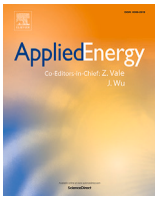
Applied Energy 415 (2026) 127889

Contents lists available at ScienceDirect

Applied Energy

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



Probabilistic electrochemical impedance spectroscopy-based fault diagnosis of solid oxide electrolysis cell systems under variable operating conditions

Luka Žnidarič [a,*], Đani Juričić [a], Bertrand Morel [b], Žiga Gradišar [a,c]

[a] Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
[b] Univ Grenoble Alpes, CEA, LITEN, DTCH F-38000 Grenoble, France
[c] Jožef Stefan Postgraduate School, Jamova 39, 1000 Ljubljana, Slovenia

HIGHLIGHTS

- Fault diagnosis under varying conditions using impedance spectroscopy.
- Probabilistic model captures operating point effects on system behaviour.
- Distribution comparison enables robust detection of system deviations.
- Achieves 97% fault detection and 96% fault isolation accuracy.
- Validated on 90-day electrolysis stack experiment with 600+ spectra.

ARTICLE INFO

Keywords:
Solid oxide electrolysis cell
Electrochemical impedance spectroscopy
Gaussian process
Wasserstein distance
Support vector classifier
Fault diagnosis

ABSTRACT

This work presents a probabilistic diagnostic framework for solid oxide electrolyser cell systems based on electrochemical impedance spectroscopy. Instead of assuming a fixed operating point, the framework explicitly accounts for variations in operating conditions. Spectral data are deconvoluted using an equivalent circuit model whose parameters are obtained via variational Bayes inference, resulting in probabilistic estimates. Parameters inferred under nominal (healthy) operating conditions are then described by a Gaussian process model and subsequently used to check for faults. Specifically, deviations from nominal behaviour are quantified by means of the Wasserstein distance, which measures the discrepancy between the predicted and experimentally obtained parameter distributions. These distance-based residuals form a set of features that enable fault detection and fault isolation using a support vector classifier. The complete framework is validated on data from a 6-cell solid oxide electrolyser short stack collected over a 90-day experimental campaign comprising more than 600 electrochemical impedance spectroscopy spectra. On the test set, the proposed approach achieved 97% accuracy for fault detection and 96% accuracy for multiclass fault isolation under variable operating conditions, demonstrating robust diagnostic performance in dynamically operated solid oxide electrolysis cell systems.

1. Introduction

Solid oxide electrolysis cell (SOEC) systems have emerged as a promising technology for efficient hydrogen production with low or potentially zero carbon emissions. Operating at high temperatures (700–1000 °C), they offer excellent electrical efficiency and potential for integration with renewable energy sources [1,2]. Despite these advantages, the long-term durability and reliability of solid oxide electrolysis cells (SOECs) remain major obstacles to large-scale deployment [3,4]. High operating temperatures accelerate multiple degradation mechanisms such as material decomposition reactivity of the O_2 electrode with the electrolyte, microstructural evolution of the fuel electrode, and contact loss and degradation of the protective coating layer of the interconnect [5]. These degradation processes significantly impact system performance over time, as observed in long-term operation studies [6]. This makes health monitoring, early fault detection, and degradation-aware operation critical for enabling field deployment of solid oxide electrolysis cell (SOEC) technology. Recent work has further emphasised that SOEC degradation is governed by cross-scale coupling mechanisms, linking microstructural evolution with macroscopic system

* Corresponding author.
Email address: luka.znidaric@ijs.si (L. Žnidarič).

https://doi.org/10.1016/j.apenergy.2026.127889
Received 30 January 2026; Received in revised form 28 March 2026; Accepted 10 April 2026
Available online 15 April 2026

Acronyms

ARD	automatic relevance determination
BIC	Bayesian information criterion
DRBS	discrete random binary signal
DRT	distribution of relaxation times
ECM	equivalent circuit model
EIS	electrochemical impedance spectroscopy
ELBO	evidence lower bound
GP	Gaussian process
PEM	proton exchange membrane
SOEC	solid oxide electrolysis cell
SVC	support vector classifier
VB	variational Bayes
WD	Wasserstein distance

behaviour, thereby motivating the integration of multiscale modelling and data-driven approaches for improved diagnosis and mitigation [7].

Inferring the origin of faults and degradation strongly depends on the quality of *features*, i.e., quantities (directly measurable or indirectly evaluated from available process data) that should be sensitive only to the presence of faults in the system, while being insensitive to external disturbances.

Feature generation can be categorised into three main approaches [8]: the passive approach (using only available sensor data), the active approach (using additional perturbations applied to the operational signals), and a combination of both. The choice of approach depends on several factors, in particular the required diagnostic performance indicators and the cost of implementation.

Several review papers highlight the taxonomy of passive diagnostic approaches in hydrogen electrochemical devices in general, and in solid oxide fuel cell and electrolyser systems in particular, as discussed in the early paper by Barelli et al. [9]. A plethora of model-based approaches for passive diagnosis has been recently reviewed in Yang et al. [10]. They span physical models (0D to 3D) as well as data-driven models [11].

The problem with passive approaches, in general, is twofold. First, they can cope only with the slowest dynamic modes of the stack, i.e., modes visible at the system level. Second, many different faults and degradation modes may result in similar feature signatures, rendering passive approaches insufficient for inferring the full range of possible fault modes. This is consistent with observations, reported independently by several authors, that passive approaches primarily cope with stack faults affecting the area-specific resistance as reported in Gallo et al. [12] and more recently Yazbeck et al. [13].

To achieve better discrimination between faults, insight into the fast dynamic modes of the stack is required. This implies the need for additional, deliberate perturbations of the stack that can elicit the response of its higher dynamic modes. This is the aim of *active diagnosis*. The most widely used active approach relies on electrochemical impedance spectroscopy (EIS).

EIS is widely recognised as a powerful tool for analysing the dynamic processes within electrochemical devices [14,15]. It reveals characteristic signatures of charge transfer, gas diffusion, and mass transport phenomena that can be associated with underlying physical degradation modes. To derive useful diagnostic information, electrochemical impedance spectroscopy (EIS) spectra are typically interpreted using equivalent circuit models (ECMs), whose parameters reflect ohmic resistance, electrode polarisation, diffusion effects, and related mechanisms [16]. In recent years, fast EIS methods based on discrete random binary signal (DRBS) have enabled the acquisition of high-quality impedance spectra under realistic dynamic operation [17,18], thus making EIS-based diagnostics feasible in varying environments.

A central challenge, however, remains unresolved. Namely, conventional EIS-based diagnosis is mostly applied under laboratory conditions and assumes fixed operating conditions. How to conduct consistent EIS interpretation under variable operating conditions *in situ* and *in operando* is rarely discussed. There are only a few papers that point out shortcomings of the conventional EIS in practical applications, e.g., [19,20]. Under this assumption, deviations in equivalent circuit model (ECM) parameters are attributed to faults or degradation. In real SOEC operation, where current, temperature, gas composition, and flows could vary, this assumption does not hold any longer. As demonstrated in Dolenc et al. [21], even healthy SOECs exhibit substantial variations in EIS spectra solely due to changes in operating conditions. Consequently, ECM parameters cannot be treated as fixed but must instead be modelled as functions of the process variables. Failing to account for this dependency leads to a high rate of false alarms and unreliable diagnostics.

Surprisingly, the attempts to account for the operating conditions in the online diagnostic applications are rather limited. An attempt in the domain of batteries can be found in Cho et al. [22] where the authors explicitly model the ECM parameters as a function of operating temperature and state of charge as two most significant process variables that define the operating point.

In the context of proton exchange membrane (PEM) fuel cell system in Martín et al. [23] the authors analyse the interaction between electrical and thermal models to simulate the system behaviour at different operating regimes. They model the dependence of the ECM parameters on thermal conditions to capture the dynamic behaviour of the system. The entire model is used only for simulation and not for diagnosis.

An approach addressing variable operating conditions is presented in Maradesa et al. [24], where the authors developed an EIS deconvolution method that can take into account experimental conditions of the observed device. The approach builds on the deconvolution by means of distribution of relaxation times (DRT), whereas a stochastic Gaussian process (GP) model is developed to relate DRT characteristics with the operating parameters in nominal healthy conditions. In this work, diagnosis is not considered.

Despite these efforts, a comprehensive diagnostic framework that explicitly accounts for varying operating conditions remains lacking. To address this challenge, we build on recent advances in probabilistic modelling for EIS-based diagnosis. First, we estimate ECM parameters from each EIS measurement using a variational Bayes (VB) approach Žnidarič et al. [25]. It provides full posterior distributions rather than point estimates. These distributions naturally quantify the uncertainty in the impedance fit and form the basis for principled residual computation. Second, we model the dependency of ECM parameter distributions on operating conditions using GP regression. GP models have previously demonstrated strong performance in energy and electrochemical systems due to their ability to capture nonlinear relationships and deliver calibrated uncertainty estimates [26]. In this work, GP models are trained solely on healthy data to represent the nominal behaviour under varying operating points.

The discrepancy between predicted (GP-based) and measured (VB-based) ECM parameter distributions is quantified using the Wasserstein distance (WD), which measures the dissimilarity between probability distributions. Finally, support vector classifier (SVC), coupled with Platt calibration for probabilistic outputs [27], is used to detect and isolate faults based on the WD-derived features. Classification methods have already shown strong potential in fuel-cell diagnostics [28,29].

The objective of this paper is to demonstrate that such a probabilistic, set-point-aware framework enables accurate fault detection and isolation in dynamically operated SOEC systems. Using experimental data from a 6-cell short stack subjected to both induced and spontaneous

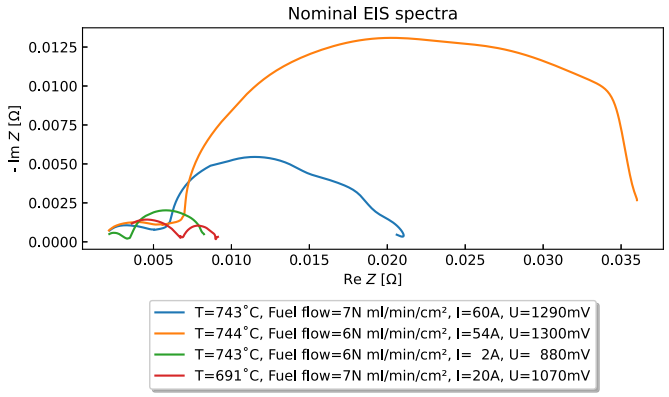


Fig. 1. Comparison of EIS spectra measured at different operating conditions (current, temperature, and fuel flow). Variations in the operating point lead to significant changes in the impedance, highlighting the need for diagnostic methods that take such variations into account.

faults, we show that the proposed method reliably detects and isolates faults while properly accounting for variations in operating conditions. The results confirm that set-point-aware modelling is essential for reliable EIS-based diagnosis and that the integration of VB, GP modelling, WD residuals, and SVC classification yields a robust diagnostic pipeline suitable for in-field operating conditions.

The remainder of this paper is organised as follows. Section 2 formulates the diagnostic challenges that arise from variations in operating conditions. Section 3 presents the proposed modelling and diagnostic framework. Section 4 provides a detailed explanation of the underlying methodology. Section 5 describes the experimental design. Section 6 presents the results obtained from applying the algorithm to the experimental data. The paper closes with conclusions and perspectives for the follow-up.

2. Problem statement

The conventional EIS-based diagnosis procedure tacitly assumes that the process operates at constant operating conditions. During in-field operation, however, it is not always possible to ensure identical operating conditions for EIS evaluation and, consequently, consistent comparison of the obtained spectra.

If spectra are evaluated under *equal operating conditions*, one can relatively unambiguously associate changes in the spectra with faults or degradation modes. If, on the other hand, EIS spectra are taken at different operating conditions, these variations can significantly impact the spectra and the associated ECM parameters.

A natural question, therefore, arises: how can one consistently infer about the SOEC system from EIS curves measured at different operating conditions? To better demonstrate the problem, Fig. 1 shows EIS curves taken at different current, temperature, and fuel flow conditions. The differences can be rather substantial, for instance, in cases where higher steam conversion is reached.

3. The proposed diagnostic framework

The key idea of the proposed framework is to take into account the variability of operating conditions when performing fault diagnosis. Instead of treating nominal ECM parameters as fixed, they are modelled as functions of the process variables Fig. 2.

Specifically, each ECM parameter θ_i is expressed as a function of the operating conditions

$$\theta_i = f_i(\mathbf{x}),$$

where process variable $\mathbf{x}$ includes conditions such as temperature, current density, fuel flow, and voltage.

Fig. 2. The idea of modelling the ECMs as a function of the vector of process variables.

The model is static because it is assumed that during the EIS session the process variables can be considered constant. Moreover, to cope with uncertainties in the process, we use GP regression $f_i(\mathbf{x}) \sim \mathcal{GP}(m_i(\mathbf{x}), k_i(\mathbf{x}, \mathbf{x}'))$, where $m_i(\mathbf{x})$ and $k_i(\mathbf{x}, \mathbf{x}')$ represent mean and covariance kernel functions, respectively. Both are selected and optimised during preprocessing.

The outline of the probabilistic diagnostic procedure, tailored for non-stationary operating conditions, is shown on the right side of Fig. 3. The conventional EIS-based diagnostic procedure applicable in stationary operating conditions is provided on the left side of Fig. 3. Both compare the measured ECM parameters to the predicted, but only the former can confidently attribute the resulting difference to an occurring fault in case of variable operating conditions.

The procedure proposed in this paper consists of the following steps:

1. *EIS session.* Both current and voltage are sampled under perturbation of the current. To reduce the duration of EIS session and minimize the influence of noise and fluctuations, especially in fuel feed, the stack is perturbed by DRBS, see Appendix E for details.
2. *Probabilistic deconvolution of EIS.* EIS is interpreted by means of ECM. To account for inevitable uncertainties present in the process, we use the probabilistic VB procedure proposed by Žnidarič et al. [25]. As a result of the procedure, we get the probability density distributions of the ECM parameters.
3. *Predicting ECM parameters from process variables.* A GP model is applied to predict the distributions of ECM parameters as the outcome of a new EIS session, given that the values of process variables are known. In case of no fault and no degradation, the predicted and the measured values are nearly the same, otherwise they should differ.
4. *Fault detection.* To quantify the discrepancy between the two resulting distributions (the measured and the predicted) WD is utilised. The greater the dissimilarity between the two, the higher the WD, which makes WD act as a residual in a fault detection setting. Moreover, due to taking into account information about operating conditions, residuals become insensitive to variations of the set-point of the device.
5. *Fault isolation.* It is done by SVC trained on the WD-derived features, enabling identification of specific fault types under varying conditions.

4. Methodology

A detailed background on VB, GP, WD, and SVC is provided in Appendix A–D, respectively. The source code for results replication along with a demo case can be found at https://repo.ijs.si/lznidaric/fault-diagnosis.

(i) Probabilistic EIS deconvolution via VB

We interpret the measured impedance spectrum $Z(j\omega)$ using an ECM built from a sum of a resistor and two RQ elements

$$Z(j\omega) = R_s + \frac{R_1}{1 + (j\omega)^{\alpha_1} R_1 Q_1} + \frac{R_2}{1 + (j\omega)^{\alpha_2} R_2 Q_2} + j\omega L, \tag{1}$$

where R_s, R_1, R_2 are resistances, Q_1, Q_2 are constant phase elements, α_1, α_2 are fractional exponents ($0 < \alpha \leq 1$) and L is inductance, which we fixed to a constant term using high frequency data (10^{-9}). This topology

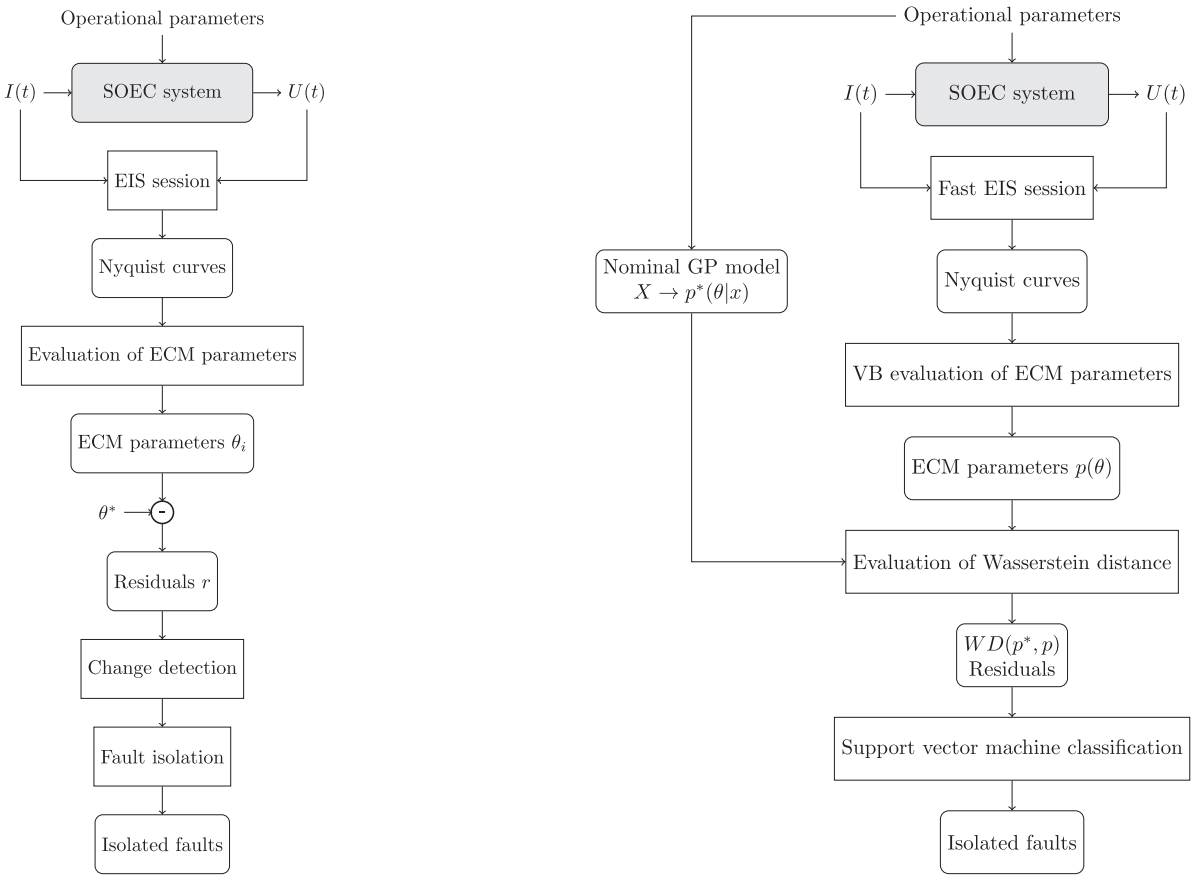


Fig. 3. Comparison between the conventional EIS procedure applicable in stationary operating conditions (left) and the novel EIS-based diagnostic procedure tailored for the non-stationary operating conditions (right).

corresponds to a commonly used extension of the Randles circuit, where multiple RQ elements represent distinct electrochemical relaxation processes [30]. The constant phase elements account for the depressed semicircles typically observed in impedance spectra. Alternative circuit topologies were also considered in earlier work, but the chosen ECM provided the best balance between model simplicity and quality of fit. RQ elements are widely used to model a range of electrochemical systems, from fuel cells [31] and batteries [32] to corrosion phenomena [33].

Let

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots]^\top = [R_s, R_1, R_2, Q_1, Q_2, \alpha_1, \alpha_2]^\top$$

denote the model parameters and $\mathcal{D} = \{Z(j\omega_k)\}_{k=1}^N$ the measurement data. Exact Bayesian inference of $\boldsymbol{\theta}$ turns out to be computationally prohibitive, so we instead opt for an approximate method to bypass these bottlenecks.

Variational Bayes (VB) facilitates inference by approximating posterior probability distribution $p(\boldsymbol{\theta} \mid \mathcal{D})$ with a parametrised distribution $q_\phi(\boldsymbol{\theta})$. A parametrised distribution is limited to a fixed family of distributions, i.e., log-normal distribution for $R_i, i \in \{s, 1, 2\}$ and $Q_i, i \in \{1, 2\}$, and beta distributions for $\alpha_i \in \{1, 2\}$, which transforms inference into a tractable optimisation task.

To allow direct comparison with the GP model outputs, we applied a logarithmic transformation to log-normal posteriors, which converts them into an equivalent normal distribution.

(ii) Modelling the nominal ECM parameters with GP

A GP model is derived for each ECM parameter. It defines a distribution over functions

$$f_i(\mathbf{x}) \sim \mathcal{GP}(m_i(\mathbf{x}), k_i(\mathbf{x}, \mathbf{x}')), \tag{2}$$

where $m_i(\mathbf{x})$ is the mean function (here assumed to be zero) and $k_i(\mathbf{x}, \mathbf{x}')$ is the covariance or kernel function. The kernel function determines the smoothness and correlation structure of the modelled function. For common kernel function options see Appendix B.

Operating conditions influence ECM parameters. Therefore, we model them as functions of process variables $\mathbf{x}$ (temperature, current density, flows, etc.)

$$\hat{\boldsymbol{\theta}}_i = \mathbf{f}_i(\mathbf{X}) + \boldsymbol{\varepsilon}_i, \tag{3}$$

where $\hat{\boldsymbol{\theta}}_i = [\hat{\theta}_i^1, \hat{\theta}_i^2, \dots, \hat{\theta}_i^n]^\top$ denotes n (log transformed) *noisy* measurements of i-th ECM. $\boldsymbol{\varepsilon}_i$ denotes white measurement noise with corresponding matrix of variances $\mathbf{D}_i = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2)$. Both $\hat{\boldsymbol{\theta}}_i$ and $\mathbf{D}_i$ are provided by running VB on the impedance data. $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^\top$ denotes the matrix of our inputs, which is provided by measuring the process variables.

The set of kernel function hyperparameters, denoted by $\boldsymbol{\vartheta}_i$, is learned by maximising the log marginal likelihood

$$\log p(\hat{\boldsymbol{\theta}}_i \mid \mathbf{X}, \boldsymbol{\vartheta}_i) = -\tfrac{1}{2}\hat{\boldsymbol{\theta}}_i^\top \left(\mathbf{K}_i(\mathbf{X}, \mathbf{X}) + \mathbf{D}_i\right)^{-1}\hat{\boldsymbol{\theta}}_i - \tfrac{1}{2}\log\left|\mathbf{K}_i(\mathbf{X}, \mathbf{X}) + \mathbf{D}_i\right| - \tfrac{n}{2}\log(2\pi),$$

where entries of matrix $\mathbf{K}_i$ are given by the chosen kernel function k_i

$$[\mathbf{K}_i(\mathbf{X}, \mathbf{X})]_{j,k} = k_i(\mathbf{x}_j, \mathbf{x}_k), \quad k_i'\text{s hyperparameters given by } \boldsymbol{\vartheta}_i. \tag{4}$$

Optimisation is done using ADAM optimizer [34]. From measured $\hat{\theta}_i$ and optimised ϑ_i, we infer the posterior distribution of the latent function values θ_i^* at unseen inputs $\mathbf{X}^*$

$$\theta_i^* | \hat{\theta}_i, \vartheta_i \sim \mathcal{N}\left(\mathbf{m}_i^*, \mathbf{\Sigma}_i^*\right), \tag{5}$$

where

$$\mathbf{m}_i^* = \mathbf{K}_i(\mathbf{X}^*, \mathbf{X})[\mathbf{K}_i(\mathbf{X}, \mathbf{X}) + \mathbf{D}_i]^{-1}\hat{\theta}_i, \tag{6}$$
$$\mathbf{\Sigma}_i^* = \mathbf{K}_i(\mathbf{X}^*, \mathbf{X}^*) - \mathbf{K}_i(\mathbf{X}^*, \mathbf{X})[\mathbf{K}_i(\mathbf{X}, \mathbf{X}) + \mathbf{D}_i]^{-1}\mathbf{K}_i(\mathbf{X}, \mathbf{X}^*).$$

To select the most appropriate kernel function we applied Bayesian Information Criterion (BIC) [35]. Bayesian Information Criterion (BIC) was chosen due to its sound theoretical background, simplicity and wide applicability in the Bayesian setting. In our case, the best trade-off between bias and variance was obtained by Matérn-3/2 kernel with a white-noise term (see Appendix B for details).

(iii) Wasserstein distance between the distributions

To detect changes in i-th ECM parameter we compare new measurements (obtained as VB posterior) $\hat{\theta}_i^*$ with the nominal GP predictive distribution $\theta_i^* | \hat{\theta}_i, \vartheta_i$ using squared WD of order two, denoted as W_2^2.

Since the compared distributions are both normal $\hat{\theta}_i^* \sim \mathcal{N}(\hat{\boldsymbol{m}}_i^*, \hat{\mathbf{\Sigma}}_i^*)$, $\theta_i^* | \hat{\theta}_i, \vartheta_i \sim \mathcal{N}(\boldsymbol{m}_i^*, \mathbf{\Sigma}_i^*)$, in that case W_2^2 simplifies to

$$W_2^2\left(\hat{\theta}_i^* \parallel \theta_i^* | \hat{\theta}_i, \vartheta_i\right) = \|\hat{\mathbf{m}}_i^* - \mathbf{m}_i^*\|_2^2 + \mathrm{tr}\left(\hat{\mathbf{\Sigma}}_i^* + \mathbf{\Sigma}_i^* - 2\left(\left(\mathbf{\Sigma}_i^*\right)^{1/2}\hat{\mathbf{\Sigma}}_i^*\left(\mathbf{\Sigma}_i^*\right)^{1/2}\right)^{1/2}\right) \tag{7}$$

see Appendix C for further details. This simple closed-form solution makes implementation in real applications computationally feasible and affordable on standard control platforms.

The WD is treated as a residual, which is close to zero if the two distributions are close and different from zero otherwise. For the specific application below, the residual feature vector is defined as

$$\mathbf{w}_d = \left[W_{2,R_1}^2, W_{2,R_2}^2, W_{2,Q_1}^2, W_{2,Q_2}^2\right]^\top,$$

and serves as the basis for the subsequent fault-detection and fault-isolation classifiers. It turned out that limiting the number of free ECM parameters to be identified alleviates the problem of excessive fluctuations in the parameter estimates. As a result of the compromise, R_s was not used in regression but was fixed to 10^{-5}, similarly L was fixed to 10^{-9}. We decided not to regress on the remaining ECM parameters α_1, α_2, since their inclusion in the learning process did not significantly improve the results.

Redundancy in α suggests that shape changes are likely aligned with variations in other features.

(iv) Calibrated classification with SVC

We trained an SVC on a set of $\mathbf{w}_d$ for fault detection and isolation. To obtain reliable class probabilities, we applied Platt scaling to the SVC decision scores, with parameters (A, B) fitted by cross-validated logistic loss. This supports thresholding and decision policies under uncertainty (see Appendix D).

Algorithm implementation

The algorithm was implemented in Python. The variational Bayes component was developed using the PyTorch framework [36], whereas the Gaussian process regression and support vector classification were implemented using scikit-learn [37].

5. Experiment design

5.1. System and instrumentation

The experimental study was carried out on a 6-cell anode-supported short stack manufactured by SolydEra and operated at the CEA laboratory within the H2020 REACTT project (https://www.reactt-project.eu/). Each cell has an active area of 80 cm^2 and is arranged in a cross-flow configuration, with steam–hydrogen supplied to the cathode inlet and air supplied to the anode inlet. The short stack is housed inside a furnace, whose temperature governs the overall stack thermal environment.

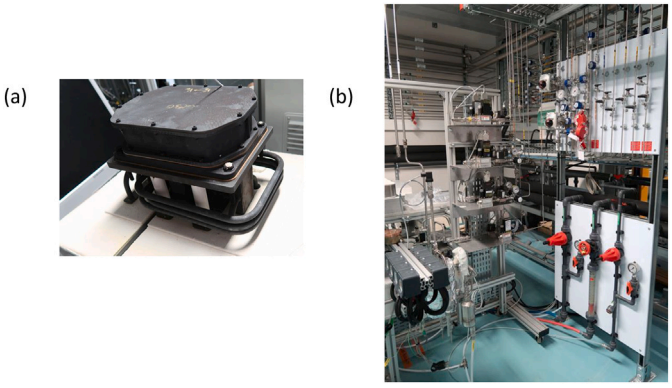


Fig. 4. (a) The 6-cell short stack from SolydEra and (b) the associated test bench at CEA used for thermal management, reactant delivery, electrical operation and data acquisition.

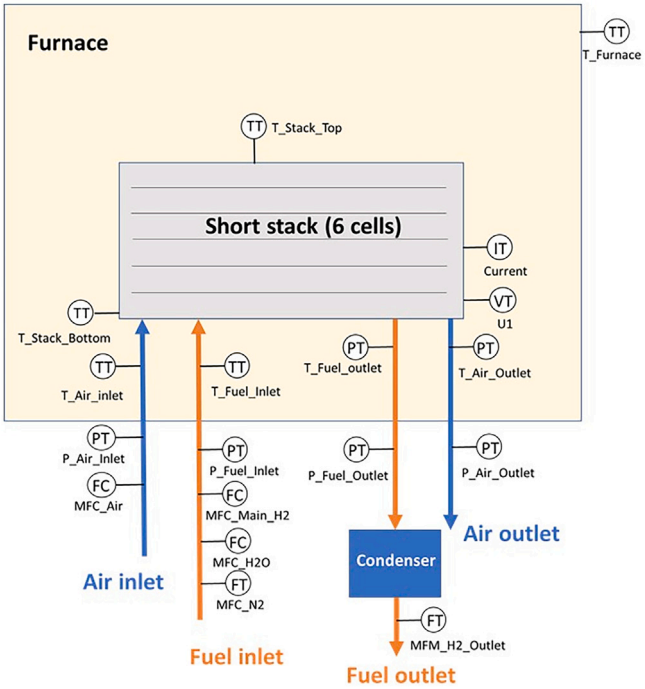


Fig. 5. Simplified flowsheet of the SOEC system with the 6-cell short stack and the principal instrumentation used for monitoring process variables and defining operating points.

Fig. 4 shows the short stack and the corresponding test bench. A simplified process flowsheet with the main instrumentation is presented in Fig. 5.

5.2. Measured signals

During the 90-day (~ 2000h) campaign, a comprehensive set of process variables was recorded with a sampling interval of 1 s. These measurements include the stack current, individual cell voltages, furnace and stack temperatures, inlet and outlet temperatures on both the air and fuel sides, inlet/outlet pressures, reactant flow rates (air, steam–hydrogen mixture), and the hydrogen outlet flow.

Table 1
Descriptions of the measured signals used in the diagnostic framework.

Name	Description
Voltage	Cell voltage [V]
Current	Stack current [A]
P_Air_Inlet	Air inlet pressure [mbar]
P_Air_Outlet	Air outlet pressure [mbar]
P_Fuel_Inlet	Fuel inlet pressure [mbar]
P_Fuel_Outlet	Fuel outlet pressure [mbar]
MFC_Air	Air inlet flow rate [slph]
MFC_Auxiliary_H2	Auxiliary hydrogen inlet flow rate [slph]
MFC_Main_H2	Main hydrogen inlet flow rate [slph]
MFC_N2	Nitrogen inlet flow rate [slph]
MFC_H2O	Water vapour inlet flow rate [slph]
MFM_H2_Outlet	Hydrogen outlet volumetric flow [slph]
T_Stack_Bottom	Temperature at stack bottom [°C]
T_Stack_Top	Temperature at stack top [°C]
T_Air_Inlet	Temperature at air inlet [°C]
T_Air_Outlet	Temperature at air outlet [°C]
T_Fuel_Inlet	Temperature at fuel inlet [°C]
T_Fuel_Outlet	Temperature at fuel outlet [°C]
T_Furnace	Furnace temperature [°C]

The complete list of measured quantities, along with their identifiers used throughout this paper, is provided in Table 1. These variables form the operating-condition inputs to the GP models and are used to characterise the system state during EIS acquisition.

We did not pick the minimal adequate set of input variables by hand, since during the optimisation an unnecessary input can be rendered irrelevant by the optimizer inflating the length scale of the pertaining GP kernel [38]. We did, however, focus on only one electrolyser cell out of six to limit the scope of our paper. Cell 1 was selected arbitrarily, solely because it was the first in the sequence. The analysis in this study is performed on a single cell to enable a controlled evaluation of the proposed diagnostic framework. This choice isolates the relationship between measured features and fault conditions without the additional complexity introduced by stack-level gradients. The methodology itself is not limited to single-cell measurements and can be extended to multiple cells or the entire stack. However, due to thermal and gas distribution gradients, measurements from a single cell are primarily sensitive to local faults, while faults in distant cells may be less pronounced. In practice, the framework can be applied flexibly, i.e., to individual cells, selected subsets, or aggregated stack signals, allowing for a trade-off between diagnostic resolution and system complexity.

The evolution of process variables over the course of the experiment is shown in Fig. 6. The variable names in the figure correspond directly to those listed in Table 1, ensuring consistent mapping between the raw measurements and the diagnostic framework.

5.3. Description of the experimental campaign

The stack was operated continuously for approximately 90 days under a variety of operational regimes.

Throughout the active campaign EIS measurements were acquired every 2 h. In total, more than 600 EIS spectra were recorded, each followed by ECM identification to obtain parameter posteriors for the diagnostic framework. The timeline of operational regimes and faults is summarised in the list below:

- *Fault-free nominal operation* [Day 1–31 and Day 49–68]: at 750 °C with fuel flow 6 N ml/min/cm^2. Fuel's steam/hydrogen mixture is 90/10 (a common setting, see [39]), where hydrogen is present to avoid Ni reoxidation [40];
- *Fault-free performance mapping* [Day 32–48]: changing current (~1A to ~60A), temperature (750 °C, 725 °C, 700 °C) and fuel flow (5,6,7,8,9 N ml/min/cm^2) to map set-point behaviour;
- *Humidification of O_2 electrode* [Day 69–72]: a slow increase of humidification from 0% to 12% over the course of a few days. Air that is feeding the O_2 electrode is usually dry. Humidification of the air is suspected to accelerate the performance degradation of the electrode [41]. Post-mortem analysis of cells exposed to humid air revealed the formation of SrO [41–43], $Sr(OH)_2$[44], $SrSO_4$ [45], and $SrCO_3$ in the presence of CO_2 [46], for instance;
- *Hydrogen starvation* [Day 74–78]: changing the composition of fuel from 90/10 to 97/3, 99/1, 100/0, where fractions indicate steam/hydrogen ratios, to try to observe a change in performance, similarly to [47,48];
- *Steam pump failure* [Day 78–79]: fuel flow was impeded by faulty pump behaviour, which resulted in rising steam conversion rates. After the failure, the pump was overhauled and restored to function;
- *High steam conversion* [Day 82–84]: fuel flow was lowered (4,5,6 N ml/min/cm^2) so the device could achieve a high conversion rate within the operating constraints. High rates of utilisation can allow for higher efficiency of SOEC [49], but they may also introduce higher degradation rates [50].

5.4. Dataset partitioning

For model development and evaluation, the available dataset was divided into training, validation, and test subsets. To preserve the temporal integrity of the experiment and prevent leakage of fault information, the splits were performed per operating regime rather than randomly. The resulting partition is illustrated in Fig. 7. The data from the nominal regime is separated into training and validation sets, and the test set is constructed from the remaining testing period data. The test, therefore, contains nominal operation data as well as the measurements obtained during induced and spontaneous faults.

6. Results on experimental data

The objective of this section is to evaluate the performance of the proposed diagnostic pipeline across the experiment.

6.1. Gaussian process modelling of ECM parameters

The GP models were trained exclusively on healthy operating data and were used to predict the expected values of the ECM parameters under varying operating conditions. Specifically, they predicted the (natural) logarithm of ECM parameters, since the parameters are non-negative numbers and Gaussian processes (GPs) are in general not bounded.

The ECM parameters selected for the final diagnosis were R_1, R_2, Q_1, and Q_2. The parameters α_1 and α_2 were excluded, as they did not contribute significantly to improving diagnostic performance. The parameters R_s and L were fixed to 10^{-5} and 10^{-9}, respectively, since their estimated values were nearly constant, and fixing them resulted in a more stable EIS deconvolution. The proposed diagnostic algorithm does allow the addition of extra parameters if the specific case calls for it, i.e., if the added parameters increase the quality of the diagnosis.

During nominal operation, the GP predictions closely match the VB estimated parameter distributions, while clear deviations emerge during fault intervals, as shown in Fig. 8. For clarity, a close-up of the test regime is provided in Fig. 9, highlighting the model behaviour under fault conditions.

We observe that the algorithm responds clearly to the investigated fault scenarios. For the first fault (*humidification*), the variance of the Q_1 and R_2 parameters increases rapidly, while the mean predictions remain relatively stable. *H_2 starvation* also induces a pronounced increase in the prediction variance of Q_1 and R_2, accompanied by a substantial rise in the variance of the R_1 parameter. In this case, the mean predictions deviate significantly from their nominal values as well. The spontaneous *pump failure* leads to a change in system behaviour, which is reflected

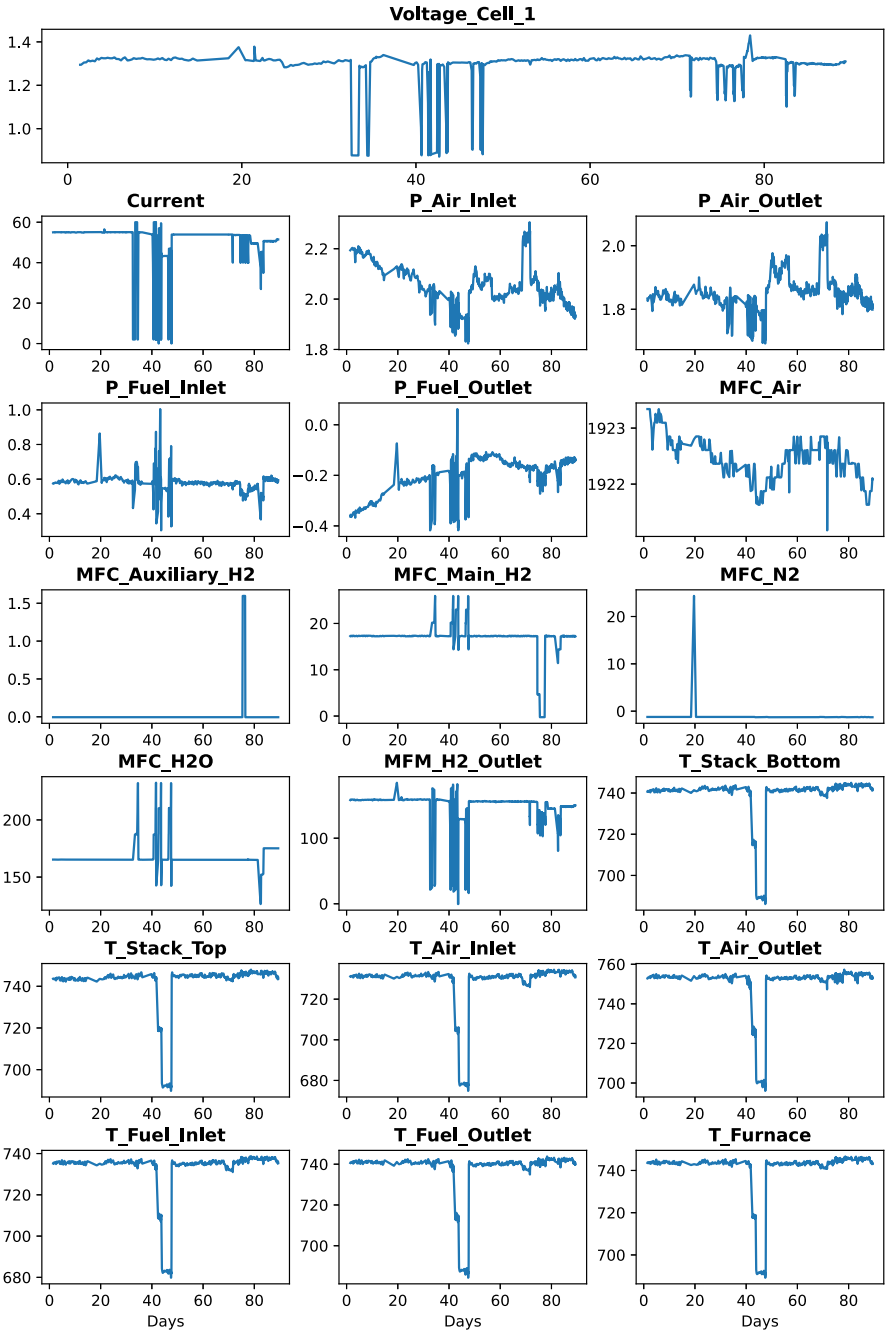


Fig. 6. Measured process variables recorded during the 90–day REACTT experiment. These signals define the operating condition inputs to the GP models. A full description of each variable is provided in Table 1.

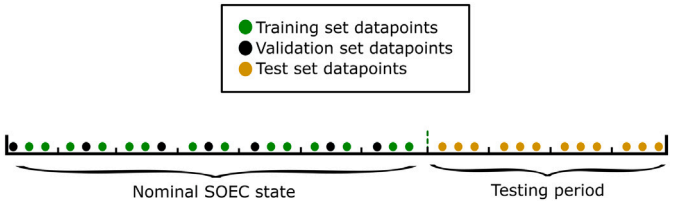


Fig. 7. Schematic representation of the dataset split into learning, validation, and test subsets. The partition respects the temporal structure of the 90–day campaign and prevents fault-event leakage into the training data.

in the steadily increasing prediction error throughout the fault interval. The prediction variance remains relatively low, indicating that the operating-condition inputs remain largely nominal, while the system itself has changed. After the pump replacement, the prediction variance increases, as expected, since the post-replacement data were not part of the model's training set. Finally, the *high steam conversion* fault primarily manifests as increased prediction error in the R_1 and R_2 parameters, while the remaining parameters remain comparatively consistent.

In the training and validation subsets, prediction variance remains low and deviations are limited. In contrast, the test subset (containing fault events) exhibits pronounced departures from the nominal GP predictions.

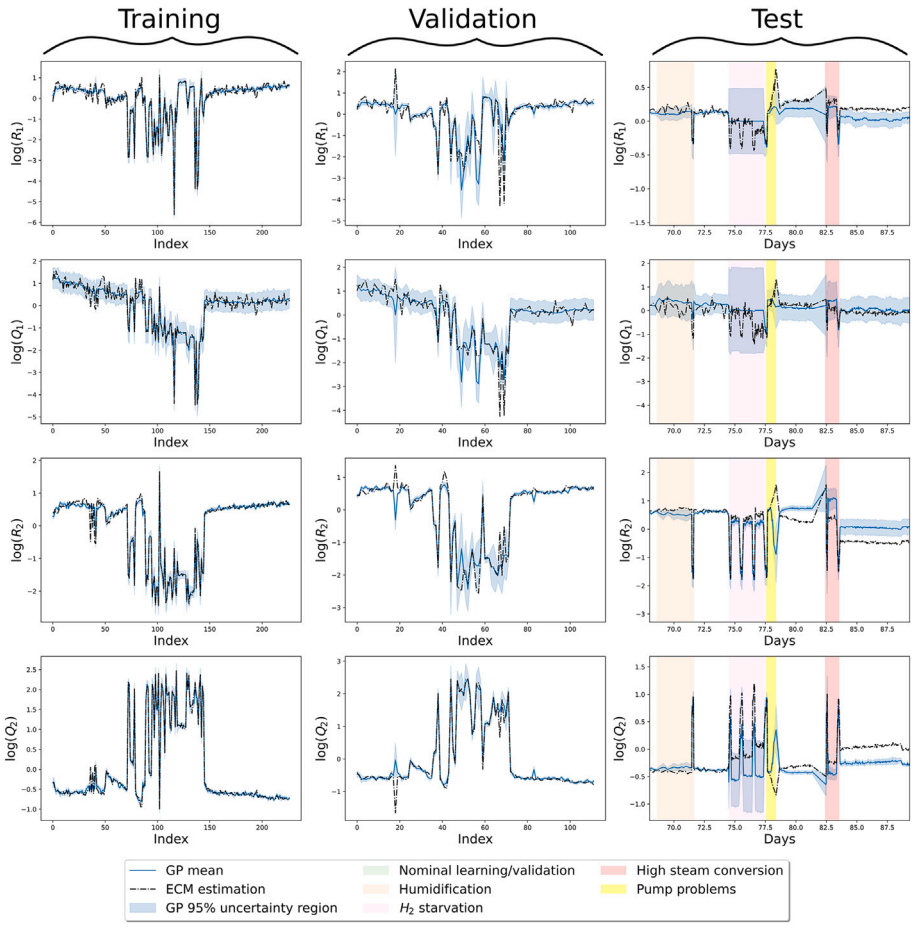


Fig. 8. GP predictions for the logarithm of ECM parameters across training, validation, and test datasets. The model accurately follows nominal behaviour and exhibits clear deviations during induced and spontaneous fault periods.

6.2. Wasserstein distance-based residuals

To quantify deviations between predicted and measured ECM parameter distributions, the squared WD was employed. This metric remained stable across varying operating conditions while remaining sensitive to deviations caused by faults. For a quantitative assessment of predictive performance, we additionally report dataset-level summary metrics. In particular, the fit between the GP predicted and VB derived distributions is evaluated using the averaged WD, while the relative root mean square error (rRMSE)[1] is used to assess the accuracy of the predictive means. Table 2 summarises these metrics across the training, validation, and test datasets. The results show low errors on the training set, moderate degradation on validation data, and a more pronounced increase on test data, indicating reduced accuracy under distribution shift while maintaining consistent qualitative behaviour.

Fig. 10 shows the WD values for the four observed ECM parameters. Induced fault periods show significant increases in the WD, and the spontaneous fault is also clearly reflected in the residual behaviour. Occasional spikes in the WD values are observed during the nominal operating regime. These primarily occur during performance mapping and other tests involving rapid changes of the operating point. Since such transient operating conditions are only sparsely represented in the training data, the GP model is not well constrained in these regions, which can lead to temporary increases in the residual values. These results confirm that the WD serves as an effective and interpretable indicator of system deviations under varying operating conditions. We note that after the pump failure, the replacement of the pump introduced a constant offset in the WD values due to a permanent change in the system configuration. This offset was estimated and subtracted from the affected WD values to restore comparability across the dataset and preserve continuity of the analysis.

6.3. Fault detection results

The WD-based residuals were used as inputs to a binary SVC classifier for fault detection. The classifier was trained to distinguish nominal from faulty operation across the full range of operating conditions.

The model achieved an accuracy of 97% on the 137 test samples (76 nominal, 61 faulty). Both classes reached a precision, recall, and F1-score of 97%, demonstrating balanced discrimination performance under varying and noisy conditions. The classification report is provided in Table 3.

Fig. 11 shows the confusion matrix. Only two false positives and two false negatives occurred, likely due to transient noise in the VB estimates near fault transitions.

Fig. 12 displays the calibrated probabilities obtained via Platt scaling. Elevated fault probabilities closely coincide with induced and spontaneous fault intervals.

6.4. Fault isolation results

To identify the type of deviation, the same WD-based feature vector was used as input to a multiclass SVC. The classifier assigned each

[1] Defined as the RMSE normalised by the standard deviation of the reference values.

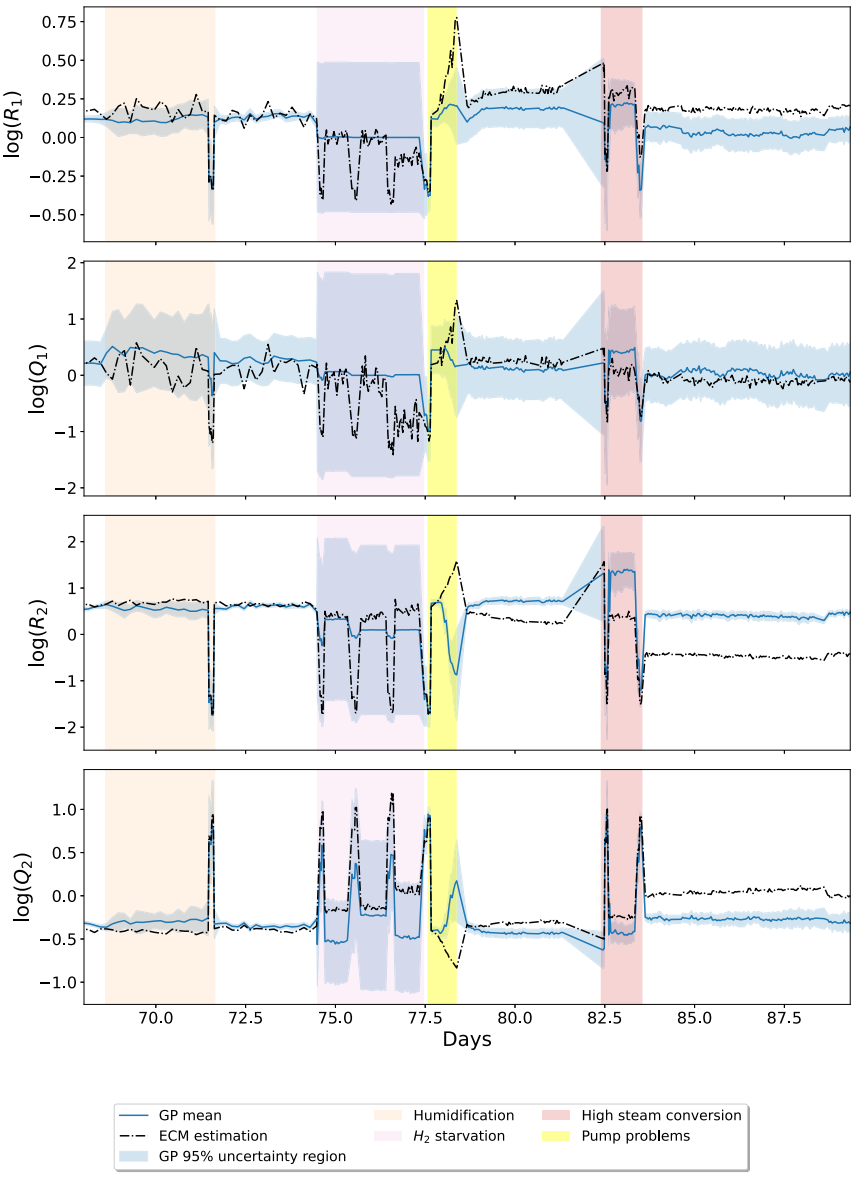


Fig. 9. Close up of the GP predictions for the logarithm of ECM parameters in the test dataset. Deviations between predicted and reference behaviour are clearly visible during fault conditions.

Table 2
Quantitative performance metrics for the GP model. Wasserstein distances are reported between the GP-predicted and VB-derived distributions, while rRMSE is computed between the GP predictive means and the Corresponding VB reference means.

Metric	Parameter	Training	Validation	Test
Wasserstein	R_1	0.0357	0.0779	0.1715
	Q_1	0.2628	0.3419	0.5090
	R_2	0.0903	0.1606	0.7420
	Q_2	0.0310	0.0532	0.2900
	Mean over 4 parameters	0.1049	0.1584	0.4281
rRMSE	R_1	0.1591	0.5886	0.7861
	Q_1	0.2159	0.4002	0.9121
	R_2	0.1346	0.2690	1.0465
	Q_2	0.0774	0.1898	0.7943
	Mean over 4 parameters	0.1468	0.3619	0.8848

sample to one of five classes: No Fault, or Faults 1–4 corresponding to the induced fault scenarios.

The overall accuracy reached 96%. Nominal data were classified with 99% precision and 95% recall. Faults 2 and 4 were identified perfectly, while Fault 1 showed minor confusion with nominal samples. Fault 3 was the most challenging due to partial overlap of its WD signature with borderline nominal behaviour.

The classification report is given in Table 4. It is evident that some faults were heavily underrepresented, therefore, we applied the synthetic minority oversampling technique (SMOTE) [51] to ensure a more balanced representation of all fault classes.

Fig. 13 shows the confusion matrix obtained for the multiclass isolation task. The matrix highlights that the classifier achieves strong separability across most fault types, with the majority of samples lying on the diagonal. Only a small number of misclassifications occur, primarily between the No-Fault class and Fault 1 or Fault 3. These errors

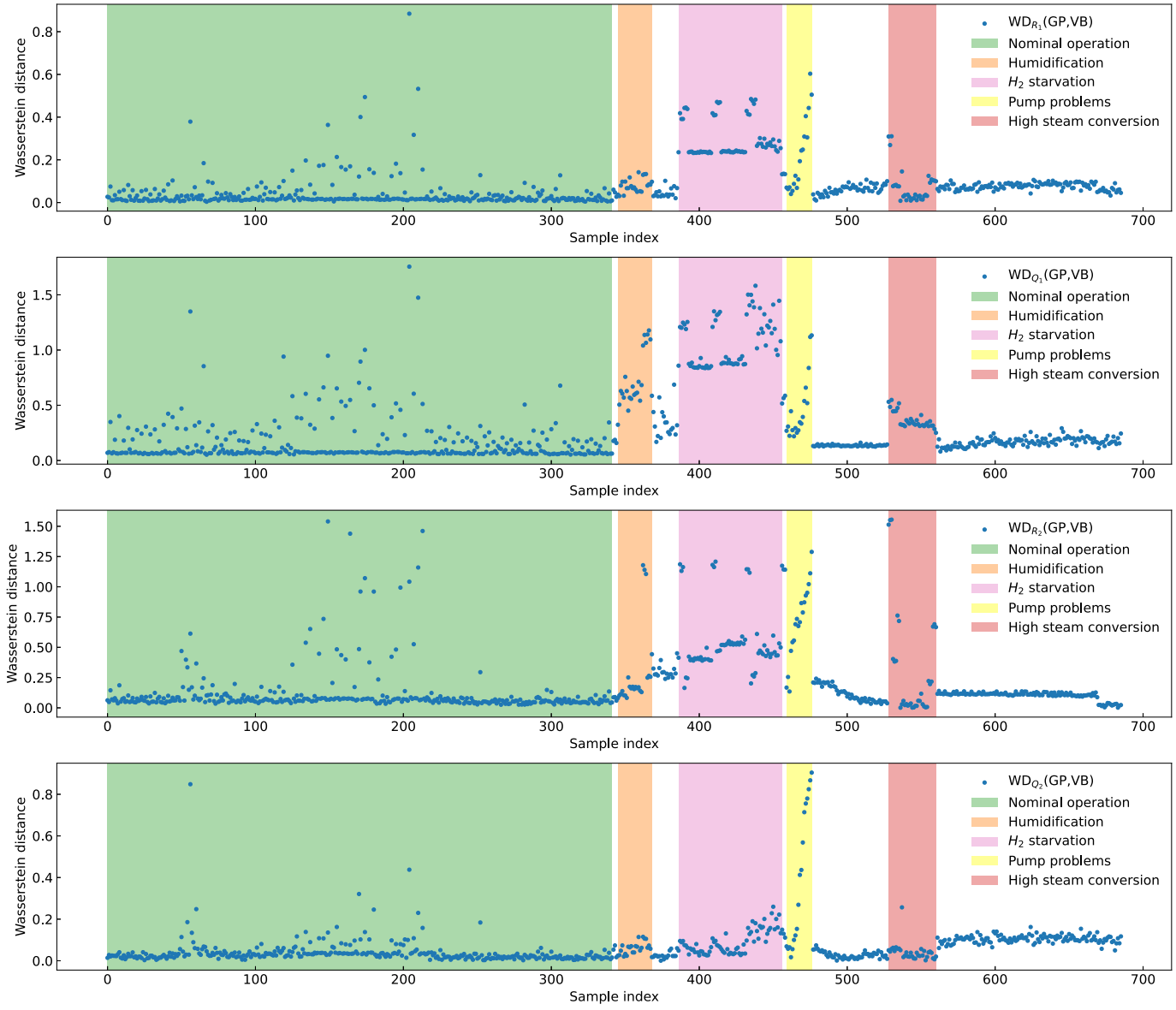


Fig. 10. Squared Wasserstein distances (blue) for all ECM parameters across the entire experiment. Periods Corresponding to induced faults are marked, demonstrating clear separation between nominal and faulty operation.

Table 3
Classification report for fault detection.

Class	Precision	Recall	F1-Score	Support
No-Fault	0.97	0.97	0.97	76
Fault	0.97	0.97	0.97	61
Accuracy	0.97			137

are consistent with the partial overlap observed in the corresponding Wasserstein-residual patterns. Overall, the structure of the confusion matrix demonstrates that the classifier preserves high accuracy even when the different fault signatures exhibit subtle differences.

Fig. 14 provides further insight into the classifier behaviour by plotting the calibrated class probabilities obtained through Platt scaling. For each sample in the test set, the figure illustrates how the classifier distributes probability mass across the five possible classes. During fault intervals, the probability mass concentrates sharply around the

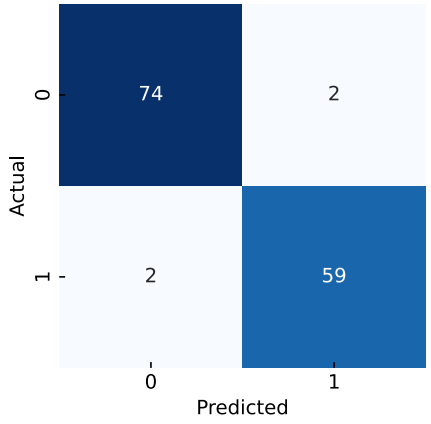


Fig. 11. Confusion matrix for the fault detection task.

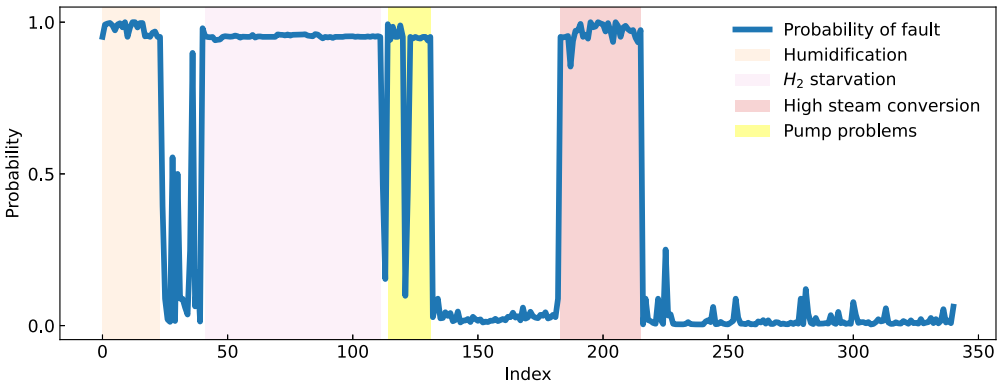


Fig. 12. Probability of fault occurrence for each sample in the test dataset.

Table 4
Classification report for fault isolation.

Class	Precision	Recall	F1-Score	Support
No-Fault	0.99	0.95	0.97	76
Fault 1	0.89	1.00	0.94	8
Fault 2	0.97	1.00	0.98	31
Fault 3	0.75	0.86	0.80	7
Fault 4	1.00	1.00	1.00	15
Accuracy	0.96			137

corresponding fault type, indicating strong discriminative confidence. Conversely, during nominal operation, the No-Fault probability remains dominant and stable. This probabilistic view confirms that the classifier does not rely on hard thresholds but instead produces meaningful confidence scores that track the underlying system behaviour.

Finally, Fig. 15 presents the isolation results over the entire experimental dataset, including training, validation, and test sequences. This timeline-level view allows direct comparison between predicted fault classes, induced operational disturbances, and the corresponding Wasserstein-residual trajectories. The figure clearly shows that the classifier assigns the correct class labels consistently across extended periods, with only short-lived misclassifications occurring near transitions or during abrupt residual spikes caused by measurement noise. Importantly, all induced faults appear distinctly separated and correctly isolated, while the spontaneous fault is also captured without prior explicit training. This confirms the robustness of the proposed framework when applied to long-duration and varying SOEC operation.

6.5. Summary of experimental results

The experimental evaluation confirms that the proposed diagnostic framework performs reliably across varying operating conditions and multiple fault scenarios. The GP models effectively normalise operating-condition variability, the squared WD provides a consistent and informative residual, and the SVC models deliver accurate detection and isolation across both induced and spontaneous faults.

Overall, the results demonstrate that the proposed approach:

- reduces sensitivity of ECM parameters to operating-condition drift;
- provides robust and interpretable residuals via WD suitable for early anomaly detection;
- achieves 97% accuracy in fault detection and 96% accuracy in multiclass fault isolation;
- and offers reliable performance suitable for real-time implementation in SOEC monitoring applications.

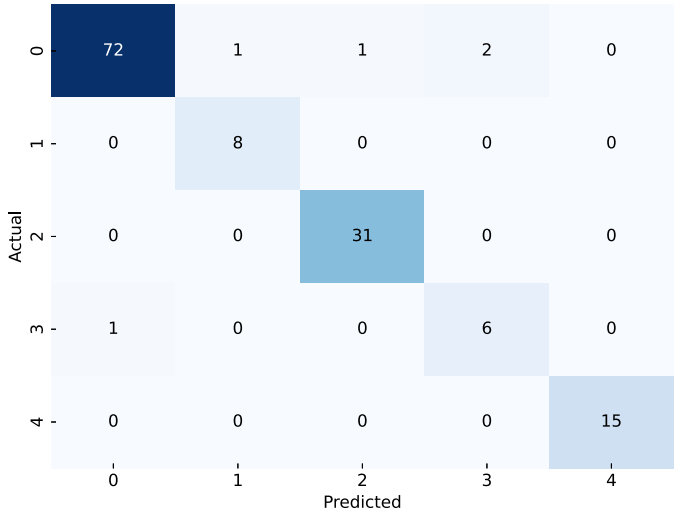


Fig. 13. Confusion matrix for multiclass fault isolation.

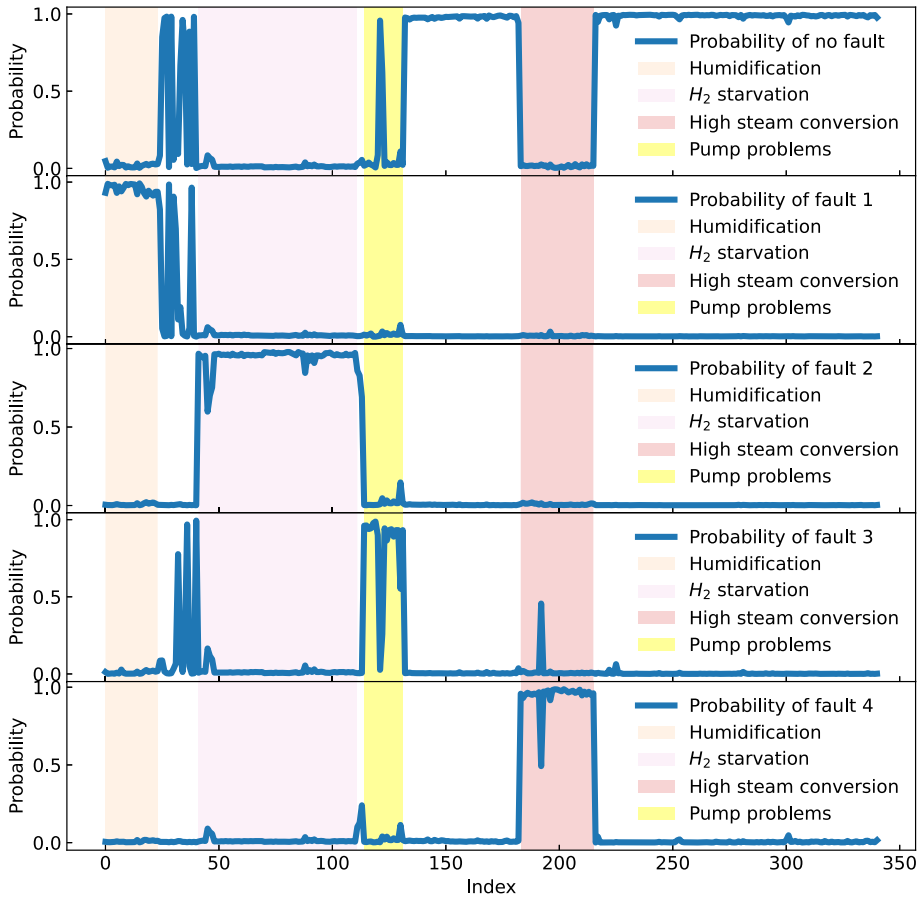


Fig. 14. Calibrated class probabilities for fault isolation in the test dataset.

7. Conclusions

We presented a diagnostic framework for SOEC systems operating under variable operating conditions, a case that remains insufficiently addressed in existing SOEC diagnostic studies. The framework combines four main steps: (i) probabilistic VB deconvolution of EIS, (ii) GP modelling of nominal ECM parameters as functions of process variables, (iii) feature extraction using WD-based residuals, and (iv) fault detection and isolation using a calibrated SVC.

The framework was validated on experimental data from a 6-cell short-stack SOEC collected during a 90-day campaign with more than 600 EIS measurements. On the test set, it achieved 97% accuracy in fault detection and 96% accuracy in multiclass fault isolation. For fault detection, both the nominal and faulty classes reached precision, recall, and F1-scores of 0.97. In the isolation task, Faults 2 and 4 were identified with perfect recall, while the most challenging class, Fault 3, still reached a recall of 0.86. These results confirm that the proposed framework can reliably distinguish healthy from faulty behaviour and isolate specific fault modes even when the operating point varies significantly.

An important advantage of the method is its low computational demand and short training time, making it suitable for deployment on embedded monitoring platforms. The GP models are trained offline on a limited set of nominal data, while online inference is computationally lightweight and negligible compared to the EIS acquisition and processing time, enabling real-time deployment. An important limitation of the present work lies in its reliance on a representative training dataset. The experimental campaign did not encompass the full range of possible degradation modes or all nominal operating conditions that may arise in real applications. Consequently, future campaigns should be designed with explicit mapping of both nominal and faulty behaviour in mind. Future work will therefore focus on evolving diagnostic algorithms that can update themselves as new data become available and thereby mitigate performance degradation caused by slow device ageing.

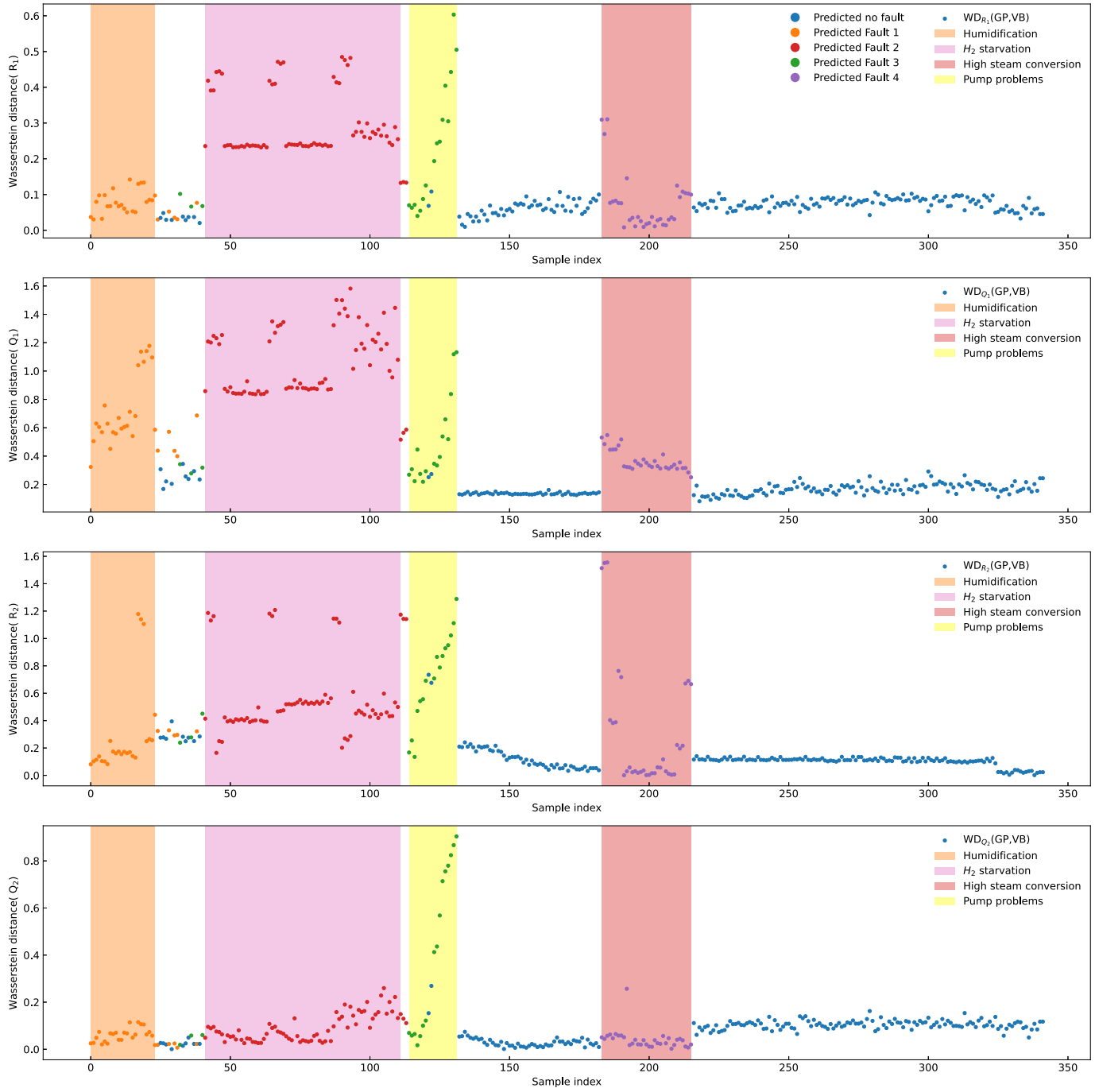


Fig. 15. Fault isolation results over the entire dataset used for SVC-based fault isolation, including samples from the learning, validation, and test subsets. Different coloured dots represent different faults, while the blue dots represent no-fault nominal operation.

CRediT authorship contribution statement

Luka Žnidarič: Writing – review & editing, Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Đani Juričić:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Bertrand Morel:** Writing – review & editing, Resources, Investigation. **Žiga Gradišar:** Writing – review & editing, Software, Methodology, Investigation, Conceptualization.

Statement

During the preparation of this work, the author(s) used the Grammarly tool to check grammar, spelling, and improve writing style. For coding purposes, the ChatGPT tool was also used to streamline the process and assist with documentation for the code attached to this journal. After using these tools/services, the author(s) reviewed and edited the content as needed and take full responsibility for the final published article.

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.

Acknowledgments

The authors would like to express their gratitude to Julien Manzi for his care about the experimental setup. This project has received funding in part from the Slovenian Research and Innovation Agency through the Research Programme P2–0001 and from the Fuel Cells and Hydrogen 2 Joint Undertaking (now Clean Hydrogen Partnership) under Grant Agreement No. 101007175. This Joint Undertaking receives support from the European Union's Horizon 2020 Research and Innovation programme, Hydrogen Europe and Hydrogen Europe Research. The authors gratefully acknowledge the anonymous reviewers for their careful reading of the manuscript and their valuable comments and suggestions.

Appendix A. VB for EIS/ECM parameter estimation

A.1. Bayesian formulation

Let θ collect the ECM parameters and $\mathcal{D} = \{Z(j\omega_k)\}_{k=1}^{N}$ the measured impedance data. The Bayesian posterior is

$$p(\theta \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \theta)\, p(\theta)}{p(\mathcal{D})}, \qquad p(\mathcal{D}) = \int p(\mathcal{D} \mid \theta)\, p(\theta)\, d\theta.$$

In our case, obtaining $p(\mathcal{D})$ directly is intractable.

A.2. Variational approximation

Choose a tractable family $\{q_\phi(\theta)\}$ (log-normal and beta in our case) and minimise the Kullback-Leibler divergence $\mathrm{KL}(q_\phi \| \; p(\theta \mid \mathcal{D}))$ or equivalently maximise the evidence lower bound (ELBO) [52]

$$\mathrm{ELBO}(\phi) = \mathbb{E}_{q_\phi}\big[\log p(\mathcal{D} \mid \theta)\big] - \mathrm{KL}\big(q_\phi(\theta) \,\|\, p(\theta)\big).$$

By minimisation we obtain optimal $q_\phi(\theta)$, which we take as the posterior; stopping is based on ELBO convergence.

A.3. Mean-field and physically consistent variational families

A mean-field factorisation $q_\phi(\theta) = \prod_i q_{\phi_i}(\theta_i)$ is adopted. To respect physical constraints:

- resistive and capacitive terms R_i, Q_i use log-normal variational factors;
- fractional exponents $\alpha_i \in (0, 1]$ use beta variational factors.

A.4. Posterior summaries and uncertainty

For each EIS session, the optimised $q_\phi(\theta)$ provides a posterior, which is put through a logarithm and then compared to the nominal GP predictions at the current operating point (Section 4).

Appendix B. GP model details

B.1. Kernel functions

GP defines a distribution over functions, which is defined through its mean function (here assumed to be zero) and its kernel function. For latter, common choices include:

- squared exponential kernel (RBF)

$$k_{\mathrm{RBF}}(x, x') = \sigma_f^2 \exp\left(-\frac{(x - x')^2}{2\, l^2}\right),$$

  where σ_f^2 is the signal variance and l is the characteristic length scale;
- Matérn kernel

$$k_{\mathrm{Matérn}}(x, x') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\,|x - x'|}{l}\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}\,|x - x'|}{l}\right),$$

  where ν controls smoothness, l is the length scale, σ_f^2 is the signal variance and $K_\nu(\cdot)$ is the modified Bessel function of the second kind;
- periodic kernel

$$k_{\mathrm{per}}(x, x') = \sigma_f^2 \exp\left(-\frac{2 \sin^2\big(\pi |x - x'|/p\big)}{l^2}\right),$$

  with period p, signal variance σ_f^2 and length scale l.

B.2. GP model (Kernel) selection via BIC

BIC was used to identify the most suitable kernel for the GP model

$$\mathrm{BIC} = -2 \log p(\mathbf{y} \mid \mathbf{X}, \theta) + c \log(n),$$

where c is the number of learned hyperparameters, and n is the number of training samples. Lower BIC values indicate a better balance between model accuracy and complexity.

Ten candidate models were evaluated, each with the general kernel structure

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \prod_{i=1}^{19} k_i(x_i, x_i'), \tag{B.1}$$

where ks are the same kernels (each for one input variable), but with different hyperparameters. The best performing model was made from Matérn 3/2 kernels, which achieved the lowest BIC value (BIC = 109.135), indicating the best trade-off between predictive performance and complexity. Other kernel candidates (RBF, linear, periodic, etc.) yielded higher BIC values, as summarised in Table B.5. The structure of the selected GP structure for the 19-dimensional input data is therefore

$$k_{\mathrm{Matérn\text{-}3/2}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \prod_{i=1}^{19} \left(1 + \frac{\sqrt{3}}{\ell_i}|x_i - x_i'|\right) \exp\left(-\frac{\sqrt{3}}{\ell_i}|x_i - x_i'|\right), \tag{B.2}$$

where l_i is the length scale for input i.

As mentioned, no input variable was deemed unimportant by hand. Since the structure (B.2) includes lengthscales l_i as hyperparameters, the optimiser can automatically exclude any irrelevant input by inflating them, see [38] for extra information on so called automatic relevance determination (ARD).

Table B.5
BIC scores obtained on the validation dataset for different kernel functions, ordered by performance.

Kernel Function	Parameters	BIC Score
Matern ($\nu = \frac{3}{2}$) + White noise	20	**109.135**
Matern ($\nu = \frac{5}{2}$) + White noise	20	110.064
Linear kernel + Matern ($\nu = \frac{5}{2}$)	20	195.925
Matern ($\nu = \frac{1}{2}$)	19	197.037
Linear kernel + Matern ($\nu = \frac{3}{2}$)	20	201.292
Linear kernel + RBF kernel	20	218.830
Matern ($\nu = \frac{3}{2}$)	19	273.772
Matern ($\nu = \frac{5}{2}$)	19	306.879
Squared Exponential (RBF) kernel	19	339.940
Periodic kernel	2	349.088
Linear kernel	1	5124.235

Appendix C. Wasserstein distance

WD, also referred to as the "earth mover's distance" [53], measures the dissimilarity between two probability distributions, and can be interpreted as the minimum cost of transportation required to transform one distribution into another.

Specifically, WD between distributions α and β is given by Séjourné et al. [54]

$$W_p^p(\alpha,\beta) = \min_{\pi\in\mathcal{U}(\alpha,\beta)} \int_{X^2} d^p(x,y)\, d\pi(x,y), \tag{C.1}$$

where $d(x,y)$ denotes the cost of transportation (generally the Euclidean distance between the transportation sites), p is the order of the method, and $\mathcal{U}(\alpha,\beta)$ denotes all possible transport plans from α to β.

(C.1) can be solved using a linear programming algorithm, like Simplex. However, for $p = 2$ there is a closed-form solution available if α and β are both normal distributions, see (7) for the definition.

WD was chosen over other distance measures as it provides a robust and truly metric distance, see [[55], p.11] for a discussion.

Appendix D. SVC

In this work, we employ SVC to discriminate between nominal and faulty operating conditions, as well as among specific fault modes, using features extracted from electrochemical diagnostics (e.g., EIS/ECM-derived features). SVC learns a separating surface with maximum margin to promote generalisation [56,57].

D.1. Decision boundaries and calibration

Let $\mathbf{z} \in \mathbb{R}^m$ denote a feature vector. SVC defines a decision function

$$f(\mathbf{z}) = \mathbf{w}^\top \varphi(\mathbf{z}) + b, \tag{D.1}$$

where $\varphi(\cdot)$ is an implicit feature map induced by a positive-definite kernel $k(\mathbf{z},\mathbf{z}') = \langle\varphi(\mathbf{z}),\varphi(\mathbf{z}')\rangle$ (e.g., RBF), $\mathbf{w}$ is the normal vector of the separating hyperplane, and b is the bias term. A hard yes/no decision follows from $\operatorname{sign}\big(f(\mathbf{z})\big)$, but probabilistic outputs are preferable for uncertainty-aware diagnosis and decision support.

Platt scaling (logistic calibration). We converted margins $f(\mathbf{z})$ into calibrated posteriors via a sigmoid (logistic) mapping [27,58]

$$\hat{p}(y=1 \mid \mathbf{z}) = \frac{1}{1+\exp\big(A\, f(\mathbf{z}) + B\big)}, \tag{D.2}$$

with parameters $A, B \in \mathbb{R}$ found by minimizing the regularised negative log-likelihood on a held-out validation set (or via cross-validation)

$$\min_{A,B} -\sum_{i=1}^{N_{\text{val}}} \Big[t_i \log\sigma\big(A f(\mathbf{z}_i)+B\big) + (1-t_i)\log\big(1-\sigma\big(A f(\mathbf{z}_i)+B\big)\big)\Big], \tag{D.3}$$

where $\sigma(u) = \frac{1}{1+e^{-u}}$. Following [27], the binary targets t_i were replaced by smoothed labels to avoid degenerate solutions

$$t_i = \begin{cases} \frac{N_+ +1}{N_+ +2}, & \text{if } y_i = +1, \\ \frac{1}{N_- +2}, & \text{if } y_i = -1, \end{cases} \tag{D.4}$$

where N_+ and N_- are the counts of positive and negative samples in the validation set. For multiclass problems, pairwise one-vs-one SVCs have been calibrated individually and combined via probability coupling [59].

Calibration quality can be assessed with reliability diagrams and proper scoring rules (e.g., Brier score, log-loss), ensuring probabilities are well matched to empirical frequencies [58,60].

D.2. Confusion matrix interpretation

Performance of the proposed fault isolation pipeline is summarised by the confusion matrix $\mathbf{C} \in \mathbb{N}^{K\times K}$, whose entries are given by

$$[\mathbf{C}]_{i,j} = \#\{\text{samples with true class } i \text{ predicted as class } j\}, \tag{D.5}$$

where diagonal entries quantify correct classifications, and off-diagonals indicate specific confusion between classes.

From $\mathbf{C}$ one derives evaluation metrics for each class k:

- recall = $\mathrm{TP}_k/(\mathrm{TP}_k + \mathrm{FN}_k)$, where TP denotes true positives and FN denotes false negatives;
- precision = $\mathrm{TP}_k/(\mathrm{TP}_k + \mathrm{FP}_k)$, where FP denotes false positives;
- accuracy score for imbalanced sets F1 = $2\frac{\text{precision}\times\text{recall}}{\text{precision}+\text{recall}}$;
- support = how many instances of each class were present in the true labels;

In the diagnostic context, these metrics indicate the reliability of:

- fault detection (nominal vs. faulty classes);
- fault isolation (assignment to specific faulty class);
- risk assessment (especially when combined with Platt posteriors from (D.2)–(D.3)).

Appendix E. Identification of SOEC

To successfully perform an EIS measurement the investigated system must be linear, causal, time-invariant [[61], p. 13-14] and stable [[61], p. 21]. Linearity was achieved by limiting the amplitude of the perturbation, resulting in an approximately linear response. The data acquisition (DAQ) equipment was properly calibrated, so causality was not broken due to mistimed samplings of the voltage and current, for example. Control elements of SOEC and balance-of-plant (BoP) ensured that the device was approximately time-invariant and stable during the EIS session. Any change in operating point was followed by a brief period (less than 10 minutes) of transient behaviour, which was waited out before continuing with the experiment.

To ensure the above conditions were met, the system was also tested using the Kramers-Kronig relations [62]. Furthermore, successful fitting of ECMs to the obtained data confirmed that the measurements were valid [63].

For perturbation DRBS was used [17]. It is a simple signal which randomly switches between values of $\pm a$. Moreover, it is a wide-sense random stationary signal designed to match the statistical properties of white noise [[64], p. 826]. Its power spectral density is given by [Isermann et al. [65] p. 163]

$$\Phi(\omega) = a^2\lambda \left| \frac{\sin\left(\frac{\omega\lambda}{2}\right)}{\frac{\omega\lambda}{2}} \right|^2, \tag{E.1}$$

where λ denotes the minimal time between two switches. DRBS can be thought of as white noise filtered by a low-pass filter, with its effective bandwidth given by $f_\mathrm{B} = \frac{1}{3\lambda}$ [[66], p. 68].

For the purpose of this work, six different discrete random binary signals (DRBSs) were used. They provided information on six partitioned spectrum bands:

- 0.1 – 1 Hz (duration 200s, f_B = 1Hz);
- 1 – 10 Hz (duration 30s, f_B = 10Hz);
- 10 – 100 Hz (duration 3.3s, f_B = 100Hz);
- 100 – 1000 Hz (duration 0.33s, f_B = 1000Hz);
- 1000 – 6000 Hz (duration 0.060s, f_B = 6000Hz);
- 1000 – 10,000 Hz (duration 0.035s, f_B = 10000Hz).

Six different input signals were needed due to memory limitations of the DAQ equipment. Thus, each signal could have a custom sampling rate, tailored to both its frequency range and to avoid memory limitations. The total duration of a full EIS sequence is on the order of a few minutes, which is significantly shorter than the characteristic time scales of the thermal and flow dynamics of the SOEC system, typically ranging from several minutes to hours. During this time, the process variables remain nearly constant, exhibiting only minor fluctuations without observable drift. This justifies the assumption that the operating conditions can be considered quasi-stationary during the EIS measurement. An example of the measured signals during an EIS sequence is shown in Fig. E.16.

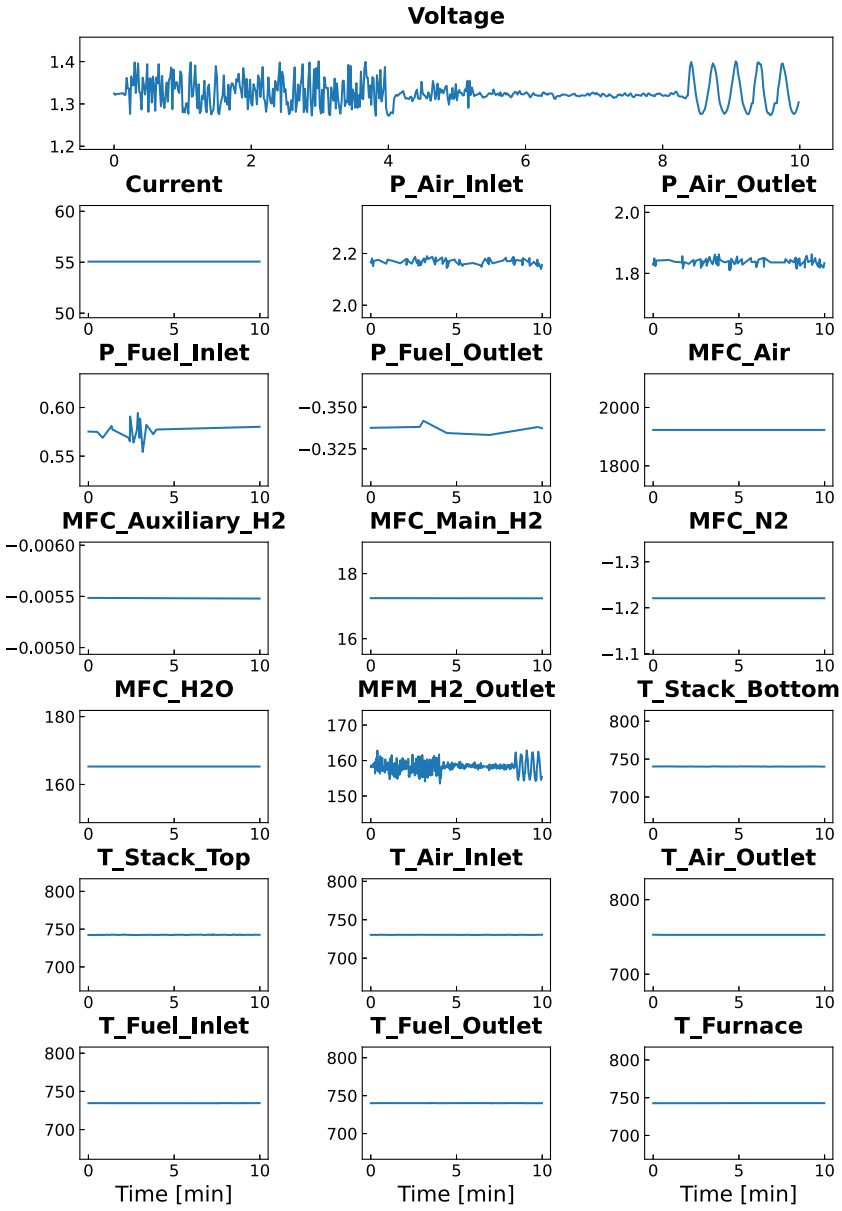


Fig. E.16. Time-domain signals recorded during a representative EIS sequence. The applied DRBSs excite the system across multiple frequency bands, while the process variables remain nearly constant, supporting the assumption of quasi-stationary operating conditions during the measurement.

Data availability

Data will be made available on request.

References

[1] Laguna-Bercero M. Recent advances in high temperature electrolysis using solid oxide fuel cells: a review. J Power Sources 2012;203:4–16.

[2] Mougin J. Hydrogen production by high-temperature steam electrolysis. Springer; 2015. 225–53. https://doi.org/10.1016/B978-1-78242-361-4.00008-X

[3] Moçoteguy P, Brisse A. A review and comprehensive analysis of degradation mechanisms of solid oxide electrolysis cells. Int J Hydrog Energy 2013;38(36):15887–902. https://doi.org/10.1016/j.ijhydene.2013.09.045

[4] Chen W, Sun C. Recent advances in high temperature solid oxide electrolytic cells. Energy Materials 2025;5(5). https://doi.org/10.20517/energymater.2024.144. https://www.oaepublish.com/articles/energymater.2024.144.

[5] McPhail SJ, Frangini S, Laurencin J, Effori E, Abaza A, Padinjarethil AK, Hagen A, Léon A, Brisse A, Vladikova D, Burdin B, Bianchi FR, Bosio B, Piccardo P, Spotorno R, Uchida H, Polverino P, Adinolfi EA, Postiglione F, Lee J-H, Moussaoui H, Van herle J. Addressing planar solid oxide cell degradation mechanisms: a critical review of selected components. Electrochemical Science Advances 2022;2(5):e2100024. https://doi.org/10.1002/elsa.202100024. https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/elsa.202100024.

[6] Hauch A, Ploner A, Pylypko S, Cubizolles G, Mougin J. Test and characterization of reversible solid oxide cells and stacks for innovative renewable energy storage. Fuel Cells 2021;21(5):467–76. https://doi.org/10.1002/fuce.202100046

[7] Fan G, Chen H, Wu T, Wang L, Xu X. Towards longevity in solid oxide electrolysis cells: multi-scale modeling and machine learning for degradation diagnosis and mitigation. J Mater Chem A Jan 2025;13. https://doi.org/10.1039/D5TA03711E

[8] Zhou J, Zhu Y, Shao Z. A combined passive-active method for diagnosing multiplicative fault. Process Saf Environ Prot 2023;178:959–75. https://doi.org/10.1016/j.psep.2023.08.045. https://www.sciencedirect.com/science/article/pii/S0957582023007401.

[9] Barelli L, Barluzzi E, Bidini G. Diagnosis methodology and technique for solid oxide fuel cells: a review. Int J Hydrog Energy 2013;38(12):5060–74. https://doi.org/10.1016/j.ijhydene.2013.02.024. https://www.sciencedirect.com/science/article/pii/S036031991300390X.

[10] Yang B, Guo Z, Wang J, Wang J, Zhu T, Shu H, Qiu G, Chen J, Zhang J. Solid oxide fuel cell systems fault diagnosis: critical summarization, classification, and perspectives. J Energy Storage 2021;34:102153. https://doi.org/10.1016/j.est.2020.102153. https://www.sciencedirect.com/science/article/pii/S2352152X20319794.

[11] Franić N, Pivac I, Barbir F. A review of machine learning applications in hydrogen electrochemical devices. Int J Hydrog Energy 2025;102:523–44. https://doi.org/10.1016/j.ijhydene.2025.01.070. https://www.sciencedirect.com/science/article/pii/S0360319925000758.

[12] Gallo M, Polverino P, Mougin J, Morel B, Pianese C. Coupling electrochemical impedance spectroscopy and model-based aging estimation for solid oxide fuel cell stacks lifetime prediction. Appl Energy 2020;279:115718. https://doi.org/10.1016/j.apenergy.2020.115718. https://www.sciencedirect.com/science/article/pii/S0306261920312113.

[13] Yazbeck Z, Bribiesca-Argomedo F, Pham MT, Morel B, Dimitriou V. System-level monitoring and diagnosis of starvation faults in solid oxide electrolyzers. Int J Hydrog Energy 2026;198:152808. https://doi.org/10.1016/j.ijhydene.2025.152808. https://www.sciencedirect.com/science/article/pii/S0360319925058112.

[14] Huang Q-A, Hui SR, Wang B, Zhang J. A review of AC impedance modeling and validation in SOFC diagnosis. Electrochim Acta 2007;52:8144–64. https://doi.org/10.1016/j.electacta.2007.05.071

[15] Nechache A, Cassir M, Ringuedé A. Solid oxide electrolysis cell analysis by means of electrochemical impedance spectroscopy: a review. J Power Sources 2014;258:164–81. https://doi.org/10.1016/j.jpowsour.2014.01.110

[16] Leonide A. SOFC modelling and parameter identification by means of impedance spectroscopy [Ph.D. thesis]. 2010. https://doi.org/10.5445/KSP/1000019173

[17] Boškoski P, Debenjak A, Mileva Boshkoska B. Fast electrochemical impedance spectroscopy. Springer International Publishing; 2017. https://doi.org/10.1007/978-3-319-53390-2

[18] Königshofer B, Höber M, Boškoski P, Nusev G, Juričic Đ, Hochenauer C, Subotic V. Performance investigation and optimization of an SOEC stack operated under industrially relevant conditions. In: 17th international symposium on solid oxide fuel cells, SOFC 2021, 1st edition, no. 1 in ECS transactions. United Kingdom: IOP Publishing Ltd.; 2021. p. 519–28. https://doi.org/10.1149/10301.0519ecst

[19] Rezaei Niya SM, Hoorfar M. Study of proton exchange membrane fuel cells using electrochemical impedance spectroscopy technique – a review. J Power Sources 2013;240:281–93. https://doi.org/10.1016/j.jpowsour.2013.04.011. https://www.sciencedirect.com/science/article/pii/S0378775313006022.

[20] Zhang X, Jiang Y, Huang L, Chen W, Brett D. Inconsistent responses of cells on operating conditions in a 5 kw proton exchange membrane fuel cell stack. Electrochim Acta 2021;391:138925. https://doi.org/10.1016/j.electacta.2021.138925. https://www.sciencedirect.com/science/article/pii/S0013468621012159.

[21] Dolenc B, Nusev G, Boškoski P, Morel B, Mougin J, Juričić Đ. Probabilistic deconvolution of solid oxide fuel cell impedance spectra. Elecetrimacs 2019 Salerno 2019:53–4. http://www.electrimacs2019.unisa.it/files/ELECTRIMACS2019_Booklet_20190514.pdf.

[22] Cho H, Hong C, Hong D, Oh S-K, Kim Y. Thermal equivalent circuit model and parameter estimation for high-capacity li-ion cell. J Electrochem Soc 2023;170(8):080520.

[23] Martín IS, Ursúa A, Sanchis P. Modelling of PEM fuel cell performance: steady-state and dynamic experimental validation. Energies 2014;7(2):670–700.

[24] Maradesa A, Py B, Ciucci F. Probabilistic deconvolution of the distribution of relaxation times from multiple electrochemical impedance spectra. J Power Sources 2024;621:235236. https://doi.org/10.1016/j.jpowsour.2024.235236. https://www.sciencedirect.com/science/article/pii/S0378775324011881.

[25] Žnidarič L, Nusev G, Morel B, Mougin J, Juričić Đ, Boškoski P. Evaluating uncertainties in electrochemical impedance spectra of solid oxide fuel cells. Appl Energy 2021;298:117101.

[26] Kocijan J. Modelling and control of dynamic systems using Gaussian process models. Springer; 2016. https://doi.org/10.1007/978-3-319-21021-6

[27] Platt J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Smola AJ, Bartlett P, Schölkopf B, Schuurmans D, editors. Advances in large margin classifiers. Cambridge: MIT Press; 2000. p. 61–74.

[28] Costamagna P, De Giorgi A, Magistri L, Moser G, Pellaco L, Trucco A. A classification approach for model-based fault diagnosis in power generation systems based on solid oxide fuel cells. IEEE Trans Energy Convers 2015;30:1–12. https://doi.org/10.1109/TEC.2015.2492938

[29] Chanal D, Steiner NY, Chamagne D, Pera M-C. Lt-pem fuel cells diagnosis based on EIS, clustering, and automatic parameter selection. IEEE Trans Veh Technol 2025;74(2):2513–26. https://doi.org/10.1109/TVT.2023.3273084

[30] Lasia A. Electrochemical impedance spectroscopy and its applications, book. Springer; 2014. https://doi.org/10.1007/978-1-4614-8933-7

[31] Leonide A, Apel Y, Ivers-Tiffee E. SOFC modeling and parameter identification by means of impedance spectroscopy. ECS Trans 2009;19(20):81. https://doi.org/10.1149/1.3247567

[32] Fleischer C, Waag W, Heyn H-M, Sauer DU. On-line adaptive battery impedance parameter and state estimation considering physical principles in reduced order equivalent circuit battery models: part 1. Requirements, critical review of methods and modeling. J Power Sources 2014;260:276–91. https://doi.org/10.1016/j.jpowsour.2014.01.129. https://www.sciencedirect.com/science/article/pii/S0378775314002249.

[33] Zhang R, Sur D, Li K, Witt J, Black R, Whittingham A, Scully JR, Hattrick-Simpers J. Bayesian assessment of commonly used equivalent circuit models for corrosion analysis in electrochemical impedance spectroscopy. NPJ Mater Degrad 2024;8(1):120. https://doi.org/10.1038/s41529-024-00537-8

[34] Kingma DP, Ba J. Adam: A method for stochastic optimization. [arXiv preprint]. 2017 arXiv:1412.6980.

[35] Stoica P, Selen Y. Model-order selection: a review of information criterion rules. IEEE Signal Process Mag 2004;21(4):36–47. https://doi.org/10.1109/MSP.2004.1311138

[36] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Köpf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S. Pytorch: An imperative style, high-performance deep learning library. [arXiv preprint]. 2019 arXiv:1912.01703.

[37] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825–30.

[38] Sebenius I, Paananen T, Vehtari A. feature collapsing for Gaussian process variable ranking. In: Camps-Valls G, Ruiz FJR, Valera I, editors. Proceedings of the 25th international conference on artificial intelligence and statistics, vol. 151 of proceedings of machine learning research. PMLR; 2022. p. 11341–55. https://proceedings.mlr.press/v151/sebenius22a.html.

[39] Qiu P, Li C, Liu B, Yan D, Li J, Jia L. Materials of solid oxide electrolysis cells for h2o and co2 electrolysis: a review. J Adv Ceram 2023;12(8):1463–510. https://doi.org/10.26599/JAC.2023.9220767. https://www.sciopen.com/article/10.26599/JAC.2023.9220767.

[40] Shao X, Budiman RA, Sato T, Yamaguchi M, Kawada T, Yashiro K. Review of factors affecting the performance degradation of ni-ysz fuel electrodes in solid oxide electrolyzer cells. J Power Sources 2024;609:234651. https://doi.org/10.1016/j.jpowsour.2024.234651. https://www.sciencedirect.com/science/article/pii/S0378775324006037.

[41] Sassone G, Celikbilek O, Hubert M, Develos-Bagarinao K, Benayad A, Morel B, Léon A, Laurencin J. Impact of AIR moisture and operating mode on the degradation of solid oxide cells. J Electrochem Soc 2024;171(12):124507.

[42] Liu R, Kim SH, Taniguchi S, Oshima T, Shiratori Y, Ito K, Sasaki K. Influence of water vapor on long-term performance and accelerated degradation of solid oxide fuel cell cathodes. J Power Sources 2011;196(17):7090–6, proceedings of 2010 European Solid Oxide Fuel Cell Forum. doi:https://doi.org/10.1016/j.jpowsour.2010.08.014. https://www.sciencedirect.com/science/article/pii/S0378775310013327.

[43] Kim M, Muroyama H, Matsui T, Eguchi K. Influence of water vapor on performance degradation and microstructural change of (LA, SR)(CO, FE)o3-δ cathode. J Electrochem Soc 2019;166(16):F1269. https://doi.org/10.1149/2.0641915jes

[44] Liu S-S, Develos-Bagarinao K, Budiman RA, Ishiyama T, Kishimoto H, Yamaji K. Towards an atomic scale understanding of the early-stage deterioration mechanism of LSCF. J Mater Chem A 2023;11(40):21983–2000.

[45] Zhang W, Barfod R. Investigation of degradation mechanisms of LSCF based SOFC cathodes — by calphad modeling and experiments. Research Portal Denmark 2012:231.

[46] Zhao Z, Liu L, Zhang X, Wu W, Tu B, Cui D, Ou D, Cheng M. High- and low- temperature behaviors of la0.6sr0.4co0.2fe0.8o3-δ cathode operating under co2/h2o-containing atmosphere. Int J Hydrog Energy 2013;38(35):15361–70. https://doi.org/10.1016/j.ijhydene.2013.09.089. https://www.sciencedirect.com/science/article/pii/S0360319913023124.

[47] Subotić V, Königshofer B, Juričić Đ, Kusnezoff M, Schröttner H, Hochenauer C, Boškoski P. Detailed insight into processes of reversible solid oxide cells and stacks using drt analysis. Energy Convers Manag 2020;226:113509. https://doi.org/10.1016/j.enconman.2020.113509. https://www.sciencedirect.com/science/article/pii/S0196890420310402.

[48] Pan H, Wu A, Au SF, Yang Y, Song Z, Liu Z, Gong X, Guan W. Effect of the steam/hydrogen ratio on the performance of flat-tube solid oxide electrolysis cells for seawater. Sustainable Energy Fuels 2023;7:3333–41. https://doi.org/10.1039/D3SE00351E

[49] Shimada H, Yamaguchi T, Suzuki T, Sumi H, Hamamoto K, Fujishiro Y. High steam utilization operation with high current density in solid oxide electrolysis cells. J Ceram Soc Jpn 2016;124(3):213–7. https://doi.org/10.2109/jcersj2.15274

[50] Königshofer B, Höber M, Nusev G, Boškoski P, Hochenauer C, Subotić V. Accelerated degradation for solid oxide electrolysers: analysis and prediction of performance for varying operating environments. J Power Sources 2022;523:230982. https://doi.org/10.1016/j.jpowsour.2022.230982. https://www.sciencedirect.com/science/article/pii/S0378775322000076.

[51] Bowyer KW, Chawla NV, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. [arXiv preprint]. 2011 arXiv:1106.1813.

[52] Kingma DP, Welling M. Auto-encoding variational bayes. [arXiv preprint]. 2022 arXiv:1312.6114.

[53] Villani C. The wasserstein distances. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 93–111. https://doi.org/10.1007/978-3-540-71050-9_6

[54] Séjourné T, Peyré G, Vialard F-X. Unbalanced optimal transport, from theory to numerics. [arXiv preprint]. 2023 arXiv:2211.08775.

[55] Chen R, Paschalidis IC. Distributionally robust learning. Foundations and Trends® in Optimization 2020;4(1–2):1–243. https://doi.org/10.1561/2400000026

[56] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97. https://doi.org/10.1007/BF00994018

[57] Bishop C. Pattern recognition and machine learning, information science and statistics. Springer New York; 2016. https://books.google.si/books?id=kOXDtAEACAAJ.

[58] Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. Proceedings of ICML 2005:625–32. https://doi.org/10.1145/1102351.1102430
[59] Wu T-Y, Lin C-J, Weng RC. Probability estimates for multiclass classification by pairwise coupling. J Mach Learn Res 2004;5:975–1005.
[60] DeGroot MH, Fienberg SE. The comparison and evaluation of forecasters. Statistician 1983;32(1):12–22.
[61] Ljung L. System identification: theory for the user. 2nd edn. Upper Saddle River: Prentice-Hall; 1999.
[62] Lazanas AC, Prodromidis MI. Electrochemical impedance spectroscopy - a tutorial. ACS Meas Sci Au 2023;3(3):162–93.
[63] Boukamp BA. A linear kronig-kramers transform test for immittance data validation. J Electrochem Soc 1995;142(6):1885–94.
[64] Proakis JG. Digital signal processing: principles, algorithms, and applications, 4/e. Pearson Education India; 2007.
[65] Isermann R, Münchhof M. Identification of dynamic systems: an introduction with applications, vol. 85. Springer; 2011.
[66] Davies W. System identification for self-adaptive control; 1970.