

Bayesian Optimization of Electrode Conditioning of Ni(-Fe) Electrodes for the Alkaline Oxygen Evolution Reaction

J. Raphael Seidenberg, Clara Gohlke, Raphael Diebold, Vera Seidl, Anna K. Mechler, Alexander Mitsos, and Dominik Bongartz*

Efficient alkaline water electrolysis requires highly active electrodes for the sluggish oxygen evolution reaction (OER). NiFeO_xH_y materials are among the most active OER electrocatalysts, and their activity can be tailored, among other methods, by electrochemical conditioning. However, there is a lack of systematic approaches to optimize the conditioning process to achieve the best electrode activation. A promising way to develop such approaches is to use mathematical models. While mechanistic models are not readily available and hard to develop, data-driven models might offer a straightforward alternative. The use of Bayesian optimization (BO) with Gaussian processes to improve the electrode conditioning process of a Ni-Fe bulk electrode is

proposed. With this approach, an electrode conditioning process yielding a stronger activity enhancement compared to previous manual optimization is identified; at the same time, fewer experiments are also required. It is further shown that this also allows to transfer knowledge to new materials: transfer learning starting from the experimental data for the Ni-Fe electrode allows optimization of the conditioning of a Ni electrode with fewer experiments than applying BO to the Ni electrode from scratch. Overall, the potential of using data-driven numerical optimization in a hardware-in-the-loop approach for electrode conditioning is highlighted.

1. Introduction

The efficiency of alkaline water electrolysis is limited by the significant kinetic overpotentials of the oxygen evolution reaction (OER) at the anode.^[1] Therefore, highly active electrode materials that enable lower kinetic overpotentials for the OER are a key to achieving efficient water electrolysis. One promising catalyst material class is NiFeO_xH_y .^[2–6] The activity of NiFeO_xH_y electrodes was shown to be strongly improved by electrode conditioning using potential cycling.^[7] Such a conditioning could enable high electrode activity while reducing the cost and labor of electrode fabrication by replacing other synthesis steps.^[7]

Electrode conditioning using potential cycling to increase the electrode activity has been extensively investigated and applied to, e.g., steels and nickel-based materials.^[8–15] Possible explanations in

terms of underlying mechanisms often rely on the duplex layer model proposed by Burke et al.,^[11] in which both an anhydrous and a hydrous oxide layer form on the metal surface. The formation of the created layers is assumed to depend on the parameters of the potential cycling process.^[7,8,16] E.g., cathodic potentials have been associated with a reduction and rearrangement of the anhydrous layer that facilitates the growth of the hydrous layer, and anodic potentials with a thickening of the anhydrous layer. Lower scan rates have been associated with a more efficient oxide layer growth per cycle due to sufficient time for mass transport, and higher scan rates with more growth cycles in a fixed time. Despite many works aimed at understanding the underlying processes, procedures for efficiently finding the best electrode conditioning parameters for a given electrode material remain unclear. In the literature, the electrochemical conditioning has been studied

J. R. Seidenberg, A. Mitsos
Process Systems Engineering (AVT.SVT)
RWTH Aachen University
52074 Aachen, Germany

J. R. Seidenberg, D. Bongartz
Department of Chemical Engineering
KU Leuven
3001 Leuven, Belgium
E-mail: dominikbongartz@alum.mit.edu


C. Gohlke, R. Diebold, V. Seidl, A. K. Mechler
Electrochemical Reaction Engineering (AVT.ERT)
RWTH Aachen University
52074 Aachen, Germany


A. K. Mechler, A. Mitsos
JARA-ENERGY
52074 Aachen, Germany

A. K. Mechler
Electrochemical Process Engineering (IET-4)
Forschungszentrum Jülich
52425 Jülich, Germany

A. Mitsos
Energy Systems Engineering (ICE-1)
Forschungszentrum Jülich
52425 Jülich, Germany

D. Bongartz
EnergyVille
3600 Genk, Belgium

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/celc.202500284>

 © 2025 The Author(s). ChemElectroChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

by experimentally testing different combinations of parameter values or by varying one or two parameters at a time.^[8,9,13]

Similarly, in our previous work,^[7] we performed a manual optimization based on trend identification of the electrode conditioning of a Ni-Fe electrode by varying different parameters of the potential cycling such as lower and upper potential limit as well as scan rate. A 47 ± 6 mV decrease of the potential at 10 mA cm^{-2} was achieved, which could be maintained during 100h of operation at 100 mA cm^{-2} . The correlation between activation and the respective protocol parameters was investigated for different Ni-Fe alloys. Similar activation trends were found for all materials. The trends were explained with the tailored formation of the duplex layer oxide. For the trend identification, the parameters were mostly varied independently, i.e., in a one-factor-at-a-time approach, to assess the influence of each parameter. One disadvantage of such an approach is that interdependencies between parameters are neglected. As the optimal lower potential limit is known to depend on the upper potential limit,^[1,10] both were also altered simultaneously. However, other interdependencies that are not taken into account might exist. Additionally, as the method is prone to noisy experimental observations and a significant standard deviation was observed when testing for reproducibility, a threefold repetition of each experiment was performed. This led to a total of 24 experiments (eight experiments with three repetitions each) with a conditioning duration of 30 min (excl. preparation time). This presents a considerable experimental effort, especially since longer conditioning durations need to be considered in the future as they appear to lead to a higher increase in activity,^[7,8] and it might be desirable to optimize the conditioning for several materials.

Model-based approaches could potentially enable faster optimization of electrode conditioning in terms of fewer required experiments. Though, as the underlying mechanisms in electrode conditioning are complex and have not been fully understood, mechanistic models describing all the effects of the electrode conditioning process on electrode activity are not readily available. Their development is likely to require years of modeling and numerous experiments. Instead, data-driven models might be a straightforward alternative. Bayesian optimization (BO) with Gaussian processes (GPs) is a promising data-driven approach for optimization with few data points and costly experiments^[17] and has a wide range of applications including reaction and catalyst optimization^[18–20] as well as optimization based on expensive simulations, e.g., computational fluid dynamics.^[21] Being a data-driven approach, BO with GPs does not require any preexisting model or knowledge about the system to be optimized. Compared to other data-driven approaches (e.g., reinforcement

learning), BO with GPs has the advantages of quantifying uncertainty and considering both expectation and uncertainty of the model prediction for efficient exploration of the search space.^[17,22,23]

Here, we apply a BO approach to electrode conditioning. Using a GP model allows us to take into account interdependencies between parameters as expected to exist in electrode conditioning. We parameterize a cyclic voltammetry electrode conditioning process and iteratively optimize this process for a Ni-Fe electrode with 30 wt% of Fe to maximize its activity. We compare the results with the results obtained for the same $\text{Ni}_{70}\text{Fe}_{30}$ material and experimental setup in our previous work.^[7] Moreover, we investigate if the BO approach gives better activity at a similar budget of experiments and appears to be a suitable method to optimize electrode conditioning.

Another open question is how knowledge on the electrode conditioning could be transferred between different materials. Transfer learning is a promising strategy to transfer knowledge between different but related tasks.^[24,25] To investigate this possibility, we apply a rather simple transfer learning BO approach to optimize the electrode conditioning of $\text{Ni}_{99.99}$ by making use of data gathered for $\text{Ni}_{70}\text{Fe}_{30}$ and compare it to a plain BO approach.

2. Experimental Section

We employed BO using a hardware-in-the-loop approach (Figure 1). As the goal is to compare not only the achieved activation but also the number of required experiments to our previous work,^[7] we did not use any of the previously collected data herein. Instead, we started with a set of initial experiments determined by an adapted Latin hypercube sampling. Based on these initial experiments, we trained the GP model. As output of the GP model, we used the achieved difference in potential (ΔE_{10}) at 10 mA cm^{-2} . It was calculated as

$$\Delta E_{10} = \bar{E}_{10,\text{after}} - \bar{E}_{10,\text{before}} \quad (1)$$

where $\bar{E}_{10,\text{before}}$ and $\bar{E}_{10,\text{after}}$ are the average potentials of a 10 min chronopotentiometry experiment at 10 mA cm^{-2} before and after the conditioning, respectively. This relatively low current density of 10 mA cm^{-2} was chosen due to the experimental setup's limited capability to reliably measure larger current densities ($>100 \text{ mA cm}^{-2}$). The chosen value also allows a straightforward comparison with the results of our previous work.^[7] To obtain an indication whether the observed trends also hold at larger current densities, which are closer to industrial conditions, we also have included the achieved potential difference at 100 mA cm^{-2}

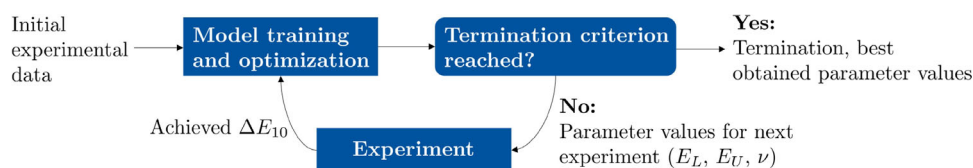


Figure 1. Overview of iterative hardware-in-the-loop approach.

(ΔE_{100}) obtained from a 60s chronopotentiometry measurement in Appendix C. As input of the GP model and parameter values to be determined by optimization, we used three parameters describing the cyclic voltammetry conditioning profile. Using the GP model, the parameter values of the next experiment were determined by optimizing the expected improvement in terms of minimizing ΔE_{10} . Based on this, the next conditioning experiment was conducted in the lab, always using a new electrode for each iteration with the determined parameters, and the result was used to train an updated GP model. Model training, optimization, and experiment were carried out in a loop until the termination criterion was fulfilled.

In the following, we will elucidate the different parts of the iterative optimization procedure in more detail and explain how the approach was adapted to a transfer learning approach.

2.1. Experimental Setup and Technique

The electrochemical experiments were conducted according to the procedure described in Gohlke et al.^[7], where the utilized electrochemical flow cell can also be seen. It features 1 cm x 1 cm parallel electrodes with an electrode spacing of 3 mm. A Hg/HgO reference electrode (ALS), a glassy carbon counter electrode (SIGRADUR G, Glassy Carbon, HTW), and a Ni or Ni-Fe plate (Ni_{99.99}, Ni₇₀Fe₃₀, HMW Hauner) working electrode was used. The composition of the working electrode material is given in wt%, as indicated in the subscript. For electrode preparation, the electrode was ground (7000 grit size), polished with an alumina-water-slurry (MicroPolish Suspensions, Buehler; polishing machine LaboPol-20, Struers) with decreasing grain size (1 μm , 0.3 μm , 0.05 μm), ultrasonicated in ultrapure water for 5 min, and dried. Experiments were conducted at room temperature (RT), ambient pressure, and in 1 M KOH, prepared from ultrapure water and KOH pellets (min. 85.0% KOH, CHEMSOLUTE). The Fe concentration was adjusted to 110 ± 10 ppb by adding Fe in diluted HNO₃ (ICP-026 calibration standard, Agilent Technologies). For the experiment, the flow cell, sealed with a torque of 0.8 Nm (MicroClick MC2, PROXXON), was aligned vertically for efficient gas removal from the electrodes. From the PTFE reservoir, the electrolyte ran through the flow cell in a single-pass with a volume flow of 3 mL/min (Ismatec Reglo Digital Pump, 4 channels à 12 rolls).

The electrochemical protocol consisted of an activity measurement before and after the conditioning. The activity was determined by first performing electrochemical impedance spectroscopy at open circuit potential (100kHz–10 Hz, 10 mV rms), followed by cyclic voltammetry (1.0–1.6V, 3 cycles) at 10 and 100 mV s⁻¹ and a chronopotentiometry at 10 mA cm⁻² for 10 min and at 100 mA cm⁻² for 60s. The conditioning was performed for 30 min (see next subsection for details on the applied conditioning). A schematic overview of the electrochemical protocol can be seen in the Supporting Information of the previous work.^[7] The potentials are reported vs. RHE (see Supporting Information of the previous work^[7] for conversion) and are

100% iR-corrected with the resistance from the respective EIS measurement. The raw data can be found on Zenodo.^[26]

When conducting the experiments, each parameter combination is measured once by conditioning a new electrode and then returned to the model to keep the experimental effort low. This is favored by the GP model considering the data not as ground truth but as potentially noisy. The noise level of the training data is learned by the GP during the training process.

2.2. Parametrization of Cyclic Voltammetry Profile for Conditioning

The cyclic voltammetry profile for conditioning was parameterized using three parameters (see Figure 2): the lower potential limit E_L , the upper potential limit E_U , and the scan rate ν . Values for both E_L and E_U are given versus RHE. The duration of the conditioning Δt was fixed at 30 min.

For the optimization, bounds on the parameters had to be defined (see Table 1). The bounds for E_L and E_U were chosen to reduce intense bubble formation caused by hydrogen and oxygen evolution, which would be difficult for the experimental setup to handle and would be potentially difficult to translate to an industrial scale. We chose a maximum scan rate of 500 mV s⁻¹ to limit the parameter space, while including commonly investigated scan rate values.^[8,10,27] Additionally, a constraint was imposed ensuring a difference of at least 0.15V between the lower and upper potential limit

$$E_L + 0.15\text{V} \leq E_U. \quad (2)$$

2.3. Determination of Initial Dataset

An initial dataset, consisting of four data points, was selected based on an adapted Latin hypercube sampling using the pyDOE Python package.^[28] The number of data points was chosen based on in-silico trials, in which artificial test data sets were

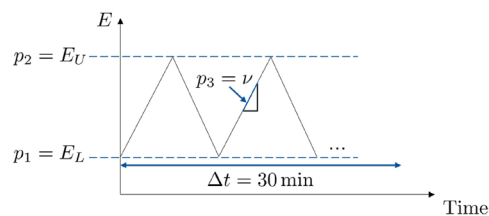


Figure 2. Parametrization of cyclic voltammetry profile for electrode conditioning with parameters p_1 , p_2 , and p_3 .

Table 1. Bounds on the parameters defining the cyclic voltammetry profile for conditioning.

Parameter	Lower bound	Upper bound	Unit
E_L	-0.35	1.45	V
E_U	-0.2	1.6	V
ν	10	500	mV s ⁻¹

used to observe that the model only starts to predict trends with at least four data points. With fewer artificial data points, the model predicted constant values for all the parameter space besides neighborhoods of the training data points. This might be related to having three parameters as inputs of the GP model. As Latin hypercube sampling returns data points from a hypercube, it does not automatically respect the constraint given by Equation (2). Although there is literature on modified Latin hypercube sampling algorithms that can handle constraints,^[29] no readily available implementation was found. Therefore, an adapted procedure was employed instead: The four data points were determined by first sampling eight data points by Latin hypercube sampling and then discarding all data points that did not fulfill Equation (2). The required number of eight data points to result in four feasible data points was identified by trial and error. For each of the four data points, an experiment was conducted, and the resulting ΔE_{10} was recorded to complete the initial dataset.

2.4. Bayesian Optimization with Gaussian Process Model

At each iteration of the hardware-in-the-loop approach shown in Figure 1, the GP model was trained based on the experimental data obtained in the last iteration and all iterations before. The GP model predicts ΔE_{10} as a function of E_L , E_U , and v . Additionally, it provides an estimate of the uncertainty of the prediction. To create the GP model, we used our open-source software package MeLOn.^[30] MeLOn uses the GP model from the gpytorch Python package^[31] and allows to translate the GP model into an optimization model compatible with our open-source deterministic global optimizer MAiNGO.^[32]

When training the GP model, the input values were scaled such that the lower and upper bounds given in Table 1 correspond to 0 and 1, respectively. The output values were standardized to a mean of 0 and a standard deviation of 1. The scaling and standardization were performed using the MinMaxScaler and the StandardScaler from the sklearn Python package.^[33] As mentioned above, the measurements are known to be noisy. However, no repeated experiments were performed in this work, and, for a fair comparison, we did not want to include any noise estimate based on our previous work.^[7] Therefore, we gave no noise estimate to the model. Instead, we allowed for homoscedastic noise to be learned to account for potentially noisy measurements.^[17,31] For the prior noise, we chose a mean of 0 and a variance of 0.5. The variance was chosen based on the assumption that the variance in the objective should be higher than the variance in the noise, such that the trends in the data are significant compared to the noise. Finally, a covariance function with a matern value of three was chosen, and 1000 training iterations (i.e., steps performed to determine the GP model for the current iteration of the hardware-in-the-loop approach shown in Figure 1) were performed.

During BO, the prediction of the GP model as well as the estimated uncertainty of the prediction were considered. In our case, we did this by using the commonly used expected

improvement acquisition function^[34] as the optimization objective. Therefore, our optimization problem formulation can be stated as

$$\begin{aligned} \max_{E_L, E_U, v} \quad & \text{El}(\mu_{\text{GP}}, \sigma_{\text{GP}}, \Delta E_{10, \min}) \\ \text{s.t.} \quad & \mu_{\text{GP}} = \mu_{\text{GP}}(E_L, E_U, v), \\ & \sigma_{\text{GP}} = \sigma_{\text{GP}}(E_L, E_U, v), \\ & E_L + 0.15 \leq E_U \end{aligned} \quad (3)$$

where El represents the expected improvement acquisition function, μ_{GP} represents the expected ΔE_{10} predicted by the GP, σ_{GP} represents the estimated standard deviation of the predicted ΔE_{10} given by the GP, and ΔE_{10} represents the minimal ΔE_{10} found among the available data points. At each iteration of the hardware-in-the-loop approach in Figure 1, this optimization problem was solved with MAiNGO^[32] until a global optimum of the acquisition function was achieved within a relative and absolute tolerance of 0.001. All other settings of MAiNGO were left at their default value.

As a termination criterion, we decided to terminate based on visual inspection of the results in terms of model prediction and achieved improvement. After eleven conducted experiments, the BO approach had explored the most promising region with several experiments and had also explored some adjacent corners without finding any improvement. Therefore, we decided to terminate. Note that further experiments could be used, ensuring a more complete exploration of the search space and reducing the probability of missing potentially better points; this would, however, ultimately increase the experimental effort. Initial attempts to terminate when reaching a pre-defined minimum value of the expected improvement were unsuccessful because this caused the BO to terminate prematurely during the first iterations.

2.5. Transfer Learning for Optimization of Electrode Conditioning of Similar Material

Finally, the transferability of existing knowledge from one material to another was investigated. We made use of the data points collected for $\text{Ni}_{70}\text{Fe}_{30}$ in the first part of this BO study to optimize the conditioning for $\text{Ni}_{99.99}$ via transfer learning.^[24,25]

Different approaches for transfer learning for BO have been proposed in literature.^[35,36] A common strategy is to use a single GP model for both the source and target task.^[36] Many works have examined dedicated approaches for constructing the kernel function, setting the GP prior, and accounting for heterogeneous data scales and noise levels between tasks.^[36] Here, we adopted a rather simple transfer learning approach following the common strategy of using a single GP model for both source and target tasks: We added a fourth input m to the GP model that reflects the material and is 0 for $\text{Ni}_{70}\text{Fe}_{30}$ and 1 for $\text{Ni}_{99.99}$. For reasons of simplicity and compatibility with the already applied GP training and optimization process using MeLOn^[30] and MAiNGO,^[32] we did not examine the mentioned dedicated approaches in this work, but note that they could present potential future improvements.

This resulting optimization problem is

$$\begin{aligned} \max_{E_L, E_U, v} \quad & \text{El}(\mu_{\text{GP}}, \sigma_{\text{GP}}, \Delta E_{10, \min}) \\ \text{s.t.} \quad & \mu_{\text{GP}} = \mu_{\text{GP}}(E_L, E_U, v, m = 1), \\ & \sigma_{\text{GP}} = \sigma_{\text{GP}}(E_L, E_U, v, m = 1), \\ & E_L + 0.15 \leq E_U \end{aligned} \quad (4)$$

where for the GP model evaluation m was kept at the value of 1 to optimize the conditioning for $\text{Ni}_{99,99}$, and ΔE_{10} was taken to be the lowest observed ΔE_{10} for $\text{Ni}_{99,99}$. For the first iteration, $\Delta E_{10, \min}$ was taken to be 0 mV as no conditioning of $\text{Ni}_{99,99}$ has been performed, yet. The initial GP model was trained purely with the eleven experimental points obtained for $\text{Ni}_{70}\text{Fe}_{30}$ (for which m is set to 0). Afterward, at each iteration, an additional point for $\text{Ni}_{70}\text{Fe}_{30}$ (for which m is set to 1) was determined by solving the above optimization problem, the point was conducted experimentally, and the GP model was retrained on all available data for both $\text{Ni}_{70}\text{Fe}_{30}$ and $\text{Ni}_{99,99}$.

For the transfer learning approach, we wanted to investigate how well a new material can be conditioned when only doing a few experiments but the training data for the GP is augmented by the eleven $\text{Ni}_{70}\text{Fe}_{30}$ data points gathered in the first part of this BO study. Therefore, we decided to limit ourselves to four experiments for the transfer learning approach. To evaluate the transfer learning approach, we compared it to starting separately with a new GP model only trained on data obtained for $\text{Ni}_{99,99}$. In that case, we decided to do more experiments, as first an initial set of experiments for $\text{Ni}_{99,99}$ is gathered to train the GP. Seven experiments were performed, consisting of four initial Latin Hypercube sampling experiments and three experiments chosen by BO.

3. Results and Discussion

3.1. Bayesian Optimization of $\text{Ni}_{70}\text{Fe}_{30}$ Conditioning

The chosen points (in terms of E_L , E_U , and v) and the corresponding experimentally achieved ΔE_{10} using the BO approach are

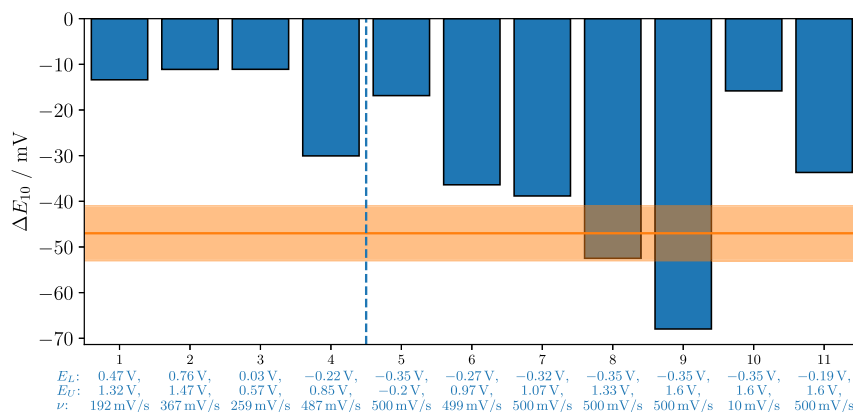


Figure 3. Chosen points (given by the values of E_L , E_U , and v below the x-axis) and experimentally achieved ΔE_{10} for the 30-min cyclic voltammetry conditioning of $\text{Ni}_{70}\text{Fe}_{30}$ at RT in 1 M KOH (blue bars). The orange line represents the best achieved ΔE_{10} of our previous work^[7] (−47 mV) and the orange shaded area represents the respective experimentally observed standard deviation (± 6 mV). Points 1–4 (left of the blue dashed line) were determined using Latin hypercube sampling, and Points 5–11 were determined using BO.

shown in **Figure 3**. A visualization of an example model iteration as well as the parameter values and the achieved ΔE_{10} and ΔE_{100} of all the points measured experimentally can be found in Appendices A and C (**Figure A1** and **Table C1–C4**). Although the values for ΔE_{100} are generally more negative than those obtained for ΔE_{10} , they show a very similar trend. The first four points to the left in **Figure 3** are based on Latin hypercube sampling; it can be observed that the achieved ΔE_{10} is in the range of −10 mV to −30 mV. The first BO experiment (Point 5) did not result in an improved electrode activation. However, throughout the next experiments (Points 6–9), a more negative ΔE_{10} and, therefore, better electrode activation was achieved step by step. The last two experiments (Points 10–11) could not improve the best-found activation, but, as will be described in the following, can be seen as an exploration of areas with high uncertainty.

The exploration of the parameter space by the model is visualized by the blue markers in **Figure 4** (the order in which the points were explored can be inferred from **Figure 3**). After the initial Latin hypercube points (blue crosses, Points 1–4), the BO approach first explored the top right corner (Point 5), and then step-by-step explored the front upper corner (Points 6–9) of the parameter space, where it found a significant improvement in electrode activation. After exploring that region, other areas of the search space were examined. Point 10 was chosen at the bottom front corner, while Point 11 was selected near the top front corner again.

The visualization of the GP model prediction in terms of both predicted ΔE_{10} (**Figure 5a**) and estimated standard deviation of the prediction (**Figure 5b**) shows what the model learned. The front upper corner is where the model learned to expect the best electrode activation (**Figure 5a**). However, after it has been explored significantly, the estimated standard deviation of the prediction is low in that region (**Figure 5b**). Therefore, the expected improvement, which was used as acquisition function in the BO as explained above, became low in this area; accordingly, other areas were subsequently explored.

While **Figure 5** gives insights into the model predictions on the three sides of the parameter space, it does not allow to

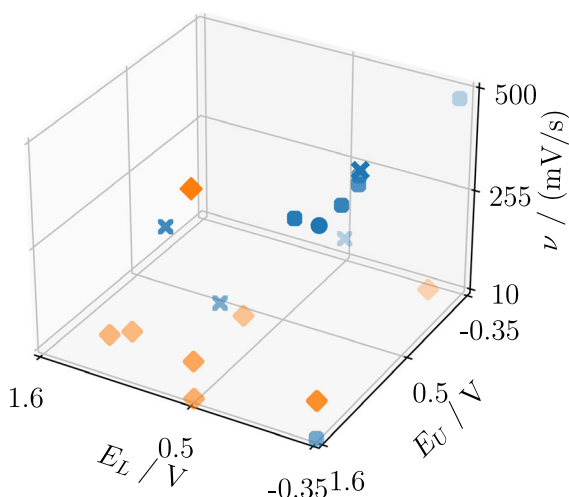


Figure 4. Chosen points of this work (Latin hypercube sampling: blue crosses, BO: blue circles, data in Table C1) and our previous work^[7] (orange diamonds) for the 30-min cyclic voltammetry conditioning of Ni₇₀Fe₃₀ at RT in 1 M KOH. More transparent points are positioned more towards the back in this two-dimensional projection of the three-dimensional parameter space.

see what is predicted inside. For that purpose, we visualized the model prediction for different scan rates (Figure 6). At high scan rates (Figure 6a), there is a strong dependency of the activation on the applied potential window. This can be seen by the rapid decrease of the achieved ΔE_{10} (increase of activation) when

approaching the potential combination of $E_L = -0.35$ V and $E_U = 1.6$ V. At lower scan rates (Figure 6b,c), this dependency in the potentials is not as observable. However, at lower scan rates, the confidence intervals are relatively large as not many experimental points were measured. Overall, we see that there are still large areas with high uncertainty. At the same time, predicted trends mainly occur in the most explored corner.

3.2. Comparison of Ni₇₀Fe₃₀ Results to our Previous Work

Using the BO approach, a significantly improved ΔE_{10} of -68 mV was achieved experimentally (Point 9, i.e., $E_L = -0.35$ V, $E_U = 1.6$ V, $\nu = 500$ mV/s) compared to our previous work, where -47 ± 6 mV (at $E_L = -0.35$ V, $E_U = 1.6$ V, $\nu = 100$ mV/s) was the best reported electrode activation. While the learned noise of the final model is only around 0.3 mV, a considerable standard deviation was observed in the threefold repetitions of the best point conducted in our previous work (around 6 mV). Nevertheless, the improvement of 21 mV is significant. A small positive effect of a higher scan rate ($\Delta E_{10} = -20 \pm 3$ mV using 500 mV/s and $\Delta E_{10} = -19 \pm 2$ mV using 100 mV/s) could also be observed for a potential window of 0.5 to 1.6 V in our previous work^[7] but was not experimentally investigated for a larger potential window as the increase of 1 mV was low compared to the noise level of $\pm 2-3$ mV. When additionally considering the data obtained in this work, a stronger effect of the scan rate

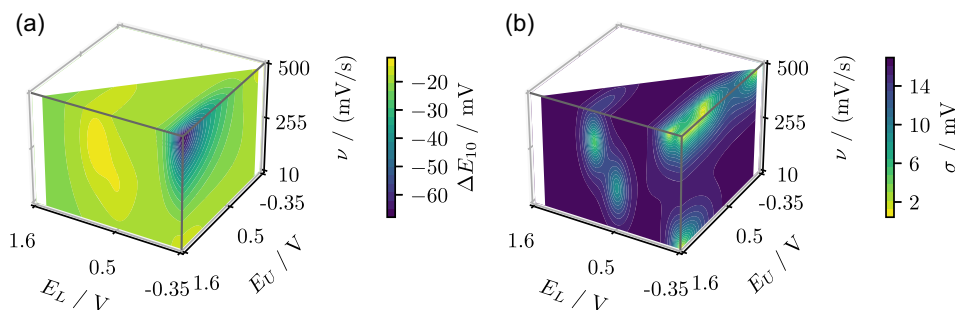


Figure 5. Overview of the GP model prediction for the 30-min cyclic voltammetry conditioning of Ni₇₀Fe₃₀ at RT in 1 M KOH. a) Predicted ΔE_{10} (μ_{GP}) on the surfaces of the parameter space cube, b) predicted standard deviation σ_{GP} on the surfaces of the parameter space cube. The BO approach mostly chose points at the top front corner, where a low ΔE_{10} and in the end also a low σ_{GP} is predicted. Afterward, it started exploring other areas. Predictions are based on the final GP model (i.e., after Point 11 was included in the model training). The white surfaces represent the parts of the cube that are not part of the feasible parameter space due to Equation (2).

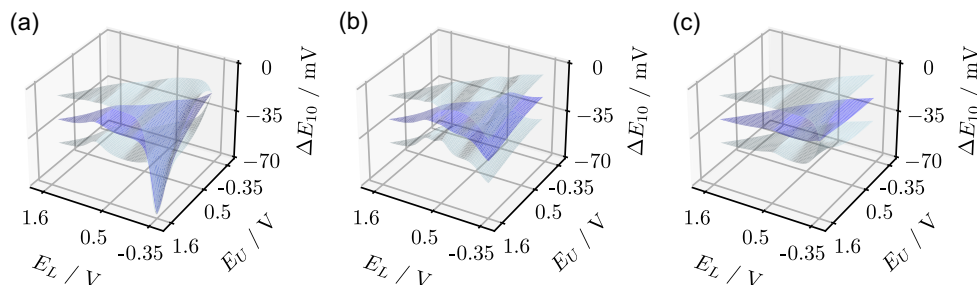


Figure 6. Mean predicted ΔE_{10} (μ_{GP}) (blue) and σ_{GP} confidence interval limits (gray) for the 30-min cyclic voltammetry conditioning of Ni₇₀Fe₃₀ at RT in 1 M KOH. The predictions are shown for different values of ν based on the final GP model. a) $\nu = 500$ mV s⁻¹; b) $\nu = 255$ mV s⁻¹; c) $\nu = 10$ mV s⁻¹.

can be observed for a potential window of -0.35 to 1.6 V (-47 ± 6 mV for 100 mV/s in our previous work^[7] and -68 mV for 500 mV/s in this work). This dependence of the effect of the scan rate on the potential window is also predicted by the GP model (Figure 5a and 6), although with high uncertainty. This example indicates that accounting for interdependencies can be crucial when optimizing electrode conditioning, potentially leading to improved outcomes.

In addition to achieving an improved electrode activation, the number of performed experiments was reduced by the BO approach: in our previous work,^[7] 24 experiments were performed, while in this work, nine experiments were performed until the best point was found, and 11 experiments were conducted overall. It should be noted that no quantitative termination criterion was used, and the termination decision was based on visual inspection. Therefore, the total number of experiments might be influenced by subjective judgment and could have been chosen slightly differently. Nevertheless, the decrease in the number of experiments is considerable, suggesting that the method generally allows for a lower number of experiments than required by our previous work based on trend identification.^[7]

One reason why the BO approach required fewer experiments could be that no repetition of experiments was performed. Performing no repetitions leads to less reliable measurements, but might be beneficial overall and is favored by the fact that the BO approach does not assume these noisy measurements to be ground truth but instead uses them to learn a homoscedastic noise during the training process.^[17,31] Although the learned noise varies a lot between the iterations and in some iterations reached negligible values, in other iterations the noise level reached an estimated standard deviation of up to 2 mV. This value is at the lower end of the standard deviations observed in our previous work.^[7] The large variation in learned noise is likely due to the low number of provided data points.

Another reason why the BO approach required fewer experiments could be that the chosen points differ significantly from those of our previous work (see Figure 4). The BO approach placed most points in the region of the parameter space predicted to have the best ΔE_{10} (top front corner of Figure 4), where the best point was indeed found. In contrast, our previous work did not explore that region at all. This can be explained by the differences of the approaches: In our previous approach based on trend identification, parameters were mostly varied separately to optimize the values for each parameter. In the BO approach, instead, all parameters were optimized simultaneously. Due to that, the interdependency of the parameters can be taken into account, which allows the discovery of additional, otherwise hidden, patterns in the data. Furthermore, in the BO approach, all available information in terms of performed experiments is used in every iteration to determine the most promising next parameter values.

We also tested how the final GP model predicts the points measured experimentally in the previous work^[7] (see Table C4). The estimated uncertainty of all predictions is considerable ($\sigma_{\text{GP}} > 10$ mV as the regions of these points have not been

explored in this work. For most points, the prediction μ_{GP} is within σ_{GP} of the average experimentally measured value ΔE_{10} . However, the model prediction for the best obtained point (-18.9 ± 11.1 mV) deviates significantly from the experimental value ($\Delta E_{10} = -47.2$ mV), showing that the model cannot reliably predict the electrode activation in every region of the search space. This is likely due to no sufficiently similar points having been explored, and in the training data, all points measured with a lower scan rate also showed a worse ΔE_{10} (see Table C1). Moreover, the optimization in this work was performed to optimize ΔE_{10} and not to reach good prediction over the whole parameter space.

3.3. Transfer Learning—Bayesian Optimization of $\text{Ni}_{99,99}$ Conditioning

To quantify the effect of transfer learning for the electrode conditioning of $\text{Ni}_{99,99}$, we compare the transfer learning approach to applying the BO approach separately for $\text{Ni}_{99,99}$.

The separately applied BO approach achieved a maximum ΔE_{10} value of -28 mV during 7 performed experiments, while the transfer learning approach achieved -34 mV during only 4 performed experiments (Figure 7). The separately applied BO approach started with the adapted Latin hypercube sampling points and did not find better points during the BO phase. Instead, the transfer learning approach chose data points close to the best parameter value combinations for $\text{Ni}_{70}\text{Fe}_{30}$. It started by choosing the best working points for $\text{Ni}_{70}\text{Fe}_{30}$ and then explored similar points by testing a lower scan rate, a lower upper potential limit and a higher lower potential limit. A visualization of the chosen points can be found in Appendix B (Figure B1).

Comparing the prediction for $\text{Ni}_{70}\text{Fe}_{30}$ (Figure 6) and for $\text{Ni}_{99,99}$ using transfer learning (Figure 8) shows that the optimal value of v seems to differ for $\text{Ni}_{99,99}$ and $\text{Ni}_{70}\text{Fe}_{30}$, which the model could already learn in the four performed iterations. Similar to the results for $\text{Ni}_{70}\text{Fe}_{30}$ (Figure 6), the predicted dependency on E_L and E_U depends on the scan rate. However, for $\text{Ni}_{99,99}$ the optimal scan rate appears to be predicted in the range of around 255 mV/s (see Figure 8) in contrast to the optimal value of 500 mV/s found for $\text{Ni}_{70}\text{Fe}_{30}$. At the value of around 255 mV/s, ΔE_{10} gets lower when approaching $E_L = -0.35$ V and $E_U = 1.6$ V. At lower (Figure 8a) and higher scan rate (Figure 8c), such a trend cannot be observed to a similar extent in the prediction. The existence of an optimal scan rate in the intermediate range has also been reported for transition metals before.^[11,16]

The results of the transfer learning approach for $\text{Ni}_{99,99}$ are in agreement with our previous work,^[7] in which a very similar best obtained ΔE_{10} of -35 ± 7 mV was found using the same potential limits of $E_L = -0.35$ V, $E_U = 1.6$ V, and a moderately lower scan rate of $v = 100$ mV/s. Additionally, both an even lower scan rate of $v = 10$ mV/s and higher lower potential limit of $E_L = 0.5$ V were separately found to result in a worse ΔE_{10} ^[7] as also predicted by the GP model.

Hence, while the employed approach for transfer learning is rather simple, it seems that transfer learning can potentially speed up the optimization of electrode activation of new materials, when

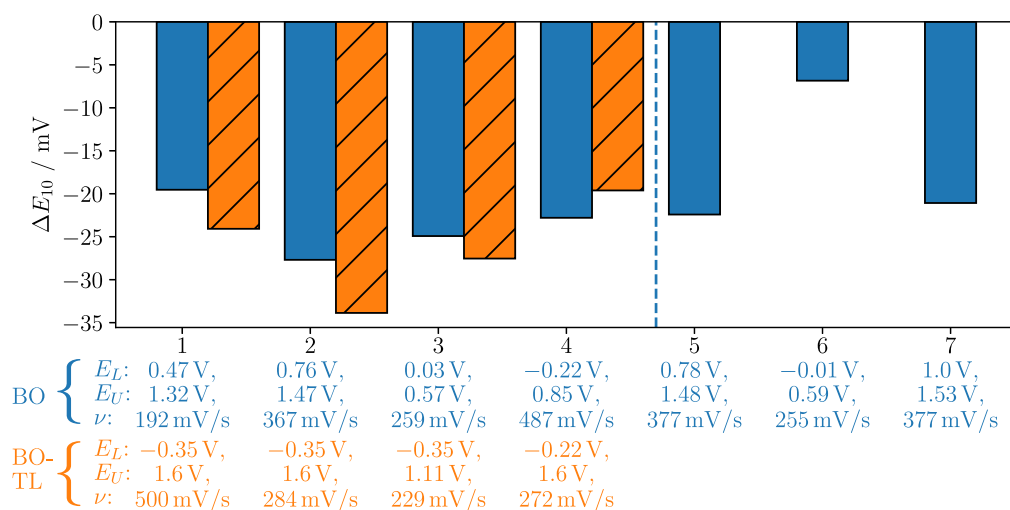


Figure 7. Chosen points (given by the values of E_L , E_U , and ν below the x-axis) and experimentally achieved ΔE_{10} for the 30-min cyclic voltammetry conditioning of $\text{Ni}_{99.99}$ at RT in 1 M KOH. The two alternatives of using a separate BO approach for $\text{Ni}_{99.99}$ (BO, blue bars, upper parameter values) and the transfer learning approach (BO-TL, orange hatched bars, lower parameter values) are shown. For the separate BO approach, Points 1–4 (left of the blue dashed line) were determined using Latin hypercube sampling, and Points 5–7 were determined using BO.

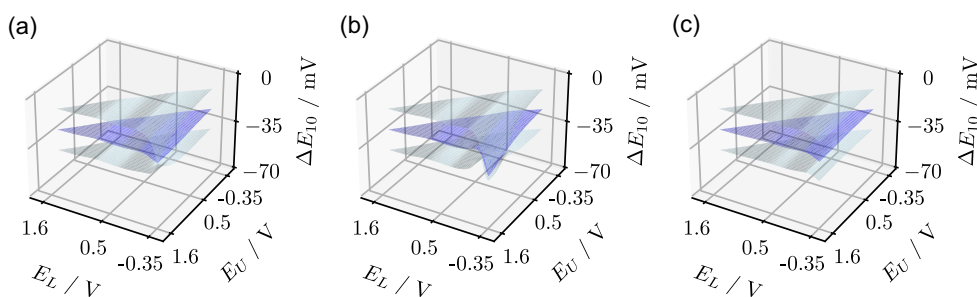


Figure 8. Mean predicted ΔE_{10} (μ_{GP}) (blue) and σ_{GP} confidence interval (gray) for the 30 min cyclic voltammetry conditioning of $\text{Ni}_{99.99}$ at room temperature in 1 M KOH. Different values of ν are shown based on the final obtained GP model when using the transfer learning approach. a) $\nu = 500 \text{ mV s}^{-1}$; b) $\nu = 255 \text{ mV s}^{-1}$; c) $\nu = 10 \text{ mV s}^{-1}$.

data for other, potentially similarly behaving materials is available. The potential similarities between the materials likely play a key role; without them, the approach might be unsuccessful.

4. Conclusion

In this work, we investigated the use of BO with GPs to optimize the electrode conditioning of a $\text{Ni}_{70}\text{Fe}_{30}$ electrode in a hardware-in-the-loop approach and compared it with prior manual experiment-based optimization. Furthermore, we compared a rather simple transfer learning approach to use knowledge of the $\text{Ni}_{70}\text{Fe}_{30}$ conditioning when optimizing the conditioning of $\text{Ni}_{99.99}$ to the usage of a separate BO approach for $\text{Ni}_{99.99}$.

For the conditioning of $\text{Ni}_{70}\text{Fe}_{30}$, the BO approach found parameter values for the cyclic voltammetry conditioning process giving better electrode activation in fewer experiments than our previous and purely experimental work.^[7] On the one hand, this could be due to no performed repetitions of experiments, which is favored by the GP learning the experimental noise. On the

other hand, the BO approach permits taking into account the interdependencies between conditioning parameters and using the knowledge of all past experiments to identify the most promising parameter values for the next experiment. While the approach allows efficient optimization, the model prediction showed significant uncertainty in large areas of the parameter space and did not reliably match all of our previous experimental results.^[7] However, a reliable prediction over the whole parameter space was also not the goal of this study since our BO acquisition function intends to find points with high activation. In the future, the procedure could be adapted to include a quantitative termination criterion reflecting the preferences of the user regarding the trade-off between the number of experiments and the exploration of areas with high uncertainty. Additionally, further constraints on the parameter values and an evaluation at higher current densities (e.g., 500–1000 mA cm^{-2}) could be included to ensure the transferability of the results to the industrial application, and the long-term stability of the achieved activation could be investigated and potentially included in the optimization objective. Concerning the modeling approach, hybrid

modeling approaches could be a promising alternative that permits the consideration of existing knowledge about mechanisms and behavior in the optimization procedure, while not having to model the complete process mechanistically. This could allow to be more strategic about the observations that i) only a comparably small region of the parameter space appears to experience large gradients with respect to ΔE_{10} and ii) after the performed iterations, a considerable uncertainty remains in the GP model prediction in large parts of the parameter space.

For the conditioning of $\text{Ni}_{99.99}$, the transfer learning approach using the data for $\text{Ni}_{70}\text{Fe}_{30}$ gave better results than a BO approach using a separate GP model for $\text{Ni}_{99.99}$ and demonstrated that transferring knowledge from one material to another similar material can be advantageous. More complex transfer learning approaches could be investigated to even better leverage prior knowledge from other materials.

Overall, the use of model-based approaches such as the one presented here provides a systematic approach for efficient optimization of electrode conditioning processes. Thus, the electrode activation can be maximized while requiring only a limited amount of time and material.

Appendix A: Example Model Iteration

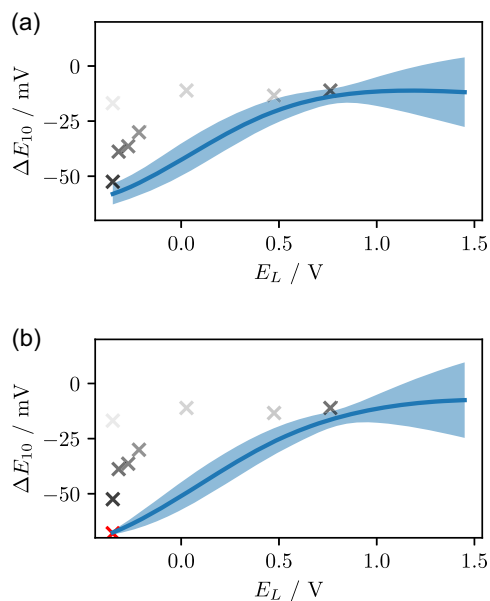


Figure A1. Visualization of Gaussian process model prediction (μ_{GP} given by the blue line) and uncertainty (σ_{GP} confidence interval marked by the blue filled area) for the 30-min cyclic voltammetry conditioning of $\text{Ni}_{70}\text{Fe}_{30}$ at room temperature in 1 M KOH. In Subfigure (a), the model was trained based on the data of the first 8 experiments. The model prediction and σ_{GP} confidence interval are shown here only as a function of E_L for fixed values of $E_U = 1.6\text{ V}$ and $\nu = 500\text{ mV/s}$. The crosses mark the experimental data and are more transparent the more they differ from those values of E_U and ν . After Point 9 was experimentally measured (Subfigure (b)), the new data point (shown as red cross) is added to the model training, and an updated model prediction is obtained. It can be seen that both model prediction and uncertainty are updated around the newly measured point.

Appendix B: Visualization of Chosen Points for $\text{Ni}_{99.99}$

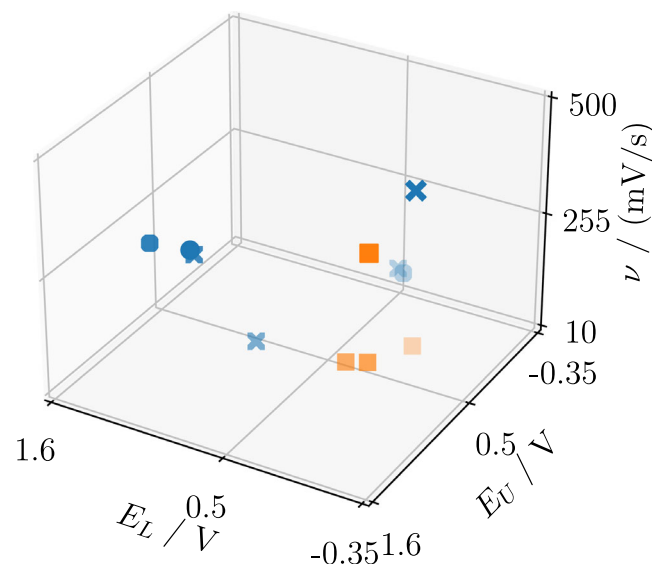


Figure B1. Chosen points for the 30 min cyclic voltammetry conditioning of $\text{Ni}_{99.99}$ at room temperature in 1 M KOH when using a separate BO approach for $\text{Ni}_{99.99}$ (blue, latin hypercube sampling; crosses, Bayesian optimization; circles, data in Table C2) and the transfer learning approach (orange squares, data in Table C3). More transparent points are positioned more towards the back in this two-dimensional projection of the three-dimensional parameter space.

Appendix C: Data of Chosen and Measured Points

Table C1. Chosen points and achieved ΔE_{10} and ΔE_{100} for the Bayesian optimization approach applied to $\text{Ni}_{70}\text{Fe}_{30}$. The point with the best achieved ΔE_{10} is marked in blue. LHS: latin hypercube sampling. BO: Bayesian optimization.

Point number	Method	E_L [V]	E_U [V]	ν [mV s ⁻¹]	ΔE_{10} [mV]	ΔE_{100} [mV]
1	LHS	0.475	1.325	192	-13.4	-41.1
2	LHS	0.763	1.469	367	-11.1	-17.5
3	LHS	0.027	0.574	259	-11.1	-13.0
4	LHS	-0.216	0.851	487	-30.1	-52.7
5	BO	-0.350	-0.200	500	-16.9	-39.8
6	BO	-0.271	0.970	499	-36.4	-81.0
7	BO	-0.320	1.070	500	-38.8	-79.4
8	BO	-0.350	1.333	500	-52.5	-85.1
9	BO	-0.350	1.600	500	-68.0	-128.5
10	BO	-0.350	1.600	10	-15.8	-27.2
11	BO	-0.190	1.600	500	-33.7	-69.9

Table C2. Chosen points and achieved ΔE_{10} and ΔE_{100} for the Bayesian optimization approach applied to $\text{Ni}_{99.99}$. The point with the best achieved ΔE_{10} is marked in blue. LHS: latin hypercube sampling. BO: Bayesian optimization.

Point number	Method	E_L [V]	E_U [V]	ν [mV s^{-1}]	ΔE_{10} [mV]	ΔE_{100} [mV]
1	LHS	0.475	1.325	192	−19.5	−32.5
2	LHS	0.763	1.469	367	−27.7	−19.9
3	LHS	0.027	0.574	259	−24.9	−24.8
4	LHS	−0.216	0.851	487	−22.8	−40.1
5	BO	0.781	1.481	377	−22.4	−19.4
6	BO	−0.012	0.587	255	−6.8	−8.9
7	BO	1.000	1.529	377	−21.1	−21.2

Table C3. Chosen points and achieved ΔE_{10} and ΔE_{100} for the transfer learning Bayesian optimization approach applied to $\text{Ni}_{99.99}$. When training the Gaussian process model in the transfer learning approach, the data for $\text{Ni}_{70}\text{Fe}_{30}$ in Table C1 as well as the data for $\text{Ni}_{99.99}$ in this table were considered. The point with the best achieved ΔE_{10} is marked in blue. BO-TL: Bayesian optimization using the transfer learning approach.

Point number	Method	E_L [V]	E_U [V]	ν [mV s^{-1}]	ΔE_{10} [mV]	ΔE_{100} [mV]
1	BO-TL	−0.350	1.600	500	−24.1	−49.5
2	BO-TL	−0.350	1.600	284	−33.9	−64.5
3	BO-TL	−0.350	1.111	229	−27.5	−59.8
4	BO-TL	−0.223	1.600	272	−19.6	−44.4

Table C4. Chosen points and average achieved ΔE_{10} in our previous work^[7] together with the respective prediction (μ_{GP} and σ_{GP}) given by the GP model trained for $\text{Ni}_{70}\text{Fe}_{30}$ in this work (see Table C1 for training data). The point with the best achieved ΔE_{10} is marked in blue. BO: Bayesian optimization.

E_L [V]	E_U [V]	ν [mV s^{-1}]	ΔE_{10} [mV]	μ_{GP} [mV]	σ_{GP} [mV]
0.500	1.600	10	−13.5	−18.2	15.6
0.500	1.600	500	−20.1	−18.1	16.4
0.500	1.600	100	−19.6	−16.0	12.7
−0.350	0.150	100	−24.9	−20.0	16.7
0.500	1.000	100	−8.4	−16.2	13.0
1.000	1.500	100	−14.5	−19.9	16.8
1.100	1.600	100	−17.9	−20.5	16.9
−0.350	1.600	100	−47.2	−18.9	11.1

Supporting Information

The code used in this work is available at <http://permalink.avt.rwth-aachen.de/?id=868717>. The raw data are made available as a Zenodo repository^[26] and can be accessed here: <https://doi.org/10.5281/zenodo.15877929>.

Acknowledgements

The authors thank Daniel Jungen for feedback on the description of the methodology, Janis Schmitt for support during the review

process, and acknowledge financial support from the German Federal Ministry of Research, Technology and Space (BMFTR project “PrometH2eus”, FKZ 03HY105A).

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

J. Raphael Seidenberg: writing—original draft (besides experimental procedure), conceptualization, methodology (Bayesian optimization), formal analysis (Bayesian optimization), visualization. **Clara Gohlke:** writing—original draft (experimental procedure), writing—review & editing, conceptualization, methodology (experimental procedure), data curation. **Raphael Diebold:** experimental measurements, data curation. **Vera Seidl:** writing—review & editing, funding acquisition, project administration. **Anna K. Mechler:** writing—review & editing, conceptualization, funding acquisition, resources, supervision, project administration. **Alexander Mitsos:** writing—review & editing, conceptualization, funding acquisition, supervision, resources, project administration. **Dominik Bongartz:** writing—review & editing, conceptualization, funding acquisition, supervision.

Data Availability Statement

The raw data are made available as a Zenodo repository²⁶ and can be accessed here: <https://doi.org/10.5281/zenodo.15877929>.

Keywords: conditioning • optimization • pretreatment • redox chemistry • water splitting

- [1] R. L. Doyle, I. J. Godwin, M. P. Brandon, M. E. G. Lyons, *Phys. Chem. Chem. Phys.* **2013**, *15*, 13737.
- [2] J. R. Swierk, S. Klaus, L. Trotochaud, A. T. Bell, T. D. Tilley, *J. Phys. Chem. C* **2015**, *119*, 19022.
- [3] S. Klaus, M. W. Louie, L. Trotochaud, A. T. Bell, *J. Phys. Chem. C* **2015**, *119*, 18303.
- [4] R. Wang, C. Wang, S. Yin, Y. Peng, J. Chen, Y. Deng, J. Li, *Catal. Today* **2021**, *364*, 140.
- [5] J. Wang, Y. Gao, H. Kong, J. Kim, S. Choi, F. Ciucci, Y. Hao, S. Yang, Z. Shao, J. Lim, *Chem. Soc. Rev.* **2020**, *49*, 9154.
- [6] M. S. Burke, L. J. Enman, A. S. Batchellor, S. Zou, S. W. Boettcher, *Chem. Mater.* **2015**, *27*, 7549.
- [7] C. Gohlke, J. Gallenberger, N. Niederprüm, H. Ingendae, J. Kautz, J. P. Hofmann, A. K. Mechler, *ChemElectroChem* **2024**, *11*, e202400318.
- [8] M. E. G. Lyons, R. L. Doyle, I. Godwin, M. O'Brien, L. Russell, *J. Electrochem. Soc.* **2012**, *159*, H932.
- [9] Y. J. Son, S. Kim, V. Leung, K. Kawashima, J. Noh, K. Kim, R. A. Marquez, O. A. Carrasco-Jaim, L. A. Smith, H. Celio, D. J. Milliron, B. A. Korgel, C. B. Mullins, *ACS Catal.* **2022**, *12*, 10384.
- [10] R. L. Doyle, M. E. G. Lyons, *J. Electrochem. Soc.* **2013**, *160*, H142.
- [11] L. D. Burke, M. E. Lyons, *J. Electroanal. Chem. Interfacial Electrochem.* **1986**, *198*, 347.
- [12] N. M. Kubo, F. Ketter, S. Palkovits, R. Palkovits, *ChemElectroChem* **2024**, *11*, e202300460.
- [13] Y. Zuo, V. Mastronardi, A. Gamberini, M. I. Zappia, T.-H.-H. Le, M. Prato, S. Dante, S. Bellani, L. Manna, *Adv. Mater.* **2024**, *36*, e2312071.

- [14] H. R. Zamanizadeh, A. O. Barnett, S. Sunde, B. G. Pollet, F. Seland, *J. Power Sources* **2023**, 564, 232828.
- [15] N. Todoroki, T. Wadayama, *Electrochem. Commun.* **2021**, 122, 106902.
- [16] M. E. G. Lyons, R. L. Doyle, M. P. Brandon, *Phys. Chem. Chem. Phys.* **2011**, 13, 21530.
- [17] C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning* 3 print edition, MIT Press, print edition, Cambridge Mass **2008**.
- [18] B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, A. G. Doyle, *Nature* **2021**, 590, 89.
- [19] V. A. Mints, J. K. Pedersen, A. Bagger, J. Quinson, A. S. Anker, K. M. Ø. Jensen, J. Rossmeisl, M. Arenz, *Acs Catal.* **2022**, 12, 11263.
- [20] J. Guo, B. Ranković, P. Schwaller, *CHIMIA* **2023**, 77, 31.
- [21] M. J. Begall, A. M. Schweidtmann, A. Mhamdi, A. Mitsos, *Comput. Chem. Eng.* **2023**, 171, 108140.
- [22] E. Schulz, M. Speekenbrink, A. Krause, *J. Math. Psychol.* **2018**, 85, 1.
- [23] R. Garnett, *Bayesian Optimization*, Cambridge University Press **2023**, <https://www.cambridge.org/core/books/bayesian-optimization/11AED383B208E7F22A4CE1B5BCBADB44>.
- [24] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, *Proc. IEEE* **2021**, 109, 43.
- [25] K. Weiss, T. M. Khoshgoftaar, D. Wang, *J. Big Data* **2016**, 3, 1.
- [26] J. R. Seidenberg, C. Gohlke, R. Diebold, V. Seidl, A. K. Mechler, A. Mitsos, D. Bongartz, *Dataset for Publication: Bayesian Optimization of Electrode Conditioning of Ni(-Fe) Electrodes for the Alkaline Oxygen Evolution Reaction* **2025**, (Accessed: October, 2025) <https://doi.org/10.5281/zenodo.15877929>.
- [27] R. L. Doyle, M. E. G. Lyons, *Phys. Chem. Chem. Phys.* **2013**, 15, 5224.
- [28] pyDOE: The experimental design package for python 2017.
- [29] M. Petelet, B. Iooss, O. Asserin, A. Lored, *AStA Adv. Stat. Anal.* **2009**, 94, 325.
- [30] A. M. Schweidtmann, D. Bongartz, D. Grothe, T. Kerkenhoff, X. Lin, J. Najman, A. Mitsos, *Math. Program. Comput.* **2021**, 13, 553.
- [31] J. Gardner, G. Pleiss, K. Q. Weinberger, D. Bindel, A. G. Wilson, in *Advances in Neural Information Processing Systems*, (Eds: S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett), 31, Curran Associates, Inc **2018**.
- [32] D. Bongartz, J. Najman, S. Sass, A. Mitsos, *MAiNGO—McCormick-Based Algorithm for Mixed-Integer Nonlinear Global Optimization*, *Process Systems Engineering (AVT.SVT)*, RWTH Aachen University. *The Open-Source Version* **2018**, (Accessed: October, 2025) <https://git.rwth-aachen.de/avt-svt/public/maingo>.
- [33] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Machine Learn. Res.* **2011**, 12, 2825.
- [34] P. I. Frazier, in *Recent Advances in Optimization and Modeling of Contemporary Problems*, (Eds: E. Gel, L. Ntamo, D. Shier, H. J. Greenberg), INFORMS **2018**, pp. 255–278, <https://pubsonline.informs.org/doi/10.1287/educ.2018.0188>.
- [35] P. Tighineanu, K. Skubch, P. Baireuther, A. Reiss, F. Berkenkamp, J. Vinogradsk, *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, (Eds: G. Camps-Valls, F. J. R. Ruiz, I. Valera), Vol. 151, PMLR **2022**, pp. 6152–6181.
- [36] T. Bai, Y. Li, Y. Shen, X. Zhang, W. Zhang, B. Cui, arXiv:2302.05927, (Accessed: October, 2025) **2023**.

Manuscript received: July 14, 2025

Revised manuscript received: September 10, 2025

Version of record online: November 10, 2025