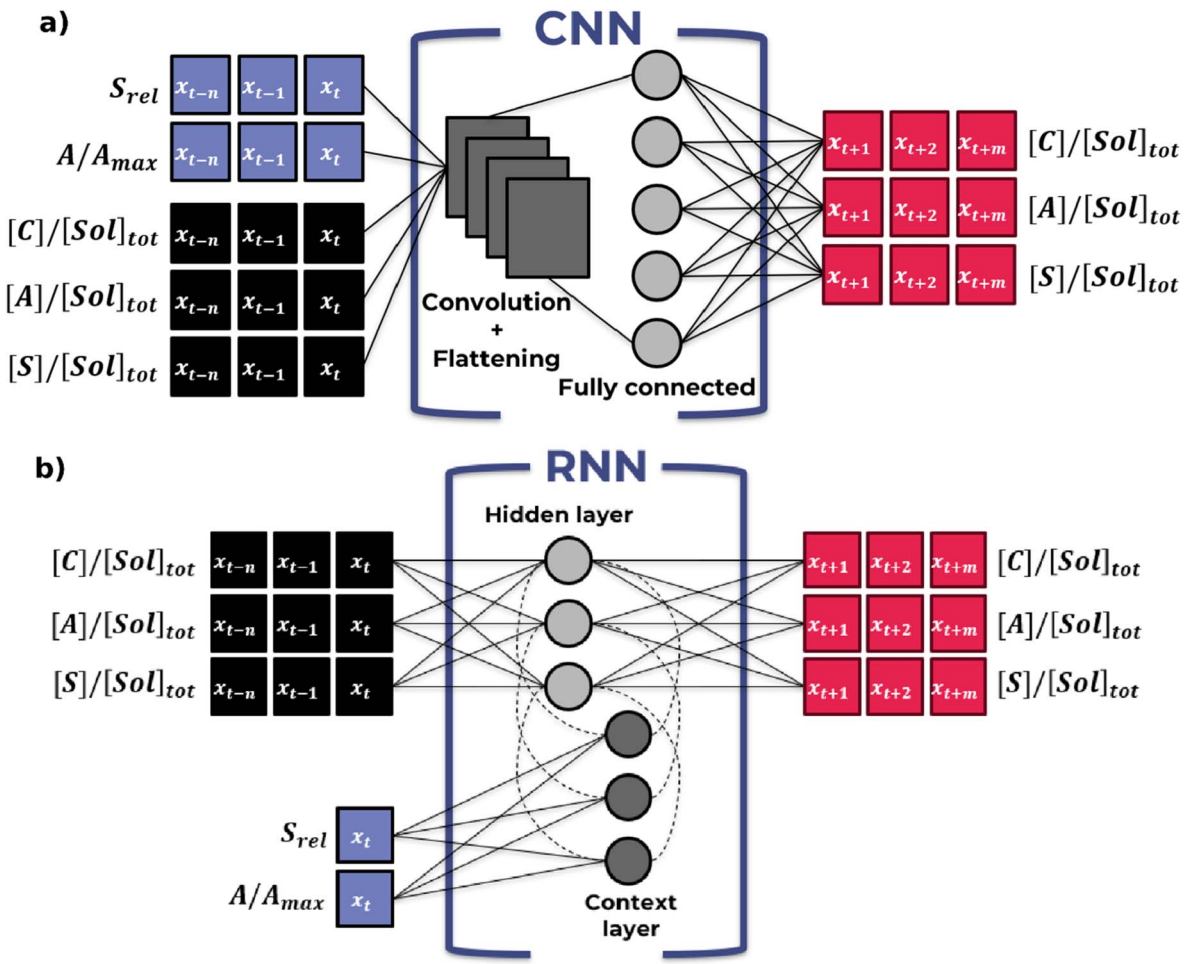

Fig. 7 Schematic structure of the neural network types investigated. (a) shows the principle of CNN-type models, while (b) displays the principle of RNN-type models.

additional features. In CNNs the applied current and the relative corrosion value can simply be treated as additional time-series inputs (see Fig. 7a). For LSTMs, these additional features are considered as conditional inputs[51] and not as time-series data, as they are not part of the forecast (see Fig. 7b). This means that the states of RNN-type architectures were initialized with a learned representation of the additional features.

In summary the model inputs are:

- $[C]/[Sol]_{tot}$, $[A]/[Sol]_{tot}$ and $[S]/[Sol]_{tot}$ being the relative cortisone, adrenosterone and side product concentrations with input sequence length n.
- A/A_{max} the normalized applied current value, with input sequence length n for CNNs and at time t for RNNs.

• S_{rel} the relative corrosion value, with input sequence length n for CNNs and at time t for RNNs.

The model output is always:

• $[C]/[Sol]_{tot}$, $[A]/[Sol]_{tot}$ and $[S]/[Sol]_{tot}$ with output sequence length m.

Details about the architecture of each model type can be found in Part B of the SI.

During the process, the model is used to forecast the concentrations until the cortisone conversion is finished. The models operate in an autoregressive feedback loop, enabling them to predict concentrations until the end of the process at any given time. In contrast to CNNs, RNNs can be explicitly designed as autoregressive models.

This means that the model decomposes its predictions into individual time steps. Then, the model's output can be fed back to itself at each step and predictions can be made conditioned on the previous one. With this structure, models can be set up to produce output with a varying length.

To explore this capability, two different variants of the LSTM (RNN) network were examined: one for use in a feedback loop and another built in an autoregressive manner (AR-LSTM). The searched model and training parameters are summarized in Table 1. CNN-type models require some additional restrictions in terms of valid parameter combinations, *e.g.* the kernel size (*i.e.* the size of the window that slides over the time-series data to extract features) must be smaller than the number of input steps, which were integrated into the search.

To ensure effective optimization the optuna framework[52] (version 3.5.0) was employed. Each model option was subjected to 1000 trials, and less promising parameter combinations were terminated early using the basic pruner class. Optuna's multi-objective optimization capability was utilized to minimize both sequence loss and long-term forecast error on the training and validation datasets. In addition to model parameters, training parameters were also optimized. Model implementation was carried out in TensorFlow (version 2.10.0). To prevent overfitting, early stopping with a patience value of one was applied in all training runs. The Adam[53] optimizer (with default settings except for the learning rate) was used as the standard optimization algorithm.

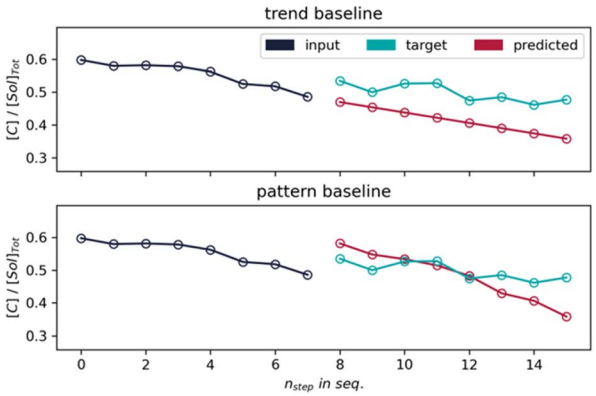

Fig. 8 Example of a sequence for one concentration. The top panel shows the linear trend baseline, while the bottom displays the pattern-based baseline.

As basis of the cost function for model training, the mean squared error (MSE) was selected. In addition to the data loss, a domain-inspired constraint was added to ensure that the sum of concentrations remains constant (for details see Chapter 2.1.1). An example of the input, target and predicted sequence can be seen in Fig. 8.

The data loss is then shown in eqn (1):

$$\mathrm{MSE}_{\mathrm{data}} = \frac{1}{B \cdot T \cdot F} \sum_{i=1}^{B} \sum_{j=1}^{m} \sum_{f=1}^{F} \left(z_{\mathrm{target},b,t,f} - z_{\mathrm{pred},b,t,f} \right)^2 \quad (1)$$

where B is the batch size (see Table 1), m is the length of the output sequence (*e.g.* m is equal to 8 in Fig. 8) and F is the number of features. The additional constraint is then defined in eqn (2):

$$\mathrm{MSE}_{\mathrm{conc}} = \frac{1}{B \cdot T} \sum_{i=1}^{B} \sum_{j=1}^{m} \left(1 - \sum_{f=1}^{F} z_{\mathrm{pred},b,t,f} \right)^2 \quad (2)$$

The combined loss is then simply the sum (see eqn (3)):

$$\mathrm{MSE}_{\mathrm{comb}} = \lambda_{\mathrm{data}} \, \mathrm{MSE}_{\mathrm{data}} + \lambda_{\mathrm{conc}} \, \mathrm{MSE}_{\mathrm{conc}} \quad (3)$$

Table 1 Searched model and training parameters

Model parameter	CNN	LSTM	AR-LSTM
Number of filters	1–100	—	—
Kernel size	1–15	—	—
Number of hidden nodes/ units	10–100	10–100	10–100
Input sequence length	5–15	5–15	5–15
Output sequence length	5–15	5–15	5–15
Output activation function	“Linear”, “Sigmoid”, “ReLU”	“Linear”, “Sigmoid”, “ReLU”	“Linear”, “Sigmoid”, “ReLU”

Training parameter	Options
Batch size	10–40
Epochs	5–40
Learning rate	1×10^{-5}–1×10^{-1}

where λ_{data} and λ_{conc} are the respective tuning factors, that enable prioritization of one component if wanted. Due to the transformation *via* the PLS model and the measurement fluctuations of the FT-IR, the sum of the concentrations at each timestep is not strictly at 50 mM. Therefore, it cannot be expected that the model output is more accurate than the data it is developed from.

4.3. Model baseline

In time-series forecasting, common baseline models for generating multistep outputs are either repeating the last observed input time step for the desired number of output steps or simply replicating the previous sequence. As one might anticipate, given trends of increase or decrease, both variants tend to yield inaccurate results for long-term forecasts. Therefore, we propose two trend-replicating baseline models as a lightweight and interpretable reference for multivariate time-series forecasting. The first baseline (Fig. 8, top) assumes a consistent trend (*i.e.*, constant rate of change) across recent time steps and extrapolates linearly into the future based on the last observed value. The second baseline (Fig. 8, bottom) captures the short-term temporal dynamics by repeating the recent input pattern and linearly adjusts the repetition based on the average trend observed in the input.

4.4. Model evaluation

When evaluating the model outside of training, only the data loss is considered. Instead of averaging it per batch it is then averaged over the number of sequences in the test dataset. Beside the metrics per sequence, the final model must be accurate and reliable on long-term predictions. Therefore, two additional metrics will be used in the evaluation of the models. The most obvious metric is the average long-term forecast error (exemplarily shown in Fig. 9). This forecast emulates the actual use in the process, where the concentrations for the rest of the experiment are predicted after a certain warm-up time. The long-term forecast error is then calculated for each of three species over all samples (indicated by s) in eqn (4) like:

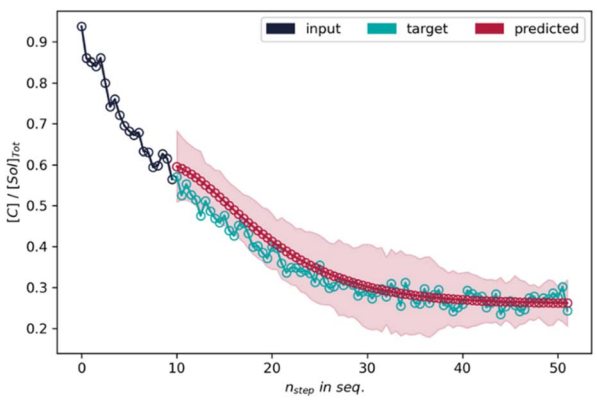

Fig. 9 Long-term forecast of a random experiment from the test dataset, for a single concentration.

$$\mathrm{MSE}_{\text{long-term},f} = \frac{1}{S}\sum_{s}^{S}\left(z_{\text{target},s,f} - z_{\text{pred},s,f}\right)^2 \quad (4)$$

where f indicates the species and z is one relative concentration at a certain timestep. One sample represents one predicted time step and S is the total number of samples in the test dataset. The average long-term forecast error is then simply the mean over the number of species.

The predicted concentrations in Fig. 9 also show prediction boundaries. These boundaries reflect the model's average error dependent on the number of forecasted time steps. These boundaries are calculated on the test dataset, in the following way: a long-term forecast is made as shown in Fig. 9, then the RMSE over all test sequences for each forecast step is calculated. The boundaries are then plotted by adding and subtracting the RMSE per forecast step to the actual forecast.

Next to the mean squared error of the sequence and of the long-term forecast we introduce the rolling window analysis to quantify the reliability of the models. If a model is used to regularly update the forecast in an ongoing process, multiple predictions at the same time can occur. The change of the input sequence will lead to a different output sequence, even for prediction at the same time in the output sequence. Predictions of a reliable and robust model will be within a small range, while predictions of an unreliable model will deviate significantly.

To quantify the reliability, the mean value of each rolling window standard deviation is calculated. The rolling window standard deviation is calculated for each experiment in the test dataset (for details see Part B of the SI).

5. Results and discussion

Fig. 10 shows an example forecast for the AR-LSTM model using data from the test dataset. As can be seen, the model demonstrates robust capability in predicting various chemical concentrations throughout the experiment, which was conducted with a constant applied current. Additionally, the model effectively maintains the chemical balance. In this test experiment, the electrode exhibited a relative corrosion level between 0 and 0.5, as quantified by the OCT-based corrosion feature. This indicates that the electrode surface was already partially degraded, providing a realistic scenario for evaluating the model's robustness to corrosion effects.

The development of the model, as well as the accuracy it can achieve, is significantly dependent on the amount and quality of the data available. The chemical concentrations derived from the PLS model exhibit natural fluctuations inherent in FT-IR measurements, which are challenging to capture accurately with a time-series model, if capturing them is even wanted. These fluctuations are also evident in the chemical balance (see Fig. 10), presenting additional complexity. Such random fluctuations pose a substantial challenge to modeling efforts, inherently limiting the achievable accuracy. However, this represents only a snapshot of the evaluation. The long-term forecast error also resembles the most challenging scenario in the application: predicting the entire experiment from the first

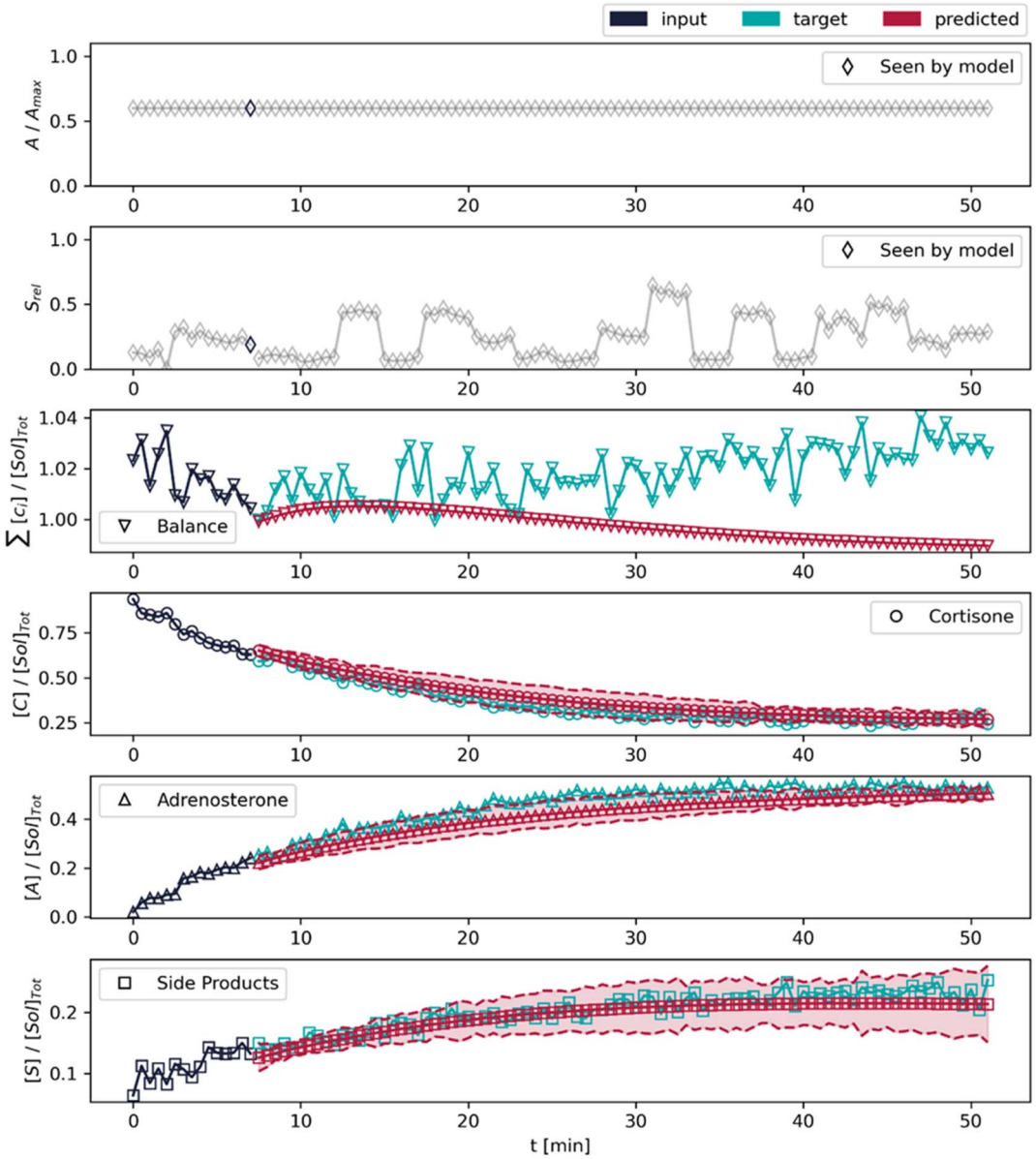

Fig. 10 Long-term forecast of a random experiment from the test dataset. The top row shows the applied current relative to the maximum. The second panel from the top shows the relative corrosion feature. The third panel from the top displays the conserved concentration. The panels below show the concentration for each species.

sequence alone. In real-world applications, forecasts are updated with every new measurement and further refined with each change in the applied current, providing more data to the model. This iterative updating process significantly enhances the model's predictive capability over time.

In the evaluation of model performance, two distinct setups were analyzed: one excluding the corrosion feature (Table 2) and one incorporating it (Table 3). In the absence of the corrosion feature, the AR-LSTM model outperformed the baseline, the CNN and the LSTM on the sequence. Specifically, the AR-LSTM model exhibited the lowest MSE and MAE for sequential data. For the long-term forecast the CNN achieved the lowest errors, with the LSTM and the AR-LSTM being close. Furthermore, it can be observed that the baseline models, which were assessed using 15 input and 9 output steps, corresponding to the mean values identified for other models, achieved sufficient results on the sequence. However, they exhibit significant limitations in forecasting accuracy and reliability. This outcome supports the hypothesis that a more sophisticated model is necessary for improved predictive performance. For all found model and training parameters see Part C of the SI.

Regarding reliability, the AR-LSTM performs significantly better than all others. However, one must consider that the reliability value is significantly influenced by the output sequence length. Specifically, the significantly smaller reliability STDs of 4.528×10^{-3} (Table 2) and 3.535×10^{-3} (Table 3), without and with the corrosion feature, respectively, confirm

Table 2 Model performance without corrosion feature

Metric	CNN	LSTM	AR-LSTM	Linear baseline	Pattern baseline
Sequence MSE	1.323×10^{-3}	9.645×10^{-4}	7.763×10^{-4}	1.538×10^{-3}	3.416×10^{-3}
Sequence MAE	2.724×10^{-2}	2.218×10^{-2}	2.009×10^{-2}	2.804×10^{-2}	4.376×10^{-2}
Forecast MSE	3.405×10^{-3}	4.462×10^{-3}	4.409×10^{-3}	2.633×10^{-1}	2.408×10^{-2}
Forecast MAE	4.759×10^{-2}	5.300×10^{-2}	5.324×10^{-2}	3.842×10^{-1}	1.251×10^{-1}
Reliability STD	1.282×10^{-2}	9.536×10^{-3}	4.528×10^{-3}	2.682×10^{-2}	1.700×10^{-2}

Table 3 Model performance including the corrosion feature

Metric	CNN	LSTM	AR-LSTM
Sequence MSE	1.179×10^{-3}	9.591×10^{-4}	8.095×10^{-4}
Sequence MAE	2.510×10^{-2}	2.259×10^{-2}	2.002×10^{-2}
Forecast MSE	3.463×10^{-3}	5.161×10^{-3}	3.736×10^{-3}
Forecast MAE	4.762×10^{-2}	5.692×10^{-2}	4.879×10^{-2}
Reliability STD	1.332×10^{-2}	9.902×10^{-3}	3.535×10^{-3}

that the AR-LSTM architecture is advantageous in terms of reliability over the other architectures.

Upon the inclusion of the corrosion feature, a notable improvement was observed in the sequence metrics for the CNN model, while the forecast metrics and the reliability remained at the same level. For the basic LSTM no significant change can be observed on the sequence task, the forecasting error and the reliability deviation increased. For the AR-LSTM the sequence errors slightly increased, while the forecast errors and the reliability deviation decreased, which is still a positive effect for the usage in the process.

Overall, both the CNN and AR-LSTM models exhibit the lowest sequence and forecasting errors, demonstrating nearly equivalent accuracy. However, the AR-LSTM model shows significantly lower deviation in its predictions over time. The inclusion of the corrosion feature enhances the performance of the AR-LSTM, whereas it does not necessarily benefit the CNN. In summary, the AR-LSTM emerges as the superior model for the process application, with the corrosion feature further improving its accuracy. Despite the challenges, the developed model can be described as accurate and valuable to the overall process monitoring approach. Its ability to adapt and improve with continuous data input underscores its potential utility in dynamic and complex experimental settings.

6. Conclusions

A real-time process monitoring strategy was presented to address issues like electrode corrosion and process deviations. The strategy was developed on the electrochemical synthesis of steroids, specifically the conversion of cortisone to adrenosterone. The chosen reaction is suitable because it proceeds in a single step, is safe to perform, and is relatively inexpensive. A total of 40 experiments were conducted. The experimental setup includes FT-IR spectroscopy, combined with PLS modeling for rapid quantification of chemical species, and OCT measurements to monitor the electrode surface.

The first step of the process monitoring strategy involves the processing of OCT images to quantify electrode corrosion. A preprocessing pipeline was introduced that can automatically and dynamically mask the area of interest on images of the electrode surface. This mask is recalculated for each scan and allows for the correlation of the number of visible pixels with the corrosion of the electrode surface. This novel method of dynamic mask generation and pixel quantification offers a robust and computationally efficient approach to monitoring electrode corrosion in real-time, representing a significant advancement.

In the second step, the paper introduces a machine learning-based time-series forecasting model that incorporates real-time sensor data, including applied current and (optionally) corrosion measurements, to predict reactant and product concentrations. Various neural network architectures, such CNNs and LSTMs, were optimized and evaluated for their forecasting capabilities. In addition to the corrosion detection itself, it was assessed whether an incorporation of the relative corrosion value into the forecasting model is beneficial. The models were evaluated on the sequence as well as on long-term forecasts. To enhance accuracy and integrate chemical reasoning, domain-inspired neural networks were employed. These networks incorporate a concentration conservation constraint into the loss function, ensuring the relative sum of concentrations remains constant over time.

The accuracy and development of the model are significantly influenced by the quantity and quality of the data. Inherent fluctuations in FT-IR measurements and deviations introduced by PLS modeling when converting spectra to actual concentration values, present challenges that limit the precision and achievable accuracy of the time-series model. Despite these challenges, the identified models prove to be accurate and valuable for the overall process monitoring and prediction strategy. The consideration of the corrosion feature was found to be beneficial for long-term forecasts.

Overall, the paper underscores the potential of integrating real-time monitoring with machine learning to develop robust and scalable electrochemical processes. This integration paves the way for more sustainable and efficient pharmaceutical production. In future applications, the developed model and corrosion detection system can be utilized as a framework for predictive process control. This optimization involves balancing productivity and electrode condition by automatically adjusting the applied current.

Author contributions

Josef Tausendschön: conceptualization, methodology, software, validation, formal analysis, data curation, writing – original draft, visualization. Michael Poelzl: methodology, software, validation, investigation, data curation, writing – original draft. Nikola Petrovic: investigation, writing – original draft, visualization. Jason D. Williams: methodology, software, writing – review & editing. Elisabeth Fink: methodology, software, writing – review & editing, project administration, funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Data availability

Experimental data for this article, are available at zenodo at https://doi.org/10.5281/zenodo.17158303. The code for preparation of the experimental data can be found at https://doi.org/10.5281/zenodo.17285101. The version of the code employed for this study is version 3. The code for training all described neural networks can be found at https://doi.org/10.5281/zenodo.17285101. The version of the code employed for this study is version 3. The trained as well as baseline models including code for can be found at https://doi.org/10.5281/zenodo.17285101. The version of the code employed for this study is version 3. A list of all performed experiments as well as statistics about parabola heights can be found in the supplementary information (SI). Supplementary information is available. See DOI: https://doi.org/10.1039/d5dd00458f.

Acknowledgements

This work was funded by the European Union under GA number 101058715. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them. Further, this work was funded through the Austrian Research Promotion Agency (FFG) as part of the "Twin4Pharma" project within the COMET Module program. The Research Center Pharmaceutical Engineering (RCPE) is funded within the framework of COMET – Competence Centers for Excellent Technologies by BMIMI, BMWET, Land Steiermark, and SFG. The COMET program is managed by the FFG.

Notes and references

1 M. Yan, Y. Kawamata and P. S. Baran, *Chem. Rev.*, 2017, **117**, 13230–13319.
2 D. Pollok and S. R. Waldvogel, *Chem. Sci.*, 2020, **11**, 12386–12400.
3 S. Garcia-Segura, J. D. Ocon and M. N. Chong, *Process Saf. Environ. Prot.*, 2018, **113**, 48–67.
4 C. Friebe, A. Lex-Balducci and U. S. Schubert, *ChemSusChem*, 2019, **12**, 4093–4115.
5 R. D. Little and K. D. Moeller, *Chem. Rev.*, 2018, **118**, 4483–4484.
6 D. Cantillo, *Chem. Commun.*, 2022, **58**, 619–628.
7 N. Petrović, B. K. Malviya, C. O. Kappe and D. Cantillo, *Org. Process Res. Dev.*, 2023, **27**, 2072–2081.
8 J. Luo, B. Çıtmacı, J. B. Jang, F. Abdullah, C. G. Morales-Guio and P. D. Christofides, *Chem. Eng. Res. Des.*, 2023, **197**, 721–737.
9 D. Manson, *Handbook of Steroids*, HAYLE MEDICAL, New York City, 2019.
10 F. Sommer, C. O. Kappe and D. Cantillo, *Chem.–Eur. J.*, 2021, **27**, 6044–6049.
11 S. Ke, *Mini-Rev. Med. Chem.*, 2018, **18**, 745–775.
12 G. Sedelmeier and J. Sedelmeier, *Chimia*, 2017, **71**, 730.
13 L. Reguera, C. I. Attorresi, J. A. Ramírez and D. G. Rivera, *Beilstein J. Org. Chem.*, 2019, **15**, 1236–1256.
14 S. Shaikh, H. Verma, N. Yadav, M. Jauhari and J. Bullangowda, *ISRN Anesthesiol.*, 2012, **2012**, 1–11.
15 P. Gupta and A. Mahajan, *Environ. Chem. Lett.*, 2019, **17**, 879–895.
16 H. R. Khatri, N. Carney, R. Rutkoski, B. Bhattarai and P. Nagorny, *Eur. J. Org. Chem.*, 2020, **2020**, 755–776.
17 L. Michel and M. Ueli Angst, *Measurement*, 2022, **201**, 111713.
18 J. Menzel, A. Slesinski, P. Galek, P. Bujewska, A. Kachmar, E. Frąckowiak, A. Washio, H. Yamamoto, M. Ishikawa and K. Fic, *Energy Storage Mater.*, 2022, **49**, 518–528.
19 R. Wang, H. Zhang, Q. Liu, F. Liu, X. Han, X. Liu, K. Li, G. Xiao, J. Albert, X. Lu and T. Guo, *Nat. Commun.*, 2022, **13**, 547.
20 M. He, K. Fic, E. Frackowiak, P. Novák and E. J. Berg, *Energy Environ. Sci.*, 2016, **9**, 623–633.
21 A. G. Star and T. F. Fuller, *J. Electrochem. Soc.*, 2017, **164**, F901–F907.
22 X. Shan, U. Patel, S. Wang, R. Iglesias and N. Tao, *Science*, 2010, **327**, 1363–1366.
23 M. A. Morin, W. Zhang, D. Mallik and M. G. Organ, *Angew. Chem., Int. Ed.*, 2021, **60**, 20606–20626.
24 P. Sagmeister, R. Lebl, I. Castillo, J. Rehrl, M. Horn, J. Kruisz, M. Sipek, S. Sacher, D. Cantillo, J. Williams and C. O. Kappe, *Angew. Chem., Int. Ed.*, 2021, **60**, 8139–8148.
25 E. Fink, M. Y. S. Ibrahim, M. Wolfgang, C. O. Kappe and D. Cantillo, *ACS Electrochem.*, 2025, **1**(10), DOI: 10.1021/acselectrochem.5c00249.
26 L. I. Stephens and J. Mauzeroll, *J. Chem. Educ.*, 2019, **96**, 2217–2224.
27 F. F. Rivera, T. Pérez, L. F. Castañeda and J. L. Nava, *Chem. Eng. Sci.*, 2021, **239**, DOI: 10.1016/j.ces.2021.116622.
28 M. Vasanthapandiyan, S. Singh, F. Bononi, O. Andreussi and N. Karmodak, *J. Chem. Phys.*, 2023, **159**, DOI: 10.1063/5.0165835.
29 L. F. Catañeda, F. F. Rivera, T. Pérez and J. L. Nava, *Curr. Opin. Electrochem.*, 2019, **16**, 75–82.
30 H. T. Moreno and A. Schaum, in *IFAC-PapersOnLine*, Elsevier B.V., 2022, vol. 55, pp. 103–108.

31 K. Liu, Y. Gao, C. Zhu, K. Li, M. Fei, C. Peng, X. Zhang and Q. L. Han, *Control Eng. Pract.*, 2022, **124**, DOI: 10.1016/j.conengprac.2022.105176.
32 Y. Mukouyama and S. Nakanishi, *Front. Energy Res.*, 2020, **8**, 111713.
33 R. Chandra, S. Goyal and R. Gupta, *IEEE Access*, 2021, **9**, 83105–83123.
34 B. Lim and S. Zohren, *Philos. Trans. R. Soc., A*, 2021, **379**(2194), DOI: 10.1098/rsta.2020.0209.
35 A. Vaswani, G. Brain, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, *arXiv*, 2023, preprint, arXiv.1706.03762, DOI: 10.48550/arXiv.1706.03762.
36 D. Salinas, M. Bohlke-Schneider, L. Callot, R. Medico and J. Gasthaus, 2019, *arXiv*, preprint, arXiv.1910.03002, DOI: 10.48550/arXiv.1910.03002.
37 R. Sen, H.-F. Yu and I. Dhillon, *arXiv*, 2019, preprint, arXiv.1905.03806, DOI: 10.48550/arXiv.1905.03806.
38 J. L. Hudson, A. Kube, R. A. Adomaitis, I. G. Kevrekidis, A. S. Lapedes and R. M. Farber, *Chem. Eng. Sci.*, 1990, **45**, 2075–2081.
39 Y. Zhang, Q. Tang, Y. Zhang, J. Wang, U. Stimming and A. A. Lee, *Nat. Commun.*, 2020, **11**, 1706.
40 J. C. Davies, D. Pattison and J. D. Hirst, *J. Mol. Graphics Modell.*, 2023, 108356.
41 H. Ji, J. C. Liu and H. Ye, *Nondestr. Test. Eval.*, 2025, **40**(12), DOI: 10.1080/10589759.2025.2456668.
42 L. Chen, Y. Wu, Y. Liu, T. Liu and X. Sheng, *IEEE Access*, 2021, **9**, 10699–10710.
43 N. Petrović, B. K. Malviya, C. O. Kappe and D. Cantillo, *Org. Process Res. Dev.*, 2023, **27**, 2072–2081.
44 N. Petrović, G. R. Cumming, C. A. Hone, M. J. Nieves-Remacha, P. García-Losada, Ó. de Frutos, C. O. Kappe and D. Cantillo, *Org. Process Res. Dev.*, 2024, **28**, 2928–2934.
45 S. Schippling, in *Optical Coherence Tomography in Neurologic Diseases*, Cambridge University Press, 2015, pp. 4–13.
46 J. M. Schmitt, S. H. Xiang and K. M. Yung, *J. Biomed. Opt.*, 1999, 95–105.
47 Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang and F. Shi, *Biomed. Opt. Express*, 2018, **9**, 5129.
48 S. V. Mohd Sagheer and S. N. George, *Biomed. Signal Process. Control*, 2020, 102036.
49 G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang and L. Yang, *Nat. Rev. Phys.*, 2021, **3**, 422–440.
50 J. Tausendschön, E. Fink, N. Petrovic and J. Williams, *s-x-AIPI Pharma Use Case Experimental Data (Dataset)*, 2025, DOI: 10.5281/zenodo.10663999.
51 P. Remy, *GitHub repository*, 2020, https://github.com/philipperemy/cond_rnn/.
52 T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama, in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, 2019, pp. 2623–2631.
53 D. P. Kingma and J. Ba, *arXiv*, 2017, preprint, arXiv.1412.6980, DOI: 10.48550/arXiv.1412.6980.