Version of Record: https://www.sciencedirect.com/science/article/pii/S0263876223000667
Manuscript_e7815c19463f83be934b02faa23a5b7c

Machine learning-based ethylene and carbon monoxide estimation, real-time optimization, and multivariable feedback control of an experimental electrochemical reactor

Berkay Çıtmacı[a], Junwei Luo[a], Joon Baek Jang[a], Carlos G. Morales-Guio*[,a], Panagiotis D. Christofides*[,a,b]

[a]*Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA, 90095-1592, USA.*
[b]*Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095-1592, USA.*

Abstract

Electrochemical reduction of CO_2 gas is a novel CO_2 utilization technique that has the potential to mitigate the global climate crisis caused by anthropogenic CO_2 emissions, and enable the large-scale storage of energy generated from renewable sources in the form of carbon-based chemicals and fuels. However, due to the complexity of the electrochemical reactions, the explicit first-principle models for CO_2 reduction are not available yet, and there has been a limited effort to develop process modeling, optimization and control of CO_2 electrochemical reactors. To this end, a rotating cylinder electrode (RCE) reactor has been constructed at UCLA to understand the mass transfer and reaction kinetics effects separately on the productivity. In the RCE reactor, the applied potential strongly influences the reaction energetics and the electrode rotation speed affects the hydrodynamic boundary layer and modifies the film mass transfer coefficient, which involves convective and diffusive transport. The present work aims to develop a multi-input multi-output (MIMO) control scheme for the RCE reactor that integrates techniques from artificial and recurrent neural network modeling, nonlinear optimization, and process controller design. Specifically, production rates of two products from the experimental reactor, ethylene and carbon monoxide, are controlled by manipulating two inputs, applied potential and catalyst rotation speed. Process dynamics and controllability are analyzed, a feedback control strategy is designed and the controllers are tuned accordingly. The experimental electrochemical cell is employed to gather data for process modeling

*Co-corresponding authors; E-mails: moralesguio@ucla.edu, pdc@seas.ucla.edu

and implement the multivariable control system. Finally, the experimental results are presented which demonstrate excellent closed-loop performance by the control system and regulation of the outputs at three different set-points including an economically-optimal set-point.

Key words: Electrochemical CO_2 reduction; Multi-input multiple-output control; Experimental data modeling; Real-time optimization; Neural network modeling.

1. Introduction

The percentage of renewable energy in the electricity grid has increased as decarbonization efforts have gained momentum against the detrimental effects of global warming. With current advances and the increasing popularity of sustainability, falling cost and increasing availability of renewable electricity generation, electrochemical methods have become an attractive alternative for transforming CO_2 gas into organic chemicals and synthetic fuels (De Luna et al. (2019)). However, the overall reaction mechanisms of this transformation have not been fully understood and limited efforts have been made to build dynamic models to understand and control this electrochemical process. Moreover, the applications of CO_2 reduction (CO_2R) have not gone beyond the bench scale (e.g., Jang et al. (2022)). The major bottleneck of the industrial implementation of electrochemical CO_2 reduction technology is the deconvolution of intrinsic kinetics from mass, heat, and charge transport effects, which has prevented the development of accurate reaction mechanisms (Jin et al., 2021). To further explore the fundamentals of electrochemical CO_2 reduction, a gastight rotating cylinder electrode (RCE) cell was recently developed, which can decouple the effects of mass transfer from surface reaction kinetics (Jang et al., 2022). This novel electrochemical reactor shown in Fig. 1 has demonstrated that mass transport phenomena and intrinsic reaction kinetics can independently affect the productivity and selectivity of electrochemical CO_2R, which implies the potential to control the product distribution of the reaction by manipulating multiple inputs. To understand the hydrodynamics effects on the very thin (μm scale) catalyst surface boundary layer, which is ultimately crucial, Richard et al. (2023) simulated the gastight RCE reactor using computational fluid dynamics (CFD) software. Even though there are such endeavors to capture the mechanism of CO_2R in RCE cells in detail, there is no dynamic model available yet.

In the absence of steady state or dynamic process models such as first principles-based models that rely on known physical relations, it is possible to build models using data-driven approaches, such as gray box or black box models that give an output for a corresponding input without exposing correlations. The involvement of machine learning (ML) models in electrolyzers has attracted attention over the last decade due to their capability to approximate nonlinearities with no prior physical information of the system (Corriou, 2004). The use of ML in electrochemistry has recently received attention and has been used in predicting the next generation of catalysts without earmarking major budgets and time for experiments with different material combinations. For instance, Timoshenko et al. (2020) used a neural network trained with X-ray absorption fine structure spectroscopy data to reproduce the rate and time based structural changes of the catalyst under CO_2 reduction. In another work, Chen et al. (2020) built a catalyst database using density functional theory (DTF) simulations and used these data to build an extreme gradient boosting (XGBoost) regression model to predict the change in Gibbs free energies in CO adsorption to find the most feasible CO_2 reduction electrocatalyst among more than 1000 combinations of metals and nonmetals. However, there have been limited efforts to dynamically model the electrochemical reactions and advanced ML methods such as recurrent neural networks (RNN) should be used in the electrolyzer context to capture time-dependent process relations.

The use of RNNs has been becoming widely common as they are very promising for leveraging the process data for various applications (Hussain, 1999). RNNs have been used successfully in modeling various processes, and they can be very efficient in modeling the dynamic behavior of electrochemical reactions. The ability of RNNs to learn time dependencies makes this approach an alternative to first-principle models, as they can capture the trends in data emerging from the behavior of the process with respect to variation in inputs. This enables RNNs to be used in process control tasks; especially highly nonlinear aspects can be learned due to the nonlinear activation functions in the hidden layers of the neural network structure (Wu et al. (2020)). This would also be extremely valuable to control a process that does not have a dynamic first-principle model. Cheng et al. (1995) is one of the early examples of using neural networks before RNN architectures such as long short-term memory (LSTM) became widely popular to model dynamic processes with

long and variable dead times. The study successfully models a pH neutralization process and uses internal recurrent neural networks (IRN) with variable dead times, which feed back the calculation from the hidden node as an input, behaving like a one time step delay. Tian et al. (2001) modeled the dynamic behavior of a batch methyl methacrylate (MMA) polymerization reactor with a hybrid stack of RNN models. Using this model, an effective feedback control scheme was implemented to regulate the temperature, and in turn, key process variables like monomer conversion. In general, ML models that can represent transient behavior can be used for process control system design and implementation. In this direction, Ren et al. (2022) summarized how to incorporate various ML models into a model predictive controller (MPC). For example, Wu et al. (2021) introduced Monte-Carlo dropout method to the LSTM training case to improve the modeling performance. In addition to that, the co-teaching method was employed to include ideal first-principles model data in the training for a better performance. Then, these RNN models were incorporated into an MPC to simulate the control performance of a CSTR. Khalid et al. (1993) used RNNs to create a multi-input-multi-output (MIMO) control scheme for a testbed furnace temperatures in which the weights of the RNN model are adopted as the operation proceeds. This neuro-controller was shown to be successful in setpoint tracking and against disturbances. With the advancements in sensor technologies and tools to digitalize experimental systems, feedback control with ML models can be implemented in a smarter manner.

Our previous work on the CO_2 reduction process incorporates smart manufacturing techniques into the experimental field to fully automate and digitalize the setup to leverage the potential for voluminous data production from multiple sensors to accelerate the experimental procedures and contribute to the scale-up efforts. Specifically, Çıtmacı et al. (2022b) summarized the efforts to connect UCLA's experimental RCE reactor to the Clean Energy Smart Manufacturing Institute's (CESMII) Smart Manufacturing Innovation Platform (SMIP) to securely store, organize and contextualize data generated during the experiments as well as meta information of the setup components. SMIP can also be used as a deployment environment for data-driven models and control, to monitor real-time data, and to extract correlations between experimental parameters. In addition to that, Çıtmacı et al. (2022b) elaborates on the RCE's sensors and how these are upgraded to smart sen-

sors using the available data. One example is the automated gas chromatogram (GC) code, which fully automates a manual procedure by imitating the steps followed by the experiment supervisor. Consequently, the experimental setup becomes more compact and efficient.

There are previous data modeling and control efforts for the RCE setup with a different catalyst electrode, which are used as a base for this work. Luo et al. (2022) explains in detail how statistical feedforward neural network (FNN) architectures can be used to model the steady state operation points given the initial experimental input parameters. This model is trained with a database of open-loop steady state experiments conducted over the past years, accounting for the experimental uncertainties. This work also presents a method to reciprocally use empirical first-principal models with the developed ML models to improve the experimentally extracted correlations. Finally, Çıtmacı et al. (2022a) used support vector regression (SVR) to model real-time reaction rates for ethylene and inserted the SVR model into dynamic mass balance equations to implement a feedback control scheme at economically optimized setpoints. The SVR model constructed in this work also accounts for a fast catalyst decay and uses delayed feedback concentrations from the GC to update the model. Thus, this work has been the first successful single-input-single-output (SISO) control instance of the complex electrochemical CO_2 reduction process using ML methods.

Motivated by the above considerations, this work proposes an ML-based scheme to implement real-time optimization (RTO) and multivariable feedback control in an experimental electrochemical reactor for CO_2 reduction. Specifically, two RNN models are developed on the basis of existing experimental data to estimate the dynamic response of the reactor operation. Subsequently, the information from the sensors and RNN models are integrated and used by two Proportional-Integral (PI) controllers that manipulate two inputs, applied voltage and electrode rotation speed, to the electrochemical cell, constructing a MIMO control scheme. In addition, an operational steady state model and an RTO are developed to calculate the economically optimum setpoints for the ethylene and carbon monoxide production rates by integrating market information. The proposed control and optimization scheme is demonstrated by a series of experiments that control the production of ethylene and carbon monoxide from the RCE cell.

We believe the demonstration of a MIMO feedback control and RTO using this bench-scale

reactor will be critical for the realization and operation of an industrial-scale CO_2 electrolyzer for the following reasons. First, moving forward from SISO to MIMO control and generalizing the approach to account for different combinations of inputs and outputs is essential for reactions involving multiple products such as the CO_2R. In addition, due to the non-selective nature of the reaction, there is a limit to solving this problem solely through the development of catalysts or the design of single-unit reactors, and efforts should be made to integrate multiple processes. Such process integration approaches include not only the connection of upstream/downstream separation or conversion processes, but also the design of a multi-stage electrolyzer sequence (Ramdin et al., 2021). Real-time feedback control of a single-unit reactor demonstrated in this work acts as a building block for constructing a process control network where its communication with neighboring units is extremely important. Finally, the implementation of RTO which optimizes the overall process but can be applied to the actuation of individual control units is critical considering how tightly the control of a single-unit reactor is tied to other processes and the electric grid infrastructure.

The rest of this manuscript is organized as follows. In the section entitled "Preliminaries", the experimental reactor setup and database generation for ML model training are described. In the next section, entitled "Modeling and Optimization of the Experimental Electrochemical Process", the construction of a steady model using FNNs and the real-time optimization of this model are discussed. In the following section entitled "Dynamic Modeling of the Experimental Electrochemical Process", the methodology for enhancing our dataset to train a dynamic RNN model is elucidated. The section entitled "MIMO Control Architecture and Controller Tuning" elaborates on the implementation of the two PI controllers, their integration with the RNN models and the tuning of the controllers. Finally, the performances of the models and the controllers are evaluated.

2. Preliminaries

This section presents the details of the experimental setup used in this work. All of the experimental devices apart from nuclear magnetic resonance (NMR) are digitized via Laboratory Virtual Instrument Engineering Workbench (*LabVIEW*) software. The experimental process employs a code for processing GC signals, controllers and actuators, and it is fully automated with Python

scripts that are integrated into a *LabVIEW* interface, in addition to the ML models that predict the concentrations at each second. Fig. 1 shows the experimental RCE reactor at UCLA. The specific feedback controller parameters are determined in advance. The control system implementations include closed-loop experiments with arbitrary starting concentrations, which are driven to the setpoints. The process data flow is connected to the database of *SMIP* provided by *CESMII*.

2.1. Process Overview

Electrochemical CO_2 reduction on copper is a complex process. There are 17 chemicals produced and their reaction pathways are complicated because processes of different time scales, including mass and charge transfer, adsorption and desorption, and surface reaction, are convoluted involving multiple reaction intermediates. Mass transport characteristics of an electrochemical system affect the transfer of reactant to the catalyst surface as well as the removal of intermediates and products away from the surface. The relative time scales of different processes in the overall reaction can be realized and controlled systematically in our RCE reactor (Jang et al. (2022)). Among various products generated from this electrochemical reactor from CO_2R on polycrystalline Cu, hydrogen (H_2), carbon monoxide (CO), methane (CH_4), and ethylene (C_2H_4) are in the gas phase and can be detected using GC. The relevant reactions for these products are shown below:

$$2CO_2 + 8H_2O + 12e^- \rightarrow C_2H_4 + 12OH^- \tag{1a}$$

$$CO_2 + H_2O + 2e^- \rightarrow CO + 2OH^- \tag{1b}$$

$$CO_2 + 6H_2O + 8e^- \rightarrow CH_4 + 8OH^- \tag{1c}$$

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \tag{1d}$$

In our previous works (Çıtmacı et al. (2022a) and Luo et al. (2022)), we used electropolished atomically-flat polycrystalline copper cylinder electrodes as CO_2R catalyst. The smooth nature of the exposed electrode surface caused a fast catalyst deactivation throughout the experiments, and various modeling and control approaches were applied to handle this challenge. In this study,

we use the same experimental setup with nanoporous copper cylinder electrodes. Nanoporous structure directly synthesized on Cu cylinders increases the roughness of the surface with higher electrochemically active surface area (ECSA) inside the pores (Roberts et al. (2015)) and is more resistive to catalyst deactivation. Inside the pores, internal pore diffusion is the dominant mode of mass transport where all species have a long residence time and the electrochemical environment (e.g., concentration, pH, and electrical potential) becomes highly localized. The high ECSA as well as the longer residence time of intermediates due to internal pore diffusion shift the selectivity towards producing more multicarbon (C_{2+}) products. To account for this change in selectivity, new modeling and control approaches are applied in this work.

2.2. Experimental Setup

The RCE system consists of two electrode chambers divided by an anion-exchange membrane, a mass flow controller (MFC), a potentiostat, a temperature control block, and a modulated speed rotator (MSR). During the experiment, pure CO_2 gas is fed at a fixed mass flowrate at 20 $mL \cdot min^{-1}$ into both the cathode chamber, where nanoporous Cu cylindrical electrode is rotating in 0.2 M $KHCO_3$ electrolyte solution, and the anode (Pt foil) chamber. CO_2 and H_2O molecules are transformed into 12 liquid-phase and 5 gas-phase (H_2, CO, CH_4, C_2H_4, and C_2H_6) products. Hydrodynamics and convective mass transport can be regulated systemically through the control of electrode rotation speed actuated by the MSR. Furthermore, the potentiostat can set the applied potential on the working electrode, and measurements are taken using Ag/AgCl as a reference electrode. Thus, the reaction kinetics and diffusion effects can be deconvoluted by running experiments at multiple applied potential and electrode rotation speeds. Finally, the electrochemical cell is hermetic so that gas phase products can be quantified by a gas chromatogram (GC) in real time. An automated GC code is written for triggering injections, peak detection, baselining, and calculation of the areas under the peaks to quantify the gas phase concentrations in ppm using available calibration data, as explained in Çıtmacı et al. (2022b). One GC injection takes 14.3 minutes to complete, and is followed by 6 minutes of cool down before the following GC injection. Thus, when a GC measurement is obtained, it is delayed and is related to the reactor overhead gas concentrations from 14.3 minutes ago. Liquid phase products accumulate in the electrolyte solution and are measured by

NMR at the end of the experiment.

In this work, the main output of the reactor is the production rates, denoted by $r_{C_2H_4\,|\,CO}$ for C_2H_4 and CO. The GC measures the concentrations in ppm and these concentrations are converted to production rates via the following equation.

$$r_{C_2H_4\,|\,CO} = \frac{C^{ppm}_{C_2H_4\,|\,CO}}{10^6} \times \frac{\dot{V}_{\mathrm{CO_2}}}{60 \times 10^6} \times \frac{P}{RT} \tag{2}$$

where $C^{ppm}_{C_2H_4\,|\,CO}$ is the concentration of C_2H_4 or CO measured by the GC in ppm, $\dot{V}_{\mathrm{CO_2}}$ is the CO_2 inlet flowrate in $mL \cdot min^{-1}$ at standard temperature and pressure (STP), P is the standard pressure at 1 atm, R is the universal gas constant and T is the standard temperature at 0 $°C$. The GC takes a fixed volume of gas (for example, 1 mL) at atmospheric pressure. Since all the terms except the concentration on the right hand side of Eq. 2 are constants, the production rates are proportional to the concentration in ppm.

This experimental setup is automated and digitalized as explained in detail in our previous work (Çıtmacı et al. (2022b)). Before the digitalization and automation efforts started, there were already accumulated open-loop steady state experimental data obtained under different input parameters, which were also sent to the SMIP. These open-loop steady state experiments were conducted under a fixed applied potential (V vs Ag/AgCl) and catalyst rotation speed (RPM) and the setup was operated until the system reached a steady state. During the experiments, the gas phase concentrations are measured via GC at 15^{th}, 35^{th}, 55^{th}, and 75^{th} minute, and the resulting current (A) and a calculated variable surface potential (V vs SHE) is sensed and recorded each second. The surface potential is the remaining potential across the surface of the catalyst electrode after accounting for the Ohmic drop in the electrolyte due to solution resistance and it is the more relevant type of potential parameter as it affects the charge transfer on the surface of the catalyst electrode. The surface potential (V vs SHE) is calculated as follows:

$$E_{\mathrm{surface}} = E_{\mathrm{applied}} - i \times R + E_0 \tag{3}$$

where E_{surface} is the surface potential, E_{applied} is the applied potential measured against the reference

electrode, i is the electrical current, and R is the solution resistance between the working electrode and the reference electrode measured by electrochemical impedance spectroscopy (EIS) (Jang et al. (2022)). E_0 is the standard reduction potential of the reference electrode used (Ag/AgCl/1 M KCl), so Eq. 3 removes the potential drop across the solution due to the resistance to ion transport in these systems.

Remark 1. Experimental conditions may cause uncertainty, therefore experimental data modeling brings some unique challenges compared to well-structured data, such as data generated by simulations or obtained from industrial facilities. In the RCE setup, despite the electrolyte resistance being kept as steady as possible ($6.2 \pm 0.2\ \Omega$), the resistance values are measured a priori, and may vary for each experiment based on stock solution preparation, environmental temperature, connectivity of inner electrical circuit, etc. Also, the differences in roughness factors that emerge from doing cyclic voltammetry during the preparation of the porous electrode may affect the catalyst activity, especially at lower overpotentials. Consequently, it is possible to observe different product concentrations at the same input conditions. Therefore, the averages and standard deviations from the steady state open-loop experimental data are shown in Table 1 and are used in the ML model construction. Also, a plot is shown in Fig. 3 for three experiments conducted at 100 RPM and in the close vicinity of -1.317 V vs SHE that have nearly overlapping current flows but produce C_2H_4 concentrations within a standard deviation of 43.4 ppm.

The existing experimental data were collected under various applied potentials at the electrode rotation speed of 100 and 800 RPMs as shown in Table 1. Also, an empirical correlation for estimating the current value at a specific surface potential and electrode rotation speed is developed using the steady state experiments and is as follows:

$$i = i_0 \Omega^{0.203} e^{kE_{surface}} \tag{4}$$

where i_0 and k are constants and Ω is the rotation speed. In order to find i_0 and k, Eq. 4 is linearized into the following form.

$$\ln \frac{i}{\Omega^{0.203}} = kE_{surface} + ln(i_0) \tag{5}$$

Average values of the experimentally measured current and surface potential and rotation speeds are linearly fitted into Eq. 5 to find the constants. The distribution of the fitted data is shown in Fig. 2. The following open- and closed-loop experiments, and data shifts can be explained, and this experimental system can be simulated using this empirical correlation.

2.3. Open-Loop Step Change Experiments

The modeling objectives of this study are C_2H_4 and CO concentrations; therefore, it is important to mention the trend with respect to the changes in one of the parameters while keeping the other one constant, where the parameters are the applied potential and the rotation speed of the catalyst. C_2H_4 has a very strong correlation with the surface potential and as the applied/or surface potential increases, C_2H_4 increases strongly. On the other hand, CO concentrations have a very strong correlation with the rotation speed, so that increasing the rotation speed results in increasing CO concentrations. These effects can be seen from our steady state machine learning model discussed in the following section and the cross effects will be discussed in detail.

In order to control the experimental setup, dynamic data must be generated in addition to the steady state data. Following the steady state experiments, various applied potential and rotation speed step change experiments were conducted separately. The timing of the GC injections is the same as the previous open-loop steady state experiments, and it started at the 15^{th} minute with a sampling period of 20.3 minutes. Based on the steady state experiments, it takes a maximum of 5 GC injections (around 80 minutes) to reach the new steady state and stay there. However, to see the shorter-term effects, there are a few experiments in which the step change is applied at 3 injections time. One example of a step change experiment is starting the experiment under a fixed applied potential and an initial rotation (e.g., at 100 RPM) and changing the rotation speed to 200 RPM after 5 GC injections and then changing it to 800 RPM until the new steady state is reached. Then, the reverse procedure is applied such that the rotation speed is first reduced to 200 RPM from 800 RPM and then to 100 RPM. A similar procedure is applied for the surface potential, in which the applied potential is manipulated to adjust the surface potential to the desired value. The surface potential values are increased/decreased, and the changes in gas phase concentrations are recorded. One instance of rotation speed step change experiments under a constant applied potential for CO

concentration is shown in Fig. 4. It is important to note that changing the rotation also affects the current passed between electrodes and thus affects the surface potential. Consequently, when the rotation speed is changed in an open-loop step change experiment, the surface potential does not remain constant and is subject to small changes. However, this effect is relatively small.

3. Modeling and Optimization of the Experimental Electrochemical Process

3.1. FNN Modeling for Steady State Setpoints

After selecting the controlled outputs and manipulated inputs, a process model needs to be developed to capture the input-output relationship and used for the design and tuning of the control system. Ideally, a mathematical model with an explicit form (e.g., first-principles model) is the best modeling option because of its explainability and reliability. However, in this work, it is challenging to develop such a model due to the complexity of the reaction mechanisms and the lack of full understanding of the electrochemical reactions. As summarized in Nitopi et al. (2019), there are several articles proposing respective unique explanations of the reaction mechanisms for this reaction. Therefore, there is not a single solid conclusion of the reaction chemistry that can be used to develop a first-principles model.

To address this challenge, a data-driven model is developed. Specifically, a feed-forward neural network (FNN) is trained based on the experimental data collected from the steady state experiments discussed in Section 2.3. There are seventeen products coming out from the electrochemical reactor. The FNN model uses two inputs (i.e., surface potential and catalyst rotation speed) to predict the production rate of sixteen product species. Hydroxyacetone production rates are not included in the modeling phase since its production rates are either 0 or very low (under 2 ppm). The inputs are normalized with a standard scaling factor. The FNN model has a hidden layer with 64 neurons activated by a ReLu function. The Softplus function, $f(x) = ln(e^x + 1)$, is selected as the activation function of the output layer to ensure non-negative prediction since the reactor did not consume any of the product such that the output of our FNN model can not be negative. Additionally, the Softplus function predicts the output with a smoother curve which aligns the physical expectation better than other candidate activation functions, such as ReLu and Sigmoid.

The mean squared error (MSE) function is utilized as a cost function to train the FNN model as explained in Luo et al. (2022).

Although an FNN model has the potential to provide a universal approximation to any non-linear relation (Scarselli and Tsoi, 1998) and has demonstrated reliable performance in addressing chemical engineering and process control tasks (Kramer and Morgado-Dias, 2018; Mohanty, 2009), it nominally treats each training data point equally, and therefore, it can lead to relationships that are affected by data points that have significant experimental variability. To account for this issue, we calculated the coefficient of variance for each data point, based on their respective mean and variance, and used it as the weight of the specific data point in the training process to account for data uncertainty (Luo et al., 2022). Specifically, the loss function used to train the weighted-FNN model can be expressed as follows:

$$Loss = \frac{1}{d}\frac{1}{m}\sum_{i=1}^{d}\sum_{j=1}^{m}\frac{1}{v_{i,j}^2}\left|y_{i,j} - \hat{y}_{i,j}\right|^2 \tag{6}$$

where d is the number of training data points, m is the number of output states, $y_{i,j}$ is the i^{th} reference data point for the j^{th} product, and $\hat{y}_{i,j}$ is the predicted production rate for the j^{th} product under the i^{th} input combination. $v_{i,j}$ is the coefficient of variance of the i^{th} data point for the j^{th} product, which can normalize the variability of each data point and provide unbiased weight for products having production rates in different magnitudes. With the weighted loss function, the FNN model is granted more tolerance for prediction error to prevent the model from overfitting the data uncertainty when a training data point has a higher variance. The visualization of the FNN model can be seen in Fig. 5.

Remark 2. Although the production rate of the liquid phase product cannot be measured in real-time during the experiment, it can be calculated based on the result of the NMR after the experiment is complete. Therefore the steady state neural network model can predict the production rate of the liquid-phase product but cannot be implemented in real-time with the dynamic model.

Twelve averaged experimental data points representing surface potentials, rotation speeds and corresponding 16 product output concentrations from 36 steady state experiments were used for

training and testing of the model. Nine of the experiments are earmarked for training and 2 were used for testing. There are 4 experimental points for each 100 RPM and 800 RPM experiments, 2 experimental points for the 400 RPM experiments, and one experimental point for 200 and 600 RPM experiments, each. The model is trained for 3000 epochs based on a mean squared error (mse) loss function. The mse values are 0.0055 and 0.0063 for training and testing. The results for gas phase products are shown in Fig. 6. Ethane is not included in the plots since its concentration does not go higher than 3 ppm under our operation range and thus it is very low. The hydrogen concentrations in Fig. 6 (a) show a linearly increasing trend with increasing surface potential. The rotation speed has a small increasing effect on the concentrations. Methane concentrations shown in Fig. 6 (b) are exponentially increasing with increasing surface potentials at higher overpotentials, and the effect of rotation speed is more significant at higher overpotentials. CO concentrations shown in Fig. 6 (c) show a very strong proportional correlation with rotation speed. However, at lower rotation speeds, an increase in surface potential results in a decrease in CO concentrations. At higher rotation speeds, the CO concentrations are affected very weakly by the surface potentials. Finally, the ethylene concentrations shown in Fig. 6 (d) exhibit exponentially increasing trends with increasing surface potential. The effect of rotation speed on the ethylene concentrations is small, and the direction depends on the specific rotation speeds.

3.2. Real-Time Optimization

The developed FNN model based on weighted data can predict the reactor performance by mapping combinations of control actions with the production rates of each species produced by the reactor. The next step is to apply this information to our multi-variable control scheme. Specifically, the prediction from the FNN model is used to solve an optimization problem computing the optimum setpoint for the multivariable control system. The optimization problem is designed to maximize the economic benefit of operating the reactor. To simplify the optimization problem for this study, electricity consumption is assumed to be the only operational cost for our reactor and the revenue of the operation is the total value of the generated product calculated based on the prediction of the FNN model. Thus, the optimum setpoint is where the reactor profit is maximized. The mathematical expression of this optimization problem is given below and a third-party software

IPOPT (Wächter and Biegler, 2006) is utilized to solve this optimization problem. Specifically, the optimization problem has the form:

$$\mathcal{J} = \underset{\hat{x}\in\mathbf{D}}{\arg\max} \;\; R(\hat{x}, E_{surface}) - C(I, E_{surface}) \tag{7a}$$

$$s.t. \quad F_{nn}(E_{surface}, \Omega) = \hat{x} \tag{7b}$$

$$C(E_{surface}, I) = c_e \times I \times E_{surface} \tag{7c}$$

$$R(\hat{x}, E_{surface}) = \sum_{j=1}^{m} c_j \times \hat{x}_j \tag{7d}$$

$$100 \leq \Omega \leq 800 \tag{7e}$$

$$-1.3 \leq E_{surface} \leq -1.19 \tag{7f}$$

In the above equations, functions $C(E_{surface}, \Omega)$ and $R(E_{surface}, \Omega)$ calculate the cost and revenue of operating the reactor, where $E_{surface}$, Ω, I, c_i, and c_e stand for the surface potential, the rotation speed of the catalyst in the unit of rotations per minute (RPM), current (A), the market price of the j^{th} product, and the electricity price respectively. $\hat{x}$ is the prediction from the FNN model containing production rates for $m = 16$ product species. When solving for the optimum setpoint, the initial guess for control actions (e.g., $E_{surface}$ and Ω) is first made by the users and provided to the optimization problem among the product and electricity prices. Subsequently, the IPOPT will alter the control actions, which leads to a change in the energy consumption and the production rates predicted by the FNN model, to maximize the profit of the reactor. Once the optimum control actions are found, the corresponding production rate of C_2H_4 and CO, given by the FNN model, will be sent as the setpoint for the multivariable control system. As an example, we assumed that the electricity price is the only varying price parameter, and by picking a value of 0.066 $\$/kWh$, the most economically feasible setpoints are found to be 112 ppm for C_2H_4 and 490 ppm for CO. For the second case scenario, a 40% price decrease for H_2 and C_1 products (e.g., CO) and a 60% price increase in C_2 products (e.g., C_2H_4 or C_2H_5OH) are assumed. The optimization problems are solved and setpoints are calculated to be 283 and 350 ppm for C_2H_4 and CO respectively. Finally, one pair of setpoints is selected to show that selectivity can be adjusted at any desired value such

as 1:1. Thus, setpoints corresponding to 1:1 ratio are calculated to be at 200 ppm and this value is selected to be the final setpoint. This optimization design assumes that all the products are sold at market values. However, the extra costs like the separation of gas and liquid products must be included in a real case scenario.

4. Dynamic Modeling of the Experimental Electrochemical Process

In order to implement multivariable control of this experimental process, there are two dynamic models needed: one for C_2H_4 and another one for CO. The concentration of the gas products in the headspace of the RCE cell can be approximated as a mixing volume where the gases produced on the catalyst mix and are carried over by the CO_2 gas flow (Çıtmacı et al., 2022a). This approach can be improved by modeling only the dynamic reactor data, which will involve all the inherent reactor dynamics without being derived by mass balances built for other similar reactor types (e.g., CSTR). Due to the absence of first principle dynamic equations of the RCE reactor for electrochemical CO_2 reduction, data-driven models can be built to model process dynamics.

4.1. Automatic Data Fitting to GC Measurements

The experimental concentration data from the GC are few and discrete; therefore, it is challenging to apply the deep-learning method (e.g., neural network modeling) to the raw experimental data. To tackle this, the GC results are combined using polynomial best fits to approximate a probable experimental trajectory, which significantly increases the number of data points, and enables building a recurrent neural network model.

In order to connect the respective GC points, 3^{rd} order polynomials are used for 3 data points. From the experimental side, three GC measurements represent 40 minutes of the experimental time span. For each 3 respective GC measurements, 2 polynomial curves are fitted. For example, the first polynomial fit is between 2^{nd}, 3^{rd}, and 4^{th} GC measurements. The second polynomial curve fitting is between 3^{rd}, 4^{th}, and 5^{th} GC measurements. However, with this method, there are two polynomial trajectories for the interval between each respective GC point. The most convenient polynomial fit between these two trajectories can be chosen intuitively. To automate this procedure, an algorithm is used to select the lower trajectory for an increasing concentration of more than 5%

and the upper trajectory for a decreasing concentration of more than 5%. This assumes that the aforementioned increase or decrease is in an exponential trend. If the concentrations that decrease or increase are less than 5%, then opposite trajectories are selected, implying that the experimental trajectory evolved into a logarithmic change phase. Due to the combination of different polynomial fits with this method, the trajectory obtained might not be smooth. To resolve this issue, a Savitsky-Golay filter with a time window of 500 seconds is applied and the obtained trajectory is adjusted to the correct timing to compensate for the time dislocation due to the filter window length. Fig.7 demonstrated an example of using the proposed method to enhance the experimental data from one experiment. In this example, sixteen data points, represented in red circles, are collected using the GC during over five hours of experiment. This raw data provides very limited information to train a neural network model. After the data fitting, the available concentration estimations increased to 18000 for this one experiment, which is shown as the black curve in Fig.4.

Remark 3. This data fitting is not claimed to be the actual experimental trajectory or the best data interpolation method. Changing the order of polynomials or the number of fitting data points in the intermediate trajectory may result in different interpolated trajectories. Other functions (e.g., linear interpolation, sigmoid function, etc.) can be considered as additional candidates for this fitting method.

4.2. Recurrent Neural Networks and LSTMs

RNN models are proven to be effective in capturing trends from time series data, and thus, are used in this work to model the dynamic behaviors of the electrochemical reactor for feedback control purposes. The increased amount of probable trajectory time series data is used to train this RNN model. Specifically, the RNN model can learn the time series from a defined time window and the correlations between the respective data points Wu et al. (2019). The RNN architecture is depicted in Fig. 8 (a). Input parameter vectors in time series are fed into the RNN and each time series vector is subject to recurrent calculations in the hidden layers. RNNs learn the time dependencies and provide results to the output layer, which is usually a fully connected dense layer to better map the hidden states into meaningful time series outputs.

One of the most powerful RNN architectures for time series is the Long-Short-Term-Memory (LSTM) model. LSTMs are different from simple RNNs due to their resistance to exploding/vanishing gradients exhibited in other neural network models thanks to the forget, input and output gates in each recurrent unit (Wu et al. (2019)). Each gate and the LSTM recurrent unit are shown in Fig. 8 (b). Here, cell state is used to transfer useful information and past relations from the previous recurrent units to the next recurrent units, and thus, it can be considered as the memory of this sequence (Yu et al. (2019)). Cell state is able to keep all the information from the initial time step recurrent unit until the final time step recurrent unit. The more relevant relationships are kept in long-term memory, whereas the less relevant information is removed at each time step. On the other side, there is the hidden state used to keep the output from each recurrent unit and transfer it to the following LSTM layer. These hidden-state outputs are formulated in a way that ultimately needs to be fed into a dense layer outside the LSTM, as mentioned in the previous paragraph. At each recurrent LSTM unit, the previous hidden state vector is combined with the new time step input vector and fed into the gates.

Additionally, there are three gates in an LSTM layer: forget, input, and output gates. Each gate contains trainable sigmoid activated neural networks as shown in 8 (b). The combined vector of previous hidden state and new input vector is fed to the forget gate. Due to the sigmoid activation function, the output is between 0-1. If the forget gate yields 0, then the previous cell state will be forgotten. Conversely, if the output is close to 1, the previous information from the cell state is retained between the forget and input gates. Next, there is the input gate, which also has a sigmoid activated trainable neural network that decides to what degree the new input vector should be remembered. The output of the input gate is pointwise multiplied with the tanh activated neural network outputs which is trained to learn the effect of the new input vector for the current time step output. Tanh activation function yields results between [-1,1], a negative value signifies that the new input vector might have a decreasing effect. The input gate yield is added pointwise to the cell state value. Finally, the output gate, again, filters the previous hidden state and input vector with a sigmoid activated neural network and this output is pointwise multiplied by the tanh activated neural network output of the hidden state. This gives the timestep output in the hidden

state (Yu et al. (2019)). The LSTM structure can be represented mathematically as follows:

$$i(k) = \sigma(\omega_i^x x(k) + \omega_i^h h(k-1) + b_i) \tag{8a}$$

$$f(k) = \sigma(\omega_f^x x(k) + \omega_f^h h(k-1) + b_f) \tag{8b}$$

$$c(k) = f(k)c(k-1) + i(k)tanh(\omega_c^x x(k) + \omega_c^h h(k-1) + b_c) \tag{8c}$$

$$o(k) = \sigma(\omega_o^x x(k) + \omega_o^h h(k-1) + b_o) \tag{8d}$$

$$h(k) = o(k)tanh(c(k)) \tag{8e}$$

$$\hat{x}(k) = \omega_y h(k) + b_y \tag{8f}$$

where k is the time step, i is the output from the input gate, h is the hidden state, c is the cell state, f is the forget gate, and o is the output gate. Furthermore, $w^{h,x}$ is the weight matrix to the hidden state vector h and input vector x, b_i, b_f, b_c, b_o, b_y represents biases and the subscript y indicates relationship to the output (Wu et al. (2020)).

4.3. LSTM based RNN Model Architecture

Two LSTM models are trained using the data calculated through the polynomial best fits between GC measurements. The inputs for the first LSTM model are surface potential, rotation speed, and current, and the output is C_2H_4 production rate converted to ppm using Eq. 2. For the second LSTM, the inputs are surface potential and rotation speed, and the output is CO production rate in ppm. Based on the experimental observations, this process has dead times of 600 to 1500 seconds for step changes on different input parameters. It can be seen in Fig. 7 that it takes around 1 injection for the rotation speed change to show its effect on concentration. Also, it was seen from the experimental data that it might take more than 1500 seconds after the dead time for the process output to reach the new steady state. Consequently, the past time window is very important for the dynamic behavior of the process and this time window should be around one hour. From this point, the models are trained with various time windows. The best model performances are obtained for 3600 seconds of time window for C_2H_4 and 3800 seconds for CO.

If one time window consists of 3600×3 (3 is the number of inputs parameters for the C_2H_4 estimator) data points, the model would be too big, computationally expensive, and perform poorly

due to high number of training parameters. In addition to that, the predictions would be very noisy. Thus, the time window can be discretized in a way that still represents the last one hour without violating the correlations between consecutive data points. For example, the last one hour can be represented with data points at each 100 seconds. As a result, the (3600×3) time window is reduced to the (36×3) time window without losing relevant information. The same approach is applicable for reducing the model size for the CO estimator, which would have (3800×2) input parameters for a per second base estimation and has (38×2) input parameters after 100 seconds of discretization. The model will make a prediction every 100 seconds. Also, the model is trained to predict the next 800 and 600 seconds of the production rates for C_2H_4 and CO respectively, with 100 seconds of discretization, having the output shape (8×1) and (6×1) in each LSTM. This way, LSTM is seen to learn the delay behavior better compared to an output shape of (1×1). Finally, both models are built with 200 hidden nodes in one hidden recurrent layer. It was seen that the prediction accuracy decreases if we decrease the number of hidden nodes. Also, the computation time increases and erroneous nonlinear trends are predicted by the model if we increase the number of hidden nodes.

Furthermore, regularization methods, such as recurrent dropout and L2 kernel regularizer, are performed to increase the generalization performance of the models to unseen data. Specifically, the following regularization methods are tuned to the best performance at 30% for recurrent dropout and 0.08 L2 kernel regularization. Increasing or decreasing those tuned parameters may lead to divergent predictions. The recurrent layer is connected to a dense layer of 8 nodes for C_2H_4 and 6 nodes for CO, which correlates the information from the recurrent node with the production rate values. A sigmoid function is used in the dense (FNN) output node to limit the predictions by the highest and lowest values in our training set. Multiple experiments from constant applied potential and rotation and two long step change experiments were selected for the test set. The remaining experiments were used to train the models. The predictions of the LSTM models are evaluated with an unseen testing set, and one set of the testing examples is demonstrated in Fig. 9.

Remark 4. Electrochemical CO_2 reduction in RCE cell is seen to have long process delays of more than 1000 seconds. The models take around 1 hour time window to model those delays. However, 3600 seconds of time window is a very long sequence. Thus, training those models only with open-

loop experiments might cause poorer performance in a closed-loop context. These models perform very successfully when only one of the manipulated input parameters varies sharply while the other one is constant. On the other hand, when two manipulated input variables vary at the same time gradually, the large time window of the LSTMs might not catch the dynamic trends as necessary. In order to enhance the LSTMs trained with the open-loop experiments, two preliminary controlled experiment were conducted, using the initial RNN models trained only with open-loop experiments. The dynamic data obtained from this experiment are used to train both estimators, which also leveraged the models to improve themselves to adapt their weights to the dynamic changes of the feedback control context.

The models need 3600 and 3800 input data points to make predictions for C_2H_4 and CO, respectively. Thus, the first prediction is obtained at the 3600^{th} second for the C_2H_4 estimator and at the 3800^{th} second for the CO model. The RNN architecture could have been built in such a way that the concentration output from the previous time step is fed to the LSTM model to guess the next time step concentration. This approach could have been useful to model the experiment starting from the beginning of the experiment rather than 3600^{th} second. However, before each experiment, pre-experiment measurements are performed including EIS to measure the solution resistance and cyclic voltammetry (CV) to check catalytic activity which already generates some products while cycling through a potential window. Thus, depending on the time span after this CV procedure, the development of gas products in the early stage may vary. This can be seen from the brown experimental curve (Experiment 3) in Fig. 3, in which there was a longer waiting time between pre-experiment measurements and the actual open-loop experiment compared to the other two experiments. This caused the delayed development of gas concentration trajectory at the early stage with a slower accumulation of gas products in the reactor headspace. Thus, feeding back the previous time step concentration estimation to the LSTM model causes confusion for the training set. As a result, the LSTM estimators wait about an hour until the system reaches equilibrium and are not affected by previous experimental development.

Remark 5. Initially, both of the estimators were trained with three inputs: Surface potential (V vs SHE), electrode rotation speed, and current. The current is a significant indicator for the C_2H_4

concentration and it is very important to include the current for training. However, the CO concentration is more dependent on the rotation speed and the surface potential. When the models were trained with open-loop step-change experiments, it was seen that the LSTM model for CO has a stronger correlation with current rather than surface potential. On the other hand, increasing current under a fixed applied potential decreases the surface potential, as can be seen from Eq. 3. Conversely, if the applied potential is not constant, the surface potential will increase with increasing applied potential. Thus, increasing the rotation speed will have an increasing effect on the CO concentration, whereas it will also increase the current and decrease the surface potential. Consequently, the LSTM model trained with open-loop experiments will learn that increasing current means increasing CO. This provides a great training accuracy in the testing set performance of open-loop experiments. However, the model should also learn the competition between the surface potential and the rotation speed for dynamic experiments. When this model was tested, it was found that the model yields incorrect trends when the ethylene concentration is increased by adding more surface potential (thus increased current) and the rotation is decreased. From the known correlations and GC results, the CO concentrations should have decreased, whereas the estimator predicted increasing concentrations due to increasing current. After seeing this deficiency with the dynamic model prediction based on model involving three inputs, the CO estimator was retrained using two inputs, surface potential and rotation speed, and the aforementioned problem was resolved. A diagram that explains how the model training and verification were carried out using both open-loop and closed-loop experiments is shown in Fig. 10.

The C_2H_4 and CO estimators are trained with 83393 and 112789 data points obtained through probable experimental trajectories. The difference between the training test sizes is due to using different experiments for training both models. Also, the CO model is trained with 2 preliminary dynamic controlled experiments whereas the C_2H_4 model is trained with only one. The training and testing mean absolute error performances are 0.62 ppm and 3.15 ppm for C_2H_4 model and 0.63 ppm and 11.1 ppm for CO. This indicates that the models fit to the training data quite well, but it generalizes to the new data satisfactorily well within the experimental standard deviations.

4.4. Regularization Effects on Experimental Data Modeling

In 2014, Srivastava et al. (2014) proposed the dropout method that randomly eliminates a percentage of hidden units during the training to eliminate/alleviate the overfitting effects in neural networks. When a unit is randomly excluded for an epoch, it prevents the rest of the neurons from excessively co-adapting. The dropout method is also used in our neural network architecture between the LSTM output and dense output layers. The addition of dropout in between these two layers has shown considerable overfitting mitigation. Specifically, the recurrent dropout in our model trained with Tensorflow/Keras uses the approach proposed by Semeniuta et al. (2016) and handles the connection between pointwise multiplication and pointwise summation in the input gate shown in Fig. 8 (b).

Although dropout is a very strong regularizer, adding other regularization methods to our model training can further improve the generalization performance of the model. L1 and L2 regularizers can be used as a part of the model training loss function. L1 regularizer is used to keep the specified model parameter close to 0 whereas L2 regularizer is used to prevent the model parameter from having too high values (Cortes et al. (2012)). The regularizers can be applied on model weights, biases or to the output. In our model training, an L2 kernel regularizer is used on the weights. The L2 kernel regularizer in the loss function is shown in Eq. 9.

$$\hat{J}(w; X, y) = J(w; X, y) + \frac{1}{2}\lambda w^T w \tag{9}$$

where $\hat{J}$ is the modified loss function, J is the loss function (e.g., mean squared error), w is the weight matrix that will be optimized, λ is the user-defined L2 regularization parameter that determines the intensity of the regularization.

Fig. 11 illustrates the enhancement in the model predictions with the addition of L2 regularization and dropout. The experiment shown here is used to test run the *LabVIEW* script with the experimental setup under a closed-loop trial. The slateblue curve is the fitted data with the method mentioned in Section 4.1 and the model predictions are supposed to converge to this curve. Orange prediction is the model trained without any regularization method and it can be seen that

the predictions have large deviations from the fitted curve. Then, a slight L2 regularization with the λ=0.04 is shown with the dashed blue curve. Between the 7500^{th} and 10000^{th} seconds, there is a slight improvement; however, the latter stages of the experiment still have significant error margins on the order of 300 ppm. Then, the L2 regularization value is increased to λ=0.08 (dash-dotted blue curve), and even though the noise in the predictions decreases and the trends are better represented, the model overshoots the probable experimental trajectory. Then, the regularization value was increased to a relatively high value at λ=0.15 (dotted blue curve), and this resulted in a drift and increased noise in predictions at the later stages of the experiment. This demonstrates that the low and high values of L2 regularization parameter λ do not improve the model, whereas a suitable fine-tuned λ value boosts the model performance. Finally, the best model with λ=0.08 is further improved with an appropriate percentage of recurrent dropout (red curve) and the dynamic trends are captured as well as the error predictions are greatly reduced. Thus, it is shown that the application of proper regularization parameters significantly augments the model generalization to unseen data. The regularization values presented in this section also generalize well to the other experimental operations (open- and closed-loop experiments).

Remark 6. The test case in Fig. 11 is a closed-loop experiment in which both surface potential and rotation are manipulated. Thus, this case is harder to generalize than open-loop experiments for the dynamic behavior.

4.5. Implementation of the LSTM Model in Real-Time Operation

The estimators are built to give a prediction each 100 seconds. However, for efficient control, we are expecting a per-second model feedback. Also, if the models were able to yield a per-second prediction, which would require a much larger model than a time window of 36/or 38 data points every 100 seconds, it would be computationally expensive for the *LabVIEW* program and might cause a time shift for 1 second intervals since there are many calculations being made at the same time and more data being kept in the memory. Therefore, to tackle this problem, a linear regression extrapolation approach is used to estimate 99 time steps between two consecutive LSTM predictions. In this approach, the recent LSTM prediction and the previous one are kept in memory and fitted to a line via linear regression for the previous 100 points. For the next 99 seconds, this linear

regression formula is used to extrapolate between 101^{th} to 199^{th} following predictions. With this method, there is a very small loss of accuracy, but a gain of interconnecting points estimation and computational efficiency. Since the 100-second interval is quite small in the development time of the experimental trajectory, the linear nature of this guess does not disturb the modeling of the setup. With this approach, long time window processes can be modeled more accurately using smaller models for many time steps without losing important information.

5. MIMO Control Architecture and Controller Tuning

This study aims for the multivariable feedback control using two PI controllers for C_2H_4 and CO concentration by manipulating applied potential and rotation speed.

5.1. Using Data for Process Parameter Extraction

The first-order plus time delay (FOPTD) model, as shown in Eq. 10 is used to extract the process parameters for the transfer function. Various open-loop experiments were conducted by applying separate step changes to applied potential and rotation speed. After applying the step change, the changes in C_2H_4 and CO concentrations are fitted to a sigmoid function and normalized between 0-1 for easy comparison between various step changes. From those plots, a tangent line from the inflection point is drawn to intersect the X axis. The intersection value on the X axis is defined as the dead time, indicated by θ, and the time required to reach 63% of the final steady state is defined as τ. These parameters are used to extract the FOPTD model shown in Eq. 10, where K is the steady state gain and s in the Laplace domain variable of the form:

$$G(s) = \frac{-Ke^{-\theta s}}{\tau s + 1} \tag{10}$$

However, electrochemical CO_2 reduction is a highly nonlinear process; therefore, one FOPTD model will not be sufficient to model the entire operating range. Thus, various FOPTD models should be extracted for different surface potential and rotation speed ranges. Our open-loop experiment range might not include an instance of a change between the limits of the designated ranges; thus, we should check general correlations for the dead time and τ for different step changes. If there is a general correlation for the time parameters, it is possible to extract steady state gains

from the FNN model. The graphs presented in Fig. 12 are used to display the dynamic responses of C_2H_4 ((a),(b)) and CO ((c),(d)) under various rotation speeds and rotation speed changes.

Fig. 12 (a) clearly shows that the concentration change exhibits the same dynamic behavior under applied potential step regulations and constant rotation, which means that we can assume the same dead time and τ for the tuning process. After analyzing the second plot, it was seen that similar dead times and τ values are obtained for rotation speeds less than 200 RPM. These values are used in the FOPTD model for C_2H_4. A similar procedure is followed for the extraction of the CO FOPTD model parameters.

Fig. 12 shows the dynamic concentration evolutions for C_2H_4 and CO under a step change in one input parameter (applied potential or rotation speed) while keeping the other input parameter constant. For example, in Fig. 12 (a), the normalized dynamic data fittings for applied potential change under a constant rotation speed overlap with each other for each experiment and they have the same dead time. In Fig. 12 (b) and (d), the dynamic data for the change in C_2H_4 and CO concentrations with respect to a variation in electrode rotation speed under a constant applied potential is demonstrated. For C_2H_4, it was seen that for the rotation speed changes under 200 RPM, the dead times are around the same value, and this is shown in Fig. 13. For CO, the variation in τ values for different rotation speed changes are illustrated in Fig. 13 (b). For each step change experiment, θ and τ values are extracted from the data and they are averaged for a more generalized results. These values will be used to extract FOPTD models for operational regions. The extracted θ and τ values can be seen in the multivariable control array shown below in Eq. 11.

As the time constants and delays for the FOPTD model are obtained, there is now a need to obtain steady state gains. Electrochemical CO_2 reduction process is inherently highly nonlinear. This can also be seen from the exponential increase in operational steady states in C_2H_4 concentrations with increasing surface potentials shown in Fig. 6 (d) and different behaviors of CO concentrations at different rotation speeds in Fig. 6 (c). Thus, we divide the operation range into 9 regions as the following:

- Potential ranges: = (-1.19, -1.26V), (-1.26, -1.30 V), (-1.30, -1.32 V);

- Rotation ranges = (100-200 RPM), (200-400 RPM), (400-800 RPM).

The multivariable control array and relative gain array shown below have the following structure.

	$100:200\,RPM$	$200:400\,RPM$	$400:800\,RPM$
$-1.19:-1.26\,V\,vs\,SHE$	$\ddots$	$\ddots$	$\ddots$
$-1.26:-1.30\,V\,vs\,SHE$	$\ddots$	$\ddots$	$\ddots$
$-1.30:-1.32\,V\,vs\,SHE$	$\ddots$	$\ddots$	$\ddots$

These ranges are expected to satisfactorily linearize the process. The steady state gains are calculated from the FNN model. Using the FOPTD models and different operating windows, we can obtain the multivariable control arrays such as shown below in Eq. 11 for the [-1.19, -1.26 V] and [100, 200 RPM] region.

$$\begin{bmatrix} C_{C_2H_4}(s) \\ C_{CO}(s) \end{bmatrix} = \begin{bmatrix} \frac{-1427e^{-1113s}}{952s+1} & \frac{-0.030e^{-1557s}}{1395+1} \\ \frac{-919e^{-1409s}}{952s+1} & \frac{-0.64e^{-1234s}}{540+1} \end{bmatrix} \begin{bmatrix} E_{surf}(s) \\ \Omega(s) \end{bmatrix} \tag{11}$$

This array is represented in the Laplace domain and four distinct FOPTD models are used to represent the dynamic process behavior. The first element of the FOPTD matrix in the first row represents the influence of $E_{surface}$ and the second element in the same row represents the dynamic effects of rotation change on the C_2H_4 concentration. The second row represents the same impact on CO concentration for the respective inputs. The transfer function matrices in different regions are given below:

$$\begin{bmatrix} \begin{bmatrix} \frac{-1427e^{-1113s}}{952s+1} & \frac{0.030e^{-1557s}}{1395+1} \\ \frac{919e^{-671s}}{1409+1} & \frac{0.64e^{-1234s}}{540+1} \end{bmatrix} & \begin{bmatrix} \frac{-1576e^{-1113s}}{952s+1} & \frac{0.14e^{-1557s}}{1395+1} \\ \frac{708e^{-671s}}{1409+1} & \frac{0.59e^{-1234s}}{540+1} \end{bmatrix} & \begin{bmatrix} \frac{-1532e^{-1113s}}{952s+1} & \frac{0.21e^{-1557s}}{1395+1} \\ \frac{-407e^{-671s}}{1409+1} & \frac{0.23e^{-1234s}}{540+1} \end{bmatrix} \\ \begin{bmatrix} \frac{-3985e^{-1113s}}{952s+1} & \frac{0.039e^{-1557s}}{1395+1} \\ \frac{808e^{-671s}}{1409+1} & \frac{0.49e^{-1234s}}{540+1} \end{bmatrix} & \begin{bmatrix} \frac{-4413e^{-1113s}}{952s+1} & \frac{0.13e^{-1557s}}{1395+1} \\ \frac{945e^{-671s}}{1409+1} & \frac{0.72e^{-1234s}}{540+1} \end{bmatrix} & \begin{bmatrix} \frac{-5158e^{-1113s}}{952s+1} & \frac{0.13e^{-1557s}}{1395+1} \\ \frac{1171e^{-671s}}{1409+1} & \frac{0.29e^{-1234s}}{540+1} \end{bmatrix} \\ \begin{bmatrix} \frac{-5445e^{-1113s}}{952s+1} & \frac{0.016e^{-1557s}}{1395+1} \\ \frac{403e^{-671s}}{1409+1} & \frac{0.50e^{-1234s}}{540+1} \end{bmatrix} & \begin{bmatrix} \frac{-5088e^{-1113s}}{952s+1} & \frac{0.0029e^{-1557s}}{1395+1} \\ \frac{857e^{-671s}}{1409+1} & \frac{0.65e^{-1234s}}{540+1} \end{bmatrix} & \begin{bmatrix} \frac{-4780e^{-1113s}}{952s+1} & \frac{0.019e^{-1557s}}{1395+1} \\ \frac{832e^{-671s}}{1409+1} & \frac{0.35e^{-1234s}}{540+1} \end{bmatrix} \end{bmatrix}$$

5.2. Relative Gain Array

There are two general approaches for the MIMO control system design of the RCE reactor based on the manipulated inputs and controlled output relationships: employing two proportional-integral (PI) controllers or a model predictive controller (MPC). If there are specific input-output couplings that are strong (i.e., multivariable interactions are relatively weak), then it is appropriate to use two PI controllers. In the opposite case where multivariable interactions are strong, an MPC should be used to regulate the process accounting explicitly for these interactions. To evaluate the strength of the multivariable interactions, a relative gain array (RGA) approach is used. In this work, from the experimental results, we have observed that the electrode rotation speed has a strong effect on the CO concentration, while the surface potential has a very strong effect on the C_2H_4 concentration. RGA is used to quantify these impacts and is calculated as follows (Bequette (2003)):

$$\Lambda = \begin{bmatrix} \frac{k_{11}k_{22}}{k_{11}k_{22}-k_{12}k_{21}} & \frac{-k_{12}k_{21}}{k_{11}k_{22}-k_{12}k_{21}} \\ \frac{-k_{21}k_{12}}{k_{11}k_{22}-k_{12}k_{21}} & \frac{k_{11}k_{22}}{k_{11}k_{22}-k_{12}k_{21}} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix}$$

where k_{ij}, $i = 1, 2$ and $j = 1, 2$, are the steady-state gains in the matrix transfer function models; for example, referring to Eq. 11, $k_{11} = -1427$, $k_{12} = -0.030$, $k_{21} = -919$ and $k_{22} = -0.64$. The RGA only needs the steady state gains. If the diagonals of RGA are close to 1, then it means that the surface potential can be used to control the C_2H_4 production and the rotation speed can be used to control the production of CO using two single-loop proportional-integral controllers. The RGAs of the different regions are presented below:

$$\begin{bmatrix} \begin{bmatrix} 0.9704 & 0.0296 \\ 0.0296 & 0.9704 \end{bmatrix} & \begin{bmatrix} 0.9706 & 0.0294 \\ 0.0294 & 0.9706 \end{bmatrix} & \begin{bmatrix} 1.0181 & -0.018 \\ -0.018 & 1.018 \end{bmatrix} \\ \begin{bmatrix} 0.9444 & 0.0556 \\ 0.0556 & 0.9444 \end{bmatrix} & \begin{bmatrix} 0.9610 & 0.0390 \\ 0.0390 & 0.9610 \end{bmatrix} & \begin{bmatrix} 0.9977 & 0.0023 \\ 0.0023 & 0.9977 \end{bmatrix} \\ \begin{bmatrix} 0.9693 & 0.0307 \\ 0.0307 & 0.9693 \end{bmatrix} & \begin{bmatrix} 0.9669 & 0.0331 \\ 0.0331 & 0.9669 \end{bmatrix} & \begin{bmatrix} 0.9908 & 0.0092 \\ 0.0092 & 0.9908 \end{bmatrix} \end{bmatrix}$$

Since the RGA diagonal elements for each region have values very close to 1, we can infer that two PI controllers are sufficient to control this process, where one of the PI controllers will manipulate the applied potential to control C_2H_4 while the other PI controller will manipulate the

rotation speed to control CO production rate. These RGA conclusions hold true for a broad set of steady state values of the inputs in the operating region of interest; this analysis was carried out in Canuso (2022) for a similar experimental reactor setup and will not be repeated here. Overall, it is noteworthy that the RGA results based on our lab-scale experimental reactor provide a quantitative basis for the strength of these interactions in a scaled-up reactor based on the design employed in our current process.

5.3. Controller Tuning

Two PI controllers are used for the feedback control of the experimental RCE reactor. Since the FOPTD transfer function array is obtained, the controller tuning parameters can be calculated for both PI controllers using the Cohen-Coon technique. Specifically, the FOPTD models are initially used to extract 9 pairs of controller parameters with the Cohen-Coon tuning method. The details of Cohen-Coon tuning method can be found in (Hambali et al. (2014)). However, employing 9 pairs of controller parameters might be complicated, thus, proportional gain values that are in the vicinity of each other are averaged to reduce the number of controller parameters. After having the preliminary proportional gains and integral time constants, MATLAB is used to simulate a setpoint change to fine-tune the proportional gains taking the Cohen-Coon estimations as starting points. The fine-tuned control parameters are presented in Table 2 and the simulation results for different regions are presented in Fig. 14. The integral time constants (τ_i) from Cohen-Coon are 1250 and 1000 seconds, respectively.

However, there is a unit matrix decoupler defined in the MATLAB tuning procedure, meaning that the simulator does not take the cross-coupling effects into account (e.g., rotation effects on C_2H_4 concentration). At some specific operation ranges, the cross-coupling effects might affect the performance of the closed-loop system. For example, as discussed in Section 3.1, the effect of surface potential change affects the CO concentration at lower rotation speeds. When the MATLAB-tuned parameters are tested in the experimental setup, it was seen that both of the parameters have huge fluctuations around the steady state setpoints. This means the controller proportional gains are not tuned to handle the cross-parameter effects, and they should be further improved.

It is very time-consuming and costly to fine-tune the controllers on the experimental setup by

trial and error, especially when multiple pairs of control parameters are involved. To address this problem, the dynamic RNN models trained for C_2H_4 and CO dynamic evolution modeling are used as the digital model of the experimental reactor to investigate the controller performance under numerous pairs of controller parameters in different regions. To enable this simulation, the instance current is required as the input for the estimation of C_2H_4. Thus, the correlation presented in Eq. 4 is used to approximate the current at a specific surface potential. However, this case is idealized compared to the real system because the current magnitude is largely affected by the resistance of the electrolyte solution and experimental conditions might also affect the concentration outputs.

Two PI controllers are added to the simulation. The first controller manipulates the surface potential to control C_2H_4 production rates while the second controller manipulates the rotation speed to regulate the CO concentration in the reactor overhead. It was seen that if both controllers are tuned with values in the vicinity of MATLAB-tuned proportional gains, both of the setpoints suffer from big fluctuations due to cross-coupling effects and it would take a very long time for both fluctuations to subside. Thus, the case that worked best and fastest to reach the steady state is having one of the controllers with a high proportional gain and the other controller with a low proportional gain. Specifically, the C_2H_4 controller is selected to be the more aggressive one while the CO controller is adjusted to drive the process to the setpoint slowly. Therefore, the MIMO control scheme aims to drive the C_2H_4 concentration to the setpoint quickly; and after it is reached, the second PI controller drives the CO to the setpoint slowly. While C_2H_4 is going to the setpoint, CO could even go in the opposite direction with respect to its setpoint because big changes in the surface potential could negatively affect CO, especially at low rotation speeds. However, once the surface potential is stabilized at the new setpoint, the rotation speed is adjusted slowly, while the small cross rotation effects on C_2H_4 are compensated by the aggressive potential controller. A control case simulation scenario with the new control parameters is shown in Fig. 15. The final proportional gains obtained from this simulation are presented in Table 3.

5.4. Estimator Design Using GC Measurements

In Section 2.2, the standard deviation in the experimental results is discussed under nearly-identical surface potential and rotation speed for relatively short duration experiments of less than

2 hours. The experimental conditions mentioned in Remark 1 also affect the variability in the experimental performance observed for individual experiments. Therefore, the LSTM models built using the experimental training data are expected to give a prediction that is compatible with the averaged experimental output and within the standard deviations, however, this prediction still needs to be improved for feedback control purposes. To accomplish this fine-tuning of the control of the reactor, the GC sensor feedback is incorporated into the prediction model and the control scheme.

In our previous work Çıtmacı et al. (2022a), we offered an approach to introduce the GC measurements incurring a 14.3 minutes delay into the concentration estimation and feedback control scheme. However, that approach took advantage of a hybrid model and mass balance equations, calculated the error in machine learning-based reaction rate estimation and corrected the estimator predictions accounting for the possible corrections in past time instants of the experiment. That approach is not applicable here due to LSTMs being black box models. Instead, taking advantage of the ability of RNNs to learn time series and trends in complex data series, we can scale the RNN predictions with respect to the recent GC measurement and adopt our RNN model to the real experimental trajectory without violating the trends. This is a suitable solution built on the understanding that the catalyst activity may vary from experiment to experiment, while the fundamental mechanism for the transformation of CO_2 on the copper electrodes is unchanged. The RNN model predicts the concentrations within a standard deviation, and captures the dynamic timing (e.g., dead time and τ) satisfactorily well. Thus, scaling the model does not affect the dynamic trends. To scale the RNN outputs, the prediction from 14.3 minutes ago is kept in the control loop, and it is compared to the GC measurement when the injection is completed. As soon as the GC results are obtained, the GC result is proportioned to the RNN prediction from 14.3 minutes ago and a scaling factor α is obtained. Then, the current RNN predictions are multiplied by the scaling factor until a new scaling function is obtained. This is described as follows:

$$\alpha^i = \frac{GC^i}{F^0_{RNN}(t - 14.3\,min)} \tag{12a}$$

$$F_{RNN}(t) = \alpha^i \times F^0_{RNN}(t) \tag{12b}$$

where i is the number of GC injections, GC^i is the i^{th} GC injection, F^0_{RNN} is the prediction of the unscaled RNN model and F_{RNN} is the prediction of the scaled RNN model.

This GC feedback correction is not applied after each injection. Applying a scaling factor during a big setpoint change might interfere with the model dynamics, thus we scale the RNN predictions when the consecutive GC variation is small (e.g., at steady state). When the difference between the GC measurement and RNN prediction is substantial, the scaling factor is applied when the concentrations are near a steady state, the model predictions and GC measurements are scaled to be coherent, and the system is driven to the set-point after the scaling. This correction is applied if the consecutive GC measurements are within 8% or 10 ppm vicinity of each other. The application of the scaling factor is automated on the *LabVIEW* interface. The final control diagram of the process is shown in Fig. 16.

6. Closed-loop Experimental Results and Discussion

Three scenarios are considered for the demonstration of MIMO control in this study. The first scenario aims to drive the process to the most profitable set-point under the base case, which has a selectivity of 1:4 for the C_2H_4:CO concentration ratio. The second scenario drives the selectivity ratio to 1:1 to show the flexibility of the control system by increasing the C_2H_4:CO ratio by 4-fold. This versatility could be advantageous considering a potential integration of a downstream process such as the copolymerization of ethylene and carbon monoxide, where the ratio of the two monomers impacts the structural and thermophysical characteristics of the synthesized polyketone (Brubaker et al. (1952) and Soomro et al. (2014)). The third scenario simulates a case based on sensitivity analysis in which the prices of C_2 and C_{2+} products increase by 60% while other chemical prices decrease by 40%. The third scenario drives the C_2H_4:CO concentration ratio to 4:5 which is very close to the 1:1 ratio and thus serves to challenge the accuracy of the control system for close but different concentration ratios over a broad range of concentrations. The three scenarios demonstrate the ability of controllers to implement large increases and decreases in selectivity ratios, as well as precise control of product concentrations. In each experiment, the closed-loop controller is activated at 4500^{th} second and the same control parameters are used for all scenarios. Each case starts from

the same initial applied potential and electrode rotation speed of -1.67 V vs Ag/AgCl and 300 RPM, usually yielding a surface potential of around -1.27 V vs SHE. This initial point is chosen to have both the rotation and potential values close to the region where the controller parameters change to observe the effects of changes in controller gains.

6.1. Experimental Results

Set-points for C_2H_4 and CO under the first scenario are calculated from the real-time optimizer to be 112 and 490 ppm. The experimental results for the first case scenario are shown in Fig. 17. It is noteworthy to compare these experimental results to those obtained in the simulated closed-loop experiments shown in Fig. 15 that utilized identical set-points. Due to the control strategy mentioned in Section 5.3, the controller first drives C_2H_4 to the set-point while the electrode rotation speed increases gradually, which slowly drives CO to its set-point. As shown previously in Fig. 15, this strategy can effectively compensate for the rotation-driven deviation in C_2H_4 concentrations. In this case, to decrease the ethylene concentration, the applied potential is decreased. This also has the effect of increasing slightly the CO concentration at low rotation speeds (see Fig. 6 (c)). Thus, cross-coupling effects on the input-output relationship in this specific set-point are helpful in achieving the control objective. The initial increase in CO is mostly due to the decrease in surface potential. In this region, until the rotation reaches 400 RPM, the effect of rotation speed is kept minimal with the CO controller using a small proportional gain. Once the ethylene concentration reaches its set-point, the rotation speed is increased more aggressively to eventually drive the CO concentration to its set-point.

The experimental trajectory in Fig. 17 is very similar to the simulated result in Fig. 15. The main two differences are that in the simulation, the set-points are reached earlier than in the actual experiment and the estimation of a set-point overshoot for the CO concentration, which is not observed in the experiment. These differences are naturally due to the fact that experimental observations are different from model prediction and feedback corrections, and highlight the importance of coupling simulations with experiments. The simulation also assumes that the current is well-described by Eq. 4. However, experimentally, the current might deviate from this correlation extracted from open-loop experiments for catalysts operated over relatively short durations, caus-

ing slightly different product concentrations. Likewise, when a scaling factor is applied as a part of the feedback correction, the error between the model prediction and the set-point change must be recalculated and results in some delays which do not exist in the simulation and could have an effect on the experimental observations. The pure model predictions might be very close to the set-point, which would reduce the actuation. However, when the scaling factor is applied, the model receives feedback when the predictions were off (e.g., at 10 ppm) and the controller drives the corresponding input at higher values to reach the set-point. Thus, this causes a delay in the set-point tracking, which can also explain the timing difference between the experiment and the simulation. It is noted that these timing differences are still minor, and the system is still driven to the desired multiple set-points in a short period of time of just a few hours.

The objective of the second scenario is to adjust the selectivity ratio to 1:1 for a concentration of 200 ppm for each product. The experimental results for this scenario are shown in Fig. 18. Here, the initial ethylene concentration is closer to 200 ppm, and it requires only a small increase in applied potential to reach the set-point. On the other hand, CO starts from a relatively high concentration compared to the set-point at 200 ppm. The initial surface potential activates the higher K_p and the controller starts increasing the surface potential further. At the same time, the error between CO concentration and its set-point is high, thus the electrode rotation speed continuously decreases resulting in a monotonic decrease in the CO concentration. As the rotation speed further decreases below 200 RPM, the potential effects become more important for CO. When C_2H_4 reaches its set-point and the controller starts to decrease the surface potential, the decrease in CO concentration loses momentum as the potential effects kick in, increasing the production of CO. To account for this, the MIMO controller reduces the rotation speed more strongly after around 13000 seconds to overcome the effects of fluctuating surface potential until both set-points are eventually reached after around 20000 seconds of the start of the experiment.

In the third scenario, the system is driven to a selectivity ratio of around 4:5 as shown in Fig. 19, where the set-points for C_2H_4 and CO concentrations are 283 and 350 ppm, respectively. In this experiment, the initial CO concentration is already close to its set-point, whereas the C_2H_4 concentration is less than half that of its set-point. Here, the C_2H_4 controller steadily increases the

surface potential to decrease the error between the set-point and the estimator model predictions. Although the CO controller slowly increases the rotation speed to raise the CO concentration, the strong increase in surface potential decreases the CO concentration. Because this surface potential increase occurs while the rotation speed is still at relatively slow rates, the surface potential effect is influential on the CO concentration. The fast increase of the surface potential causes a small overshoot in the C_2H_4 concentration but this is then corrected by slowly decreasing the applied potential and driving C_2H_4 slowly to its set-point. After this task is achieved, the rotation is increased to drive CO to its desired set-point.

In all 3 experiments, different set-points are reached for both C_2H_4 and CO, thus experimentally demonstrating that the developed control scheme is efficient. It must not be lost in the technical description of this paper the fact that this is the first time that this level of control is achieved for multiple products in a CO_2 electrochemical RCE reactor. The MIMO controller demonstrated here delivers on the promise of RNN-based modeling frameworks for the control of advanced reactors driven purely by electricity. This should motivate further integration of RNN-based modeling in the research and development of electrified chemical and fuels manufacturing technologies.

6.2. Experimental Errors in Closed-loop Experiments and Outlook

Each closed-loop experiment started from the same initial conditions of applied potential and rotation and yielded 189, 172 and 132 ppm for C_2H_4 and 346, 328 and 335 ppm for CO after reaching the first steady state at around 5000 seconds from the start of the experiment. This highlights the experimental difference in each of the catalysts freshly prepared, and demonstrates the need for a feedback corrector to adjust the model predictions to the GC points for each individual experiment.

Furthermore, there are shifts in current. For example, if the system generates less current in a new experiment at the same surface potential compared to the base case steady state experiments, it is very probable to obtain less ethylene. These dynamics are captured by the RNN model as many case scenarios are fed to the model during the training. One of the reasons for this current shift might be the lack of Ohmic drop compensation feature on the *LabVIEW* interface. As mentioned in Remark 1, the electrolyte solution is measured in the beginning of the experiment. However, the resistance of the electrolyte solution might increase or decrease as the experiment

proceeds, and the surface potential calculations are affected by these resistance shifts. It would be necessary to program into the potentiostat controller an algorithm that, without changing the applied potential, performs electrochemical impedance spectroscopy or a current interrupt analysis during a few seconds to measure the evolving solution resistance. This will be the focus of future works as we expand the MIMO controller to account for more operation variables and products.

7. Conclusion

In summary, this work presented the development and implementation of a recurrent neural network-based modeling framework on an experimental electrochemical reactor. The RNN-based modeling framework was used to represent relationships between process variables and gas product concentrations at the outlet of the reactor. The proposed approach combined the output of LSTM model-based gas-phase ethylene and carbon monoxide concentration estimators with GC measurements to implement multivariable control of the production rate of these two products in the reactor. Steady state data was used to construct a feedforward neural network for the calculation of feasible operating points. Using this model and an optimization model, economically-optimal, energy-efficient set-points were computed. In the absence of first-principle modeling descriptions needed to implement an efficient control system for the reactor, experimental GC measurements and data regression techniques were used to construct probable experimental trajectories for experiments run over lengths of 7 hours while the GC data available is discrete in nature, and only collected every 20 minutes. The enhanced trajectory dataset was used to train the LSTM model, which is the basis for the gas concentration estimator. In the implementation of the estimator-based multi-input multi-output control system, the real-time GC measurements are fed back to the controller for which the controller parameters have been tuned in different operating regions in order to better respond to the nonlinear nature of the electrochemical process. The estimator-based multi-input multi-output control system was successfully demonstrated to be capable of driving the process outputs to a variety of optimal set-points for C_2H_4 and CO.

8. Acknowledgements

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Advanced Manufacturing Office Award Number DE-EE0007613. The authors would like to thank Derek Richard for valuable comments on the manuscript preparation.

9. Disclaimer

References

Bequette, B. W., 2003. Process control: modeling, design, and simulation. Prentice Hall Professional.

Brubaker, M. M., Coffman, D. D., Hoehn, H. H., 1952. Synthesis and characterization of ethylene/carbon monoxide copolymers, a new class of polyketones. J. Am. Chem. Soc. 74, 1509–1515.

Canuso, V., 2022. Machine Learning-Based Operational Modeling of an Electrochemical Reactor: Handling Data Variability for Experimental Data. Master's thesis. University of California, Los Angeles.

Chen, A., Zhang, X., Chen, L., Yao, S., Zhou, Z., 2020. A machine learning model on simple features for CO_2 reduction electrocatalysts. The Journal of Physical Chemistry C 124, 22471–22478.

Cheng, Y., Karjala, T. W., Himmelblau, D. M., 1995. Identification of nonlinear dynamic processes with unknown and variable dead time using an internal recurrent neural network. Industrial & Engineering Chemistry Research 34, 1735–1742.

Çıtmacı, B., Luo, J., Jang, J. B., Canuso, V., Richard, D., Ren, Y. M., Morales-Guio, C. G., Christofides, P. D., 2022a. Machine learning-based ethylene concentration estimation, real-time optimization and feedback control of an experimental electrochemical reactor. Chemical Engineering Research and Design 185, 87–107.

Çıtmacı, B., Luo, J., Jang, J. B., Korambath, P., Morales-Guio, C. G., Davis, J. F., Christofides, P. D., 2022b. Digitalization of an experimental electrochemical reactor via the smart manufacturing innovation platform. Digital Chemical Engineering 5, 100050.

Corriou, J.-P., 2004. Process Control: Theory and Applications. Second ed., Springer.

Cortes, C., Mohri, M., Rostamizadeh, A., 2012. L2 regularization for learning kernels. arXiv preprint arXiv:1205.2653 .

De Luna, P., Hahn, C., Higgins, D., Jaffer, S. A., Jaramillo, T. F., Sargent, E. H., 2019. What would it take for renewably powered electrosynthesis to displace petrochemical processes? Science 364, 3506.

Hambali, N., Masngut, A., Ishak, A. A., Janin, Z., 2014. Process controllability for flow control system using Ziegler-Nichols (ZN), Cohen-Coon (CC) and Chien-Hrones-Reswick (CHR) tuning methods, in: Proceedings of the International Conference on Smart Instrumentation, Measurement and Applications, Kuala Lumpur, Malaysia.

Hussain, M. A., 1999. Review of the applications of neural networks in chemical process control—simulation and online implementation. Artificial Intelligence in Engineering 13, 55–68.

Jang, J., Rüscher, M., Winzely, M., Morales-Guio, C. G., 2022. Gastight rotating cylinder electrode: Towards decoupling mass transport and intrinsic kinetics in electrocatalysis. AIChE Journal 65, e17605.

Jin, S., Hao, Z., Zhang, K., Yan, Z., Chen, J., 2021. Advances and challenges for the electrochemical reduction of CO_2 to CO: From fundamentals to industrialization. Angewandte Chemie International Edition 60, 20627–20648.

Khalid, M., Omatu, S., Yusof, R., 1993. Mimo furnace control with neural networks. IEEE Transactions on Control Systems Technology 1, 238–245.

Kramer, A., Morgado-Dias, F., 2018. Applications of artificial neural networks in process control applications: A review, in: Proceedings of International Conference on Biomedical Engineering and Applications, Funchal, Portugal.

Luo, J., Canuso, V., Jang, J. B., Wu, Z., Morales-Guio, C. G., Christofides, P. D., 2022. Machine learning-based operational modeling of an electrochemical reactor: Handling data variability and improving empirical models. Industrial & Engineering Chemistry Research 61, 8399–8410.

Mohanty, S., 2009. Artificial neural network based system identification and model predictive control of a flotation column. Journal of Process Control 19, 991–999.

Nitopi, S., Bertheussen, E., Scott, S., Liu, X., Engstfeld, A., Horch, S., Seger, B., Stephens, I., Chan, K., Hahn, C., et al., 2019. Progress and perspectives of electrochemical CO_2 reduction on copper in aqueous electrolyte. Chemical Reviews 119, 7610–7672.

Ramdin, M., De Mot, B., Morrison, A. R. T., Breugelmans, T., van den Broeke, L. J. P., Trusler, J. P. M., Kortlever, R., de Jong, W., Moultos, O. A., Xiao, P., Webley, P. A., Vlugt, T. J. H., 2021. Electroreduction of CO_2/CO to C_2 products: Process modeling, downstream separation, system integration, and economic analysis. Industrial & Engineering Chemistry Research 60, 17862–17880.

Ren, Y. M., Alhajeri, M. S., Luo, J., Chen, S., Abdullah, F., Wu, Z., Christofides, P. D., 2022. A tutorial review of neural network modeling approaches for model predictive control. Computers Chemical Engineering 165, 107956.

Richard, D., Tom, M., Jang, J., Yun, S., Christofides, P. D., Morales-Guio, C. G., 2023. Quantifying transport and electrocatalytic reaction processes in a gastight rotating cylinder electrode reactor via integration of computational fluid dynamics modeling and experiments. Electrochimica Acta 440, 141698.

Roberts, F. S., Kuhl, K. P., Nilsson, A., 2015. High Selectivity for Ethylene from Carbon Dioxide Reduction over Copper Nanocube Electrocatalysts. Angew. Chemie 127, 5268–5271.

Scarselli, F., Tsoi, A. C., 1998. Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results. Neural networks 11, 15–37.

Semeniuta, S., Severyn, A., Barth, E., 2016. Recurrent dropout without memory loss, in: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan.

Soomro, S. S., Cozzula, D., Leitner, W., Vogt, H., Müller, T. E., 2014. The microstructure and melt properties of CO-ethylene copolymers with remarkably low CO content. Polym. Chem. 5, 3831–3837.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 15, 1929–1958.

Tian, Y., Zhang, J., Morris, J., 2001. Modeling and optimal control of a batch polymerization reactor using a hybrid stacked recurrent neural network model. Industrial & Engineering Chemistry Research 40, 4525–4535.

Timoshenko, J., Jeon, H. S., Sinev, I., Haase, F. T., Herzog, A., Roldan Cuenya, B., 2020. Linking the evolution of catalytic properties and structural changes in copper–zinc nanocatalysts using operando EXAFS and neural-networks. Chem. Sci. 11, 3727–3736.

Wächter, A., Biegler, L. T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Mathematical Programming 106, 25–57.

Wu, Z., Rincon, D., Christofides, P. D., 2020. Process structure-based recurrent neural network modeling for model predictive control of nonlinear processes. Journal of Process Control 89, 74–84.

Wu, Z., Rincon, D., Luo, J., Christofides, P. D., 2021. Machine learning modeling and predictive control of nonlinear processes using noisy data. AIChE Journal 67, e17164.

Wu, Z., Tran, A., Rincon, D., Christofides, P. D., 2019. Machine learning-based predictive control of nonlinear processes Part I: Theory. AIChE Journal 65, e16729.

Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: LSTM cells and network architectures. Neural computation 31, 1235–1270.

List of Figures

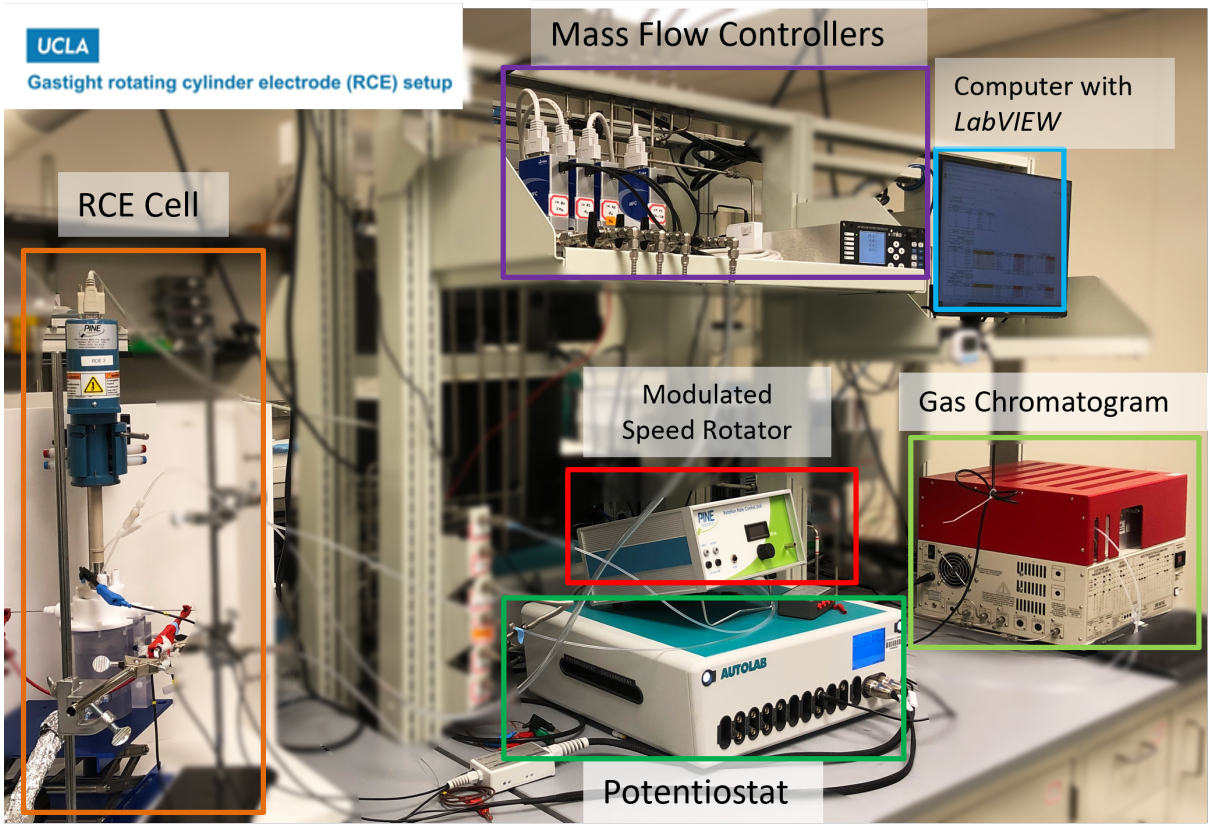


Figure 1: UCLA gastight RCE reactor setup.

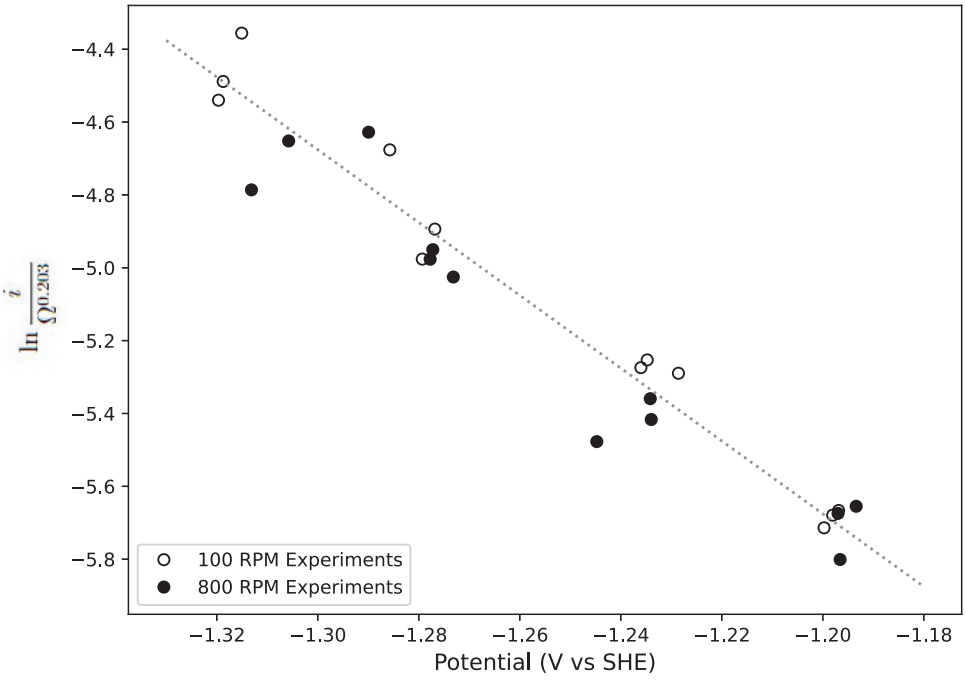


Figure 2: Empirical correlation of normalized current versus surface potential at two different electrode rotation speeds (100 RPM and 800 RPM).

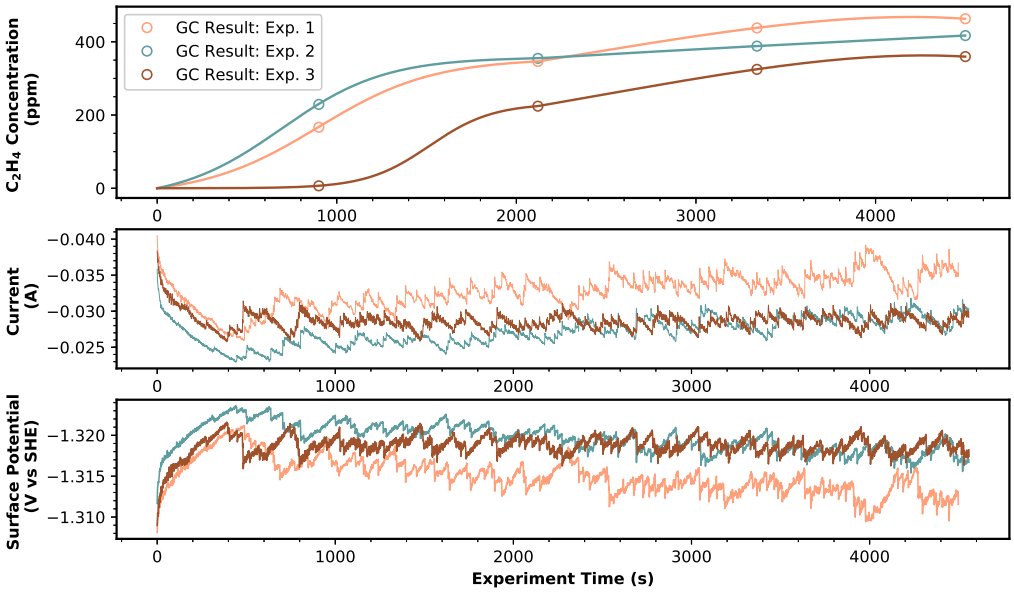


Figure 3: Comparison of 3 experimental C_2H_4 concentrations from open-loop steady state experiments conducted at nearly identical surface potentials (bottom plot) and 100 RPM. The top plot shows experimental GC results with fitted curves calculated by polynomial regression.

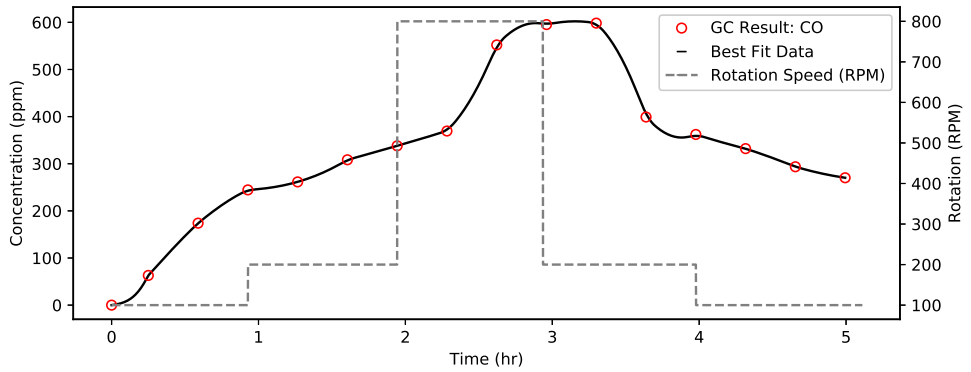


Figure 4: Probable experimental trajectory of CO concentration using polynomial fit curves for open-loop rotation speed step change experiments while keeping the applied potential constant.

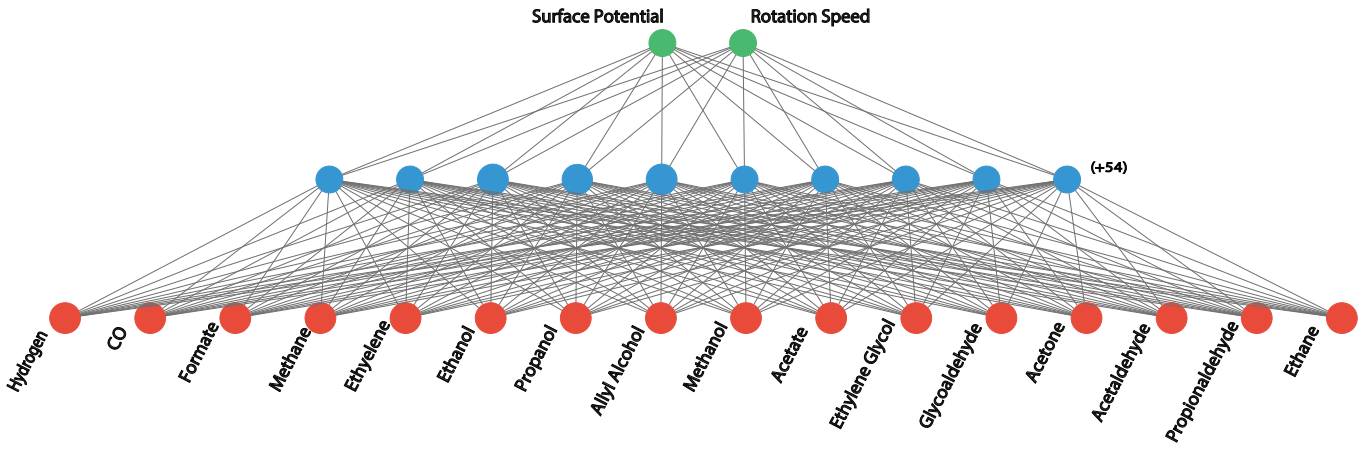
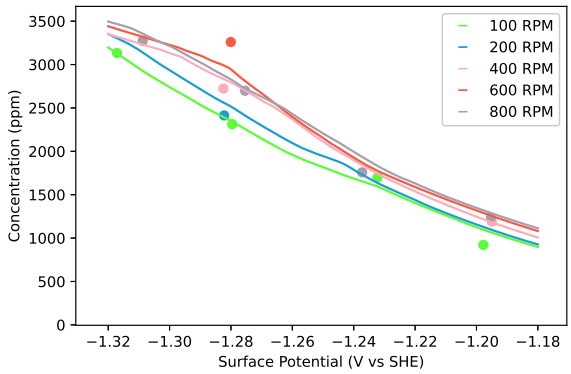
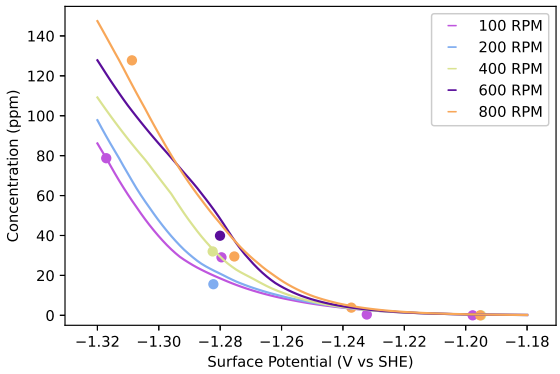


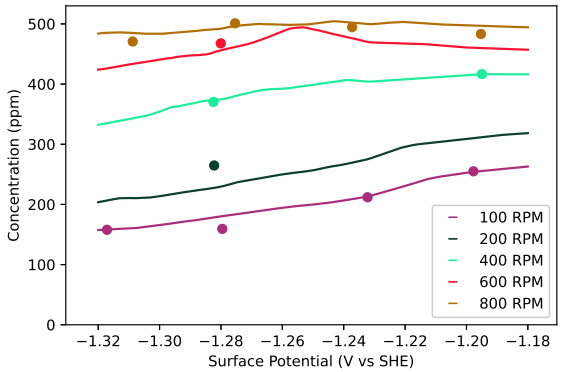
Figure 5: FNN architecture based on weighted data mapping two inputs (i.e., surface potential and rotation speed) represented in green circles to the production rates of sixteen products (outputs) represented in red circles through a densely connected hidden layer represented in blue circles. Only 10 of the 64 nodes are shown in the figure as blue circles. The model includes 54 more hidden nodes.



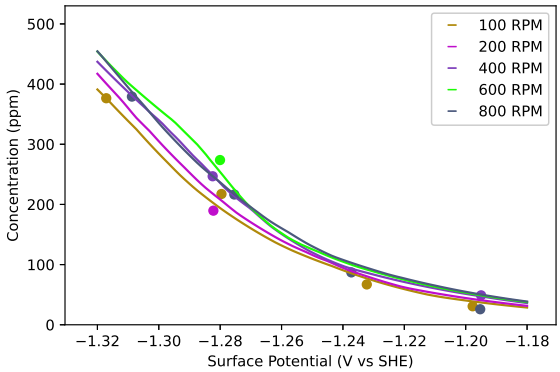
(a) H_2 concentration versus surface potential and rotation speeds.



(b) CH_4 concentration versus surface potential and rotation speed.



(c) CO concentration versus surface potential and rotation speed.



(d) C_2H_4 concentration versus surface potential and rotation speed.

Figure 6: FNN predictions for gas-phase products under various input conditions, where (a), (b), and (d) demonstrate that the production rates of H_2, CH_4, and C_2H_4 are weakly correlated to the rotation speed, and (c) demonstrates that the production rate of CO has stronger correlation with the rotation speed. The solid dots represent the experimental data which support the predicted curves calculated from the FNN model.

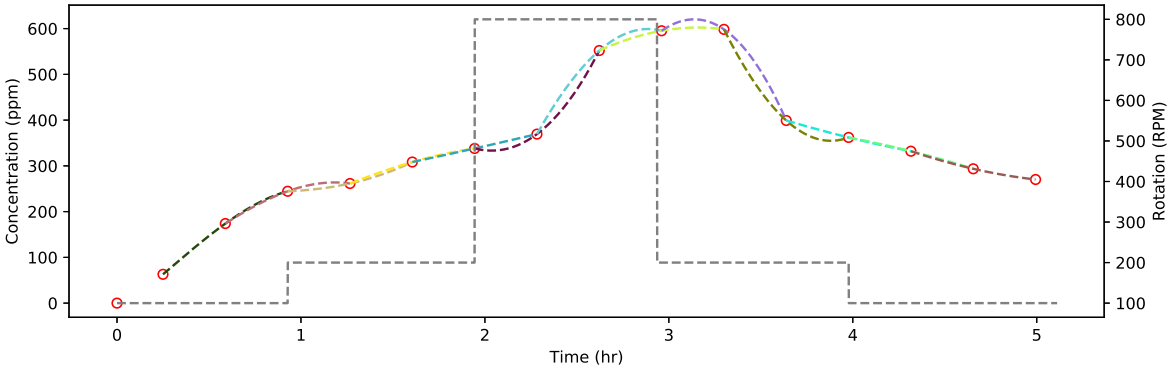
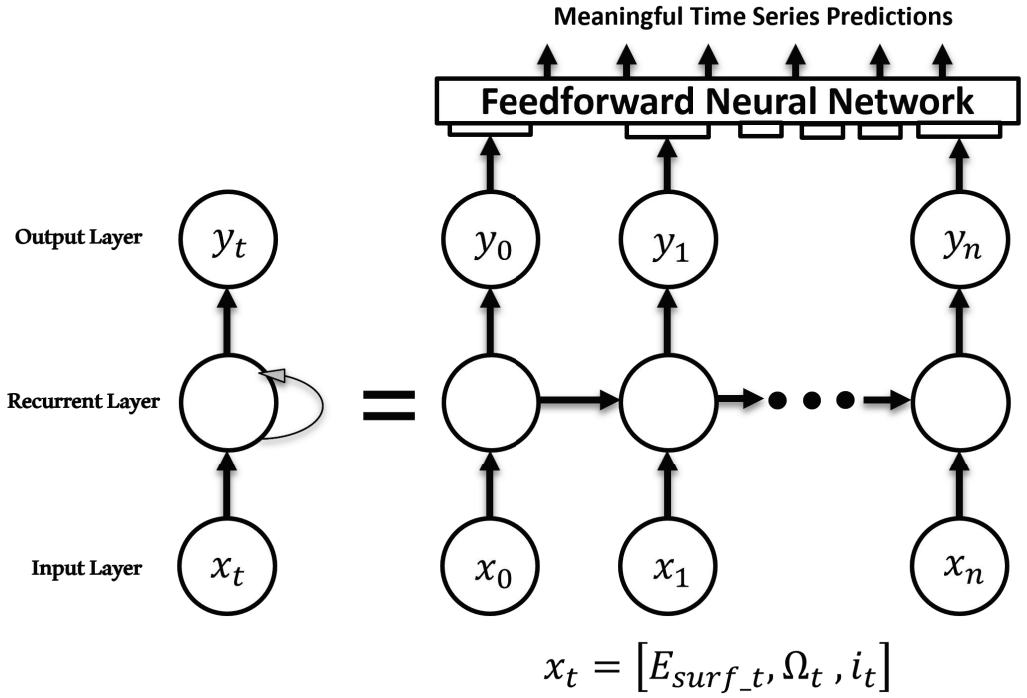
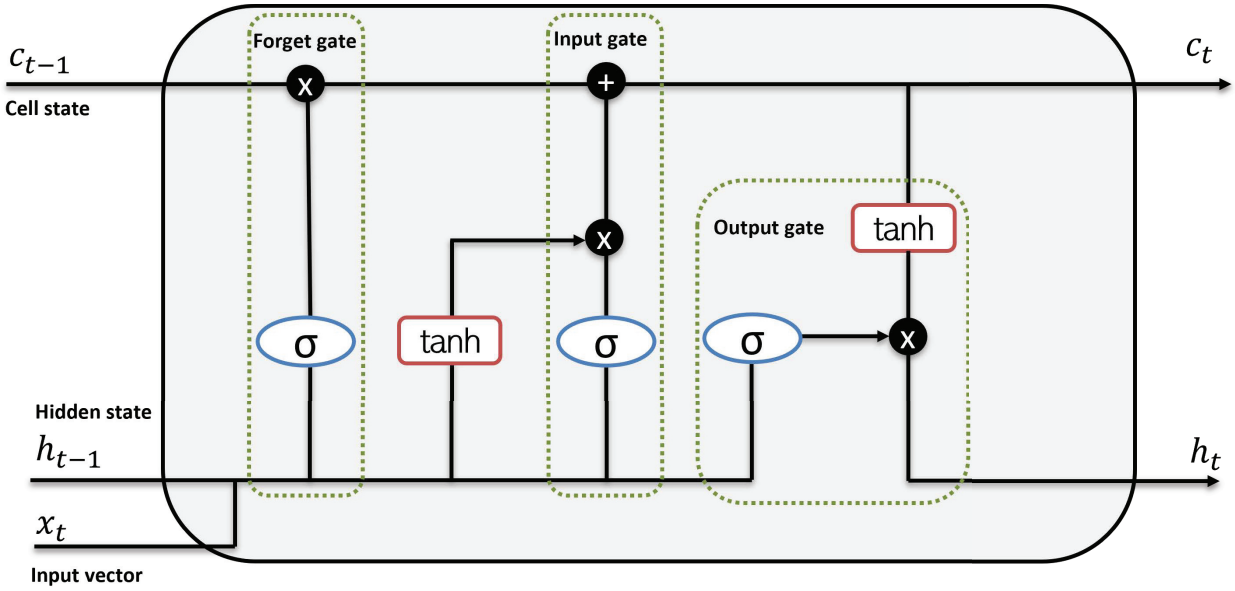


Figure 7: Visualization of auto data fitting algorithm for change in CO concentration in an open-loop step rotation speed change experiment. The best-fitted trajectory is demonstrated as the black solid curve in Fig. 4 generated by the second step of the algorithm that automatically picked the best fit from the candidate trajectories between every two points and smoothed the overall trajectory with Savitzky–Golay filter.

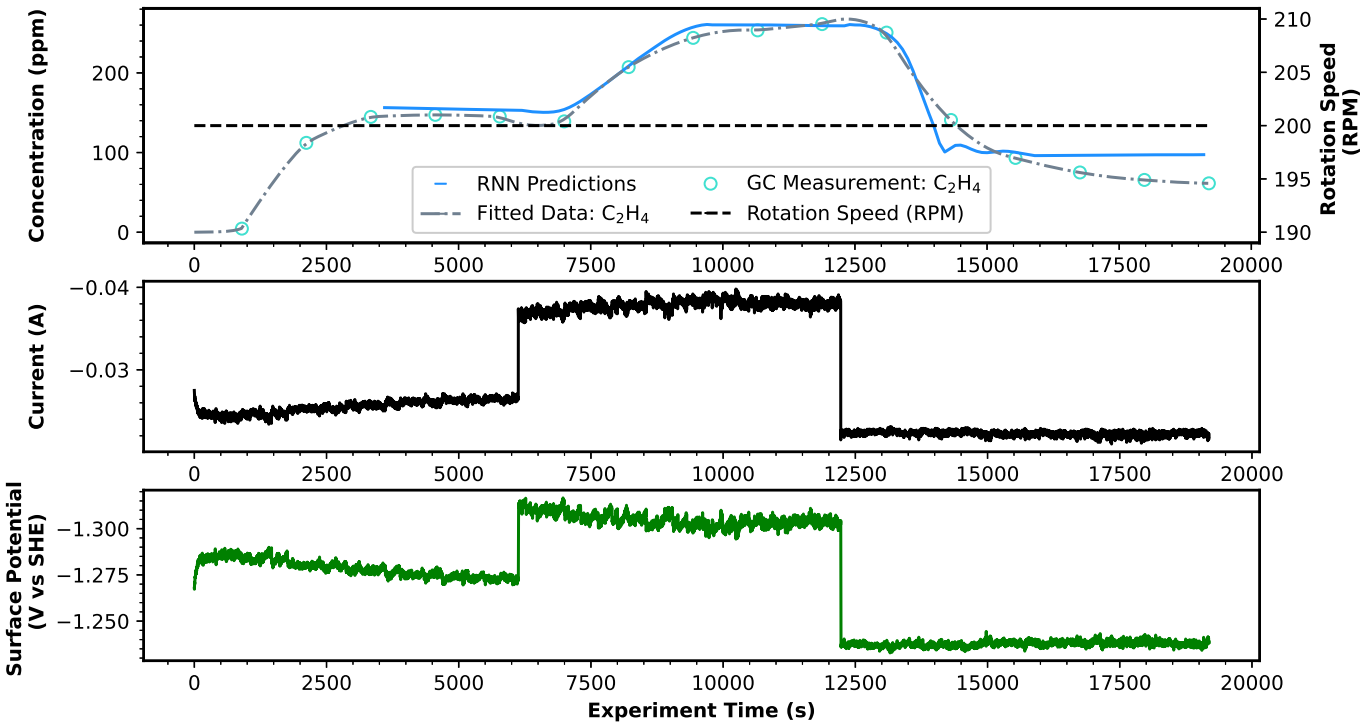


(a) Rolled and unrolled RNN architecture visualization.

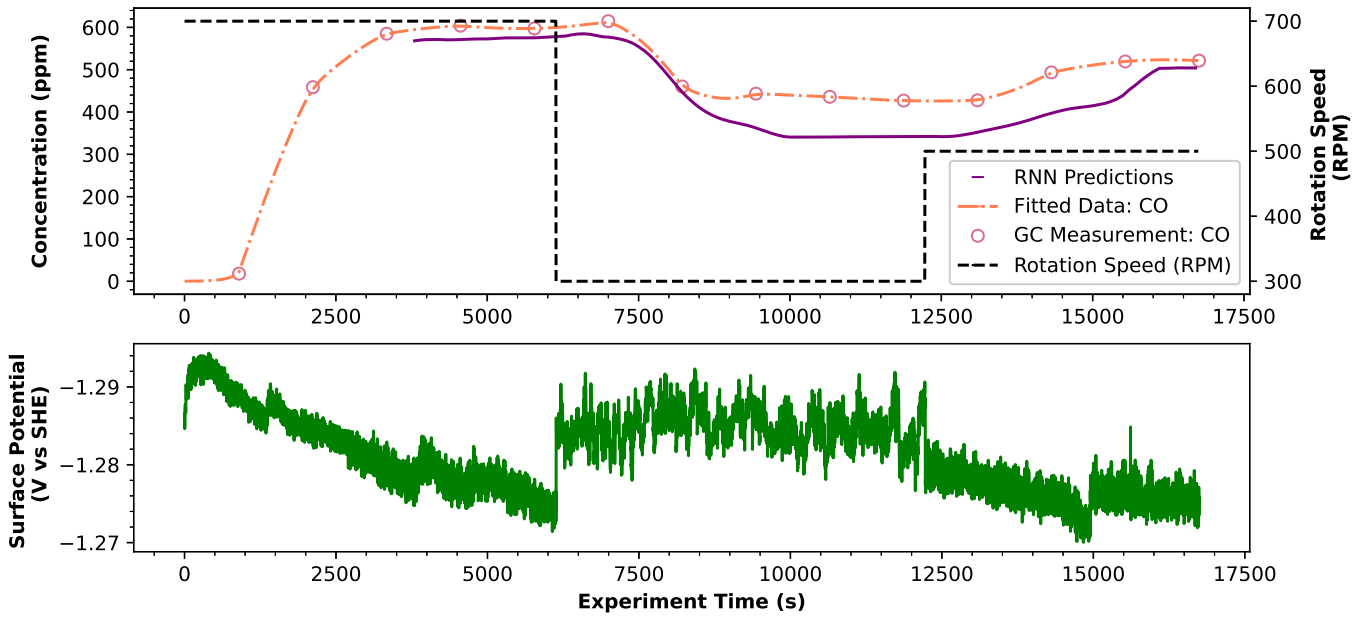


(b) LSTM layer diagram.

Figure 8: Overall structure of recurrent neural network (RNN) and long short-term memory (LSTM) network unit. The LSTM unit fits into the empty circles of the recurrent layer in the top plot.



(a) C_2H_4 concentration prediction model with surface potential (V vs SHE), rotation speed, and current as inputs for a time window of 3600 seconds.



(b) CO concentration prediction model with surface potential (V vs SHE) and rotation speed as inputs for a time window of 3800 seconds. This experiment was conducted under constant applied potential and the change in surface potential is due to the change in current caused by the electrode rotation speed variation.

Figure 9: C_2H_4 and CO RNN predictions for open-loop experiments from the testing set.

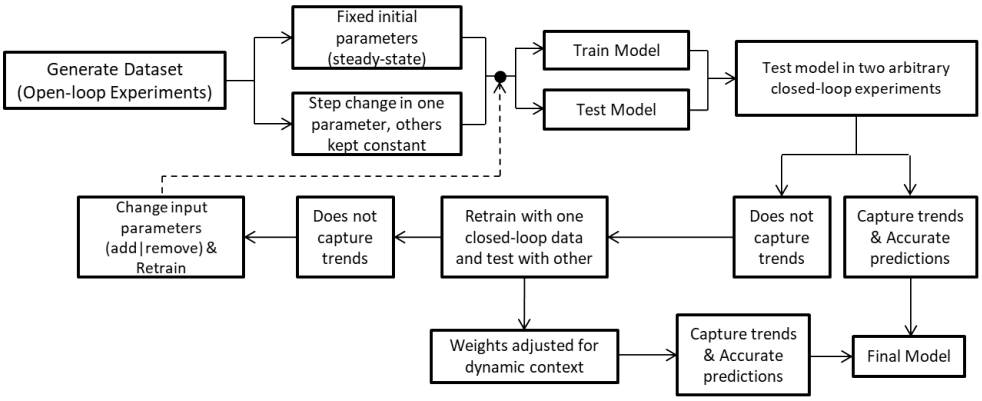


Figure 10: Procedure to optimize the LSTM model using open- and closed-loop experiments.

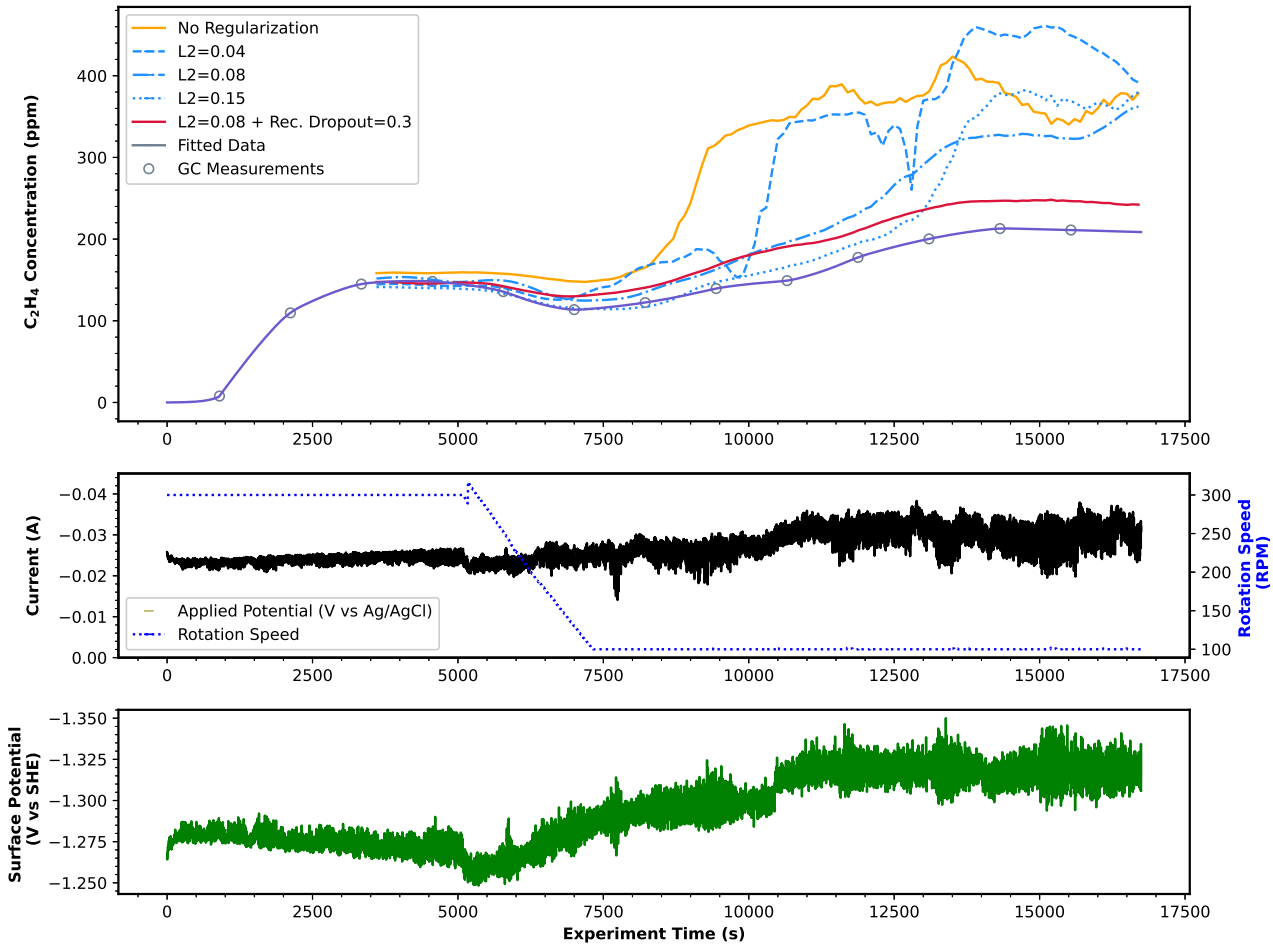
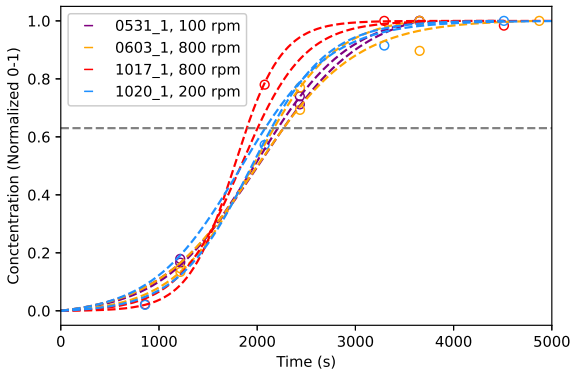
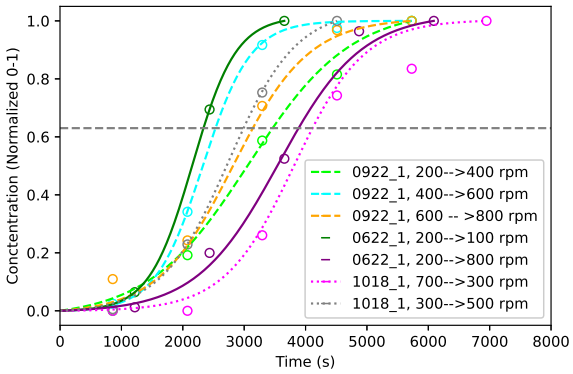


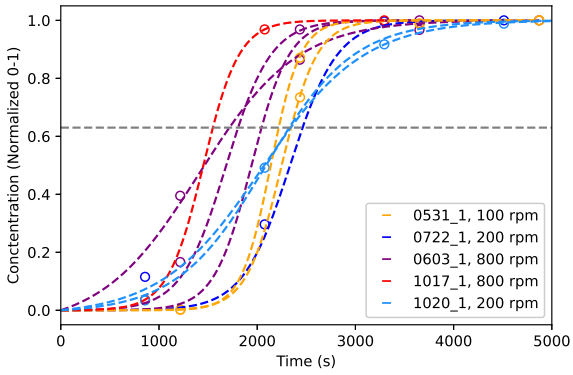
Figure 11: The effect of regularization in improving the model predictions. Multiple parameters for L2 regularization are compared to the addition of recurrent dropout and no regularization case.



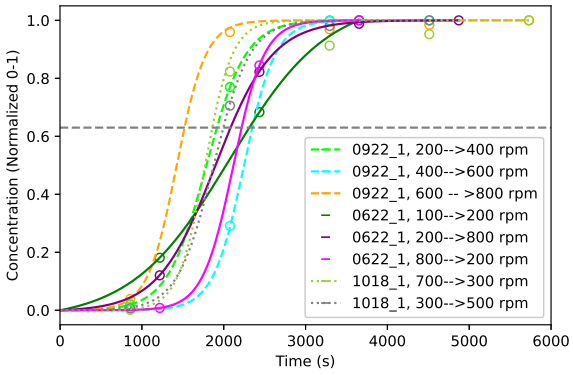
(a) C_2H_4 concentration evolution for potential change under constant rotation speed.



(b) C_2H_4 concentration evolution for rotation speed change under constant applied potential.

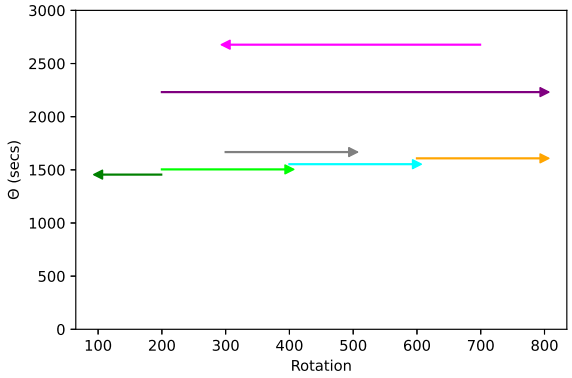


(c) CO concentration evolution for applied potential change under constant rotation speed.

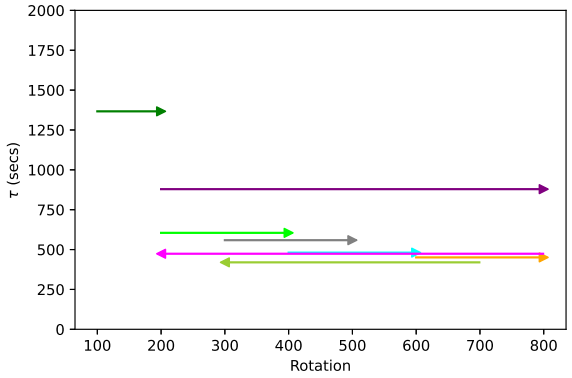


(d) CO concentration evolution for rotation speed change under constant applied potential.

Figure 12: The evolution of C_2H_4 and CO concentrations for step changes in the input parameters, fitted to a sigmoid function and normalized between 0-1.

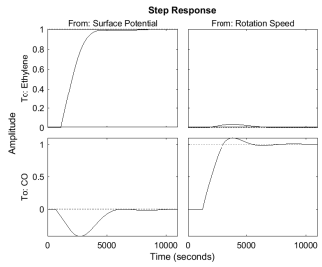


(a) Extracted dead time values for various rotation speed changes in open-loop step change experiments for fitted C_2H_4 dynamic data depicted in Fig. 12 (b).

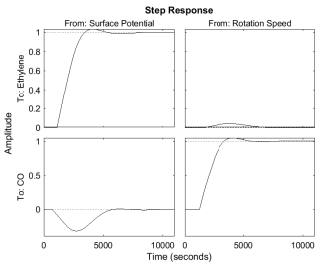


(b) Extracted τ values for various rotation speed changes in open-loop step change experiments for fitted CO dynamic data depicted in Fig. 12 (d).

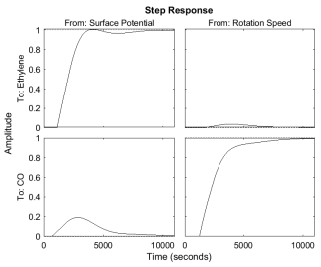
Figure 13: Process dead time (θ) and time constant (τ) distribution extracted from experimental data for dynamic changes.



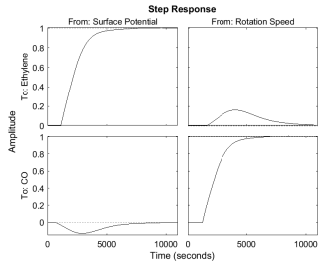
(a) -1.19 to -1.26 V vs SHE and 100-200 RPM.



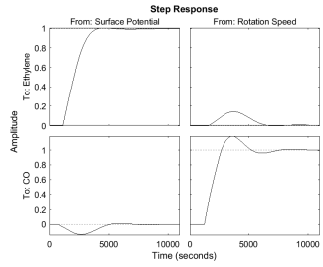
(b) -1.19 to -1.26 V vs SHE and 200-400 RPM.



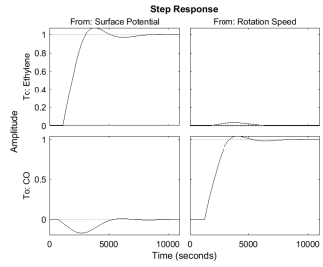
(c) -1.19 to -1.26 V vs SHE and 400-800 RPM.



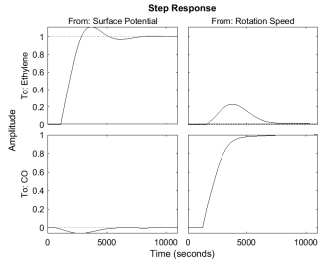
(d) -1.26 to -1.30 V vs SHE and 100-200 RPM.



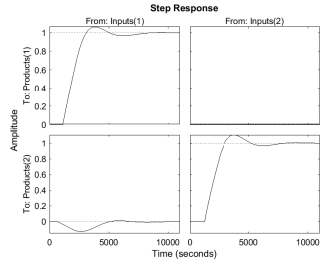
(e) -1.26 to -1.30 V vs SHE and 200-400 RPM.



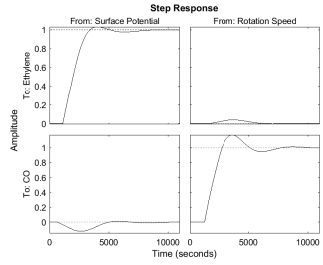
(f) -1.26 to -1.30 V vs SHE and 400-800 RPM.



(g) -1.30 to -1.32 V vs SHE and 100-200 RPM.



(h) -1.30 to -1.32 V vs SHE and 200-400 RPM.



(i) -1.30 to -1.32 V vs SHE and 400-800 RPM.

Figure 14: MATLAB-based tuning of the PI controllers using a decoupler.

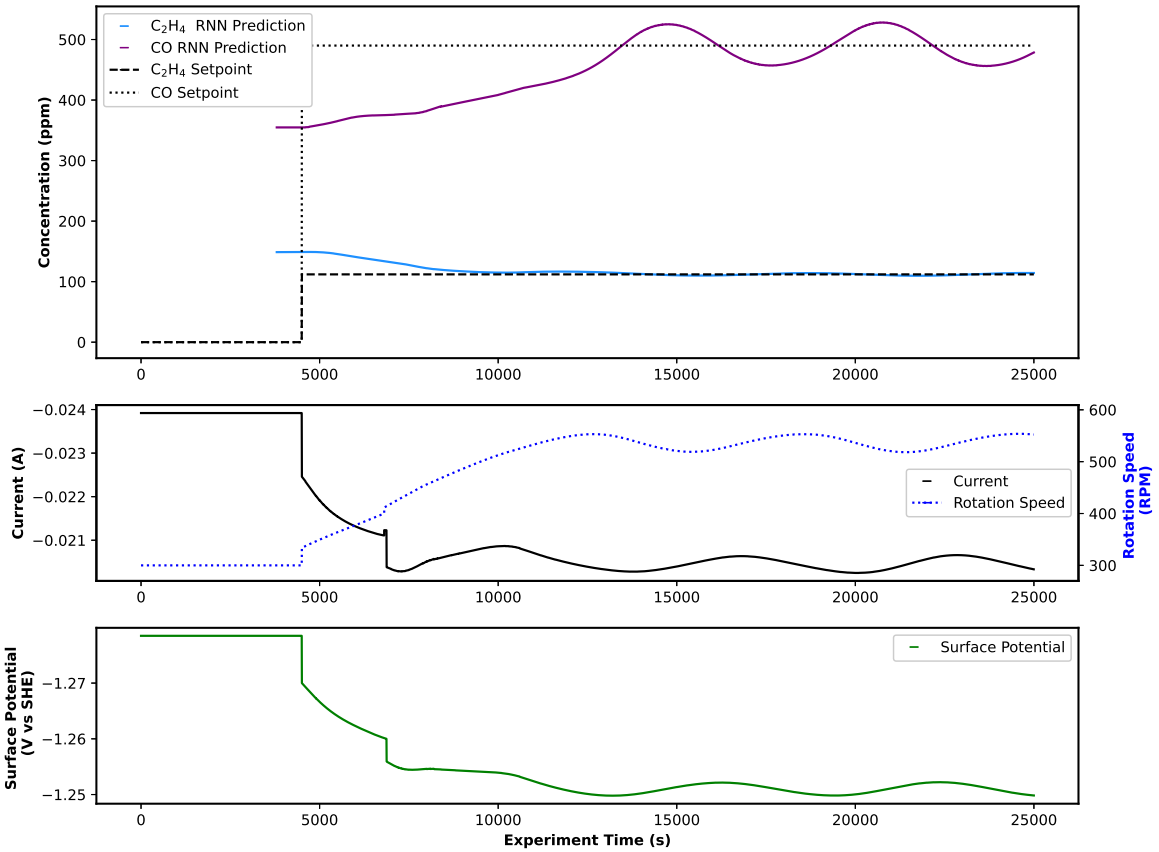


Figure 15: Closed-loop simulation using the dynamic RNN models and the empirical surface potential-current correlation used to determine the final PI controller parameters.

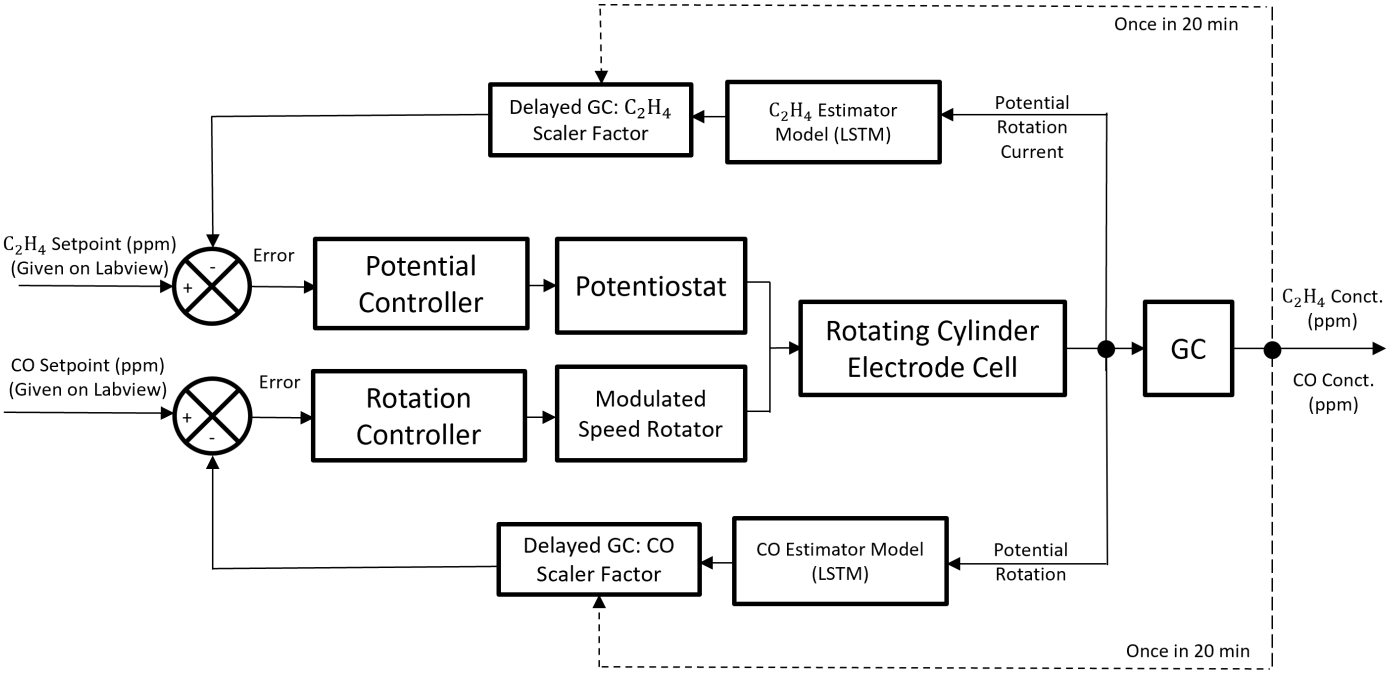


Figure 16: Closed-loop system structure using multi-input multi-output control system with ML estimators.

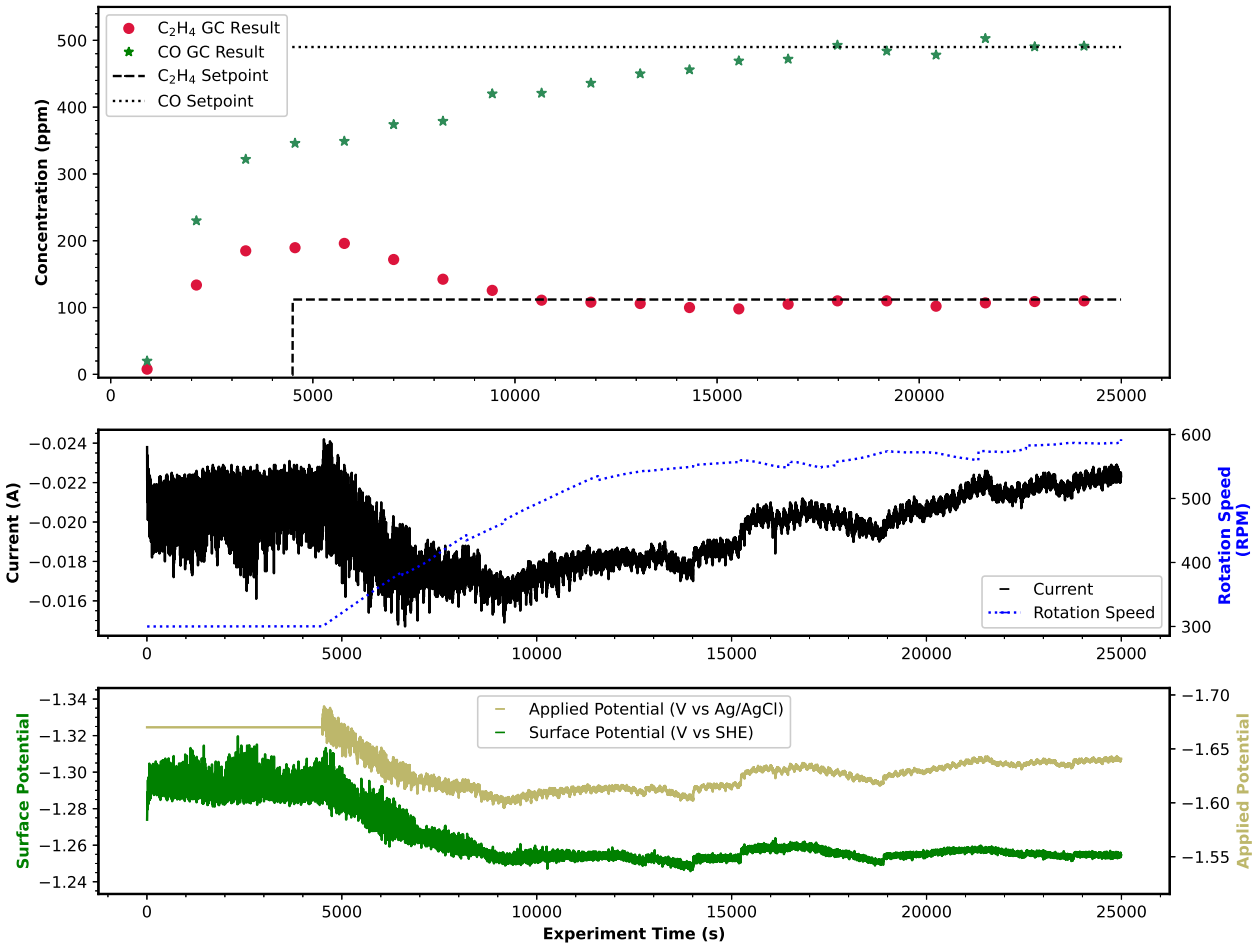


Figure 17: Closed-loop experimental results for the economically optimal set-points for which the C_2H_4:CO selectivity ratio is 1:4.

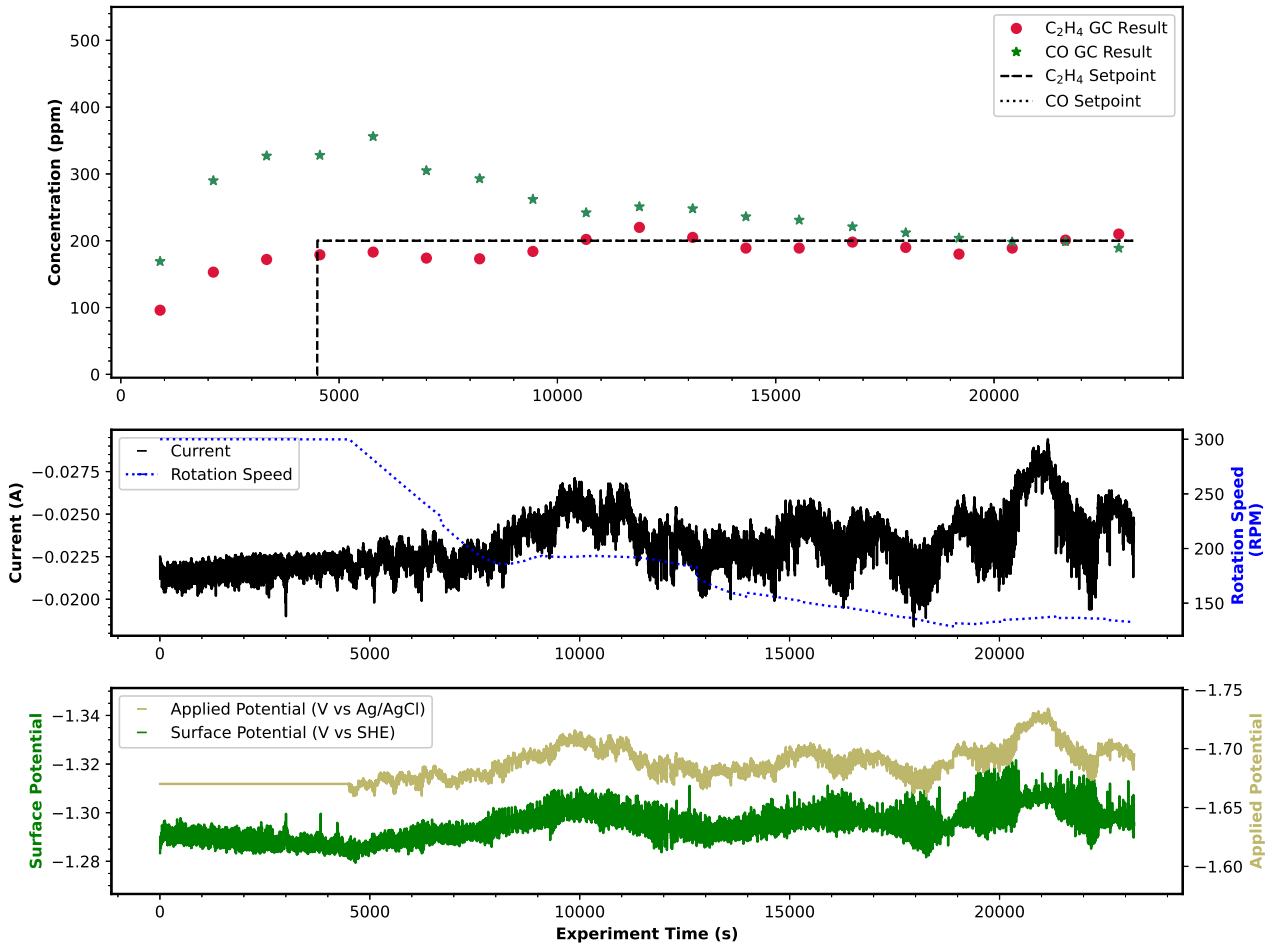


Figure 18: Closed-loop experimental results for the C_2H_4:CO selectivity ratio of 1:1 which results in a higher selectivity towards ethylene compared to the economically optimal case.

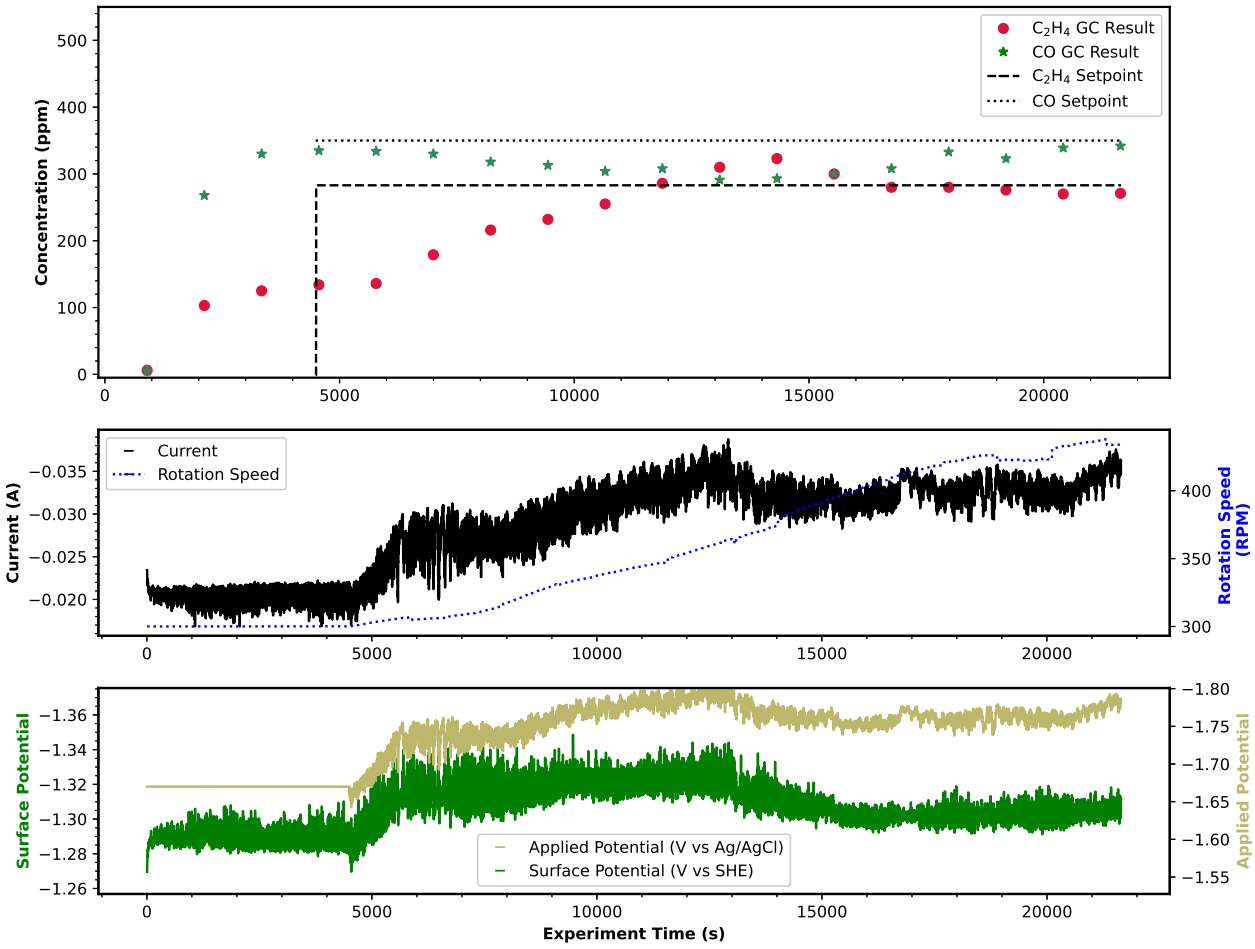


Figure 19: Closed-loop experimental results corresponding to setpoints with C_2H_4:CO selectivity ratio of 4:5.

List of Tables

Table 1: Averages and variations in surface potential, current, and C_2H_4 and CO concentrations in repeated open-loop experiments under different operating conditions.

Rotation Speed	**Potential (V vs SHE)**	**Current (A)**	**C_2H_4 (ppm)**	**CO (ppm)**
100 RPM	-1.31	-0.0296	376	157
	-1.28	-0.0201	217	159
	-1.23	-0.0131	67	211
	-1.20	-0.0086	30	255
800 RPM	-1.30	-0.0359	363	505
	-1.27	-0.0267	216	500
	-1.24	-0.0173	86	488
	-1.20	-0.0129	25	483

(a) Averages of experimental inputs and outputs in various ranges.

Rotation Speed	**Potential (V vs SHE)**	**Current (A)**	**C_2H_4 (ppm)**	**CO (ppm)**
100 RPM	0.0029	0.0029	35	14
	0.0047	0.0032	96	21
	0.0041	0.0002	43	42
	0.0015	0.0002	11	8
800 RPM	0.0128	0.0030	76	65
	0.0019	0.0010	49	56
	0.0060	0.0010	31	22
	0.00199	0.0010	9	16

(b) Standard deviations of experimental inputs and outputs in various ranges.

Table 2: Controller gains (K_p) resulting from MATLAB-based tuning.

		RPM		
		100 - 200	200 - 400	400 - 800
V vs SHE	-1.19 : -1.26	-0.00038		
	-1.26 : -1.30	-0.000126		
	-1.30 : -1.32			

(a) Applied potential controller.

		RPM		
		100 - 200	200 - 400	400 - 800
V vs SHE	-1.19 : -1.26	0.69		1.4
	-1.26 : -1.30			
	-1.30 : -1.32			

(b) Rotation speed controller.

Table 3: Final controller gains (K_p) resulting from closed-loop system simulation.

		RPM		
		100 - 200	200 - 400	400 - 800
V vs SHE	-1.19 : -1.26	-0.00042		
	-1.26 : -1.30	-0.00025		
	-1.30 : -1.32			

(a) Applied potential controller.

		RPM		
		100 - 200	200 - 400	400 - 800
V vs SHE	-1.19 : -1.26	0.25		0.35
	-1.26 : -1.30			
	-1.30 : -1.32			

(b) Rotation speed controller.