Chemical Engineering Research and Design 185 (2022) 87–107

Available online at www.sciencedirect.com

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

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

Berkay Çıtmacı[a], Junwei Luo[a], Joon Baek Jang[a], Vito Canuso[a], Derek Richard[a], Yi Ming Ren[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

ARTICLE INFO

Article history:
Received 11 May 2022
Received in revised form 26 June 2022
Accepted 28 June 2022
Available online 1 July 2022

Keywords:
Electrochemical reactor
CO_2 reduction
Support vector regression
Data-driven modeling
Gas chromatography
Real-time optimization
Feedback control

ABSTRACT

With the increase in electricity supply from clean energy sources, electrochemical reduction of carbon dioxide (CO_2) has received increasing attention as an alternative source of carbon-based fuels. As CO_2 reduction is becoming a stronger alternative for the clean production of chemicals, the need to model, optimize and control the electrochemical reduction of the CO_2 process becomes inevitable. However, on one hand, a first-principles model to represent the electrochemical CO_2 reduction has not been fully developed yet because of the complexity of its reaction mechanism, which makes it challenging to define a precise state-space model for the control system. On the other hand, the unavailability of efficient concentration measurement sensors continues to challenge our ability to develop feedback control systems. Gas chromatography (GC) is the most common equipment for monitoring the gas product composition, but it requires a period of time to analyze the sample, which means that GC can provide only delayed measurements during the operation. Moreover, the electrochemical CO_2 reduction process is catalyzed by a fast-deactivating copper catalyst and undergoes a selectivity shift from the product-of-interest at the later stages of experiments, which can pose a challenge for conventional control methods. To this end, machine learning (ML) techniques provide a potential approach to overcome those difficulties due to their demonstrated ability to capture the dynamic behavior of a chemical process from data. Motivated by the above considerations, we propose a machine learning-based modeling methodology that integrates support vector regression and first-principles modeling to capture the dynamic behavior of an experimental electrochemical reactor; this model, together with limited gas chromatography measurements, is employed to predict the evolution of gas-phase ethylene concentration. The model prediction is directly used in a proportional-integral (PI) controller that manipulates the applied potential to regulate the gas-phase ethylene concentration at energy-optimal set-point values computed by a real-time process optimizer (RTO). Specifically, the RTO calculates the operation set-point by solving an optimization problem to maximize the economic benefit of the reactor. Lastly, suitable compensation methods are introduced to further account for the experimental uncertainties and handle catalyst deactivation. The proposed modeling, optimization, and control approaches are

* Corresponding author.
** Corresponding author at: Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA.
E-mail addresses: moralesguio@ucla.edu (C.G. Morales-Guio), pdc@seas.ucla.edu (P.D. Christofides).
https://doi.org/10.1016/j.cherd.2022.06.044

the first demonstration of active control for a CO_2 electrolyzer and contribute to the automation and scale-up efforts for electrified manufacturing of fuels and chemicals starting from CO_2.

1. Introduction

Over the last decade, a great deal of interest has been elicited around the idea of using renewable-based electricity rather than fossil fuels as the energy source for large scale manufacturing of chemicals (e.g., De Luna et al., 2019). Direct electrocatalytic transformation of carbon dioxide (CO_2) to fuels and chemicals can enable global scale renewable energy storage and close the anthropogenic chemical carbon cycle. However, outside of well-established chloralkali, water electrolysis, and aluminum refining processes, most electrified chemical manufacturing processes are limited to bench-scale demonstrations. This is particularly the case for processes that use gaseous reactants and produce a complex mixture of multiple different products, such as the electrochemical reduction of CO_2 (e.g., CO_2 electrolyzers are currently limited to electrode areas of around 5 cm^2 or less) or the reduction of nitrogen to ammonia (Jeng et al., 2022; Niu et al., 2021). One of the major challenges of industrially implementing electrochemical reduction of CO_2 is the absence of a well-understood reaction mechanism (Jin et al., 2021).

To further investigate 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 and reaction kinetics as well as promote the production of multiple valuable products (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 the reactor, which implies the potential to control the product distribution of the reactor by manipulating certain inputs. The demonstration of control in production rates during CO_2 electrochemical reduction is rare, with only one example by Creel et al. (2019) demonstrating that it is possible to control the selectivity of CO by using light to illuminate the cathode in plasmonic catalyzed electrochemical CO_2 reduction reactions.

The development of chemical process models for control can be approached from two opposite directions, one using first-principles to build up a model and the other using a data-driven approach that condenses the input-output relations in a black-box model. The use of machine learning (ML) models in the context of electrochemical reactors has received attention over the past ten years due to their demonstrated ability to approximate universal linear/nonlinear relations with little a priori knowledge of the system (Corriou, 2004). ML has been used broadly from efforts geared towards gaining understanding of the role of adsorbates in changing the catalytic properties of copper electrodes to the discovery of CO_2 electrocatalysts based on a large data set of electrolyzer literature. For example, Wu et al. (2021) used a four-layer neural network as well as other ML algorithms to investigate the influence of different supplementary adsorbates on the performance of electrochemical CO_2 reduction on copper surface, and reported tens of adsorbates that have a major impact among other candidates. In another work, Malek et al. (2021) tested various ML algorithms, such as bagging regressors, regression trees, and gradient boosting, to predict the most feasible electrochemical CO_2 reduction catalyst material using an extensive database comprising academic and industrial reports. Alternatively, classic ML algorithms usually require less training data than deep learning techniques and are better suited for the analysis of electrochemical data sets acquired under more controlled and better defined conditions of transport, such as those generated in the RCE cell (Mahesh, 2020). Among ML algorithms, support vector machines (SVM) provide a complementary way to perform data-driven modelling with lower probability of over fitting (Tan et al., 2018; Ray, 2019). Proposed by Vapnik et al. (1996), support vector regression (SVR), based on SVM, can be used to perform regression for static and dynamic models by defining a margin around a proposed hyperplane for data fitting and penalizing data points beyond the margin. This method has shown strong

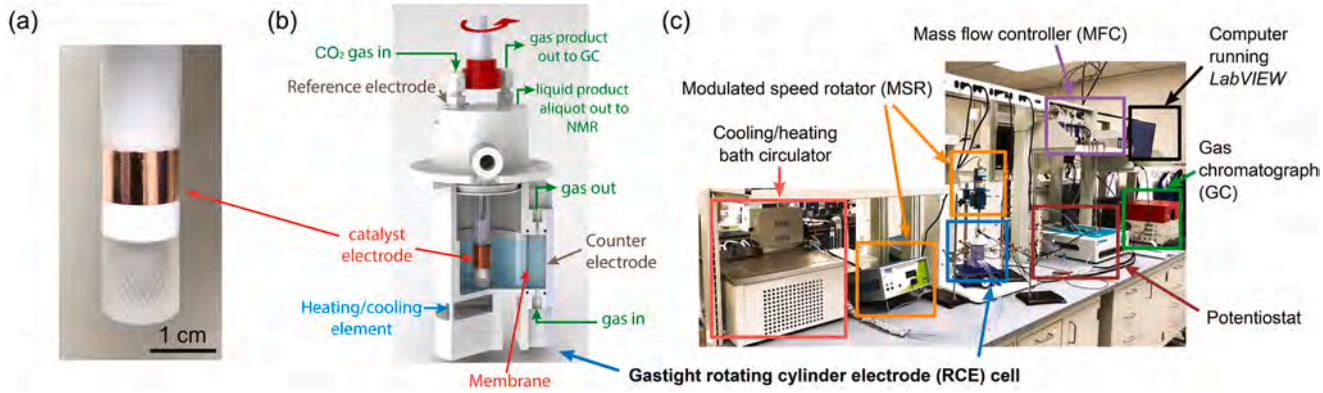


Fig. 1 – Components of the electrochemical reactor. (a) Electropolished atomically-flat polycrystalline copper cylinder electrode, (b) Schematics of gastight rotating cylinder electrode (RCE) cell, and (c) Bench-scale electrochemical reactor setup with sensors, actuators, and automation devices.

performance in representing time-series data and pattern recognition (Basak et al., 2007; Yu and Jiang, 2016; Burges, 1998). For example, Pascual et al. (2013) developed an SVR model to predict the steady-state performance of a reverse osmosis desalination plant using 3990 steady-state data points and further simulated its dynamic perturbations from steady-state operation. Shokry et al. (2015) demonstrated the ability of SVR to model a dynamic Photo-Fenton process, a photochemical oxidation reaction common in pollutant treatment, and the SVR model effectively predicted the production of OH radicals with data from a pilot plant. Tan et al. (2018) modeled NO_X emissions from coal combustion boilers based on observations of 73 h of experiment using SVR and Artificial Neural Network (ANN), which showed that SVR is a more robust option than ANN for this process.

Motivated by the above considerations, this work proposes an ML-based scheme to implement real-time optimization (RTO) and feedback control in an experimental electrochemical reactor for CO_2 reduction. Specifically, an SVR model is developed on the basis of existing experimental data to estimate the dynamic response of the reactor operation, and to account for inherent disturbances such as sensor uncertainties and catalyst degradation. Subsequently, the information from the sensors and ML model is integrated and used by a Proportional-Integral (PI) controller that manipulates the input of the reactor. In addition, a real-time optimizer (RTO) is developed to compute the optimum set-point for the reactor by integrating the steady-state prediction from a neural network model and valid market information. The proposed control and optimization scheme is demonstrated by a series of experiments that control ethylene production of the RCE reactor.

The rest of this manuscript is organized as follows. In the section entitled “Preliminaries”, the experimental reactor setup and catalyst deactivation are described. In the next section, entitled “Machine-Learning Modeling", the experimental method of extracting, formulating and building the SVR model and related model improvements are discussed. In the following section entitled “Real-Time Optimization", the methodology to calculate the most economically feasible set-points is discussed. In the section entitled “Feedback Control", the implementation of a PI controller is integrated with an estimator for the reactor overhead ethylene concentration and real-time GC measurements; then the performance of the model and the controller is evaluated.

2. Preliminaries

This section introduces the background of the experimental electrochemical reactor employed in this work. Specifically, the experimental system and methodology are discussed and important variables are explained. An overview of the catalyst deactivation is then used to explain the background of the control objective. All process equipment mentioned in this section, except nuclear magnetic resonance (NMR), is connected to a Laboratory Virtual Instrument Engineering Workbench (*LabVIEW*) interface. The operational procedures required in the experiment, including but not limited to real-time GC data processing, controller activation, and equipment actuation, are fully automated by a computer program developed with Python and implemented through a *LabVIEW* interface. Fig. 1 shows UCLA’s RCE system along with the array of sensors and actuators in the system.

The ML model is integrated into *LabVIEW* and the outputs of the ML model are calculated on a per-second basis. The parameters of the feedback controllers are assigned a priori to take specific values for specific time intervals of operation. This includes closed-loop control experiments where the set-point is changed between two time intervals. For gas product quantification, a Python script was developed to automatically process the raw GC data, detect the gas product signals, and calculate the corresponding gas phase concentrations utilizing a GC caligration file. This GC script is then used to get real-time data to correct the ML model and improve the controller performance. The process data is synchronized with an online database of the Smart Manufacturing Innovation Platform (*SMIP*) provided by The Smart Manufacturing Institute (*CESMII*). *SMIP* was used for data storage and monitoring and to facilitate data accessibility for all researchers when building models on site.

2.1. *Electrochemical reactor setup*

We have recently described the construction of an RCE reactor setup (Jang et al., 2022) which can be used to produce over 16 different gas and liquid products during the electrochemical reduction of CO_2 on a copper electrode. As shown in Fig. 1 b, this electrochemical reactor setup consists of six major components. These are: the reactor with its two chambers containing respectively the working (cathode) and counter (anode) electrodes separated by an ion-exchange membrane that prevents product cross-over between the two chambers, a potentiostat that regulates the potential applied to the working electrode, a mass flow controller to adjust the mass flowrate of the feed gas (CO_2), a modulated speed rotator (MSR) to adjust the rotation speed of the electrode, a cooling/heating block to control the temperature of the cathode compartment, and a computer. In addition to these six major components, a gas chromatograph (GC) and a nuclear magnetic resonance (NMR) spectrometer are utilized to quantify the chemical species in the gas and liquid-phase products, respectively. In the work presented here, the GC has been automated for on-line gas product quantification.

During the experiment, inlet flowrate of feed gas is fixed to 20 $mL\,min^{-1}$ by the mass flow controller and is bubbled directly through a 0.2 M $KHCO_3$ electrolyte solution in the cathode chamber. Subsequently, CO_2 is reduced to synthetic fuel and chemical products which result from proton-electron transfer processes occurring on the surface of the rotating cylinder electrode, which serves as the cathode and working electrode. In addition, the potential applied to the working electrode is continuously measured against the reference electrode, while the MSR modulates the rotation speed of the RCE through magnetic coupling and the reactor temperature is controlled through a bath circulator that flows coolant at a specified temperature through the cooling/heating block. Eventually, gas products leave the reactor headspace and travel through polytetrafluoroethylene (PTFE) tubing to be quantified by the GC once every sampling period. The liquid products accumulate in the electrolyte solution and are then quantified by NMR after electrolysis (Jang et al., 2022). Since the concentrations of the liquid products cannot be measured while the reactor is operating (NMR sample preparation, analysis and quantification take long and are infeasible to implement feedback control on this data), only the gas products are considered as the outputs of the process for the proposed control scheme.

Remark 1. In addition to working as an actuator, the potentiostat also functions as a sensor that measures and records the current density flowing between the working electrode and the counter electrode in real-time. Based on this measurement, the potential on the surface of the electrode, which is known to directly affect the electrochemical reaction, can be calculated with reference to the standard hydrogen electrode (SHE) by Eq. (1).

$$E_{surface} = E_{applied} - i \times R + E_0 \quad (1)$$

where $E_{surface}$ is the surface potential, $E_{applied}$ is the applied potential measured against the reference electrode, i is the current density (negative value for reductive currents), and R is the resistance of the solution between the working electrode and the reference electrode measured using electrochemical impedance spectroscopy (EIS) (Jang et al., 2022).

E_0 represents the standard electrode potential of the reference electrode (Ag/AgCl/1 M KCl), so that $E_{surface}$ is referenced against the standard hydrogen electrode (V vs. SHE) after correction.

2.2. Identification and quantification of the process output

Among all products generated by this electrochemical reactor, hydrogen (H_2), carbon monoxide (CO), methane (CH_4), and ethylene (C_2H_4) are in the gas phase and can be detected using gas chromatography (GC). The relevant reactions for these products are shown below:

$$2CO_2 + 8H_2O + 12e^- \rightarrow C_2H_4 + 12OH^- \quad (2a)$$

$$CO_2 + 6H_2O + 8e^- \rightarrow CH_4 + 8OH^- \quad (2b)$$

$$CO_2 + H_2O + 2e^- \rightarrow CO + 2OH^- \quad (2c)$$

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad (2d)$$

The measured production of gas products (C_2H_4, CH_4, CO, and H_2) by the GC is the target component of the feedback control scheme, which is implemented by the following procedure. Firstly, the temperature-programmed GC separates molecules injected onto the columns, which have different elution times. Subsequently, separated molecules are detected using a thermal conductivity detector (TCD) and a flame ionization detector (FID) in the GC system. In this study, the temperature program runs for 14.33 min after which the heated column oven requires 6 min of cooling time before the next injection. This means that the GC has a sampling period of 20.33 min. The delay in the GC analysis limits the possibilities for real-time control, but this can be overcome to some extend through the development of a ML-based estimator which allows for the calculation of real-time gas compositions in the cell overhead. This estimator is described in Section 3.

Remark 2. Because of the strong correlation of the ethylene productivity with respect to the surface potential and the electric current, it is chosen as the initial component of interest for modeling and control. In addition, ethylene has numerous industrial applications and is the most produced organic molecule worldwide with a current installed production capacity of over 200 million metric tons per year.

2.3. Catalyst deactivation

Catalysts play a central role in electrochemical reactors. Our experimental process uses a smooth, cylindrical, polycrystalline copper electrode shown in Fig. 1 a. Copper has previously been shown to have the highest selectivity for C_2 products, making it the catalyst of choice in CO_2 electrolysis (Kim et al., 2014). However, the copper catalyst continuously degrades as the reactions proceed, and this causes a drift in steady state, even when the input (typically applied potential which is the driving force for electrochemical reactions) is held constant. The catalyst deactivation is attributed to surface restructuring, blocking of sites by reactive carbon species, and absorption of impurities from the electrolyte (Nitopi et al., 2019). The deactivation is particularly fast on flat, non-porous catalysts such as the one used here. Electrodes with a higher porosity have a higher density of active sites and can be operated at lower overpotentials, so the deactivation processes appear to occur over longer time periods (Kas et al., 2015).

The deactivation mechanism of copper under electrochemical CO_2 reduction environments is complicated and different potential sources of deactivation have been pointed out over the last decades. Hori et al. (2005) showed that metal impurities, mainly Zn^{+2} and Fe^{+2}, and trimethylamine existing in a 0.5 M $KHCO_3$ electrolyte accumulated on the copper electrode during testing affecting product selectivity and current densities. These authors recommended the utilization of reagents of the highest purity or the use of pre-electrolysis using platinum black cathodes as a way to purify the electrolyte solution and delay catalyst deactivation (Hori et al., 2005). Another source of deactivation that has been proposed is formation of graphitic carbon species from decomposition of reactive intermediates (DeWulf et al., 1989; Shiratsuchi et al., 1993; Xie et al., 2014). Such reaction intermediates could block catalytic sites and poison the electrode surface. On the other hand, Kim et al. observed surface reconstruction of polycrystalline copper to Cu(100) facet both in KOH (Kim et al., 2014) and $KHCO_3$ (Kim et al., 2018) electrolytes under reductive potentials and associated these surface restructuring to changes in product distribution. Despite the complexity of the deactivation mechanism which can be attributed to various factors, anodic pulsing could be used to mitigate catalyst deactivation and prevent changes in activity and selectivity of copper under operation (Shiratsuchi et al., 1993; Engelbrecht et al., 2018).

Catalyst reactivation procedures such as anodic pulsing or potential sweeping could be programmed and integrated eventually into the control system of future CO_2 electrolyzers. However, catalyst oxidation procedures can also lead to excessive surface roughening, loss of electric conductivity and catalyst dissolution and must be further investigated. In this work, we have introduced an additional parameter of the cumulative integral of current to address the deactivation issue. As a general trend, either the energy required for molecules to pass the activation energy barrier increases or the number of active sites decreases as the catalyst continues to deactivate. One way to compensate for the loss of activity is to increase the applied electric potential to ensure that a similar number of reactant molecules can continue to be transformed despite the deactivation. This is accomplished by interpreting general trends of deactivation and introducing the integral of current parameter. Integration of the current passed in an electrolyzer is a simple, yet effective, way to track the degree

of use of a catalyst within an electrochemical system with broad applications beyond CO_2 electrolyzers and will be discussed in detail in Section 3.4.

Remark 3. A different catalyst morphology may considerably delay the deactivation. Specifically, for electrochemical CO_2 reduction reactions, another available catalyst morphology is copper cubes that have pores on the electrode surface, allowing both internal and external mass transfer to play a role (Roberts et al., 2015). However, this study is only focused on the reactions with flat copper and compensation of the catalyst deactivation via introduction of the cumulative integral of current parameter and feedback control. Future work will investigate control with the copper cube catalyst.

3. Machine learning modeling

One of the main objectives of this research is to control the experimental electrochemical reactor. In the absence of a first-principles model for the electrochemical system, a simple PI controller could be implemented and tuned on the basis of direct measurement feedback without the use of any model. However, this scenario exhibits many limitations in terms of control performance. To construct a closed-loop control system, the frequency of feedback from GC measurements (20.33 min) would be inadequate to provide accurate and reliable control of the reactor. An effective PI control in the presence of catalyst deactivation would benefit from feedback on a per-second basis rather than once every 20.33 min from the GC. The use of a Fourier Transform Infrared Spectrometer (FTIR), a device that can detect the concentration of gas products every second with high reliability, was considered as an alternative method of measuring and providing feedback to the control system (Ke et al., 2022). However, because of the higher volume flowrate required for FTIR, the product concentrations from the reactor would have been reduced by an order of magnitude. This reduction in concentration would have placed the product concentrations near the limit of detection, making FTIR inadequate for the reactor used in this investigation. Due to this limitation, building a dynamic machine-learning model is deemed to be a necessary step in controlling the electrochemical system.

3.1. Data collection and pre-processing

The data used to develop the ML model is collected from open-loop experiments performed by the following procedure. Mass flow controllers are arranged to maintain 20 $mL\,min^{-1}$ of CO_2 gas flow to the reactor. Specifically, before each experiment, the resistance of the buffer solution is measured, and the applied potential is adjusted to compensate for the measured solution resistance. This is done automatically by the potentiostat during the experiment. A steady rotation speed is set for the cylinder electrode so that the hydrodynamics are well developed in the cell, and then the applied potential (V vs. Ag/AgCl) is set to a desired value. Both rotation speed and applied potential are kept constant throughout the experiment. The experiment takes approximately 80 min, and 4 GC measurements are taken during the experiment at 15, 35, 55 and 75 min mark. The corresponding real-time current and applied potential values are recorded simultaneously. It is known that concentration of the gases in the reactor overhead is not equilibrated by the first GC injection, so this data point is ignored for the purposes of modeling reactor performance. As a result, each experiment produces 3 data points of concentrations for each gas product. The corresponding current densities and surface potentials for relevant GC runs are averaged in a time frame to represent the gas equilibrium in the reactor overhead. It is assumed that the average residence time of this electrochemical reactor is around 5–8 min under conditions operating close to steady state. Therefore, the surface potential and current values are averaged in a 3-minute time window, from 8 to 5 min prior to the GC injection to best represent the current density and surface potential corresponding to each GC result.

The experiments with GC measurements below the detection limit or with unusual increases in the electrolyte resistance were marked as outliers. After these data points have been eliminated, the results of 48 experiments are considered for modeling. The resulting database includes 144 data points of surface potential, current, rotation, and gas product concentrations.

Remark 4. Using small-scale experimental data to develop an ML model raises additional unique challenges in comparison to using simulation or well-structured industrial data, since it contains more experimental uncertainty. The measured resistance of the $KHCO_3$ solution can vary for each trial depending on the preparation of stock solution, temperature, conductivity of the inner electrical circuit, etc., although the values were kept as consistent as possible ($7 \pm 0.2\,\Omega$). Similarly, the electrodes are both mechanically and electrochemically polished prior to each experiments, but even the standardized polishing process can lead to different catalyst activities at low overpotentials. Furthermore, data measurements are limited by the detection range of the sensors, since the production scale is small with this experimental setup. As a result, it is possible to obtain different observations from the same input conditions. Thus, averages and standard deviations of some of the data from open-loop experiments, presented in Table 1, are utilized in the ML modeling.

3.2. Model selection and evaluation

Luo et al. (2022) have modeled the reactor setup for the estimation of the steady-state production rate using an ANN. However, this ANN model only gives the average estimation of what the reactor would produce throughout the 80-minute experiment and cannot be used for dynamic data modelling. A very effective method of constructing a dynamic data driven model involves using Recurrent Neural Networks (RNN) as this architecture accounts for the history of the process. Wu et al. (2019) examined in detail the use of RNNs in process control. One of the dynamic ML modeling techniques is Long Short Term Memory (LSTM) RNNs. LSTM networks have an underlying architecture that processes a time window and keeps the necessary information at specific time steps. Therefore, the use of LSTM to store these dependencies is justified. However, neural network architectures are very data-dependent, and 3 data points with 20-min intervals per experiment is insufficient to capture the general behavior of an 80-minute experiment. Another example is using ML techniques for parameter-based modeling (such as reaction rate at specific temperatures) and combining this ML model with first-principle time-dependent equations.

Based on the work of Luo et al. (2022), the applied potential and rotation speed are two important inputs for

Table 1 – Variations in surface potential, current, and C_2H_4 concentrations in repeated open-loop experiments at certain operating conditions.

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

Potential (*V vs SHE*)	Total Current Density (mA/cm^2)	C_2H_4 Concentration (*ppm*)
-1.447	11.25	140.61
-1.426	9.6	150.71
-1.407	6.09	139.01
-1.365	3.87	47.81
-1.318	1.7	9.62
-1.262	1.19	1.28

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

Potential (*V vs SHE*)	Total Current Density (mA/cm^2)	C_2H_4 Concentration (*ppm*)
0.010	2.360	64.38
0.001	1.687	42.40
0.005	0.960	23.397
0.013	1.059	13.93
0.010	0.214	2.64
0.003	0.082	1.82

characterizing the reactor operation. However, these two inputs are held constant throughout the experiment. To create a dynamic state estimator, the model must be trained with inputs that vary during the experiment. Thus, the state estimator model should also incorporate other inputs, such as the current and surface potential, which vary throughout the experiment with the catalyst degradation.

The available data is separated into training and testing sets. 80% of the data is destined for training and used to train various machine learning methods, including linear, ridge, lasso, polynomial regressions, support vector regression, and gradient-boosted decision tree algorithms. Five-fold cross-validation (CV) and mean square error (MSE) were used for model selection. The MSE of the different models are plotted against each other in Fig. 2.

Among the tested regression methods, two of them had remarkably higher training and testing performance, which were gradient boosting and support vector regression (SVR) with a polynomial kernel. Since decision tree-based algorithms, such as gradient boosting, do not produce a smooth output (Guelman, 2012), they present challenges when combined with control. Boosting methods are very successful at predicting the GC measurement points but can be very misleading for the areas in-between two consecutive GC points. In contrast, SVR predictions are more realistic, as they yield predictions for the operating conditions between GC points that align closely with the expected behavior of the system. Thus, the polynomial kernel SVR is chosen because of its smooth, continuous properties and ability to predict intermediate points more accurately.

Remark 5. It is important to note that the SVR model is trained only from GC measurements as described in Section 3.1, so the SVR model cannot capture the gas concentration prior to the GC measurement at 35 min. This requires the gas production to reach a pseudosteady state such that the production rate at the catalyst surface is equal to that of the gas products entering the GC. In the beginning of the experiment, the electrolyte is not saturated with the product gases such as hydrogen, ethylene, methane and carbon monoxide. As a result, the initial gases produced must dissolve in the electrolyte until the electrolyte is saturated. Then, the gases begin to accumulate in the reactor overhead and become detectable by the GC. Therefore, until equilibrium is reached, the SVR predictions are different from the actual GC readings. After the time necessary for equilibrium passes, the SVR predictions and GC measurements start to converge.

3.3. *Support Vector Regression (SVR)*

Support Vector Machine algorithms are similar to least-squares-based regression methods. However, instead of minimizing the residuals, SVR aims to optimize the generalization error bound; thus, it tries to come up with a generalized regression equation (Basak et al., 2007). SVR assumes that the support vectors have at most a deviation of ε from the proposed function and idealy intends to keep all data points within this margin. The SVR algorithm only penalizes points that are outside the support vectors (Basak et al., 2007). A visualization of non-linear regression parameters of SVR is shown in Fig. 3.

We note that ξ, ξ^* are slack variables for opposite sides of the support vectors and represent the distance between the outside points and the support vectors. For a linear SVR, the optimization problem is as follows (Basak et al., 2007):

$$\begin{aligned} \min \quad & \tfrac{1}{2}\,\|w^2\| + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \\ \text{st.} \quad & \begin{Bmatrix} y_i - \langle w, x\rangle - b \le \varepsilon + \xi_i \\ \langle w, x\rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{Bmatrix} \end{aligned} \tag{3}$$

where $\langle \cdot, \cdot \rangle$ is the inner product, w is the weight matrix, C is a parameter that decides the tolerance limit for divergences greater than ε, and l is the number of data points beyond the support vectors. The first two optimization constraints represent the support vectors. Solving the above convex optimization problem yields the linear regression function. The same procedure can be applied to a kernel function such as a polynomial. The degree of the polynomial is selected prior to the optimization.

The polynomial kernel $k(x_j, x_k)$ can be defined as follows:

$$k(x_j, x_k) = (1 + x_j^T x_k)^n \tag{4}$$

where x_j and x_k are instances of input vectors, x_j^T is the transpose of x_j, and n is the order of the polynomial.

$$k(x_j, x_k) = \langle \varphi(x_j), \varphi(x_k) \rangle \tag{5}$$

where $\varphi(x)$ is a non-linear kernel function and $\langle \cdot, \cdot \rangle$ denotes the inner dot product between the values. While our training data has 4 features, for simplicity, Eq. (6e) below only illustrates 2 features with a 2^{nd} degree polynomials for the two instances of the input vectors x_1 and x_2. Each feature contains the current, i_i, and surface potential, e_i, data of two experiments as shown in the following example:

$$x_1 = [e_1, i_1] \tag{6a}$$

$$x_2 = [e_2, i_2] \tag{6b}$$

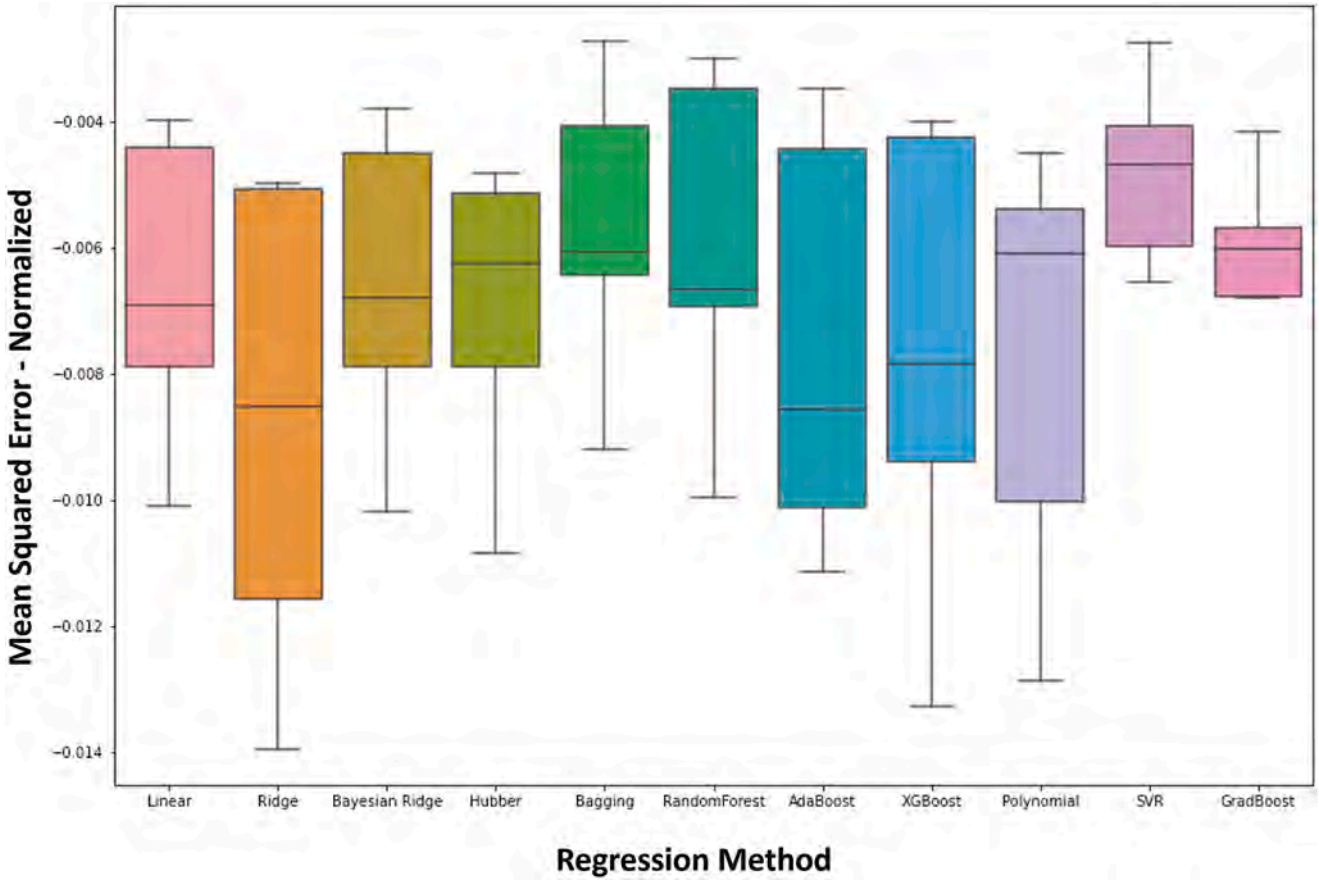


Fig. 2 – Box plots for five cross-validation mean squared errors on a normalized scale for multiple candidate ML models.

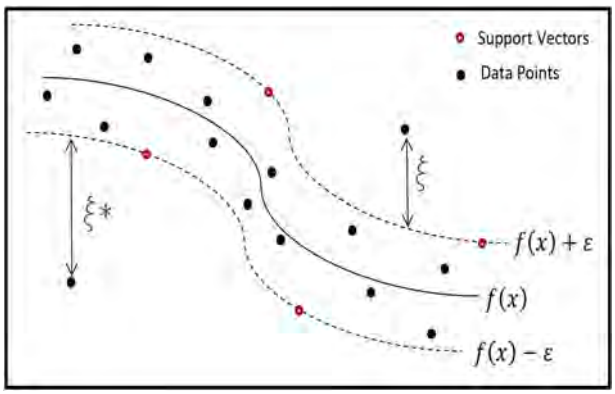


Fig. 3 – The proposed hyperplane f(x) and the margins f(x) ± ε.

$$\varphi(x_1) = [1, e_1, i_1, e_1^2, i_1^2, e_1 i_1] \tag{6c}$$

$$\varphi(x_2) = [1, e_2, i_2, e_2^2, i_2^2, e_2 i_2] \tag{6d}$$

$$k(x_1, x_2) = 1 + e_1 e_2 + i_1 i_2 + e_1^2 e_2^2 + i_1^2 i_2^2 + e_1 i_1 e_2 i_2 \tag{6e}$$

where e_1 and e_2 are surface potentials, and i_1 and i_2 are the current values of input the vectors x_1 and x_2, respectively.

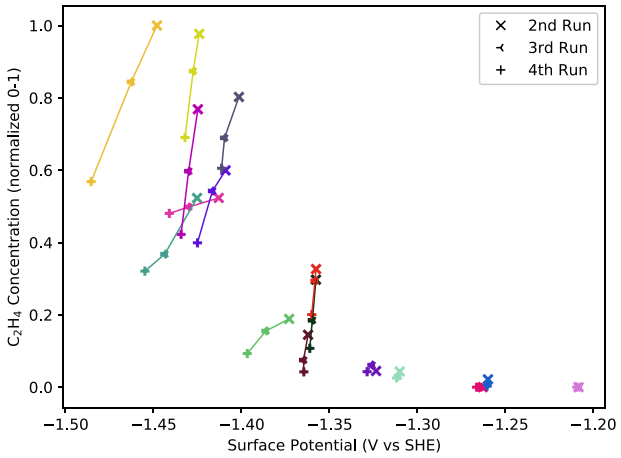
The objective function would then take the following form:

$$f(x) = \sum_{j=1}^{N} (\alpha_j^* - \alpha_j) k(x_j, x) + b \tag{7}$$

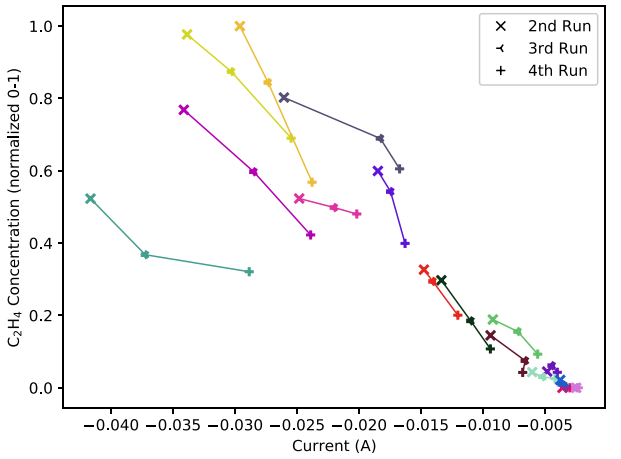
where α_j^* and α_j are the Lagrange multipliers of the optimization problem. The convex optimization problem is solved to find the coefficients of the polynomial kernel (Basak et al., 2007).

3.4. *Model training and feature engineering*

The Python package Scikit-Learn is used to fit the data to a support vector regression model. Grid search is implemented to find the best hyperparameters, by training the proposed model with all combinations of predefined sets of hyperparameters. The hyperparameters yielding the highest cross validation performance are selected to fit the final model. After grid search, the data are fit to a polynomial kernel of 5^{th} order consisting of the surface potential, rotation speed and current as input parameters and ethylene concentration as an output. Since this model has a relatively small training set, feature engineering is applied to improve the model performance. Fig. 4 is useful for interpreting general trends in the data. Throughout the open-loop experiments, although the manipulated input variables (e.g., applied potential and rotation speed) are fixed, the ethylene concentration in the product is continuously decreasing. As shown in Fig. 4, the current density generally decreases over the duration of the experiments, and therefore, it increases the surface potential according to the relationship described by Eq. (1) (since the sign in current and potential signifies the direction, increasing or decreasing remarks are made with respect to absolute values). The general decreasing trend of current is due to catalyst deactivation which will be represented better in the model by the integral of the current as explained in the following paragraphs. The model is first fit to the data set and some preliminary fitting results are presented in Fig. 5.



(a) C_2H_4 trend against current at 100 RPM in 0.2 M $KHCO_3$.



(b) C_2H_4 trend against current at 100 RPM in 0.2 M $KHCO_3$.

Fig. 4 – The trends in ethylene concentrations with increasing current and surface potential illustrated with results from various experiments. Same colors show 3 GC results from the same experiment. Ethylene concentrations are normalized between 0 and 1.

Despite having a satisfactory approximation to the GC results and rational predictions in between two consecutive GC points, the model fails to catch some behaviors in the current and surface potential. In the experimental data, there exists a natural deviation in ethylene production rates measured under the same operating conditions. Also, the SVR model predictions have a strong correlation with the current input. However, mimicking the trend in the current could result in missing some inherent behaviour in ethylene concentration, particularly when the experiment is run for longer than the 80 min used for the collection of the open-loop data used in the initial SVR model training. This motivates the pursuit of an additional input that can represent the catalyst degradation and account for the historical effect of the inputs throughout the experiment. Since the catalyst deactivates continuously, the ethylene concentration is expected to decrease monotonically over time under a constant applied potential and rotation as shown in Fig. 5. Thus, the new input must be increasing or decreasing monotonically and must be derived from the available system parameters.

A parameter that is compatible with these prerequisites is the cumulative integral of the current. The current is always negative, as defined by the convention, since a cathodic potential is applied. Therefore, the cumulative area under the current is increasing at each time step even if the current is decreasing in magnitude. Thus, a very high integrated value of the current hints that the degradation has reached a very high level and the reaction is slowing down accordingly. This is consistent with the experimental trend that experiments at higher potentials show a more significant overall catalyst deactivation, because a higher applied potential leads to a larger flow of current. Additionally, the historical effect of the current in the experiment is taken into account through integration. The integral of the current can be defined as follows:

$$P(t) = \int_0^t i_{av}(\zeta)\, d\zeta \tag{8}$$

where $P(t)$ represents the integral of the current term, t is the time, and $i_{av}(t)$ is the averaged current values that the GC measurements are based on. The integral is calculated numerically with the trapezoidal rule based on per-second current data. The impact that the integral of the current parameter makes is shown on Fig. 6. In the concentration plot of Fig. 6, the large fluctuation in SVR predictions between the second and third GC points is due to fluctuations in the surface potential, but this is reduced to a more rational trend and the third GC point is captured after the integral of the current is introduced.

Since the data set is relatively small for dynamic modelling, auxiliary performance boosting methods are explored to increase the model accuracy. One of the beneficial methods is feature engineering (Heaton, 2016). The input parameters can be augmented by applying some mathematical transformations to the existing inputs. Some of the common feature engineering methods are polynomial, logarithmic and reciprocal transformations (Heaton, 2016). Since first-principles models are in the development phase, the mathematical forms of all parameters are not exactly known. For example, in the case of a polynomial transformation, $x_1^2 x_2$ might be a more efficient parameter than just x_2, where x_i is an arbitrary input. Similarly, feeding the transformation $log(x_1)$ might be a better input than x_1 if there is a possible logarithmic correlation in the nature of the phenomenon. Feature engineering can increase the number of inputs in the model, and this causes a trade-off between computation time and accuracy. The performance of the feature engineered models is tested under polynomial degrees 2, 3, and 4, logarithmic, and reciprocal transforms, in addition to regression tree and lasso transformations. The mathematical backgrounds for decision tree and lasso can be found in (Myles et al., 2004) and (Park and Casella, 2008). The results are presented in Table 2.

Feature engineering has the potency to improve the R^2 of the model by 8% and reduce the absolute error in the test set. The best-performing feature transformation is polynomial, and the R^2 metric continues to improve as its degree increases. The best polynomial is the fourth degree polynomial, but this transformation greatly increases the number of inputs. The third degree polynomial also appears to have a good performance, and it achieves this performance with fewer inputs. This is expected as the third degree polynomial captures both the exponential increase in partial current density as a function of the applied potential and the fact that ethylene production becomes limited by mass transport at the highest overpotentials. Thus, a third degree

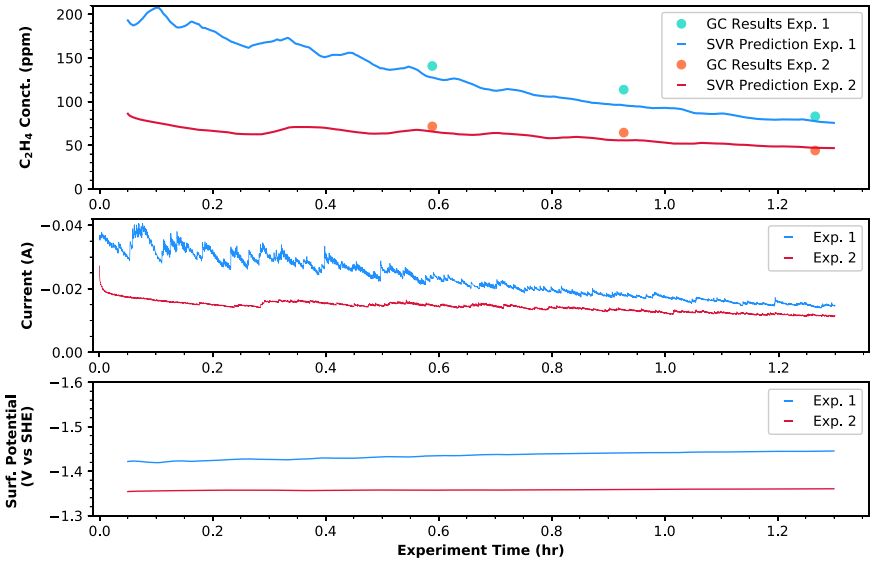


Fig. 5 – Performance of the SVR model in open-loop experiments. The concentration and surface potential plots do not start from time zero because those values are averaged over a 3 min time window in open-loop experiments.

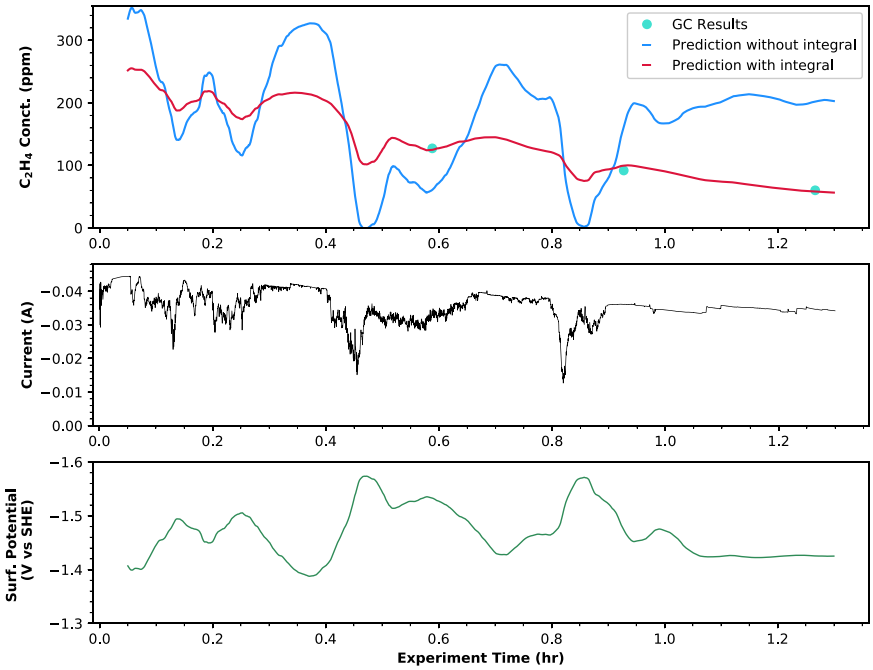


Fig. 6 – Improvement of the predictions after adding the integral of the current parameter as input. The concentration and surface potential plots do not start from time zero because those values are averaged over a 3 min time window in open-loop experiments.

polynomial feature transformation was selected to retrain the model.

Training and testing performance of the improved SVR model, after including the integral of the current as a new parameter and implementing the feature engineering, is shown in Fig. 7. The mean absolute error is 11.3 ppm and the R^2 score is 0.92.

4. Real-time optimization

This section demonstrates a real-time optimization strategy for the electrochemical reactor. Specifically, we first approximate the cost and revenue of the reactor at various set-points (i.e., ethylene concentration). Subsequently, the optimization process is developed based on the economic evaluation result of the reactor. Lastly, open source software is introduced to perform the real-time optimization for this work.

4.1. *Set-point optimization*

Set-point optimization is critical for process operation; however, the location of the optimum varies with different practical considerations. In this work, we develop a framework to decide the optimum operating conditions for the electrochemical reactor automatically, by considering the result from the economic evaluation based on the operational neural network model from Luo et al. (2022). Specifically, we approximate the energy consumption to maintain

Table 2 – Performance of the model using various feature engineering methods.

Feature Engineering Method	R^2	Mean Abs. Error
No Feature Engineering	0.842	13.24
Polynomial degree 2	0.894	10.39
Polynomial degree 3	0.922	11.31
Polynomial degree 4	0.924	10.31
Logarithmic + Reciprocal	0.853	14.59
Regression Tree	0.847	11.70
Lasso	0.841	13.24

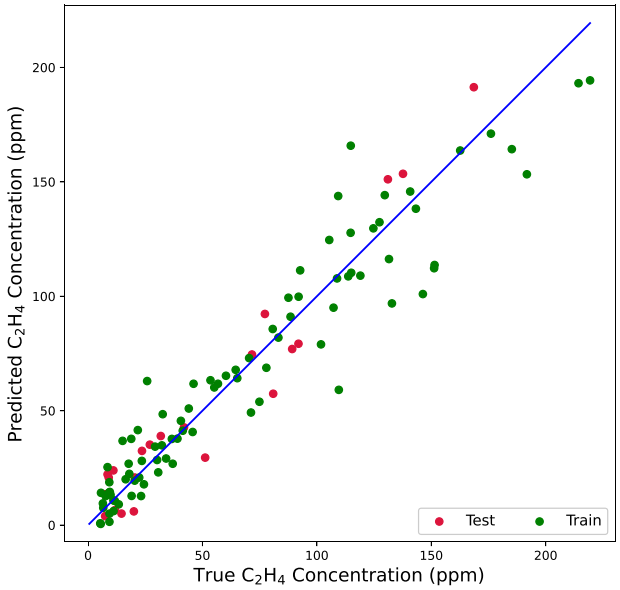


Fig. 7 – Testing and training performance of improved SVR model.

each ethylene concentration set-point at steady-state by using the following equations:

$$I = \sum_{i=1}^{m} P_i e_i A F \tag{9a}$$

$$M_i = \frac{P_i A V_g}{F_0} \tag{9b}$$

$$E = IVt \tag{9c}$$

where Eq. (9a) approximates the current density of the reactive surface by calculating the amount of transferring electrons, which is proportional to the overall production rate of the reactor. A is the area of the reactive surface, F is the Faraday constant, V_g is the standard molar volume of gases, F_0 is the feed flow rate of the CO_2 gas and e_i is the number of electrons transferred to form a molecule of the ith product. The values of these parameters are listed in Table 3. Additionally, P_i, $i = 1, \ldots, m$ is the molar production rate of the ith product, which is predicted by a statistical feed-forward neural network (FNN) model described in Luo et al. (2022). The FNN model is developed to take the surface potential (V) and rotation speed (r) of the electrode, which was fixed at 100 rpm for the control experiments in this study, as inputs to predict the production rates for all products. Fig. 8 shows the comprehensive profile of the ethylene production rate predicted by the FNN model, which implies a proportional relationship between the surface potential and the ethylene production rate, and an inversely proportional relationship with respect to the proportional speed. The reduction of ethylene production at higher rotations speeds was determined to be caused by the decrease in the residence time of carbon monoxide at the electrode/electrolyte interface by Jang et al. (2022). Furthermore, according to the balance-based equation for this reactor (Eq. (9b)), a unique concentration (M_i) for each product can be found from the corresponding production rate P_i. Therefore, the energy consumption (E) to operate the reactor at a specific set-point of ethylene concentration is calculated by adopting the equation of electric energy (Eq. (9c)).

Table 3 – Parameters of economic evaluation.

Notations	Value	Unit
A	3	cm^2
F	96,485.3	$C\,mol^{-1}$
V_g	22.4	$L\,mol^{-1}$
F_0	0.02	$L\,min^{-1}$
V	Variable	Voltage
I	Variable	Amp
E	Variable	Watt
M	Variable	Ppm

We assume the electric energy consumption is the only type of cost to operate this reactor, and we approximate the revenue of the reactor on the basis of the sale price and production rate of each product. Therefore, the optimum ethylene concentration set-point to operate the electrochemical reactor can be determined by solving the following optimization problem:

$$J = \underset{\hat{x} \in D}{\operatorname{argmax}} \left(R(\hat{x}, V) - C(I, V)\right) \tag{10a}$$

$$\text{s. t.} \qquad F_{nn}(V, r) \approx x \tag{10b}$$

$$C(V, I) = c_e \times E(V, I) \tag{10c}$$

$$R(\hat{x}, V) = \sum_{i=1}^{m} c_i \times \hat{x}_i \tag{10d}$$

$$r = 100 \tag{10e}$$

$$-1.5 \le V \le -1.27 \tag{10f}$$

where $\mathbf{D} \in \mathbb{R}^m$ is the bounded state space of the production rates, and the boundary of the space $\mathbf{D}$ is determined from the training data set used to develop the FNN model (Luo et al., 2022). The vector $x = [P_1, P_2, \ldots, P_m]$ contains the actual steady-state production rate of each product during the real-time reaction, and $\hat{x}$ denotes the predicted production rates given by the FNN model. c_e and c_i, $(i = 1, \ldots, m)$ are the price of electricity and the sale price for the i^{th} product listed in the Table. 4. In this study, the rotation speed of the working electrode is set to be constant at 100 revolutions per minute (rpm), and the surface potential is bounded from -1.5V to -1.27V *vs SHE* shown in Eq. (10e) and Eq. (10f). The FNN prediction is used in Eq. (10c) and (10d) to approximate the revenue and cost of operating this reactor and the constraint shown in Eq. (10b) ensures the FNN prediction is accurate and reliable. It is noted that the bench-scale reactor used here is 8–10 orders of magnitude smaller in ethylene production rates compared to existing commercial ethylene plants. Thus the set-point optimization is an interesting conceptual experiment as it includes operating constraints

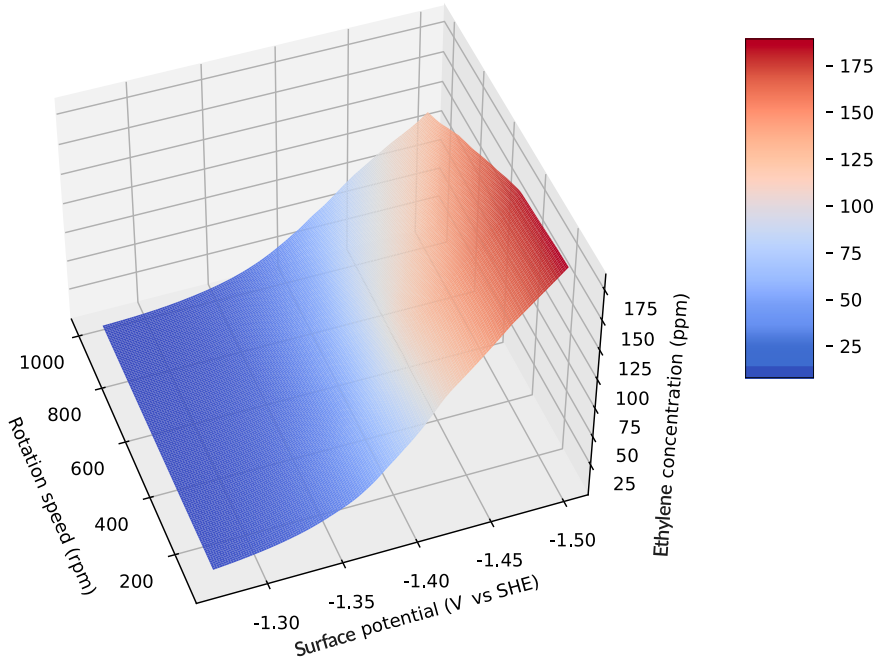


Fig. 8 – The map of FNN prediction for C_2H_4 production rates.

Table 4 – Chemical information of products.

Index	Products	Number of Transferred Electrons	**Chemical Formula**
1	methane	8	CH_4
2	ethylene	12	C_2H_4
3	methanol	6	CH_3OH
4	ethanol	12	C_2H_5OH
5	acetate	8	CH_3COO^-
6	ethylene glycol	10	$(CH_2OH)_2$
7	glycolaldehyde	8	$HOCH_2CHO$
8	acetaldehyde	10	CH_3CHO
9	n-propanol	18	C_3H_7OH
10	allyl alcohol	16	C_3H_5OH
11	acetone	16	CH_3COCH_3
12	propionaldehyde	16	C_2H_5CHO
13	carbon monoxide	2	CO

but it does not capture the complexity of future electrified large scale production systems. This approach remains valuable as it is based on experimental data and should be translatable to increasingly larger systems.

Open-source software for large-scale optimization problems, Ipopt, is utilized to address the set-point optimization problem. In this work, we use the forward finite difference method by adding small steps Δu on the optimized variables (i.e., potential and rotation speed), to approximate the first-order derivatives of the optimization problems. Additionally, second-order derivatives are approximated with the Quasi-Newton method to provide information for the calculation of search directions (Wächter and Biegler, 2006). The derivatives and constant parameters (e.g., products and electricity prices) are provided to Ipopt to optimize the operating conditions in terms of surface potential and rotation speed. Finally, the results are converted to the corresponding ethylene concentration set-point (**c**) by using Eqs. (10b) and (11), where F_o and a are the gas inlet flow rate (0.2 $L\,min^{-1}$) and a constant (1,000,000) for the unit conversion, as follows:

$$\mathbf{c} = \frac{aP_iAV_g}{F_o} \tag{11}$$

The optimization is performed for various electricity costs ranging from 0.023 to 0.03 dollars per kWh to provide optimized results over this range (Guerra et al., 2019). As shown in Fig. 9, subfigures a and b demonstrate the approximate daily profit to operate the electrochemical reactor and the ethylene concentration profile at a rotation speed of 100 rpm. In the plot, the optimal set-point shifts to a lower concentration of ethylene with increasing electricity price, implying that the optimizer is capable of making intelligent decisions to reduce the production rate as operating costs increase. As a result, the optimum operating conditions for electricity price at 0.023 and 0.03 dollars per kWh are utilized for dynamic control experiments to demonstrate the controller performance in a cost changing scenario, such as the electricity Time-of-Use (TOU) rate plan in actual operation.

Remark 6. The model is based on FNN calculations built using a statistical ML model, which generates averaged steady-state calculations taking catalyst deactivation into account. Since the catalyst deactivation cannot be defined with first-principle models in this study, the representation of deactivation is embedded into the FNN model with the statistical ML method. The FNN model is trained based on open-loop experiments and 3 data points were taken with equal intervals of 20 min, which shows the concentration decrease due to the catalyst decay. These results are averaged to give the pseudo steady-state concentration under a fixed applied potential and rotation speed. This model is optimized in our study to find economically optimal set-points while accounting for the catalyst deactivation (Luo et al., 2022).

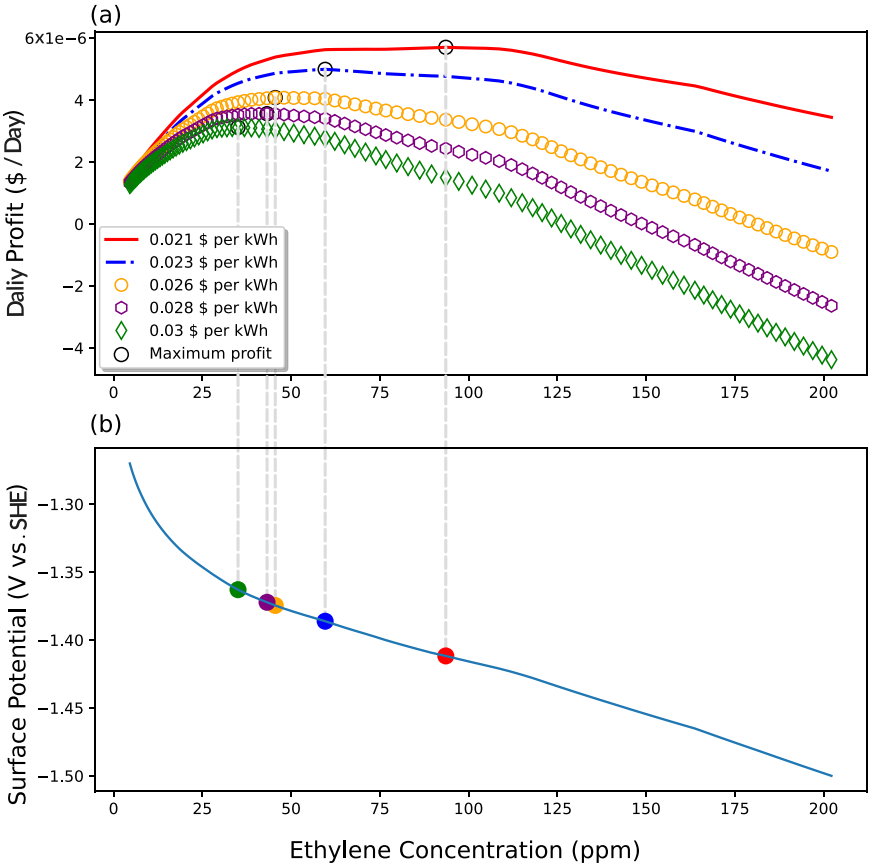


Fig. 9 – Approximated daily profit to operate the electrochemical reactor under various ethylene set-points with changing electricity price. (a) The approximated daily profit profile to operate the reactor under different electricity cost. The open black points are the maximum profits that can be obtained by operating the reactor under the respective electricity prices. (b) The ethylene concentration profile under various surface potential conditions. The solid colored points emphasize the optimum ethylene set-points that give the maximum profit.

Remark 7. The catalyst needs to be regenerated once the selectivity shift happens. Catalyst regeneration is the only way to reverse the selectivity shift. However, this study does not explore catalyst regeneration. A future process operation and control study for this reactor will be based on a more stable catalyst, which does not exhibit severe deactivation.

5. Feedback control

This section demonstrates the application of the experimental control schemes to the electrochemical reactor. Driven by the motivation mentioned in Section 2.2, the control objective is to regulate the ethylene production to energy-optimal set-point values by manipulating the input variable. Therefore, a proportional-integral (PI) control algorithm is adopted to calculate the control actions for this study, where the manipulated input is the potential applied to the working electrode and the controlled output is the ethylene concentration measured from the reactor headspace gas phase product. The PI control algorithm in discrete-time form is expressed by the following formulas (Mhaskar et al., 2004):

$$e(t_k) = y_{sp} - y \tag{12a}$$

$$u_p(t_k) = K_C e(t_k) \tag{12b}$$

$$u_I(t_k) = u_I(t_k - 1) + \frac{K_C}{\tau_i}\left(\frac{e(t_k) + e(t_k - 1)}{2}\right)\Delta t \tag{12c}$$

$$u(t_k) = u_p(t_k) + u_I(t_k) \tag{12d}$$

where y_{sp} is the desired set-point and y is the estimate of the controlled variable at time t_k. The controller error $e(t_k)$ is defined as Eq. (12a). Eq. (12b) and (12c) demonstrated the controller outputs from the proportional ($u_p(t_k)$) and integral ($u_I(t_k)$) terms, respectively, where K_C denotes the controller gain, τ_I represents the integral time constant, and Δt is the controller sampling time. The final control action ($u(t_k)$) calculated by the PI controller is shown as Eq. (12d), which is the potential on the surface of the catalyst.

Specifically, the surface potential ($E_{surface}$) required to reach or maintain the desired set-point is computed by the PI controller, and then the corresponding applied potential ($E_{applied}$) is back-calculated by the actuator based on Eq. (1), since the $E_{surface}$ compensates for the Ohmic drop from the electrolyte solution resistance. The surface potential is more relevant than the applied potential ($E_{applied}$) in describing the Tafel kinetics of electrochemical reactions involving electron transfer. It is noted that the primary and secondary current distributions in the RCE geometry are uniform, which is not the case for most existing CO_2 electrolyzers using H-type or compression cells. To further explain the control of the production rate, it is important to demonstrate how a change in $E_{applied}$ affects the current and $E_{surface}$. Intuitively, increasing the potential applied to the electrode can increase

the potential on the reacting surface. However, the current density also has a positive correlation with the applied potential, and it is competing with the potential distributed to the reacting surface. Thus, the actuator needs to adjust the applied potential with respect to the output of the PI controller and the real-time measurement of the current density.

Remark 8. The reactor may be controlled in real-time by manipulating two input variables: applied potential and rotation speed. The present study examines the control of a single gas phase product concentration (ethylene) by using a single-loop feedback controller that manipulates the applied potential. In addition to the applied potential, the rotation speed can be effectively used to control the outlet concentration of another gas product. The CO concentration is a good candidate as a second output given its strong dependence on rotation speed. Thus, multivariable control of the electrochemical reactor is a feasible task and will be discussed in a future work.

On the other hand, the controller output should be bounded by the highest and lowest surface potential values in the training set of the SVR model. The applied potential does not have a strict upper bound due to the variable electrolyte resistance which ranges between 6.8 and 7.2 ohms, but the surface potential is strictly bounded by the $E_{surface}$ range in the SVR training data set. In order to stay within the model confidence, an instant conversion from surface potential to applied potential is made during the control.

Due to the presence of a mixing volume in the gas phase of the reactor, it was previously determined that the GC concentrations during open-loop experiments were actually representative of the conditions in the reactor 5–8 min before the GC injection was taken, as discussed in Section 3.1. Thus, current and potential values were averaged over this 3-minute window for open-loop experiments used to train the SVR model. Because of the dynamics of the reactor, this method provided the best means of accurately correlating the production rate of ethylene with the conditions in the reactor. However, in the case of closed-loop experiments, the SVR model is used to predict the instantaneous production rate in real time, and the gas phase ethylene concentration model discussed in Section 5.1 accounts for the dynamic time delay in the response of the GC to the instantaneous change on the catalyst surface under a fixed rotation speed at 100 RPM. Thus, a sampling time of two seconds was used for averaging by the SVR during closed-loop control to provide real-time estimates. The real-time production rates predicted by the SVR model were then used as inputs for the gas-phase ethylene concentration model to control the reactor with a PI controller.

5.1. *Ethylene concentration estimator*

As mentioned in Remark 5, the SVR model is constructed to represent the product concentration on the catalyst surface. However, this does not represent what is measured by the GC. To have efficient control, it is necessary to estimate the gas-phase ethylene concentration in real time.

It is experimentally known that the GC and reactor overhead need a certain amount of time to reach equilibrium. This equilibrium volume is not exactly known and cannot be assumed to be equal to the reactor overhead volume, since gas bubbles adhere to the inner walls of the reactor and the surface of the rotating electrode depending on the rotation speed. To address this issue, the GC and reactor overhead are approximated as a continuously stirred tank reactor (CSTR), and the equilibrium volume and residence time are calculated using the first-principle CSTR equations, as shown in Eq. 14. It is mentioned in the previous sections that $r_A(t)$ represent the reaction rate and is calculated by the SVR model. The GC results are equalized to Eq. 14d, $r_A(t)$ is calculated by the SVR model and the τ and V parameters are varied to find the best fit. τ and V are correlated by the following equation:

$$V = \frac{60 \times \dot{V}_{CO_2} \times \tau}{1000} \tag{13}$$

where V is the equilibrium volume in L and $\dot{V}_{CO_2}$ is the volumetric flow of CO_2 into the reactor in $mL\,min^{-1}$. Since CO_2 flowrate is known to be 20 $mL\,min^{-1}$, the fitting is made by varying only the residence time τ for multiple reaction rate calculations and corresponding GC measurements.

This set of differential equations can be solved analytically to estimate the gas-phase ethylene concentration and can be modelled as follows:

$$\frac{dN_A}{dt} = F_{A0} - F_A + r_A(t) \times s \tag{14a}$$

$$F_{A0} = 0 \tag{14b}$$

$$\frac{dC_A}{dt} = -\frac{1}{\tau}C_A + \frac{r_A(t) \times s}{V} \tag{14c}$$

$$C_A(t) = e^{\frac{-t}{\tau}} \int_0^t e^{\frac{\zeta}{\tau}} \frac{r_A'(\zeta)}{V} d\zeta + C_0 e^{\frac{-t}{\tau}} \tag{14d}$$

$$r_A'(t) = r_A(t) \times s \left[\frac{mol}{s}\right] \tag{14e}$$

where N_A is the number of moles, F_A is the molar flowrate, τ is the residence time, $r_A(t)$ is SVR model calculation for instant reaction rate for ethylene production, s is the surface area of the catalyst, V is the volume of the proposed reactor and $C_0 = 0$ ppm. Several experimental data fittings with varying residence times and volumes were tested to fit the data with the model predictions. As a result, the suitable residence time and volume are selected to be 1800 s and 0.6 liters, respectively, as shown in Fig. 10. A CSTR model assumes perfect mixing; however, this reactor is neither a CSTR nor perfectly mixed. In addition, the initially produced gases dissolve into the electrolyte, so the liquid must be saturated before the gases diffuse into the vapor phase at a constant rate. The reactor headspace concentration model cannot fully capture those details, and thus it fails to capture the first GC point. However, it is seen experimentally that an ideal CSTR can satisfactorily represent the reactor overhead for the remaining GC injections.

Remark 9. A higher order model, such as multiple CSTRs in series, can be used to model the imperfect mixing better than a single CSTR. However, the development of a residence-time distribution model for this reactor is beyond the scope of this first work. Furthermore, this makes it mathematically more complex to introduce the sensor feedback data discussed in Section 5.3.

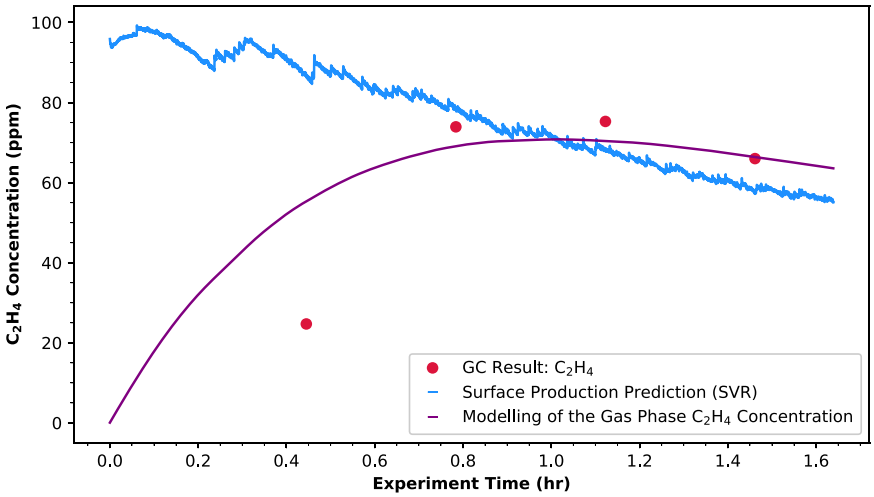


Fig. 10 – Data fitting of gas phase ethylene concentration on experimental data.

5.2. *System constraints*

Experimental systems are inherently vulnerable to varying conditions. In this electrochemical setup, the initial catalyst activity can vary each trial within some error range. That is, with the same initial fixed potential, two experiments might give different product concentrations as illustrated by the average rates and standard deviations summarized in Table 1. Although the catalysts were cleaned with the same procedure each time, the exact catalyst activity was not known prior to the experiment.

Hori proposed that methane and ethylene may share the same surface-bound CH_2 intermediate (Hori et al., 1997) with the dimerization rates of these fragments determining selectivity. In our extended feedback control experiments, after a certain surface potential threshold, it is seen that no matter how much potential is applied, there is no increase in the ethylene concentration, whereas the methane concentration keeps increasing. To prevent a selectivity shift to methane, it is crucial that the controller does not increase the surface potential to very high values. This non-linear feature of the electrochemical reactor makes this control problem different from traditional reactors. A rule of thumb would suggest increasing the manipulated variable to the limit to reach the set-point faster. However, in this electrochemical reaction case, this action would only deactivate the catalyst at such a rate that the output would never reach close to the set-point. Thus, the controller parameters must be selected attentively, so that the set-point is reached slowly without shifting the reaction selectivity away from ethylene and towards methane. The anticipated selectivity shift is seen to start from the −1.45 V surface potential. The experimental outcomes are illustrated and discussed in Section 5.5.

The total concentration of carbon atoms going into methane or ethylene is seen to be constant when multiple dynamic experiments are compared in terms of carbon conversion to hydrocarbon gas products. The formula for this comparison is given in Eq. (15), and the experimental justification is shown in Fig. 11:

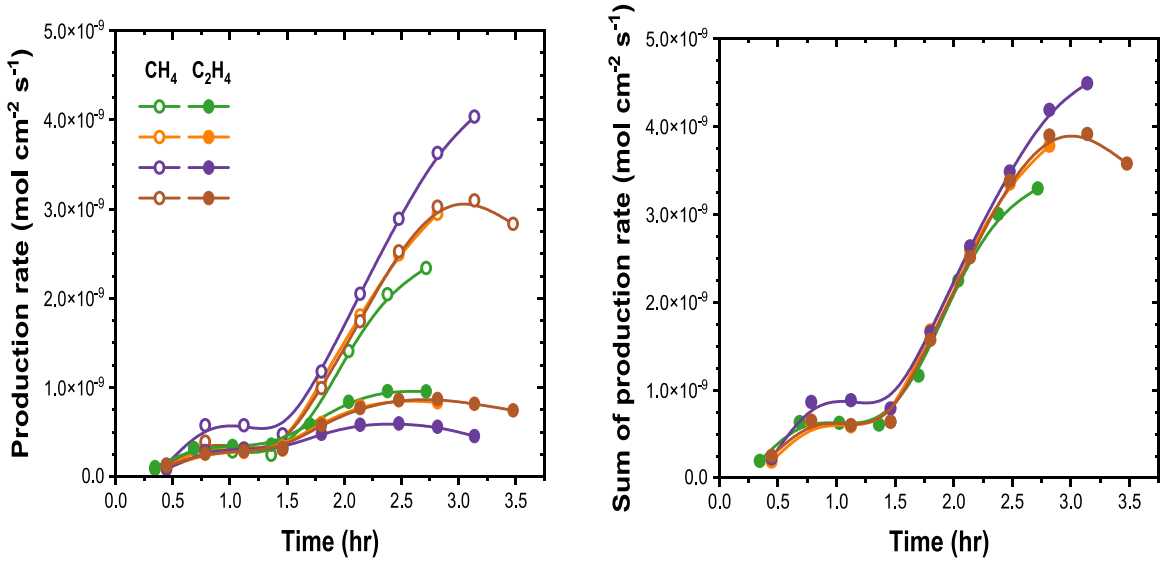
$$C_C = C_{CH_4} + 2 \times C_{C_2H_4} \tag{15}$$

where C_x is the concentration of the relevant carbonaceous gas product. The ethylene concentration is multiplied by 2 because one molecule of ethylene contains 2 carbon atoms. The repetition of the experiments demonstrated a constant concentration of carbon atoms at each GC measurement, regardless of the catalyst activity, when the surface potentials are constant before and after the set-points. Although some experiments have the same potential, rotation speed, and current, this demonstrates that the difference in catalyst activity can dictate the ethylene concentration and selectivity and bring support to Hori's idea that methane and ethylene share a common surface-bound intermediate.

5.3. *Feedback control with delayed GC measurements*

In previous sections, the modeling challenges that arose from the nature of the experimental data were emphasized. As a result of the uncertainty in experiments, the model is expected to give results within the standard deviation of repeated open-loop experiments. Implementing the control solely based on the reactor overhead estimator is expected to give close estimations to the GC results; however, this would exclude important measurement-feedback information that could be used to improve the control. Utilizing feedback from the GC measurements is a way to account for the experimental uncertainty, and this gives an opportunity for designing a complete closed-loop control, wherein the real-time sensor measurements directly impact the calculations of the manipulated inputs. Thus, the GC results obtained every 20 min are also used to improve the estimator. The control diagrams of the open-loop configuration and model-based, GC-incorporated, closed-loop configuration are shown in Fig. 12.

The purpose of the components of the control system is as follows: The PI controller node in LabVIEW calculates the control signal based on the PI control algorithm. The input to the control algorithm is the headspace ethylene concentration estimation (in ppm) and its output is the calculated surface potential. The potentiostat is both a sensor and an actuator. It senses the real-time current and can manipulate the applied potential. It changes the applied potential based on the control signal sent by the PI controller. The reactions occur in the electrochemical reactor. The electrodes and sensors are connected to the reactor. The applied potential sensed in the reactor is converted to surface potential, the current values are measured, the rotation speed and the calculated variable of cumulative integral of the current are fed to the SVR model for reaction rate estimation, which is



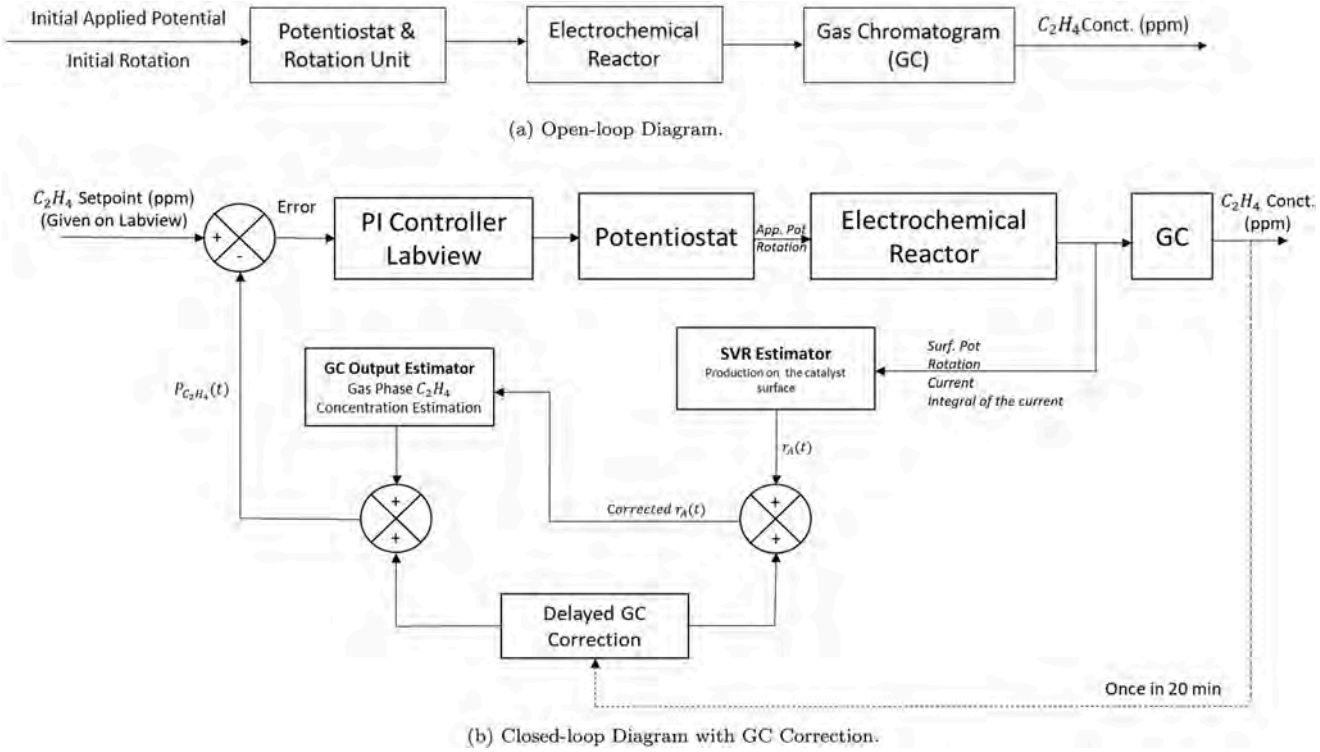
(a) Methane and ethylene production rates

(b) Sum of carbonaceous gas products

Fig. 11 – Comparison of methane and ethylene production from several experiments. Each color represents a different experiment where the set-point for ethylene was increased to 2.7-fold at the 75th minute after being kept steady for 48 min.

then used in first-principle calculations to estimate the headspace gas-phase ethylene concentration. The gases products accumulate in the reactor overhead. After the product gases equilibriate in the headspace, they travel to the GC sensor to be quantified in each 20 min. The quantification is processed automatically via an automated GC code and the concentration measurements (in ppm) are considered for feedback correction on the estimators. After the GC feedback correction, the error is calculated and sent to the PI controller for a new control signal.

In the closed-loop control, the feedback data from the sensors is introduced to the ethylene concentration estimator to realize real-time adoption of the process control curve. However, the GC provides the concentration measurement results with a 14.33 min delay from the time of the reactor overhead sampling. Thus, a correction method was developed to estimate a probable past trajectory for the process control curve. The proposed correction is activated once the GC feedback data are received. The algorithm examines the past data and calculates the cumulative correction changes in the previous 20.33 min of the predictions (the time between 2 consecutive GC injections) to better align the predicted gas concentration with the most recent GC measurement. This correction affects both the SVR model and



(a) Open-loop Diagram.

(b) Closed-loop Diagram with GC Correction.

Fig. 12 – Process control diagram for both open-loop process, and closed-loop control that incorporates the SVR model and GC feedback in the loop.

the gas-phase ethylene concentration estimation model. $r_A(t)$ term is the reaction rate with units $mol\,s^{-1}$. This correction scheme is divided into two parts. The first correction is between $t=0$ and $t=t_2$. The second correction is for after the second injection. The important abbreviations are presented below.

- n: the number of injections.
- t_i: i^{th} injection start time, $i=1\ldots n$.
- t_i^{end}: i^{th} injection end time.
- GC_i: i^{th} GC measurement result in ppm.
- $\widetilde{GC_i}$: The GC_i converted to $mol\,L^{-1}$.
- x_{i-1}: The correction applied to the SVR model (in ppm) after t_i^{end}, which should have been present during the previous injection time.
- $\tilde{x}_{i-1}$: The x_{i-1} converted to rate in $mol\,s^{-1}$ using Eq. (16).
- $P_{C_2H_4}(t_i)$: Prediction of the gas phase ethylene concentration at t_i in $mol\,L^{-1}$.
- r_A: Prediction of SVR model (in ppm) converted to $mol\,s^{-1}$ using Eq. (16).
- $r'_{A,c}$: The cumulatively corrected SVR prediction converted to $mol\,s^{-1}$ using Eq. (16).

The equation for converting the concentration to rate is as follows:

$$r_{C_2H_4} = \frac{C^{ppm}_{C_2H_4}}{10^6} \times \frac{\dot{V}_{CO2}}{60\times 10^6} \times \frac{P}{RT} \tag{16}$$

where $r_{C_2H_4}$ is the production rate in $mol\cdot s^{-1}$, $C^{ppm}_{C_2H_4}$ is the ethylene concentration in ppm, $\dot{V}_{CO_2}$ is the volumetric flow of CO_2 into the reactor in $mL\,min^{-1}$, P is the pressure in Pa, R is the universal gas constant equal to 8.314 $\frac{J}{mol\cdot K}$, and T is the ambient temperature in K.

Given in Section 5.1, Eq. (14d) shows the analytical solution for this reactor that is used to explain the mathematical background of the GC feedback correction. When the linear correction of the second GC injection $\tilde{x}_1$ is applied to the prediction of the reaction rate, the analytical solution takes the form shown in (17a) and is equal to the result of the GC measurement. Eqs. (17b), (17c), (17d) are the solution steps for (17a). The corrections between $t=0$ and t_2 and the detailed solution are calculated as follows to find the linear correction term $\tilde{x}_1$:

$$\widetilde{GC_2} = e^{\frac{-t_2}{\tau}}\int_0^{t_2} e^{\frac{\zeta}{\tau}}\left(\frac{r_A(\zeta)}{V}+\frac{\tilde{x}_1}{V}\right)d\zeta \tag{17a}$$

$$\widetilde{GC_2} = e^{\frac{-t_2}{\tau}}\int_0^{t_2} e^{\frac{\zeta}{\tau}}\frac{r_A(\zeta)}{V}d\zeta + e^{\frac{-t_2}{\tau}}\times\frac{\tilde{x}_1}{V}\times\int_0^{t_2} e^{\frac{\zeta}{\tau}}d\zeta \tag{17b}$$

$$\widetilde{GC_2} = P_{C_2H_4}(t_2) + \frac{\tilde{x}_1}{V}\times\tau\times\left(1-e^{\frac{-t_2}{\tau}}\right) \tag{17c}$$

$$\tilde{x}_1 = \frac{\widetilde{GC_2} - P_{C_2H_4}(t_2)}{\tau\times\left(1-e^{\frac{-t_2}{\tau}}\right)}\times V \tag{17d}$$

where $\tilde{x}_1$ is converted to concentration by rearranging the Eq. (16). The correction after t_2 is calculated as follows:

$$\widetilde{GC_i} = e^{\frac{-t_i}{\tau}}\int_0^{t_i} e^{\frac{\zeta}{\tau}} r_{A,c}(\zeta)\,d\zeta \tag{18a}$$

$$\widetilde{GC_i} = e^{\frac{-t_i}{\tau}}\left(\int_0^{t_{i-1}} e^{\frac{\zeta}{\tau}}\frac{r_{A,c}\hat{}'(\zeta)}{V}d\zeta + \int_{t_{i-1}}^{t_i} e^{\frac{\zeta}{\tau}}\left(\frac{r_{A,c}\hat{}'(\zeta)}{V}+\frac{\tilde{x}_{i-1}}{V}\right)d\zeta\right) \tag{18b}$$

$$\widetilde{GC_i} = \frac{e^{\frac{-t_i}{\tau}}}{e^{\frac{-t_{i-1}}{\tau}}}\widetilde{GC_{i-1}} + e^{\frac{-t_i}{\tau}}\left(\int_{t_{i-1}}^{t_i} e^{\frac{\zeta}{\tau}}\frac{r_{A,c}\hat{}'(\zeta)}{V}d\zeta + \frac{\tilde{x}_{i-1}}{V}\int_{t_{i-1}}^{t_i} e^{\frac{\zeta}{\tau}}d\zeta\right) \tag{18c}$$

$$\tilde{x}_{i-1} = \frac{\left(\widetilde{GC_i} - e^{\frac{(t_{i-1}-t_i)}{\tau}}\widetilde{GC_{i-1}} - e^{\frac{-t_i}{\tau}}T\right)\times V}{\tau\left(1-e^{\frac{(t_{i-1}-t_i)}{\tau}}\right)} \tag{18d}$$

where $T = \int_{t_{i-1}}^{t_i} e^{\frac{\zeta}{\tau}}\frac{r'_{A,c}(\zeta)}{V}d\zeta$ is calculated numerically using the trapezoidal integration rule on *LabVIEW*. We note that in Eq. (18a), the $r_{A,c}$ term accounts for all corrections. Eq. (18b) divides the calculations into two. The first integral term assumes that the previous correction was adjusted with the linear correction so that the relevant prediction overlaps with the previous GC measurement. The second integral term in Eq. (18b) is structured to find the next correction that should have been applied between $t=t_{i-1}$ and $t=t_i$.

It is important to note that this calculation assumes the equalities shown in Eq. (19) below:

$$P_{C_2H_4}(t_2) = e^{\frac{-t_2}{\tau}}\int_0^{t_2} e^{\frac{\zeta}{\tau}}\frac{r_A(\zeta)}{V}d\zeta \tag{19a}$$

$$\widetilde{GC_2} = e^{\frac{-t_2}{\tau}}\int_0^{t_2} e^{\frac{\zeta}{\tau}}\frac{r'_{A,c}(\zeta)}{V}d\zeta \tag{19b}$$

The difference between the two terms is the correction in r_A.

The objective of the calculations so far is to calculate the SVR correction term $\tilde{x}$. Following this, the correction made on the surface production rate should be cumulatively reflected in the headspace ethylene predictions. This requires a compensating factor for the ethylene production on the basis of the previously assumed surface production difference. Since GC corrections begin at t_i^{end}, the integral time interval is calculated between $t=t_{i-1}$ and $t=t_i^{end}$. This requires recalculation of $P_{C_2H_4}(t_i^{end})$ with the corresponding linear correction and would correspond to step changes in the headspace ethylene concentration prediction curve at $t=t_i^{end}$. The application of this correction to the reactor overhead ethylene prediction control curve and the relevant analytical solution are shown in Eq. (20) below:

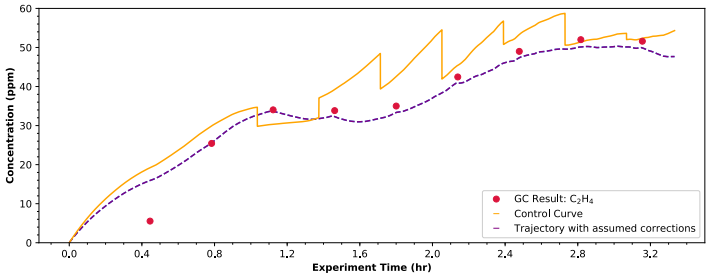
$$P_{C_2H_4}(t_i^{end}) = e^{\frac{-t_i^{end}}{\tau}}\int_0^{t_i^{end}} e^{\frac{\zeta}{\tau}}\left(\frac{r_{A,c}(\zeta)}{V}\right)d\zeta \tag{20a}$$

$$P_{C_2H_4}(t_i^{end}) = e^{\frac{-t_i^{end}}{\tau}}\left(\int_0^{t_{i-1}^{end}} e^{\frac{\zeta}{\tau}}\frac{r_{A,c}\hat{}'(\zeta)}{V}d\zeta + \int_{t_{i-1}}^{t_i^{end}} e^{\frac{\zeta}{\tau}}\frac{r_{A,c}\hat{}'(\zeta)}{V}d\zeta + \frac{\tilde{x}_{i-1}}{V}\int_{t_{i-1}}^{t_i^{end}} e^{\frac{\zeta}{\tau}}d\zeta\right) \tag{20b}$$

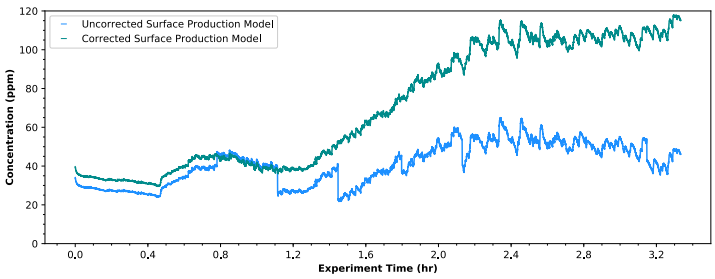
$$P_{C_2H_4}(t_i^{end}) = e^{\frac{-t_i^{end}}{\tau}}\left(\int_0^{t_i^{end}} e^{\frac{\zeta}{\tau}}\frac{r_{A,c}\hat{}'(\zeta)}{V}d\zeta + \frac{\tilde{x}_{i-1}}{V}\times\tau\times\left(e^{\frac{t_i^{end}}{\tau}} - e^{\frac{t_{i-1}}{\tau}}\right)\right) \tag{20c}$$

From Eq. (20c), it can be seen that the concentration correction that should be made to the reactor headspace concentration estimator is proportional to $\frac{\tilde{x}_{i-1}}{V}\times\tau\times\left(e^{\frac{t_i^{end}}{\tau}} - e^{\frac{t_{i-1}}{\tau}}\right)$ term. This is easy to implement when integrated in *LabVIEW* as it is added to the integration term of the analytical solution in the real-time concentration calculations as shown in Eq. (20c).

The integral of time from $t=0$ to t in the analytical solution of this headspace ethylene concentration model accounts for the past of the reaction rate. With each feedback from the GC measurement, to adjust the GC to instant



(a) The red dots represent the start of the corresponding injection t_i. The jumps in the control curve kick in at t_i^{end}. The purple dashed line is the assumed curve at the end of the 9^{th} injection as the result of the back-calculation algorithm.



(b) The surface production correction has the same timing. The cumulative correction calculated based on GC feedback is seen when the two curves are compared, especially through the end of the experiment.

Fig. 13 – The cumulative corrections based on past data and its adjustment to real-time trajectories.

production, it is assumed that the past of the experiment can be accounted for more accurately. For time interval $0 \le t < t_{i+1}^{end}$, the SVR is also adjusted cumulatively as follows:

$$r'_{A,c} = r_A(t) + \sum_{k=1}^{i} x_k \tag{21}$$

and an example of this for $0 \le t < t_5^{end}$ is like the following:

$$\begin{cases} 0 \le t \le t_2 & r'_{A,c} = r_A(t) + x_1 \\ t_2 < t \le t_3 & r'_{A,c} = r_A(t) + x_1 + x_2 \\ t_3 < t \le t_4 & r'_{A,c} = r_A(t) + x_1 + x_2 + x_3 \\ t_4 < t < t_5^{end} & r'_{A,c} = r_A(t) + x_1 + x_2 + x_3 + x_4 \end{cases}$$

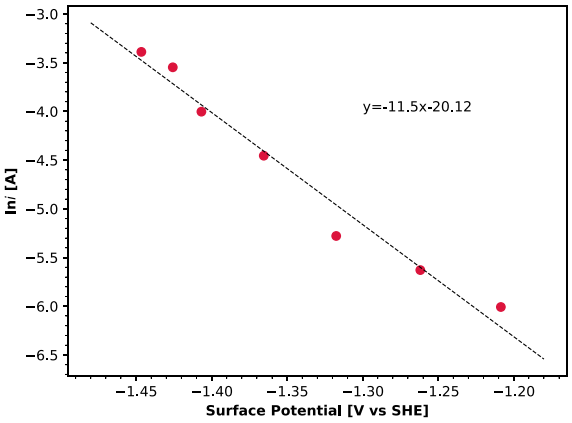
Corrections begin at t_i^{end} and assume that this correction was applied between t_{i-1} and t_i, as shown in Fig. 13. This methodology assumes a linear correction for a non-linear process. Thus, the correction is most effective when K_c is not high and when the concentration change between two consecutive GC measurements is not very large.

5.4. *Simulation and tuning*

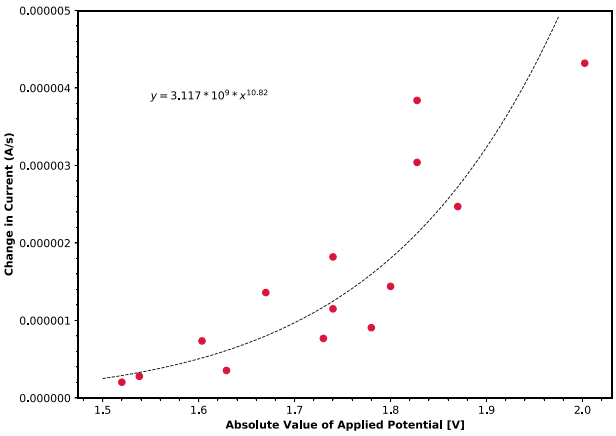
It is not practical to run experiments for several hours to determine the control parameters. Instead, a reactor simulator is built to find the controller tuning parameters. The simulator uses the data-driven SVR estimator, calculations of the first-principles modeling of the gas phase ethylene concentration, and a data-based correlation between the surface potential and current shown in Eq. (22). Using the available data, the optimal values i_0 and k are found from the data shown in Fig. 14a using the following equations:

$$i = i_0 e^{k \times (E_{Surface})} \tag{22}$$

$$i = i_0 e^{k \times (E_{Applied} - i \times R + E_{Ag/AgCl})} \tag{23}$$



(a) Current-Surface Potential Linear Correlation



(b) Current-Applied Potential Power Correlation

Fig. 14 – Regression based on averaged experimental data to find current-potential correlation parameters.

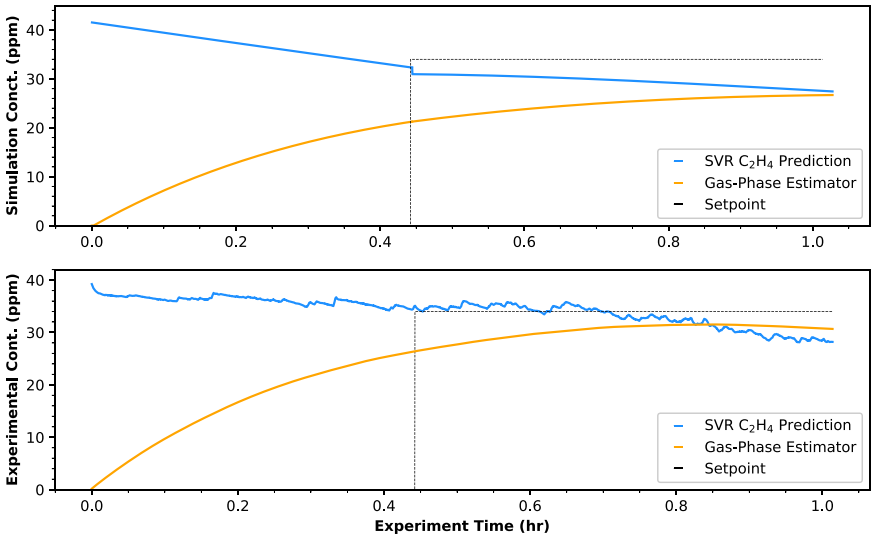


Fig. 15 – Comparison of simulated and experimental data for the set-point of 34 *ppm*.

The starting fixed potential and rotation speed are defined at – 1.65 V and 100 RPMs respectively and the corresponding current is calculated by solving Eq. (23). Then, the PI controller code calculates the necessary surface potential, and this value is fed to the SVR model. The current at this surface potential is found using Eq. (22).

In addition to the current prediction at a specific surface reaction, an approximate relation for the catalyst deactivation is extracted using the available open-loop experiments. Specifically, open-loop experiments are conducted at a fixed surface potential, and catalyst deactivation manifests itself as a reduction in current throughout the experiment. The rate of reduction in current in open-loop experiments was found to be approximately linear with time at a given potential, and a correlation between applied potential and the rate of current loss was approximated as shown in Fig. 14b by a power relation of the form:

$$\frac{di}{dt} = 3.17 \times 10^{-9} \left|E_{\text{applied}}\right|^{10.82} \tag{24}$$

It is important to note that the developed simulator does not perfectly represent the reactor. Firstly, the activity of the catalyst is variable, and the simulator uses a model derived from the open-loop experiments. Due to this, experimental proof is needed to show that the controller parameters are suitable to compensate for the inherent decrease in the current. However, the simulator is expected to produce a good estimate of the controller parameters, K_c and τ_i. Second, the SVR model does not properly represent the selectivity shift that occurs at high surface potentials because it is not trained on experiments that are long enough for the selectivity shift to occur. As a result, the controller has no knowledge as to whether the selectivity shift has occurred or not, and it assumes that increasing the potential will increase the ethylene concentration. This issue should be considered when selecting control parameters. Finally, the simulator does not use feedback data from the GC. Thus, the feedback correction mechanism for the GC discussed in Section 5.3 cannot be used during the simulation.

Before the simulator was built, an initial experiment was conducted using the control parameters, $K_c = -2.45 \times 10^{-5}$ and $\tau_i = 40$ s. However, it was seen that the catalyst deactivates faster than the controller compensates for it. After the simulator was coded, the simulation and the experimental trajectories were compared, and they are shown in Fig. 15. The experimental trajectory shown in Fig. 15 is compatible with the GC results. Both the simulation and the experiment show that this proportional controller gain is not strong enough to compensate for the catalyst decay. As a result of more simulations, after the second set-point change, a single pair of controller parameters were found not to be appropriate to control the process. This is an indication of the high non-linearity of the process and should be tackled by gain scheduling. As a result, two different gains were used for two different set-points considered in our experiments.

Multiple combinations of controller parameters were tested with the simulator. The desired parameters should not drive the process output to the controller surface potential limits to avoid an early selectivity shift to methane. Also, the controller parameters should not lead to a very sluggish response where the catalyst deactivation overpowers the system before the set-point is reached. The desired parameters in the simulation should allow some overshoot since in reality the expected concentrations are lower. Fig. 16 shows the controller simulation for two set-points, 34 and 60 *ppm*, respectively. The control parameters are $K_{c1} = -0.00011025$, $K_{c2} = -5.39 \times 10^{-5}$ and $\tau_i = 40$ s

5.5. *Closed-loop performance of set-point tracking*

In an effective control system, the controller should be able to drive the system to a new set-point in case of a set-point change. In a hypothetical scenario, the electricity price can change (due to TOU pricing, etc.), and the controller must adapt to keep the most economical or energy efficient production. In order to prove the efficacy of the estimator-based PI controller on this setup, a set-point change is introduced between two economically optimal points. The set-points are chosen to reflect a sufficiently high change in production over a reasonable period of time. Thus, the PI controller is tested for set-points of 34 and 60 *ppm*.

In an open-loop setting, while the catalyst degradation decreases the number of available catalyst sites, the activation energy reduction caused by the catalyst slowly subsides and molecules need more energy to overcome the activation energy barrier. As a result, when the applied potential is kept constant, the number of molecules that can pass the

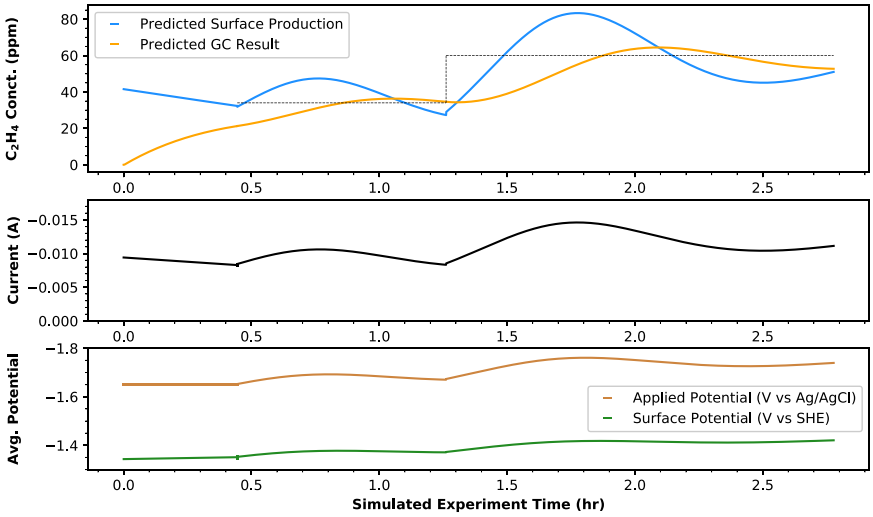


Fig. 16 – The simulation that was used to find the control parameters that were used in the closed-loop experiment.

activation energy barrier decreases as more reactive sites become unavailable. In a controlled experiment, the closed-loop potential is driven to more negative values by the PI controller to keep the energy of the molecules high enough to sustain production. The trends in current for the open-loop and closed-loop experiments are the opposite. In the open-loop experiment, the decreasing current is an indication that the ethylene concentration is decreasing as the catalyst deactivates over time. In the closed-loop experiment, the increase in the applied potential drives the current in the opposite direction with respect to the open-loop current, indicating that the manipulated input compensates for catalyst degradation.

The experiment is run for 9 injections, and it is expected that the PI controller keeps the concentration at each set-point. The controller starts at the 27^{th} minute, and the set-point change is introduced at the 76^{th} minute, after obtaining 2 GC measurements. The experimental results are presented in Fig. 17. The process succeeds in driving the gas-phase ethylene concentration to the first set-point by the third injection. Then, after the second set-point change, the controller drives the output close to 60 ppm at about 52 *ppm*. This experiment was repeated several times, with higher and lower feasible proportional controller gains, and with different catalysts. In each of the experiments, the final concentration was recorded to be between 52 and 54 *ppm*. One of the potential reasons that this system does not fully reach the higher set-point is that the dynamics of the system cause the process to approach the second set-point after 2 h. By this time, the catalyst has largely deactivated. The surface potential needed to increase the concentration of this system causes a selectivity shift. Fig. 18 shows this selectivity shift to methane more clearly. Especially, between injections 7–11, the methane concentration increases as the ethylene concentration remains constant under increasing potentials and current. This is an experimental justification that the energy given to the process was being used to produce more methane instead of producing ethylene. Thus, if the system has not reached the set-point by the time the selectivity shifts to favor methane, it is not possible to increase the concentration of ethylene to the set-point by increasing the potential. It takes around the time of 3 GC measurements to reach the

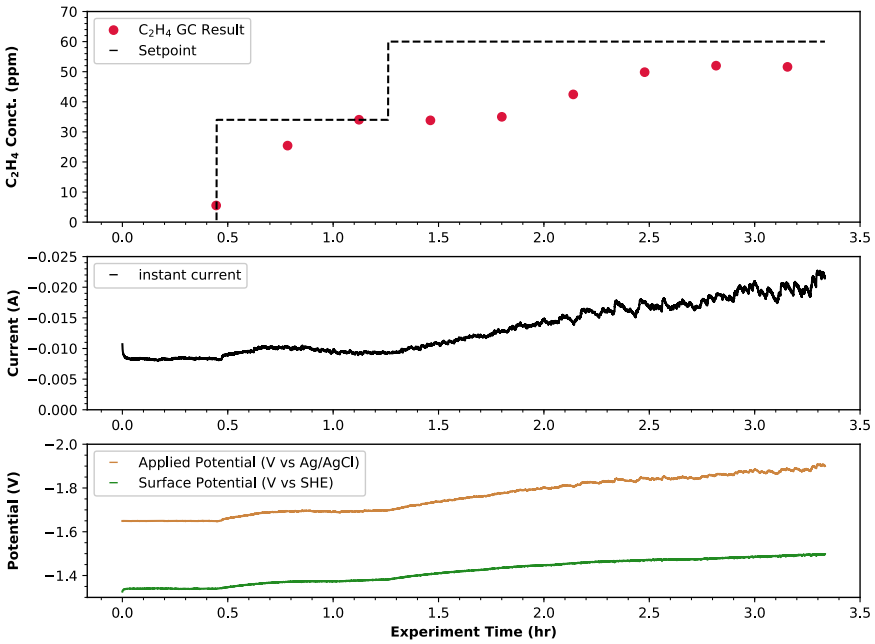


Fig. 17 – Closed-loop experiment with set-point changes.

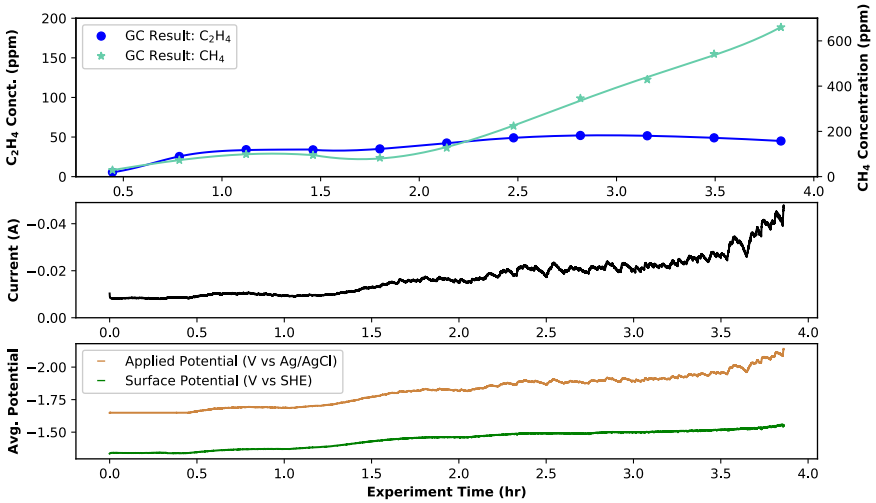


Fig. 18 – Ethylene and methane concentration comparison. The selectivity shift to methane can be seen starting from the 7th injection.The surface potential and current continue to increase, while the trend of ethylene concentration remains steady and the methane production increases sharply.

neighborhood of a new set-point. The results show that the controller can drive the system to the neighborhood of the set-points within the error margins of 10–12%.

Remark 10. The set-point change would not be possible without the integral of the current input. The current and surface potential values at 100 RPM recorded through the end of the repetition experiments were seen to yield much higher concentrations in the early GC measurements of related open-loop experiments. However, the cumulative integral of the current brings the concentration closer to the GC results by capturing appropriately the catalyst deactivation over time.

6. Conclusion and overlook

This work presented the implementation of support vector regression in an electrochemical reactor to represent process variables and concentration relationships and proposed an approach to combine the output of the SVR model-based gas-phase C_2H_4 concentration estimator with GC measurements to control the concentration of ethylene in the reactor. As the electrochemical CO_2 reduction and the reaction pathways are not fully known, current first-principal approximations are inadequate to set up an efficient control scheme. Furthermore, catalyst deactivation in the reactor was an inherent disturbance that increased process variability. To address this issue, the SVR model was built to model the experimental concentration and catalyst deactivation over an extensive window of operating conditions and was combined with first-principles modeling to predict the gas-phase ethylene concentration. The integral of the current was introduced as an input to contribute to the representation of the degradation of the catalyst, and the overall accuracy of the model was increased with feature engineering. In addition, the GC measurements were introduced to the controller to achieve a fully-closed-loop control scheme. Lastly, the economically-optimized energy-efficient set-points were calculated, and the controller was shown to be successful, as it could drive the process in the neighborhood of two energy-optimal set-points.

The approach shown in this report for the automation of the UCLA electrochemical reactor could be broadly applicable. For simpler electrochemical reactions involving a single reaction pathway, there is no challenge in selectivity and such systems require only the collection of electrical current data which is generated every second from the potentiostat. When the system produces more than one product, other analytical tools to quantify products should be implemented and automated. In this work, we have demonstrated the online data processing of GC that automatically quantifies gases, which could be readily translated to other electrochemical reactions involving gas products. This approach could be expanded to other analytical and spectroscopic tools such as high-performance liquid chromatography (HPLC), ultraviolet-visible (UV-Vis) spectroscopy, and FTIR. Although the overall implementation process would be similar, it is important to note the intrinsic differences of analytical sensors. Understanding the intrinsic nature (e.g., detection limit, sampling, and response times) of sensors as detailed in this work will be the key to successfully apply the developed automation and control approach to other electrochemical reactor systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to gratefully acknowledge financial support 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.

References

Basak, D., Pal, S., Patranabis, D., 2007. Support vector regression. Neural Inf. Process. - Lett. Rev. 11, 203–224.

Burges, C.J., 1998. A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Discov. 2, 121–167.

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

Creel, E.B., Corson, E.R., Eichhorn, J., Kostecki, R., Urban, J.J., McCloskey, B.D., 2019. Directing selectivity of electrochemical carbon dioxide reduction using plasmonics. ACS Energy Lett. 4, 1098–1105.
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.
DeWulf, D.W., Jin, T., Bard, A.J., 1989. Electrochemical and surface studies of carbon dioxide reduction to methane and ethylene at copper electrodes in aqueous solutions. J. Electrochem. Soc. 136, 1686–1691.
Engelbrecht, A., Uhlig, C., Stark, O., Hämmerle, M., Schmid, G., Magori, E., Wiesner-Fleischer, K., Fleischer, M., Moos, R., 2018. On the electrochemical CO_2 reduction at copper sheet electrodes with enhanced long-term stability by pulsed electrolysis. J. Electrochem. Soc. 165, J3059–J3068.
Guelman, L., 2012. Gradient boosting trees for auto insurance loss cost modeling and prediction. Expert Syst. Appl. 39, 3659–3667.
Guerra, O.J., Eichman, J., Kurtz, J., Hodge, B.-M., 2019. Cost competitiveness of electrolytic hydrogen. Joule 3, 2425–2443.
Heaton, J., 2016. An empirical analysis of feature engineering for predictive modeling. In: Proceedings of SoutheastCon 2016, Norfolk, VA, USA. pp. 1–6.
Hori, Y., Konishi, H., Futamura, T., Murata, A., Koga, O., Sakurai, H., Oguma, K., 2005. Deactivation of copper electrode in electrochemical reduction of CO_2. Electrochim. Acta 50, 5354–5369.
Hori, Y., Takahashi, R., Yoshinami, Y., Murata, A., 1997. Electrochemical reduction of co at a copper electrode. J. Phys. Chem. B 101, 7075–7081.
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 J. 65, e17605.
Jeng, E., Qi, Z., Kashi, A.R., Hunegnaw, S., Huo, Z., Miller, J.S., BayuAji, L.B., Ko, B.H., Shin, H., Ma, S., Kuhl, K.P., Jiao, F., Biener, J., 2022. Scalable gas diffusion electrode fabrication for electrochemical CO_2 reduction using physical vapor deposition methods. ACS Appl. Mater. Interfaces 14, 7731–7740.
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. Angew. Chem. Int. Ed. 60, 20627–20648.
Kas, R., Kortlever, R., Yílmaz, H., Koper, M.T.M., Mul, G., 2015. Manipulating the hydrocarbon selectivity of copper nanoparticles in CO_2 electroreduction by process conditions. ChemElectroChem 2, 354–358.
Ke, J., Gao, C., Folgueiras-Amador, A.A., Jolley, K.E., de Frutos, O., Mateos, C., Rincón, J.A., Brown, R.C., Poliakoff, M., George, M.W., 2022. Self-optimization of continuous flow electrochemical synthesis using fourier transform infrared spectroscopy and gas chromatography. Appl. Spectrosc. 76, 38–50.
Kim, Y., Baricuatro, J.H., Javier, A., Gregoire, J.M., Soriaga, M.P., 2014. The evolution of the polycrystalline copper surface, first to Cu (111) and then to Cu (100), at a fixed CO_2 RR potential: a study by operando EC-STM. Langmuir 30, 15053–15056.
Kim, Y., Baricuatro, J.H., Soriaga, M.P., 2018. Surface reconstruction of polycrystalline Cu electrodes in aqueous $KHCO_3$ electrolyte at potentials in the early stages of CO_2 reduction. Electrocatalysis 9, 526–530.
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. Ind. Eng. Chem. Res (in press).
Mahesh, B., 2020. Machine learning algorithms - a review. Int. J. Sci. Res. 9, 381–386.
Malek, A., Wang, Q., Baumann, S., Guillon, O., Eikerling, M., Malek, K., 2021. A data-driven framework for the accelerated discovery of CO_2 reduction electrocatalysts. Front. Energy Res. 9, 52.
Mhaskar, P., El-Farra, N. H., Christofides, P. D., 2004. A method for PID controller tuning using nonlinear control techniques. In: Proceedings of the 2004 American Control Conference, Boston, Massachusetts, pp. 2925–2930.
Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., Brown, S.D., 2004. An introduction to decision tree modeling. J. Chemom. 18, 275–285.
Nitopi, S., Bertheussen, E., Scott, S.B., Liu, X., Engstfeld, A.K., Horch, S., Seger, B., Stephens, I.E.L., Chan, K., Hahn, C., Nørskov, J.K., Jaramillo, T.F., Chorkendorff, I., 2019. Progress and perspectives of electrochemical CO_2 reduction on copper in aqueous electrolyte. Chem. Rev. 119, 7610–7672.
Niu, L., An, L., Wang, X., Sun, Z., 2021. Effect on electrochemical reduction of nitrogen to ammonia under ambient conditions: Challenges and opportunities for chemical fuels. J. Energy Chem. 61, 304–318.
Park, T., Casella, G., 2008. The bayesian lasso. J. Am. Stat. Assoc. 103, 681–686.
Pascual, X., Gu, H., Bartman, A.R., Zhu, A., Rahardianto, A., Giralt, J., Rallo, R., Christofides, P.D., Cohen, Y., 2013. Data-driven models of steady state and transient operations of spiral-wound ro plant. Desalination 316, 154–161.
Ray, S., 2019. A quick review of machine learning algorithms. In: Proceedings of 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing, Faridabad, India, pp. 35–39.
Roberts, F.S., Kuhl, K.P., Nilsson, A., 2015. High selectivity for ethylene from carbon dioxide reduction over copper nanocube electrocatalysts. Angew. Chem. 127, 5268–5271.
Shiratsuchi, R., Aikoh, Y., Nogami, G., 1993. Pulsed electroreduction of CO2 on copper electrodes. J. Electrochem. Soc. 140, 3479–3482.
Shokry, A., Audino, F., Vicente, P., Escudero, G., Moya, M.P., Graells, M., Espuña, A., 2015. Modeling and simulation of complex nonlinear dynamic processes using data based models: application to Photo-Fenton process. Comput. Aid. Chem. Eng. 37, 191–196.
Tan, P., Zhang, C., Xia, J., Fang, Q., Chen, G., 2018. NO_X emission model for coal-fired boilers using principle component analysis and support vector regression. J. Chem. Eng. Jpn. 49, 211–216.
Vapnik, V., Golowich, S.E., Smola, A., 1996. Support vector method for function approximation, regression estimation and signal processing. Adv. Neural Inf. Process. Syst. 9, 281–287.
Wächter, A., Biegler, L.T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math. Program. 106, 25–57.
Wu, D., Zhang, J., Cheng, M.-J., Lu, Q., Zhang, H., 2021. Machine learning investigation of supplementary adsorbate influence on copper for enhanced electrochemical CO_2 reduction performance. J. Phys. Chem. C 125, 15363–15372.
Wu, Z., Tran, A., Rincon, D., Christofides, P.D., 2019. Machine learning-based predictive control of nonlinear processes Part I: theory. AIChE J. 65, e16729.
Xie, J.-F., Huang, Y.-X., Li, W.-W., Song, X.-N., Xiong, L., Yu, H.-Q., 2014. Efficient electrochemical CO_2 reduction on a unique chrysanthemum-like Cu nanoflower electrode and direct observation of carbon deposite. Electrochim. Acta 139, 137–144.
Yu, Y., Jiang, L., 2016. Dynamic modeling on the MIMO system with linear programming support vector regression and its application. In: Proceedings of 2016 6th International Workshop on Computer Science and Engineering. Morgan Kaufmann Publishers, Tokyo, pp. 710–714.