DIAGNOSIS AND PROGNOSIS OF SOLID OXIDE CELL SYSTEMS: A STATISTICAL APPROACH

Luka Žnidarič

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Supervisor: Prof. Dr. Đani Juričić, Jožef Stefan institute, Ljubljana, Slovenia

Evaluation Board:
Assoc. Prof. Dr. Pavle Boškoski, Chair, Jožef Stefan Institute, Ljubljana, Slovenia
Prof. Dr. Juš Kocijan, Member, Jožef Stefan Institute, Ljubljana, Slovenia
Asst. Prof. Dr. Pierpaolo Polverino, Member, University of Salerno, Salerno, Italy

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL

Luka Žnidarič

DIAGNOSIS AND PROGNOSIS OF SOLID OXIDE CELL SYSTEMS: A STATISTICAL APPROACH

Doctoral Dissertation

SPREMLJANJE STANJA IN NAPOVEDOVANJE ŽIVLJENJSKE DOBE SISTEMOV TRDNO-OKSIDNIH CELIC: STATISTIČNI PRISTOP

Doktorska disertacija

Supervisor: Prof. Dr. Đani Juričić

Ljubljana, Slovenia, July 2024

Acknowledgments

I would like to express my deepest gratitude to all those who have supported and guided me throughout my PhD journey.

First and foremost, I would like to thank my girlfriend, Ana, for her unwavering love, patience, and encouragement. Your support has been a constant source of strength and motivation.

I am deeply indebted to my mentor, Đani, for his invaluable guidance, insightful advice, and continuous encouragement. Your expertise and dedication have been crucial to the completion of this work.

I would like to extend special thanks to my coworker, Pavle, whose guidance and support have been instrumental throughout my research. I am also grateful to my "roommates," Žiga, Martin and Žan, for creating a positive and collaborative office environment that has made this journey enjoyable.

I extend my heartfelt thanks to my family for their unconditional love, support, and belief in me. Your sacrifices and encouragement have made this achievement possible.

I would like to extend my appreciation to all of my coworkers at the E2 Department of the Jožef Stefan Institute. Your support, collaboration, and friendship have greatly enriched my experience and contributed to the success of this work.

I am sincerely grateful to the members of my PhD committee for their time, dedication, and valuable insights. Your thoughtful feedback, constructive criticism, and encouragement have greatly enriched this thesis. The rigour and expertise with which you approached my work have helped me refine and strengthen my research, and I am truly grateful for your role in this journey.

I wish to acknowledge the support of the Young Researchers Programme, grant no. PR-08986, and the Research Programme, no. P2-0001, of the Slovenian Research and Innovation Agency. Part of the results have been obtained in the frame of projects receiving funding from the Fuel Cells and Hydrogen 2 Joint Undertaking (now Clean Hydrogen Partnership) under Grant Agreements No. 101007175 and No. 875047.

Thank you to everyone who has supported me in various ways, both directly and indirectly. This thesis was only possible with you.

Abstract

Hydrogen is a promising energy carrier in the emerging green energy landscape. It is environmentally friendly, transportable, and storable. The Solid Oxide Cell (SOC) system is a complex technology that enables the bidirectional conversion of hydrogen to energy. The label SOC entails solid oxide fuel cell (SOFC) systems, which produce electrical energy from hydrogen, and solid oxide electrolyzer cell (SOEC) systems, which produce hydrogen from water and electricity.

Compared to other electrochemical energy conversion systems, a SOC system offers some remarkable advantages, such as flexible use of reactants and no dependence on rare materials and catalysts (since they operate at temperatures as high as $600 - 1000K$). Despite significant developments in the last two decades and rising needs for applications in green transport and smart grids, their deployment is still relatively slow. Part of the reason is still insufficient durability and reliability, which are inherently conditioned by the inevitable degradation processes. Reliable condition monitoring, prognostics, and health management contribute to better operability and durability of SOC devices. Diagnosis and prognosis, as the key functionalities, help the operators gain insight into the two critical questions: (i) What is the internal health status of the devices and components, and if there is a problem, what exactly is it about, and (ii) how long is the system expected to operate before facing the inevitable end? This thesis focuses on these two fundamental questions.

The body of research dealing with fault diagnosis and prognosis of SOC systems is relatively small compared to other electrochemical conversion systems (batteries in particular). Most works concern SOFC systems while the methodology is still emerging for SOEC systems. The approaches range from those based on first principles to entirely data-driven ones. A common denominator of all of them is that they build on some form of the *d*eterministic model of the underlying processes running within the stack and the balance of the plant. Hence, the impact of process disturbances and uncertainties on the final result is largely overlooked. To the best of our knowledge, this thesis provides a rare attempt to systematically quantify the influence of random noise and disturbances on the diagnostic inference and prediction of the remaining useful life. Indeed, uncertainties in the process cannot result in specific reasoning outcomes. Therefore, it is of considerable engineering and scientific interest to quantify the level of uncertainty of the results of inference, which contributes to cautious and more reliable decision-making.

The thesis first addresses the uncertainty in the deconvolution of the electrochemical impedance spectroscopy (EIS) utilizing the equivalent circuit models. It is well known that instrumentation and process noise during the EIS measurements affect the spectra, particularly in the low-frequency region. In our approach, we pursue the variational Bayes (VB) approach, which tends to be an approximation of the Bayesian approach but computationally much less overwhelming. Hence, the VB algorithm results in a code suitable for online implementation. To mitigate the inherent bias of the VB approach, mixture distributions are used as prior-posterior distributions, significantly reducing bias. The results

obtained on data gathered from EIS sessions on SOFC and SOEC stacks indicate that the serial connection of a resistance and RQ elements (parallel connection of a resistance and a constant phase element) sufficiently accurately describes the experimental data with relatively low uncertainty. That observation holds provided the EIS were taken under regular conditions, although the results can fluctuate from one EIS session to another.

The ground for fault diagnosis is the equivalent circuit model (ECM) parameters and their evaluated uncertainties. A fault is deduced from the dissimilarity between the evaluated and nominal ECM parameters. For that purpose, we introduce the Wasserstein distance (WD) between the distributions of the evaluated and the nominal ECM parameters. The conventional approaches tacitly assume that nominal ECM parameters are constant. However, nominal values of the ECM parameters vary with the operating conditions. In practice, especially in SOEC applications, the operating conditions change frequently and decisively influence the EIS spectra. Therefore, process parameters must be considered in the EIS analysis if we want consistent diagnosis resistant to changes in the operating point. To do so, we propose an innovative approach in which the nominal ECM parameters are calculated relative to the operating conditions. The data-driven Gaussian processes (GP) model relates the process variables with the ECM parameters. Based on that, each time a new EIS session is performed, the GP model predicts the ECM parameters from process variables. Then, the WD between the predicted and measured distributions of the ECM parameters is evaluated.

The problem of fault isolation is solved using the Support Vector Machine (SVM) classifier, which evaluates the probability that the system is at a specific faulty state.

Finally, the problem of predicting the remaining useful life (RUL) is approached within an entirely data-driven probabilistic framework. It is applied to a health indicator, which can be, for example, the stack voltage or the area specific resistance (ASR). The idea of the approach is simple, i.e., anticipate the end of life based on the linear trend model of the health indicator. A closed-form solution for the probability distribution of the stack's RUL emerges. That is a rather unusual result, which renders the required computational effort doable in no time. Additionally, the approach can be implemented on target platforms like industrial controllers. The idea has been validated through simulations and experimental data from an SOFC system to demonstrate its applicability to online monitoring and control of SOC systems.

In this thesis, we have been seeking solutions that will be operable and implementable on industrial control systems, which have much less computational power than PCs. The code is derived in Python and, as such, can be ported on the HW platforms developed within the framework of EU projects REACTT and RUBY, to name some examples. In both cases, the target processor is Raspberry Pi 4, which, thanks to the adequately conceptualized firmware, can include the operational code realized in Python.

The study begins by exploring Bayesian inference and introducing the Variational Bayes approach and its optimization methods. These techniques are applied to simulated and experimental SOC system data, demonstrating significant improvements in parameter estimation accuracy. The research further investigates passive diagnostic methods, employing data-driven feature extraction with Gaussian Processes to detect and isolate faults in SOC systems.

Povzetek

Vodik je pomemben element v sektorju zelene energije. Je okolju prijazen, lahko prenosljiv in primeren za dolgoročno shranjevanje. Sistemi trdooksidnih celic omogočajo pretvorbo vodika v energijo in obratno. Mednje sodijo sistemi trdno oksidnih gorivnih celic (SOC sistemi), ki proizvajajo električno energijo iz vodika, in sistemi trdooksidnih elektrolizerjev, ki proizvajajo vodik iz električne energije in vode. V primerjavi z drugimi sistemi za elektrokemijsko pretvorbo energije imajo sistemi SOC številne prednosti, kot so prilagodljiva uporaba reaktantov ter dostopnost materialov in katalizatorjev, saj delujejo pri temperaturah med 600 in 1000 ^oK. Kljub pomembnemu razvoju v zadnjih dveh desetletjih in naraščajočim potrebam po uporabi v zeleni mobilnosti ter pametnih omrežjih je njihova uvedba še vedno razmeroma počasna. Del razloga je še vedno nezadostna trajnost in zanesljivost, ki sta neločljivo povezani z neizogibnimi procesi degradacije. Za izboljšanje delovanja in trajnosti naprav SOC se kaže potreba po boljših metodah za zanesljivo spremljanje njihovega stanja in napovedovanje življenjske dobe (prognostika).

Diagnostika in prognostika sta ključni funkciji, ki pomagata operaterjem. Ponujata odgovor na dve pomembni vprašanji: (i) kakšno je notranje zdravstveno stanje naprav in komponent, ter kje je vzrok za morebitne težave, (ii) kako dolgo bo sistem še deloval, preden bo dosegel neizogiben konec?

Za diagnostiko napak in prognostiko SOC sistemov je opravljeno bistveno manj raziskav v primerjavi z drugimi sistemi za elektrokemijsko pretvorbo energije (zlasti baterijami). Na voljo je sicer kar nekaj metod, ampak predvsem za SOFC sisteme, medtem ko so SOEC sistemi še vedno v razvoju. Pristopi segajo od tistih, ki temeljijo na fizikalnih zakonitostih, do povsem podatkovno vodenih pristopih. Skupno jim je, da temeljijo na deterministčnih modelih procesov, zato večinoma zanemarjajo vpliv procesnih motenj in negotovosti na končni rezultat. Po našem vedenju, ta disertacija predstavlja edini poskus sistematičnega kvantificiranja vpliva naključnih procesov na degradacijo sklada kakor tudi na diagnostično sklepanje in predvidevanje preostale življenjske dobe. Zaradi negotovosti v procesu ne moremo deterministično sklepati o stanju sistema, zato je iz inženirskega in znanstvenega vidika kvantificiranje stopnje negotovosti rezultatov zelo zanimivo, saj na koncu prispeva k previdnemu in zanesljivejšemu sprejemanju odločitev.

Disertacija najprej obravnava negotovosti pri dekonvoluciji elektrokemijskih impedančnih spektrov z uporabo modelov nadomestnih električnih vezij. Znano je, da merilni in procesni šum med meritvami EIS vplivata na spekter, zlasti pri nizkih frekvencah. V našem pristopu uporabljamo pristop variacijske Bayesove metode (VB). Gre za približek Bayesovega pristopa, ki pa je računsko veliko manj zahteven. Zato ima VB algoritem za rezultat kodo, ki jo je mogoče implementirati na manj zmogljive računske platforme. Za ustrezno obravnavo pristranskosti VB pristopa, uporabljamo mešane porazdelitve kot a priori porazdelitve, kar znatno zmanjšuje pristranskost. Rezultati, pridobljeni na podatkih, zbranih med EIS sejami na SOFC in SOEC skladih, kažejo, da serijska povezava upornosti in RQ elementov (vzporedna povezava upornosti in elementa s konstantno fazo) dovolj natančno opisuje eksperimentalne podatke z razmeroma majhno negotovostjo. Ta ugotovitev velja v

primeru, da so bile EIS meritve, izvedene v kontroliranih pogojih, čeprav se rezultati lahko razlikujejo od ene seje EIS do druge.

Osnova za diagnostiko napak so parametri ECM in njihove ocenjene negotovosti. O napakah sklepamo na podlagi razlik med tekočimi ECM parametri in referentnimi (nominalnimi) ECM parametri. Za ta namen uporabljamo Wassersteinovo razdaljo med ocenami ECM parametrov, pridobljenimi z VB pristopom, kot osnovo za definiranje rezidualov. Wassersteinova razdalja učinkovito zazna spremembe parametrov skozi čas, kar omogoča detekcijo napak. Znano je, da analiza EIS temelji na predpostavki o stacionarnosti delovanja procesa, kar je v praksi skoraj nemogoče, še posebej pri SOEC sistemih. Ti namreč pogosto spreminjajo delovne pogoje. Delovni pogoji pa odločilno vplivajo na obliko EIS spektrov, zato jih je potrebno ustrezno upoštevati pri končnem preračunavanju spektrov. V ta namen predlagamo inovativen pristop, pri katerem se Wassersteinove razdalje napovedujejo neposredno iz procesnih parametrov. To opravimo z modelom Gaussovih procesov (GP), ki preslika procesne parametre v operativne parametre. Z uporabo Wassersteinove razdalje nato izračunamo razliko med dejanskimi vrednostmi parametrov ECM modela ter ocenami GP modela, kar služi kot temeljna informacija za diagnostiko. Problem izolacije napak rešujemo z uporabo klasifikatorja podpornih vektorjev (SVC), ki na koncu poda verjetnost, da je sistem v kateremkoli od pričakovanih napačnih ali nominalnih stanj.

Nazadnje obravnavamo problem ocenjevanja trajnosti sklada v smislu preostale uporabne dobe (RUL), v povsem podatkovno vodenem verjetnostnem okviru. Osnova za prognozo je predvidevanje razvoja enega ali več kazalnikov zdravja. Kazalnik zdravja je lahko napetost sklada ali specifična upornost (ASR). Ideja pristopa je preprosta, tj. napovedati konec življenjske dobe na podlagi linearnega trenda modela kazalnika zdravja, zato lahko dobimo rešitev za verjetnostno porazdelitev preostale življenjske dobe (RUL) za sklad v zaprti obliki. To je precej presenetljiv rezultat, ki omogoča hitro izvedbo potrebnih računalniških operacij. Po drugi strani pa se lahko pristop implementira na ciljne platforme, kot so industrijski krmilniki. Za prikaz učinkovitosti in uporabnosti ideje za direktno spremljanje in nadzor SOC sistemov smo metodo preizkusili na več simuliranih problemih in tudi na eksperimentalnih podatkih pridobljenih iz meritev na SOEC.

V raziskavi smo iskali rešitve, ki bodo uporabne in izvedljive na industrijskih krmilnih sistemih, z veliko manj računalniške moči kot osebni računalniki. Koda je napisana v programskem jeziku Python in se zlahka prenese na HW platforme, razvite v okviru projektov EU REACTT in RUBY. V obeh primerih je ciljni procesor Raspberry Pi 4, ki lahko zaradi ustrezno zasnovane programske opreme vključuje operacijsko kodo v Pythonu.

Contents

List of Figures

List of Tables

Abbreviations

ADAM ... adaptive moment estimation
APU ... auxiliary power unit
ARMA ... autoregressive moving-average
ASR ... area specific resistance
BIC ... Bayesian information criterion
BoP ... balance of plant
CPE ... constant phase element
CWT ... continuous wavelet transform
DRBS ... discrete random binary sequence
DRT ... distribution of relaxation times
ECM ... equivalent circuit model
ECMs ... equivalent circuit models
EECD ... electrochemical energy conversion devices
EIS ... electrochemical impedance spectroscopy
ELBO ... evidence lower bound
EOL ... end-of-life
FHT ... first hitting time
FOS ... fractional order system
GP ... Gaussian processes
GW ... Gauss-Wishart
HMC ... Hamiltonian Monte Carlo algorithm
IT ... information technology
KL ... Kullback-Leibler
KDE ... kernel density estimation
LLS ... Linear Least Squares
MC ... Monte Carlo
MCMC ... Markov chain Monte Carlo
NUTS ... No-U-Turn sampling
OLS ... Ordinary Least Squares
PDF ... Probability Density Function
PEM ... proton-exchange membrane fuel cell
PMF ... Probability Mass Function
RUL ... remaining useful life
SGD ... Stochastic Gradient Descent
SMR ... steam methane reforming
SOC ... Solid Oxide Cell
SOEC ... solid oxide electrolyzer cell
SOE ... Solid Oxide Electrolyzer
SOFC ... solid oxide fuel cell
SOH ... state of health
SVC ... support vector classifier

SVM ... Support Vector Machine
SVI ... stochastic variational inference
TCO ... total cost of ownership
VB ... variational Bayes
WD ... Wasserstein distance
WGS ... water-gas shift

Chapter 1

Introduction

Hydrogen is an indispensable energy carrier in the emerging green energy panorama. It is environmentally friendly, transportable, and storable. Hydrogen can be used in many branches of industry, green transportation, smart grids, and household applications. Fuel cells (e.g., low-temperature proton exchange membrane (PEM) fuel cells or high-temperature solid-oxide fuel cells (SOFC) convert hydrogen's chemical energy into electricity and heat.

Most hydrogen production methods nowadays rely on fossil fuels, such as steam reforming, partial oxidation of heavy hydrocarbons, and coal gasification [1]. Over 90% of hydrogen is produced by steam reforming of methane [2]. This process results in carbon dioxide emissions, contrasting with global decarbonisation agendas [3]. That is why eco-friendly technologies for highly efficient hydrogen production leave a minimal carbon dioxide footprint [2].

1.1 Basics of Solid Oxide Cell Systems

Solid oxide electrolysis cell (SOEC) systems. The cleanest way to produce hydrogen is by electrolysis, i.e., by splitting water into hydrogen and oxygen. Solid Oxide Electrolyzer (SOE) is one of the technologies that has attracted growing interest in the last two decades. SOEC systems operate at 700–1000°C and offer significant electrical power and higher efficiency than conventional low-temperature electrolysers [4], [5]. It gets its name due to the ceramic composition of its electrolyte.

The principle of operation of a SOEC system is as follows. Steam is supplied to the cathode, where water is subjected to the reduction reaction, resulting in a hydrogen molecule and oxygen ions thanks to the electrons supplied by the external circuit (Figure 1.1b).

$$H_2O + 2e^- \rightarrow H_2 + O^{2-} \tag{1.1}$$

Oxygen ions migrate through the electrolyte to the anode, where they are converted into oxygen gas and return the electrons to the power source [2].

$$O^{2-} \rightarrow \frac{1}{2}O_2 + 2e^- \tag{1.2}$$

Compared to other electrolysis technologies, SOECs demonstrate several strong properties:

1. Different reactants can be easily used: pure water for steam electrolysis, water with carbon dioxide to produce syngas through co-electrolysis, or only CO_2 in dry electrolysis mode.

2. SOECs do not require using expensive and rare metals such as platinum. There is a possibility of decreasing costs with increased production volume, e.g., through 3D printing. SOECs with lower electrode polarisation resistance are viewed as potentially undermining the current dominance of PEMEL systems [1].

3. Taking into account that up to 80% of the total cost of ownership (TCO) is based on the costs of electricity, SOECs show the highest potential in the long run [6].

Solid Oxide Fuel Cell (SOFC) system In SOFCs, the electrochemical oxidation of hydrogen with oxygen generates electrical and thermal energy, with water (H_2O) as the by-product, leaving no carbon dioxide (CO_2) footprint

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O. \tag{1.3}$$

In the sequel, more emphasis will be placed on SOEC devices. However, the presented methodologies apply to both types of devices.

Solid Oxide Cell (SOC) system. SOC system is an umbrella term for electrochemical energy conversion devices that operate in both fuel cell mode (SOFC) and electrolysis mode (SOEC). In a narrow sense, the label SOC is used in the literature to refer either to SOFC or SOEC, while reversible devices, capable of operating in both modes, are labelled r-SOC.

SOFCs are used to generate electricity from fuels, whereas SOECs are used to produce fuels from electricity.

The architecture from cells to systems. Without loss of generality, the point will be explained in an example of SOEC.

The power used on a solid oxide electrolysis cell depends on the operating conditions, cell design, and size, and it can reach from a few watts to several dozens of watts. The rate of the produced hydrogen can vary from a few to several dozen standard litres of produced hydrogen per hour. One has to apply more electrical energy to the input to increase the hydrogen produced on the cell input. Therefore, several single cells must be connected in series via metallic interconnects. Such a structure is referred to as a stack (Figure 1.2).

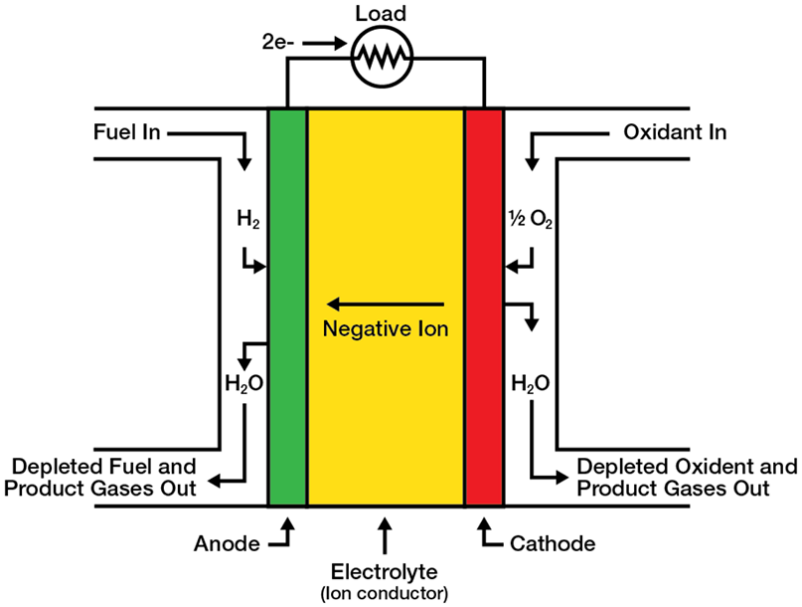
For a stack to operate correctly, the supply of the input gasses (air, steam, H_2) should be ensured. Also, their temperatures must be kept within certain limits to guarantee correct operation. Hence, besides the stack, one has to control the auxiliary units like pumps, heat exchangers, burners, etc. These units are referred to as the balance of plant (BoP). BoP and stack together are referred to as SOEC systems. An example of the SOEC system scheme is presented in Figure 1.3. Without going into the details, one can see the system is a small plant.

A SOEC system is further integrated into an electrical grid or industrial facilities (Figure 1.4).

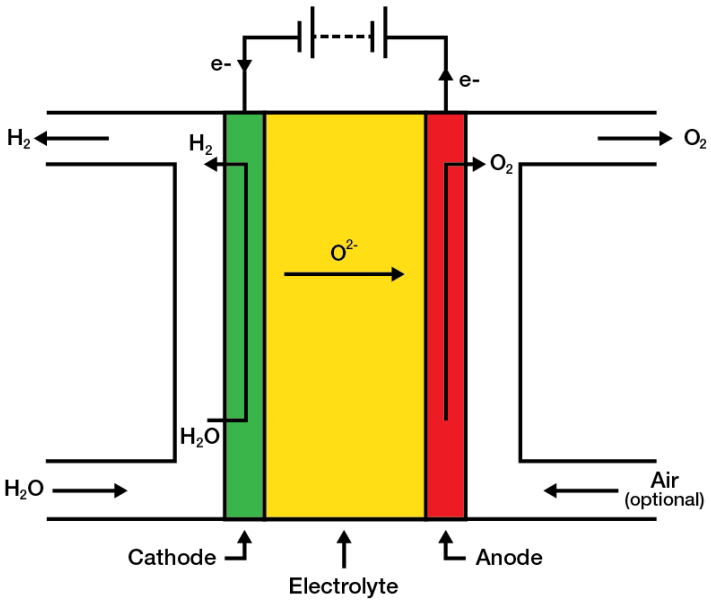
1.2 Issues in SOC Technologies

Despite the excellent potential for high conversion efficiencies, SOEC and SOFC technologies face obstacles that slow their commercial expansion. Namely, the high operating temperature leads to several degradation processes that limit lifespan. An ontology of the degradation phenomena in SOEC stacks is depicted in Figure 1.5. This issue must be addressed to prepare this technology for large-scale implementation.

Operation under high temperatures poses several challenges:



(a) The principle of operation of SOFC.



(b) The principle of operation of SOEC.

Figure 1.1: The principle of operation of SOFC and SOEC. SOFC produces while SOEC uses electrical energy. [7]

1. Scaling up the cells to larger stacks generally decreases the reliability of the actual system performance. Mechanical issues seem to be prevailing, and the root causes are different. For example, chemical reactions across the cell generate large thermal gradients that implicate material stress. Similar implications may occur due to thermal cycling during operation as well as due to fabrication errors.

2. Many physical processes occur parallel within the porous electrode and cell components. They pose difficulties in SOEC modelling because the equations form a system of coupled, nonlinear differential and algebraic equations with many parameters that

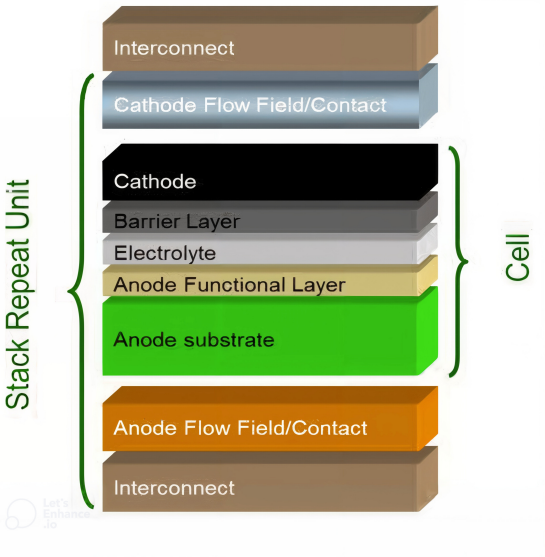


Figure 1.2: Cell and stack structure [8].

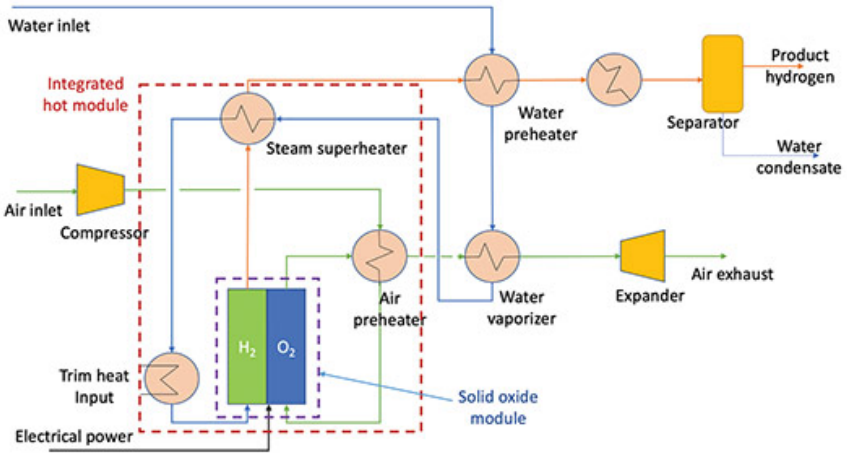


Figure 1.3: An example of the SOEC system [9].

are not independently measurable.

3. Elevated temperatures contribute to the growth of oxide layers between the interconnect plate and the electrodes, especially the cathode.

Understanding the mechanisms that lead to degradation is essential for designing, operating, and managing SOEC systems. Most of the work so far has focused on detailed modelling of the multiphysics processes in each point of the structure to understand the conditions that implicate the onset and rate of degradation and the relationship with the reliability and life span of the SOEC and SOFC systems.

An option to enforce the efficiency, reliability, and durability of SOEC systems is to implement dedicated online condition monitoring devices capable of timely detecting and localising the root cause of degradation or potential malfunction. Based on that, mitigation actions can be undertaken to avoid unexpected failures or to slow down the degradation rate whenever possible. Hence, one can extend the useful life of the equipment. On top of that, the online optimisation of the operating parameters is expected to keep the process

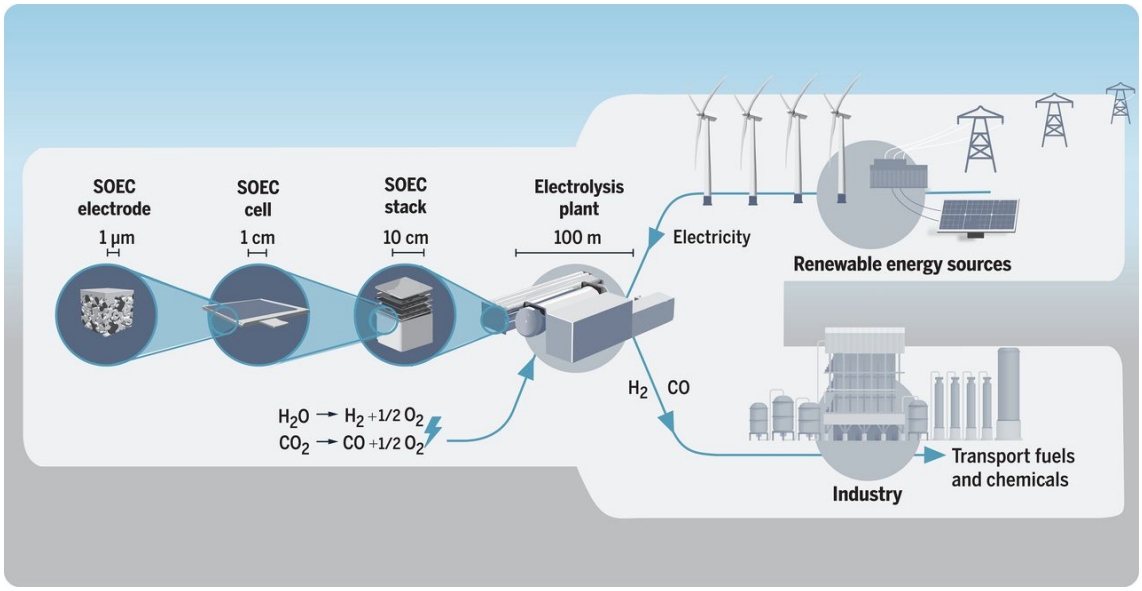


Figure 1.4: Typical SOEC system configuration from cell level to the system level [10].

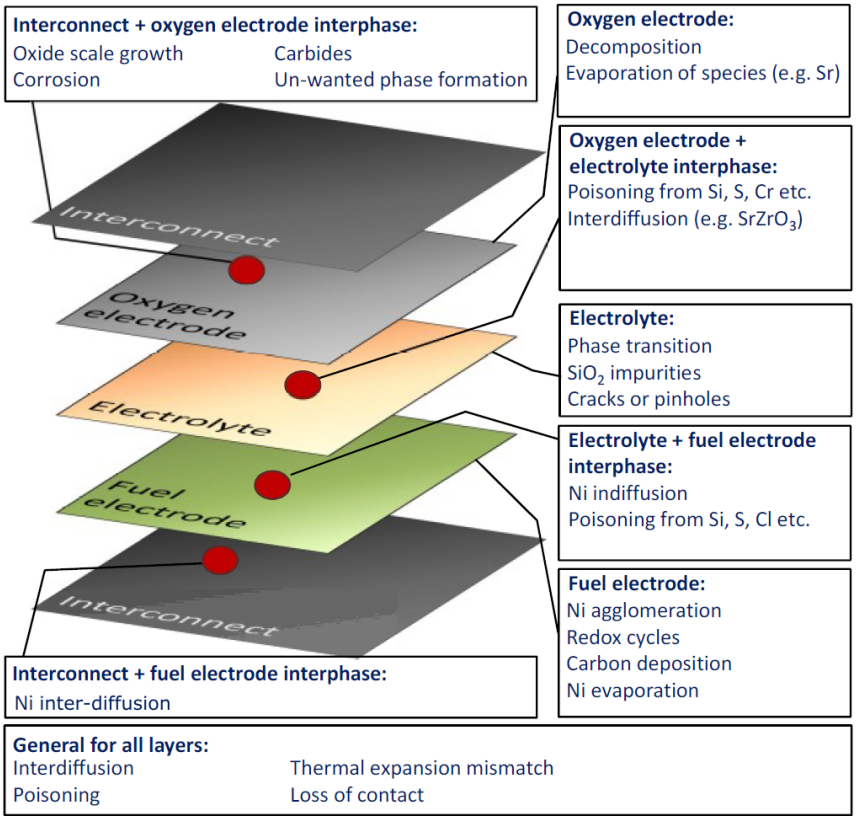


Figure 1.5: A possible ontology of the degradation mechanisms [11].

performance indicators at the best values. Thanks to such functionality, the operating conditions can also be optimised to slow down the degradation rate in periods of progressing degradation.

The whole idea of the PhD thesis goals is compactly presented in Figure 1.6.

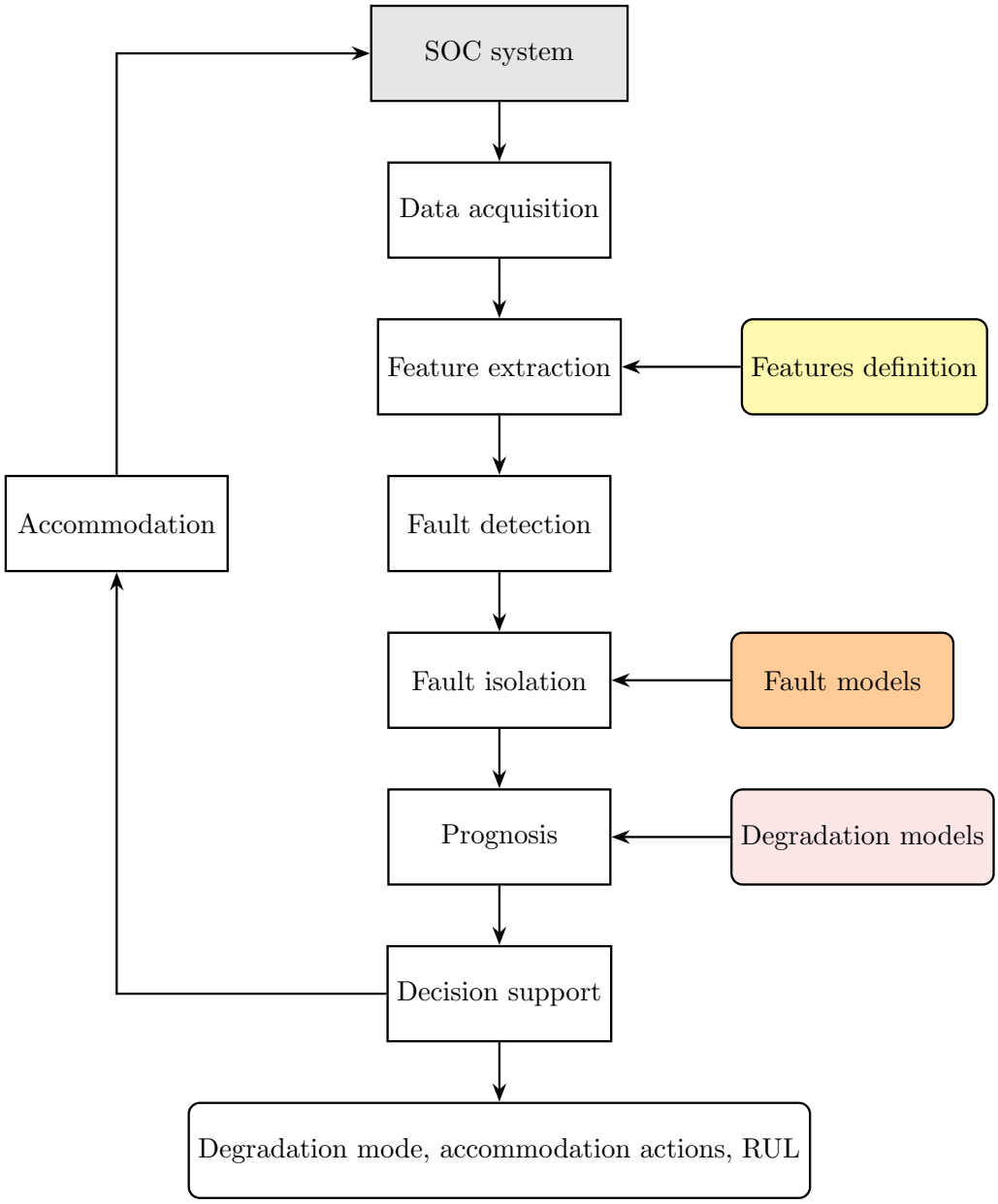


Figure 1.6: Flowchart illustrating the proposed diagnosis and prognosis framework for SOC systems, including steps for data acquisition, feature extraction, fault detection, isolation, prognosis, and decision support, with feedback loops for accommodation actions. The colored blocks are researched and presented as parts of this thesis.

1.3 Overview of the Diagnostic and Prognostic Research in SOC Systems

Systematisation of the diagnostic approaches can be done in various ways. In the sequel, we will distinguish two main approaches depending on how the diagnostic features are obtained. Diagnostic features are variables insensitive to the process operational regimes and rely only on the health condition of the system. We will consider two main diagnostic approaches: passive and active.

Passive approaches rely on collecting stack and BoP variables from the ordinary op-

eration without additional intervention into the system. On the contrary, the active approaches require external intervention for additional on-purpose perturbations [12]. In the electrochemical energy conversion systems context, the passive approach can only grasp the dominant dynamic modes of the stack (thermal modes) and not the other relevant dynamic modes related to gas diffusion and charge transport. Those dynamic modes are spread across a broad frequency scale from mHz to kHz, and to get insight, the additional active perturbation of the stack is needed.

1.3.1 Active Diagnosis

Polarization Curves

Polarisation curves provide static characteristics of the cell or stack by directly relating current density (j) and voltage (V). Some key parameters can be computed/identified from polarisation curves to summarise the change in the $j-V$ characteristics (Figure 1.7). The ASR approach can be considered among them. It relies on the linearised model that simulates the $j-V$ curve within the linear region. In this way, the ASR value change can be associated with the variation in the $j-V$ slope. For instance, an ASR approach can be considered to model SOFC voltage response over time [13].

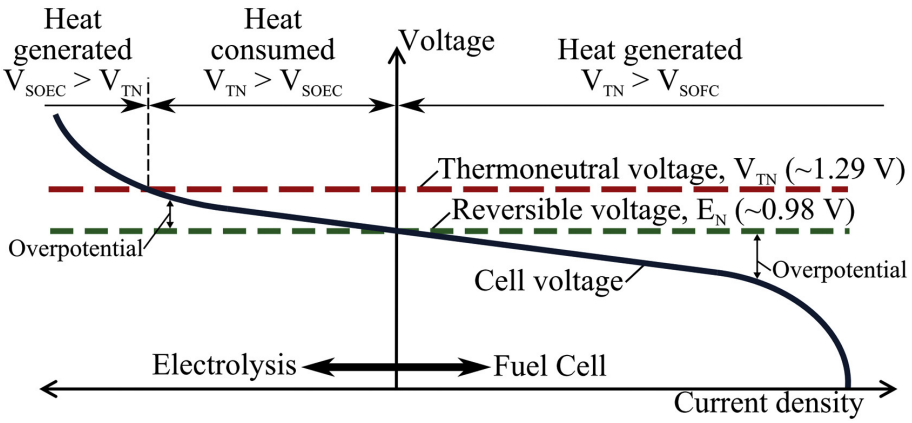


Figure 1.7: Current density versus voltage characteristic of rSOC cell/stack [14].

Electrochemical Impedance Spectroscopy (EIS)

It is widely recognised as a suitable methodology to analyse the dynamic modes of the electrochemical systems [15]. It allows the investigation of the key processes related to electrochemical reactions, charge, mass, and thermal transfer [16]. EIS employs the superposition of a perturbing stimulus (in voltage or current) and the measurements of the system response. Generally, the perturbation is a sinusoidal stimulus imposed at a given frequency and amplitude, injected during steady-state operation. Such measurement is usually performed at operating conditions within the linear region of the rSOC or SOE polarisation curve to avoid any distortion of the system response.

Many research activities have applied EIS to rSOC or SOE, although they mainly focus on cell and stack levels [16], [17]. For instance, in the work of [18], the authors showed the capability of EIS to monitor the performance of SOE systems for diagnostic purposes and to study the degradation mechanisms for different cell configurations.

Few works appeared in the last decade on evaluating EIS based on non-sinusoidal Perturbation of cells and stacks. In [19] the author proposed a fast EIS approach that

builds on using the discrete random binary signals for stack perturbation. Similar ideas have been proposed recently in [20]. The key idea is that such a signal exhibits a flat power spectrum on a particular bandwidth ω_B, which allows system excitation with all frequencies on the band $[0, \omega_B]$. That reduces the required perturbation times and increases the resolution of the evaluated impedance curve. The system transfer function is evaluated by applying the Morlet wavelet transform of the voltage and current signals. The approach has been validated in [21], [22]. Its first application to SOEC can be found in [23].

EIS Deconvolution using Equivalent Circuit Models (ECM)

The idea is to find an electrical circuit that can reproduce the EIS characteristic of the system [24]. An example of ECM structure is shown in Figure 1.8. The structure comprises a resistor-inductor R-L series connected with a resistor-capacitor R-C parallel. In [25], the capacitor in the parallel element is replaced by the constant phase element (CPE), which models the capacitance of a porous medium. The methodology presented in the work [25] for ECM features extraction is based on a patented algorithm [26], which adapts the ECM structure (i.e., the number of elements, their type, and connection) according to the shape of the Nyquist graph obtained through EIS measurements. The ECM fit can autonomously adapt the ECM without specifying technology, thus ensuring the high generality ability of the patented approach.

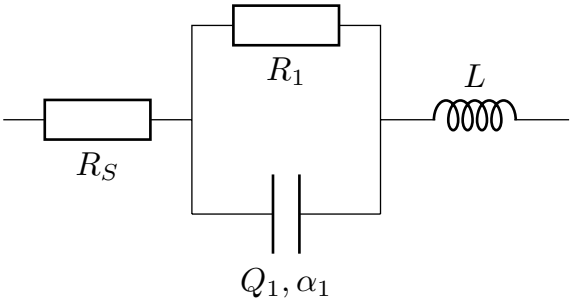


Figure 1.8: An example of an Equivalent Circuit Model (ECM) structure.

In [27], the authors propose an approach that combines EIS and polarisation curve measurements, which are further processed through ECM and deconvolution techniques to study the electrochemical features of commercial SOEs. The deconvolution techniques consist of the distribution of relaxation times (DRT), which investigates the frequency features related to EIS measurements to correlate the occurring physical phenomena to specific frequencies. Along the same line, in [28], the authors apply the EIS technique to perform an experimental characterisation of SOCs.

Different SOEC ECMs can be found in the literature. It is known that the uniqueness of the ECM model is not guaranteed [29]. Given the fact that a model does not equal true nature but serves to approximate best the measurements obtained from the system, the structural identification of the ECM depends not only on the chemical mechanisms in the cell but also on the quality of measurements and acceptable quality of the model. One finds different model structures in the literature, such as 5 RQ elements in [30] or the serial connection of 3RQ elements and Gerischer element [31].

In [32], the author performed an exhaustive in-situ stack characterisation with EIS and DRT. The DRT peak identification was performed with a SOC stack through the parameter study. According to the experimental results and previous results obtained with single cells by KIT (Karlsruhe Institute of Technology), up to five DRT peaks can be identified and

Table 1.1: Detectable SOC EIS responses according to frequency regime.

Response	**Attribution**	**Frequency peak**	**Influenced by**
P1	probably O_2 dissociation	< 1 Hz	j, T
P2	gas conversion	few Hz	j, T, p_{H_2O}, dilution
P3	O_2-electrode R_{pol}	10-100 Hz	j, T, p_{O_2}
P4	likely fuel-electrode diffusion	100-500 Hz	j, T, p_{H_2O}, dilution
P5	Fuel-electrode R_{pol} (ct)	1-4 kHz	j, T, p_{H_2O}, dilution
P6	ion transfer at O_2-electrode side	> 5 kHz	T, p_{O_2}
R_Ω	anything ohmic	intercept (corrected)	T

related to corresponding electrode processes (Table 1.1).

EIS Deconvolution using Distribution of Relaxation Times (DRT)

The DRT method relies on the idea that every impedance function that obeys the Kramers-Kronig relations can be represented as a differential sum of infinitesimal small RC elements [33]. The work of Baldinelli et al. [28] describes the performance of EIS measurements with features extracted through DRT and ECM approaches applied to rSOCs to investigate the degradation phenomena affecting such technologies [34]. An extensive experimental study was performed on electrolyte-supported SOC single-cell and stack levels. This work demonstrated the viability of DRT as an appropriate data analysis approach for the analysis and operation optimisation of SOC single cells and stacks. The DRT enabled linking single processes within complex SOC systems with their characteristic frequencies [35]. Subotić and her group conducted a suite of studies on degradation mechanisms and relation with the DRT characteristics of the stacks [23], [36].

1.3.2 Passive Diagnosis

The most common approaches to model-based diagnosis [37] build on:

1. Parameter estimation,
2. Parity relations, and
3. Observers.

The first group of approaches aims to build a physical model of the system and associate faults with one or more model parameters. Hence, model parameters play a role in diagnostic features. Parity relations are consistency relationships among the available process variables used to build features referred to as residuals. If the system is in the nominal condition, i.e., without fault, the parity relations return values near zero. A fault in the component causes certain parity relations to differ from zero. A fault is detected when at least one residual exceeds a given threshold. Fault isolation can be performed from the set of non-zero residuals. In the ideal case, each fault is associated with zero and non-zero

residuals. In the third case, the process model is the state space model, while a fault is modelled as an unknown input and parameter in the state model. Depending on the issues addressed, there are different ways to generate residuals [37].

The number of approaches for SOFCs and SOECs based on the mathematical models is somewhat limited. One of the few contributions to the residual-based diagnosis is available in [38]. The residuals are then evaluated based on a fault-symptom matrix (FSM) presented in [39]. The approach builds on the physical model of the SOFC system. FSM was first developed using fault-trees analysis and then improved via faults simulation. It should be stressed that a critical issue in model-based diagnosis is diagnostic resolution, i.e., the ability to distinguish unambiguously among different faults. This property significantly depends on the selection and placement of the sensors. A profound study on the optimal choice of sensors in SOFC to achieve the best resolution was performed by Polverino et al. [40].

The authors in [41] use a first-principles model to generate residuals. The model is validated in steady-state and dynamic states. It simulates faults and generates residuals, which provide training examples for a fault classification algorithm.

Papers [42] and [43] presented a black-box approach based on a neural network model of the nominal SOFC behaviour. In [44], the authors present a selection procedure for determining the most descriptive variables in the model.

The theory of fault detection and isolation is very well developed, and a range of tools have already been used in many fields of applications [37]. In the detection step, one has to reveal the eventual departure of the current system condition from the nominal condition. Statistical decision-making can assess change; see [45] for an overview. In [44], the authors demonstrated an application of the statistical evaluation of residuals in SOFCs.

To isolate a fault or a degradation mechanism, some discrimination in the feature space is required.—A rich arsenal of tools for fault isolation is widely used in other communities. One can refer to [37] for a comprehensive overview.

There are two major classes of approaches to fault isolation: inference methods and based on classification methods.

Classification Methods

Classification methods come into play when sufficient experimental case studies are available. A range of algorithms at disposal aim to discriminate between clusters of points in the feature space that belong to the exact degradation mechanism. Again, no references can be found related to SOEC systems and very little has been done in the SOFC domain. An approach employing the support vector machine has been proposed in the context of SOFCs by Costamagna et al. [46]. Experiments show that incipient faults can be distinguished quite reliably. More specifically, the minimum size of the fault corresponds to a 5% deviation of the associated variable from its nominal value.

A classification approach to SOFC diagnosis based on fuzzy clustering is proposed by Wang et al. [47]. The main drawback of the classification approach is that its successful operation requires a considerable set of case examples. The experiments must also contain examples from all the degradation modes of interest, which requires substantial experimental effort and cost.

Inference Methods

Inference methods incorporate prior or heuristic knowledge about the features and faults. This relationship can be expressed in Boolean IF-THEN-ELSE rules or their fuzzy version. In the SOFC context, only very few works address isolation. This can be done by using

Boolean reasoning [44]. The fault-symptom matrix corresponds to a set of logical rules, where it is usually assumed that only one fault is acting in the system at the time. If FSM is fully isolated, multiple faults can also be unambiguously revealed.

A systematic summarisation and classification of various SOFC systems fault diagnosis approaches can be found in the recent paper [48] related to the SOFC. Eighteen specific faults are addressed along with fifteen fault detection and isolation (FDI) methods, categorised into four groups, i.e., EIS-based, model-based, data-driven, and hybrid FDI. Besides, practical applications, diagnosis accuracy, diagnostic diversity, complexity, advantages, and disadvantages of all methods are thoroughly compared and evaluated.

1.3.3 Prognostics

Prognostics anticipates the system's health status unfolding and estimates its remaining useful life. Hence, prognostics can improve system reliability while reducing maintenance costs and downtime. This has been a topic of interest in different disciplines, from mechanical systems [49] to electrochemical ones [50]. The prognostics takes two phases: a learning phase and the prediction phase. During the learning phase, the predictive model is learned from the system behaviour up to time t_p (see Figure 1.9). During the prediction phase, the prognostics gives RUL predictions of the system and determine when the critical threshold is reached. The time length between the predicted time of failure (we define it as $t_{failure}$) and the starting point of prognostics t_p is the RUL.

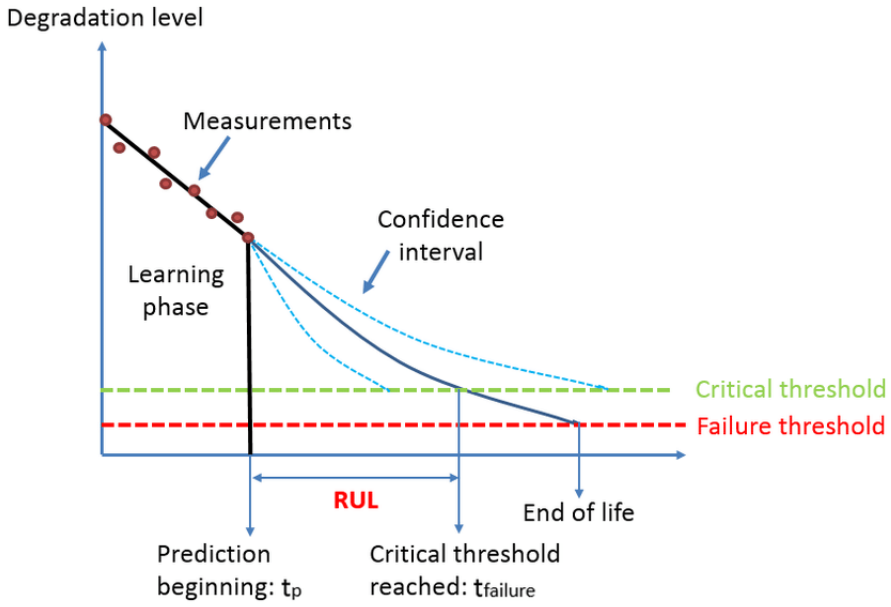


Figure 1.9: The prognostics process [51].

So far, almost no attention has been devoted to developing algorithms for the online prediction of RUL in SOFC, particularly in SOEC systems. Only a few papers address the topic. In [52], the authors adopt a neural network model in which time is used as an explicit independent variable, hence allowing voltage decay to be described as a function of time and operational parameters. To train the model, the data from run-to-failure tests are required. The approach is shown to predict the evolution of the stack voltage decay under the implicit assumption that the degradation mechanisms present during the validation phase are similar to those in the learning stage. A combination of modelling

with electrochemical impedance spectrum information for the prediction of the remaining useful life of SOFC stacks is proposed by Gallo et al. [13].

Wu and Ye [53] study the RUL prediction under two degradation mode types, i.e., anode poisoning and cathode humidification. They apply a qualitative model for the state of health and probabilistic transition between the discrete instances of the state of health. The whole study is performed in a simulation environment. From a series of simulated life-long tests, they learn the parameters of the transition model by the data-driven method (Least Squares – Support Vector Machine). The model shows a high prediction accuracy in the validation stage with 20% absolute error. Cell voltage is utilised in this work as the state of health indicator. Again, stationary operating conditions are assumed.

Unfortunately, the output voltage is directly affected by operating conditions imposed on the system. Therefore, it is difficult to distinguish whether the voltage drop is caused by the degradation of the fuel cells, a change in load condition, or if it is a response of the system due to the controller. In[54], the first attempt at RUL prediction of SOFC stacks seems to be the first attempt at RUL prediction that allows for rather general non-stationary operating conditions, thus rendering the need for constant operation unnecessary. The idea is to estimate the parameter ASR online by employing a nonlinear lumped process model. The temporal evolution of the ASR is then described by another model whose parameters are identified online based on sensor records. The RUL predictions are evaluated using Monte Carlo simulation. Figure 1.10 shows an excerpt of the algorithm performance. In [55], Dolenc et al. present an advanced algorithm version. Several papers have emerged in recent years dealing with SOFC prognostics, e.g., [56], [57].

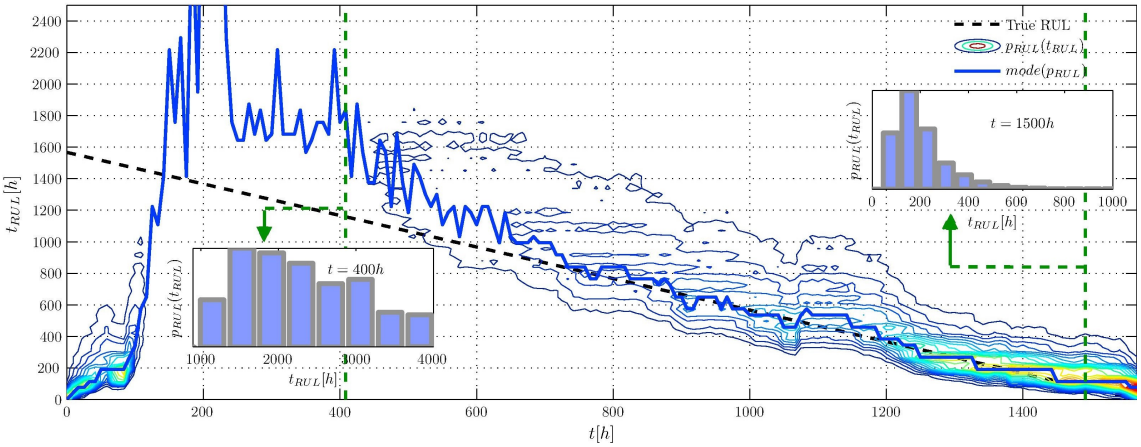


Figure 1.10: The evolution of RUL prediction during operation of a SOFC stack [55]. The dashed line is true RUL. The graph indicates the prediction is highly uncertain at the time instance $t = 400$ since the expected end-of-life spread is 1000-4000h. At $t = 1500h$, the spread narrows to [0-500] with the most expected RUL 150h, which fits the true time remaining till the end [58].

As stated before, this widens the scope of usable methods for the remaining research topics surrounding Hydrogen Fuel Cells (fault detection, fault isolation, prognosis, etc.). Furthermore, we deal with the VB approach's inherent bias by using mixture distributions as our prior/posterior distributions of choice, which significantly lowers the bias.

1.4 Purpose of the Dissertation

The dissertation focuses on the diagnosis and prognosis of SOFC and SOEC systems. The scope is to address uncertainties that are inherently present, such as disturbances in operational conditions and noise in process instrumentation. The known diagnostic approaches in the literature range from those based on first principles to the entirely data-driven ones. A common denominator of all of them is that they build on some form of a deterministic model of the processes running within the stack and the balance of the plant, hence largely overlooking the impact of process disturbances and uncertainties on the final result. That is the crucial motivation for our approach, and the backing idea is explained in Figure (1.11). Namely, from a mathematical point of view, a SOC system can be viewed as mapping a vector of input variables to a vector of output variables. Since the inputs are random variables, outputs are also random variables. Any further transformation of the output variables, what the diagnostic reasoning is, generates results, which are also random variables. In such a setup, all information is expressed by the probability density function and not by a single number. The uncertainties in process inputs have different origins:

- Random noise in the sensor readings;
- Disturbances in the BoP components, particularly in the flow of fuel. A good example is steam flow, in which random fluctuations are inevitable. Those disturbances extend in the lower frequency range therefore they can impact the results of the electrochemical impedance spectroscopy in the lower frequency region;
- Complex degradation phenomena rooted in the changes in the material of the electrodes due to ongoing chemical reactions. Changes on the microscopic level affect the physical parameters of the process on the macroscopic level in a not exactly predictable way.

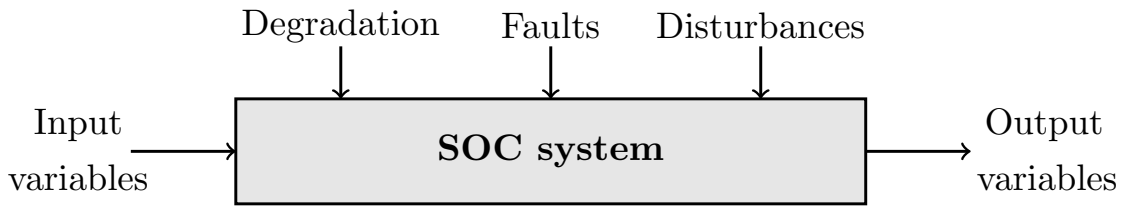


Figure 1.11: System view on the SOC system as a mapping of random variables.

Accounting for the uncertainties in the variables results in a cautious interpretation of the measurements and, ultimately, a more cautious and robust diagnosis.

To the best of our knowledge, this thesis provides a rare attempt to systematically quantify the influence of the random nature of the complex degradation processes in stacks as well as the role of fluctuations in the BoP (for example, inflows of gases) on the diagnostic inference and anticipation of the remaining useful life.

This dissertation aims to develop a unified framework for diagnosing and prognosis of SOC systems. The goal is to consider uncertainties in both models and operational data comprehensively. In contrast to pursuing fundamental methodological research, the dissertation prioritizes adopting the existing methodological tools to create algorithms with practical applicability. Recognizing the current immaturity of real-time monitoring solutions for SOC systems, the research focuses on delivering algorithms suitable for immediate

implementation. The design principles include optimizing the algorithms for portability on low-cost embedded platforms, minimizing the number of design parameters, and maintaining intuitive usability. The goal is to contribute practical tools that can significantly advance the current state of SOC system monitoring.

The focus has been put on the following points:

- ECM parameter estimation,
- fault diagnosis of the stacks and
- fast prognosis algorithm.

The parameter estimation method uses Bayesian inference to obtain the posterior distributions of the model parameters directly from EIS data. Similar results could be achieved with Markov chain Monte Carlo (MCMC) by sacrificing speed and low numerical complexity that we can achieve with the Bayesian framework. The marginal increase in speed does not significantly affect the accuracy of the estimations, so the Bayesian framework was chosen.

ECM parameter estimation is the basis for fault diagnosis. A fault can be detected by evaluating the distance of the evaluated EIS parameters from the nominal parameters using the WD. The important benefit of the approach is the automatic threshold calculation based on only one parameter - the desired probability of a false alarm.

Furthermore, we can calculate the probability of specific faults by using the distributions of the system parameters in its current state and comparing them to the system in its nominal state, provided the fault matrix is available.

There is a great need for an accurate and non-complex prognosis method that allows for a preventive servicing plan. Using ratio distribution and Linear Least Squares (LLS) posterior updates, the future trend of the desired health estimator can be simulated, and the distribution of FHT can be calculated analytically.

1.5 Goals of the Dissertation

The main goals of the dissertation are as follows.

Goal 1: The probabilistic deconvolution of electrochemical impedance spectra (EIS)

That is achieved by identifying ECM parameters from EIS data. Notably, there has been limited systematic study in this regard. The Variational Bayes approach is proposed to achieve this objective, offering a detailed exploration of underlying principles to ensure the methodology's accuracy and applicability. A comparative analysis is conducted to validate the approach, comparing the Variational Bayes approach with traditional methods such as Markov Chain Monte Carlo (MCMC). This comparative study underscores the advancements and benefits of the novel Variational Bayes approach over conventional MCMC methods.

Goal 2: Probabilistic fault detection algorithms

To design a consistent diagnostic inference under uncertainty, the variables are treated as random variables. Be it a process variable like stack voltage, a residual, or a parameter of the ECM, each of them is considered as a random variable. In contrast to many practical approaches, the output of the probabilistic diagnostic inference is not a binary decision (fault/no fault) but a probability measure indicating the presence of a fault in the system. Notably, the boundary between a normal (or nominal) system condition and a faulty condition is rarely known a priori. Typically, end-users expect the designer of a diagnostic system to define this boundary. Establishing appropriate threshold values for fault detection features can be a challenging task. However, within the probabilistic

framework, one can select the threshold values that align with a permissible false alarm rate. This approach simplifies the determination of thresholds for fault detection through straightforward statistical analysis of measurement data collected in normal conditions.

Goal 3: Probabilistic Fault Isolation

Fault isolation in diagnostic inference aims to identify the root cause of a detected fault. For this purpose, a SVM framework is proposed, utilizing WD values between GP model predictions and estimated ECM parameter distributions. This approach generates a list of suspected faults with associated likelihood measures. By considering different degradation modes and combining information from estimated parameters, the method pinpoints and isolates specific issues, providing a more nuanced and targeted approach to system maintenance.

Goal 4: An algorithm for health prognosis

While diagnosis indicates the location of the problem, prognosis is intended to estimate the remaining useful life of the system. The dissertation relies on monitoring the statistical evolution of a health indicator and estimating the time the indicator hits some prescribed boundary limit (marking the end of life). A statistical approach, referred to as the Linear First-Hit Times Prediction (LFHT) algorithm, relying on linear trend estimation, is proposed. The remarkable property of the LFHT algorithm is that the distribution of the FHT can be expressed in a closed form.

Goal 5: Validation of the proposed algorithms through simulations and experimental data

The algorithms are initially tested using simulated case studies, followed by validation of data obtained from experimental campaigns.

1.6 Hypotheses of the Dissertation

The dissertation will test the following hypotheses:

- **Hypothesis 1:** The VB approach significantly improves the efficiency of EIS deconvolution compared to traditional MCMC, thereby making the approach usable for online applications.
- **Hypothesis 2:** Probabilistic diagnostic inference, which treats all variables as random variables, provides a consistent framework allowing for systematic calibration of the diagnostic system, in particular, easy tuning of the detection thresholds.
- **Hypothesis 3:** The proposed statistical approach for estimating the remaining useful life of SOC systems offers significant benefits over existing methods.

1.7 Scientific Contributions

The scientific contributions of this thesis are the following:

- **A fully probabilistic approach to modeling SOFC and SOEC systems:** This approach uses the VB method for parameter estimation, accounting for uncertainty. Although the current application may seem marginal, the potential practical use of these systems necessitates robust parameter estimation methods that handle uncertainty, especially in further developed algorithms incorporating mixture distributions.
- **A novel approach to fault detection in electrochemical energy conversion devices (EECD):** This approach employs GP and support vector classifier (SVC),

enabling a probabilistic framework that synthesizes information from measured signals and EIS spectra.

- **A novel approach to fault isolation in EECD:** Building on the fault detection algorithms, this approach expands to allow accurate and rapid fault isolation while maintaining a probabilistic framework.
- **A closed-form solution for calculating distributions of RUL:** This method applies to systems with a linear trend in their health indicators, facilitating use in online applications on low-cost computing equipment, such as Raspberry Pi and similar devices.

1.8 Structure of the Thesis

The structure of this thesis is as follows: The first section consists of definitions of different electrochemical and mathematical terms. Following is the description and methodology of the VB approach in combination with ECMs. Furthermore, this section contains results on simulated and experimental data. Thirdly, the methodology and definitions for Wasserstein distance are given, along with an explanation of our motivation for the choice. Similar to before, the method for fault detection is tested on simulated and experimental data. Fourthly, the fault detection method is extended to fault isolation utilizing differences in distributions and their Wasserstein distances to provide a successful isolation method. Fifthly, the novel approach to the probabilistic prognosis of SOC systems is presented and extensively tested on simulated and experimental examples. Finally, in the concluding session, we point out the main aspects of this thesis and the main steps forward in SOFC/SOEC domain.

Chapter 2

Probabilistic EIS Deconvolution

2.1 The Electrochemical Impedance Spectroscopy

The impedance Z of a linear two-terminal electrical circuit relates the current $i(t)$ passing through the circuit and the voltage $u(t)$ on the terminals. In mathematical terms, impedance is defined as the ratio between the Fourier transform of voltage and the Fourier transform of current. Therefore, the impedance is a complex number and is a function of the frequency ω, i.e. $Z(j\omega) = |Z(j\omega)| \cdot e^{arg(Z(j\omega))}$. It is expressed in the same units as the resistance, i.e. in ohm Ω.

If a sinusoidal current is applied, the resulting voltage will also be sinusoidal (c.f. Figure 2.1). That means if $i(t) = I_0 \sin \omega t$ then $u(t) = U_0 \sin(\omega t + \varphi)$ leading to

$$Z(\omega) = \frac{\mathcal{F}\{u(t)\}}{\mathcal{F}\{i(t)\}} = \frac{U_0 e^{j(\omega+\varphi)}}{I_0 e^{j\omega}} = Z_0(\omega) e^{j\varphi}$$

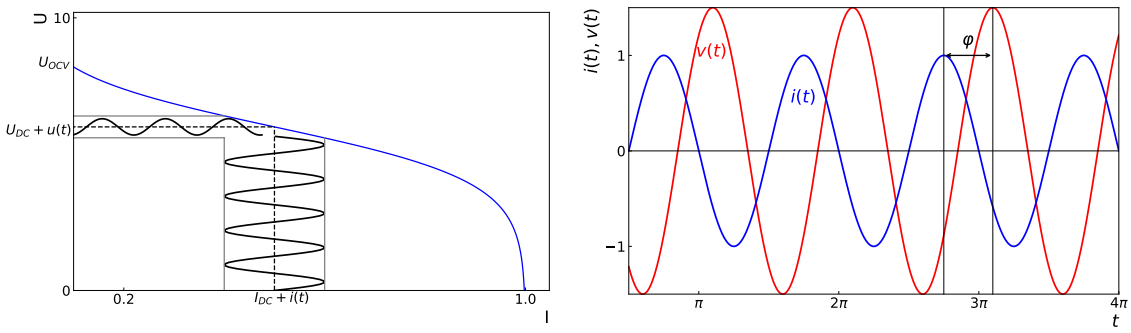


Figure 2.1: The principle of electrochemical impedance spectroscopy.

The impedance is divided into two main components: resistance, which represents the real part of the impedance and reactance, which represents the imaginary part of the impedance. The reactance is responsible for the phase shift in the response signal (linear capacitance or inductance), while the resistance is responsible for changes in the amplitude. The measuring unit of impedance is in Ohms [Ω].

$$Z = R + jX = Z' + jZ''$$

The conventional approach to the evaluation of the impedance is straightforward. It relies on applying (a small-amplitude) sinusoidal excitation $i(t)$ on system input at the

given operating point (I_{DC}) and observing the response $u(t)$ at the operating point U_{DC} (Figure 2.1).

Proper EIS stands on the three important conditions:

- *Linearity.* The system's response signal must be directly proportional to the input signal. If linear, no other harmonics are generated during the experiment. Generally, SOC systems are nonlinear. Hence to comply with the linearity conditions the amplitude of the excitation signal should be small enough in order not to excite non-linear modes of the system. This requirement can be accomplished by taking into account the information about the system's static characteristic (IV curve) and then choosing a small enough amplitude of the perturbation.

- *Stationarity.* It is mandatory that during the EIS measurement session, the system operating points I_{DC} and U_{DC} remain constant. In the SOC systems, that means that ideally, all the process variables, including input flows of gasses and temperatures in the process, should stay fixed. That is hard to ensure, particularly during the long measurement sessions. During this time, drifts and environmental disturbances may occur, which can significantly affect the accuracy of the final result. Therefore, maintaining stationarity is challenging at frequencies as low as, say $< 1mHz$. In the experiments, usually, 3 to 10 periods are required to achieve sufficient frequency resolution [59]–[61]. One way to reduce the time required for performing EIS is by using broadband excitation signals [62].

- *Causality.* That refers to the expectation that the perturbed voltage $u(t)$ on output is the result of only the perturbation in current $i(t)$. However, the process operating point, e.g. the stack temperature, can affect the voltage response. That means the same current excitation may result in different voltage responses. Overlooking the fact that the output voltage is causally related to other process conditions besides current, can lead to misleading results. That will be one of the key issues addressed below.

2.1.1 Fast EIS

The fast EIS used in this study is proposed by Boškoski et al. [62]. Conventional EIS methods employ single components of sine waves to excite the system [38]. In turn, the impedance is evaluated only for a limited set of predetermined frequencies. Multiple periods of a sinusoidal excitation signal are necessary for each desired frequency to achieve results that are precise enough.

The discrete random binary signal (DRBS), on the other hand, has emerged as a viable candidate for the excitation signal. It can take either of the two values and alternate between them at random times. A signal of this type could be thought of as being made up of an endless ensemble of sinusoids that are defined at all feasible frequencies and have their amplitudes changed correspondingly. An example of DRBS is shown in Figure 2.2 in a time domain and its power spectrum density. The effective bandwidth of the DRBS is defined as:

$$f_b = \frac{1}{3\lambda},$$

where λ is the minimal time between the two shifts [63]. The power spectrum of the signal is very flat in the interval $[0, f_b]$ and is thus quite close to the power spectrum of white noise.

To determine the gain and phase characteristics of a system at a particular frequency, a spectral decomposition of the input and output signals must be applied. As proposed

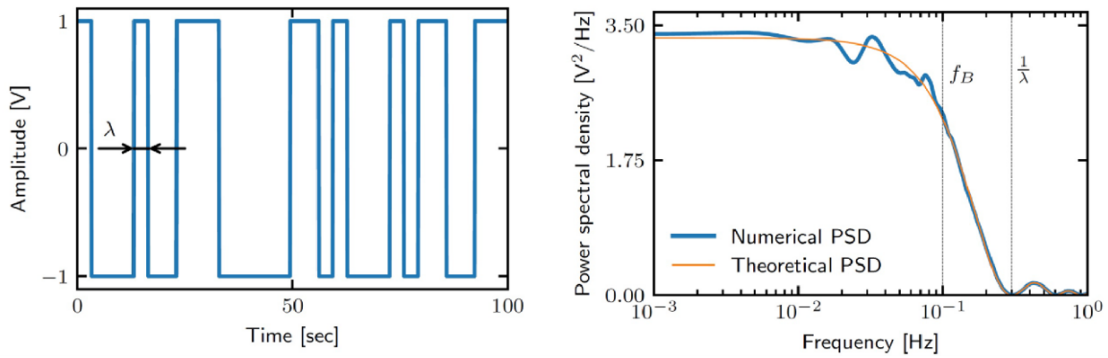


Figure 2.2: The underlying idea of the fast electrochemical impedance spectroscopy. The perturbation is performed by a random switching binary signal (the so-called discrete random binary signal - DRBS). The time domain representation of DRBS is shown on the left, while its power spectrum is on the right side [64].

by Nusev et al. [64], a continuous wavelet transform (CWT) was utilized instead of the conventional short-time Fourier transform. By scaling the mother-wavelet $\psi_{u,s}(t)$, CWT enables a flexible time-frequency resolution. It is described as [65]:

$$Wf(s,\nu) = \int_{-\infty}^{\infty} f(t)\psi^*_{\nu,s}(t)dt, \tag{2.1}$$

where the complex conjugate of the mother wavelet is denoted by $\psi^*_{\nu,s}(t)$, and ν and s are the translation and scaling parameters, respectively. Based on the characteristics of the chosen wavelet function, the ν and s parameters can be simply converted in time t and frequency f.

Morlet, Morse, or Lognormal wavelets can be utilized as a mother wavelet. It is possible to tune wavelet parameters such as duration and symmetry, minimum and maximum frequencies used to define the frequency range, and the number of frequency points from which coefficients are generated. This study implements the Morlet wavelet as in [64], defined as:

$$\psi(t) = \pi^{-\frac{1}{4}} \left(e^{-j\omega_0 t} - e^{-\frac{{\omega_0}^2}{2}} \right) e^{-\frac{t^2}{2}}$$

Details on using CWT with the Morlet wavelet in the context of EIS, as used in this study, can be found in [66].

Similarly to the Fourier transform, the impedance based on CWT can be calculated as the ratio of the wavelet coefficients of the voltage $u(t)$ and the current $i(t)$:

$$Z(t,f) = \frac{Wu_{\text{cell}}\,(t,f)}{Wi_{\text{cell}}\,(t,f)}$$

where $Wu_{\text{cell}}\,(t,f)$ and $Wi_{\text{cell}}\,(t,f)$ are the wavelet coefficients of $u(t)$ and $i(t)$ calculated using (2.1). The instantaneous transfer function $Z(t,f)$ in the CWT analysis, unlike the Fourier transform, requires both time and frequency to be defined.

2.2 Spectral Deconvolution using the Equivalent Circuit Models

The EIS spectra should be interpreted in some way in order to be used in diagnostic reasoning. The very common approach is using equivalent circuit models that mimic the

impedance and, at the same time, offer a physical interpretation of the electrochemical processes inside the stack.

The ECMs are composed of basic building blocks which result from detailed analysis and linearisation of the first principles models referring to mass, energy, and charge conservation [67], [68].

In the sequel, we will briefly review the most frequently applied building blocks (besides conventional elements like resistors, capacitors and inductors).

Constant phase element (CPE)

It is used to describe the impedance of a double-layer capacitance. It represents a general form of the impedance of the capacitor and reads:

$$Z_{CPE}(j\omega) = \frac{1}{(j\omega)^{\alpha}Q}$$

where $\alpha \in (0, 1]$ and represents the fractional order coefficient. When $\alpha = 1$ the CPE describes an ideal capacitor.

Cole-Cole (RQ) element

It is the most commonly used element in ECM and represents a parallel connection of a resistor and CPE element. Its impedance reads:

$$Z_{RQ}(j\omega) = \frac{R}{1 + (j\omega)^{\alpha}QR} = \frac{R}{1 + (j\omega\tau)^{\alpha}},$$

where $\tau = (QR)^{\frac{1}{\alpha}}$.

Gerischer element

The Gerischer can be derived from diffusion with a 'sink term' in Fick's second law. The meaning is that the electrochemically active species is reacting away in the diffusion process to an inactive complex or absorbed species [68]. The impedance is given by:

$$Z_{GER}(j\omega) = \frac{R}{\sqrt{1 + j\omega QR}}$$

A more general form is referred to as the Cole-Davidson element. Its impedance reads:

$$Z_{\text{Cole-Davidson}}\,(j\omega) = \frac{R}{(1 + (j\omega)QR)^{\alpha}}$$

where $\alpha \in (0, 1]$ and represents the fractional order coefficient.

Generalised Warburg element

It appears useful in the cases where charge transfer in the cell is influenced also by diffusion to and from the electrode. There are two important equivalent circuit elements that describe finite diffusion finite length Warburg (FLW), which is typically used to describe impedances of fuel cells and the finite space Warburg (FSW), which is typically used to describe impedances of Li-ion batteries. Sometimes, FLW and FSW are called the short

and open Warburg elements, respectively. The impedance equation of FLW and FSW is given as follows:

$$Z_{FLW}(j\omega) = \frac{Z_0}{\sqrt{j\omega\tau}} \tanh \sqrt{j\omega\tau}$$
$$Z_{FLS}(j\omega) = \frac{Z_0}{\sqrt{j\omega\tau}} \coth \sqrt{j\omega\tau}$$

2.3 Variational Bayes Approach

To keep in line with the unified probabilistic approach to the goals of the dissertation, a fast and accurate method for deriving the distributions of SOC system model parameters is needed. The primary candidates appear to be MCMC and the VB approach. Since the complexity of MCMC method increases exponentially with the amount of data, we deem it inappropriate for online implementation of the methods. Therefore, we pursued the VB approach, which is computationally much less demanding while yet sufficiently accurate.

2.3.1 Bayesian theorem and Bayesian inference

Bayes' theorem, named after an 18th-century British mathematician, Thomas Bayes, is essentially a mathematical formula for determining conditional probability. Conditional probability is the likelihood of an event occurring based on previous events under similar circumstances. It provides a way to update the existing beliefs about events or theories given new evidence. The result of the theorem is called Bayes' rule [69], [70]. The theorem can be written as

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}, \tag{2.2}$$

where $P(A|B)$ denotes the probability of event A occurred given that event B has occurred, $P(B|A)$ denotes the probability of event B occurring given that event A has occurred, $P(A)$ is the probability of event A occurring, and $P(B)$ is the probability of event B occurring.

Bayesian inference Bayesian inference is a statistical inference method that uses Bayes' theorem to update the probability of an event under certain circumstances as more information is provided. Starting with some prior knowledge that can range from being wholly uninformative or accurate, the probability is updated to acquire the posterior distribution of the event occurring under similar circumstances. The resulting posterior distribution is the derived result of the Bayesian inference.

In the context of model-based inference, we can replace A with the vector of parameters θ and B with the data D. In addition, we can include an additional assumption (knowledge) I to the inference, which can be, for example, which model to use. With that in mind, the Bayes' theorem can now be stated as follows:

$$\overbrace{P(\theta|D,I)}^{\text{posterior}} = \frac{\overbrace{P(D|\theta,I)}^{\text{likelihood}}\overbrace{P(\theta|I)}^{\text{prior}}}{\underbrace{P(D|I)}_{\text{marginal likelihood}}} \tag{2.3}$$

Before incorporating the data, the prior distribution $P(\theta|I)$ specifies our assumption regarding the parameters θ. The likelihood $P(D|\theta,I)$ represents the probability of the data given specific parameters θ. The marginal likelihood (or evidence) $P(D|I)$ reflects the data distribution D under the additional assumption I. It is the normalization factor in Bayes' rule and is crucial for model comparison. Finally, the posterior $P(\theta|D,I)$ denotes

the distribution of the parameters after considering the observed data D and our prior assumption I. We can also express the posterior as directly proportional to the likelihood and the prior.

$$posterior \propto likelihood \times prior$$

If the D and I are independent the (2.3) becomes:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{2.4}$$

2.3.2 Variational Bayes Approach

The presented inference approach revolves around the Bayes rule (2.3). The bottleneck of the Bayesian rule is the evaluation of the marginal likelihood $p(D)$, which in the case of the probability density functions is written as $p(D)$. Imagine the vector θ is inferred from the data D, where the prior is $p(\theta)$ and the likelihood $p(D|\theta)$. Then the marginal likelihood $p(D)$ is solution of the following equation:

$$p(D) = \int_{\theta} p(D|\theta)p(\theta)d\theta. \tag{2.5}$$

For multi-dimensional cases (2.5) becomes computationally intractable. Consequently, obtaining the posterior in (2.3) also becomes unfeasible, hence the need for approximative methods, where VB approach is a strong option. The VB approach does suffer from overconfidence, i.e., underestimation of confidence region, as presented in [71, (5.4)], [72]. This trade-off for the low numerical complexity needs to be taken into account, but the benefits are too great to tarnish the usability of VB methods.

2.4 Probabilistic Spectral Decomposition by Means of the Variational Bayes Approach

2.4.1 The problem setup

Let us put the equivalent circuit model parameters into a vector $\boldsymbol{\theta}$. The data D is a set of evaluated impedance measurements $D = \{Z(j\omega_1), ..., Z(j\omega_N)\}$ where $Z(\omega_i)$ is impedance evaluated at the frequency ω_i. To infer the model's latent parameters $\boldsymbol{\theta}$ from the measured impedance spectra $\mathbf{D}$ we use Bayes' rule

$$p(\boldsymbol{\theta}|\mathbf{D}) = \frac{p(\boldsymbol{\theta}, \mathbf{D})}{p(\mathbf{D})}, \tag{2.6}$$

where $p(\dots)$ denotes probability (density) distributions. VB offers an approximate solution to the problem of calculating the posterior distribution $p(\boldsymbol{\theta}|\mathbf{D})$ in which the model parameters are regarded as latent [73]. The idea is demonstrated in various examples in [74]. Since Bayes' rule (2.6) requires calculating $p(D)$, which is usually intractable, we instead seek an adequate parameterized distribution over the latent variables $q_\phi(\boldsymbol{\theta})$, which resembles the true posterior while being easier to handle due to parametrization ϕ. To minimize the discrepancy between $p(\boldsymbol{\theta}|\mathbf{D})$ and $q_\phi(\boldsymbol{\theta})$, an optimum parameter ϕ, denoted as $\hat{\phi}$, must be found. To define the optimization problem, a loss function, also called the Kullback-Leibler divergence, denoted as KL(..||..), is introduced

$$\hat{\phi} = \arg\min_{\phi} \mathrm{KL}\left(q_\phi(\boldsymbol{\theta})||p(\boldsymbol{\theta}|\mathbf{D})\right). \tag{2.7}$$

We expand the loss function

$$\begin{aligned}\mathrm{KL}\left(q_\phi(\boldsymbol{\theta})||p(\boldsymbol{\theta}|\mathbf{D})\right) &= \mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log\frac{q_\phi(\boldsymbol{\theta})}{p(\boldsymbol{\theta}|\mathbf{D})}\right] \\ &= \mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log q_\phi(\boldsymbol{\theta})\right] - \mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log p(\boldsymbol{\theta}|\mathbf{D})\right] \\ &= \mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log q_\phi(\boldsymbol{\theta})\right] - \mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log p(\boldsymbol{\theta},\mathbf{D})\right] + \log p(\mathbf{D}) \\ &= \underbrace{-\mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log\frac{p(\mathbf{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{q_\phi(\boldsymbol{\theta})}\right]}_{=\,\mathcal{L}(\phi;\,\mathbf{D})} + \log p(\mathbf{D}),\end{aligned} \tag{2.8}$$

to obtain the definition of ELBO, denoted as $\mathcal{L}(\phi;\ \mathbf{D})$. Since $\log p(\mathbf{D})$ is a constant, we can redefine the optimization as

$$\hat{\phi} = \arg\max_\phi \mathcal{L}(\phi;\ \mathbf{D}) = \arg\max_\phi \mathbb{E}_{q_\phi(\boldsymbol{\theta})}\left[\log p(\mathbf{D}|\boldsymbol{\theta})\right] - \mathrm{KL}\left(q_\phi(\boldsymbol{\theta})||p(\boldsymbol{\theta})\right). \tag{2.9}$$

It is easy from (2.9) to see that in order to maximize $\mathcal{L}(\phi;\ \mathbf{D})$, one should maximize the expected log-likelihood $\log p(\mathbf{D}|\boldsymbol{\theta})$ while keeping the posterior $q_\phi(\boldsymbol{\theta})$ as close as possible to the prior $p(\boldsymbol{\theta})$, thus imposing a regularisation-like property, see [75, (Eq. 10.15)]. The overall idea is compactly presented in Figure 2.3.

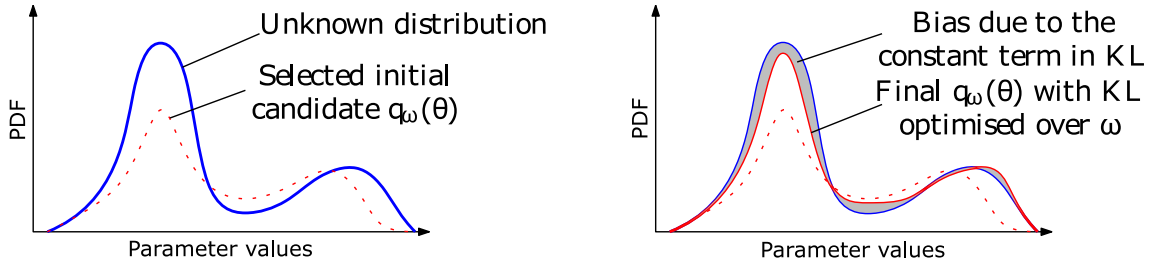


Figure 2.3: The illustration of the optimization process for finding the closest variational distribution $q_\omega(\theta)$ over the set of latent variables ω. The inherent bias of the VB approach is also presented.

2.4.2 Minimisation of the Kullback-Leibler divergence

With the help of VB, we convert the derivation of the posterior distributions from a statistical inference problem to a much simpler task of optimizing a cost function for the variational family parametrized by latent variables λ. The cost function is, in this case, ELBO. For the optimization part of our algorithm, we use adaptive moment estimation (ADAM). ADAM is one of the most prominent optimization algorithms in the research community. ADAM is a gradient descent-based optimization algorithm that was presented in 2014 by Diedrik P. Kingma and Jimmy Ba in [76]. The optimization framework of ADAM consists of adjusting the learning rate adaptively for each parameter of the model, where the adjustment is based on the history of gradients calculated for that parameter. This allows the optimizer to converge faster and produce more accurate results compared to methods with a fixed learning rate.

The ADAM algorithm ADAM consists of adjusting the learning rate adaptively for each parameter of the model, where the adjustment is based on the history of gradients

calculated for that parameter. This allows the optimizer to converge faster and produce more accurate results compared to methods with a fixed learning rate.

ADAM provides an adaptive correction of learning rates for the s^{th} element $\theta^{(s)}$ of the vector $\boldsymbol{\theta}$ during the run of the algorithm. The gradient vector $\boldsymbol{g}(t)$ is treated as a random vector. For the s^{th} element of the vector, the algorithm computes the first and second moments, denoted $m_t^{(s)}$ and $v_t^{(s)}$, respectively. From now on, m_t, v_t, and g_t refer to $m_t^{(s)}$, $v_t^{(s)}$, and $g_t^{(s)}$, respectively.

This is done by using exponentially moving averages computed on the gradient, which is evolved on a current batch:

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1-\beta_1) g_t, \\ v_t &= \beta_2 v_{t-1} + (1-\beta_2) g_t^2, \end{aligned} \tag{2.10}$$

where t denotes the number of iterations and β_1, β_2 are the parameters of the moving average. The default values for the β parameters are 0.9 and 0.999, and the initial values of the first and second moments of the gradient are both set to 0, since it turns out this does not impact the convergence rate in a significant manner. For some data window x_w, the gradients can be presented with the following vector:

$$\boldsymbol{g} = \left(g^{(1)}, \ldots, g^{(K)}\right)^T = \nabla_\theta L_{x_w}(\theta). \tag{2.11}$$

First and second moments are only estimated with m and v; therefore, we want them to satisfy the following conditions:

$$\begin{aligned} E\{m_t\} &= E\left(g_t\right), \\ E\{v_t\} &= E\left(g_t^2\right). \end{aligned} \tag{2.12}$$

The above conditions assure us that we are dealing with unbiased estimates. The first moment at step t in the recursive equation (2.10) is

$$m_t = (1-\beta_1) \sum_{i=1}^{t} \beta_1^{t-i} g_i. \tag{2.13}$$

We can see some bias still occurs in this estimate. Applying the expected value operator to equation (2.13) gives us

$$\begin{aligned} E_\theta\left[m_t\right] &= (1-\beta_1) \sum_{i=1}^{t} \beta_1^{t-i} E_\theta\left[g_i\right] \\ &\approx (1-\beta_1) \left(\sum_{i=1}^{t} \beta_1^{t-i}\right) E_\theta\left[g_t\right] \\ &= (1-\beta_1^t) E_\theta\left[g_t\right]. \end{aligned} \tag{2.14}$$

Bias correction is done automatically by ADAM during the evaluation of m and v as follows:

$$\begin{aligned} \hat{m}_t &= \frac{m_t}{1-\beta_1^t}, \\ \hat{v}_t &= \frac{v_t}{1-\beta_2^t}. \end{aligned} \tag{2.15}$$

Optimization steps are also adjusted. For each individual parameter, the algorithm updates

$$\theta_t^{(s)} = \theta_{t-1}^{(s)} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \tag{2.16}$$

where $\eta = 0.001$.

The whole idea is presented in Figure 2.4.

The use of ADAM spreads upon a wide range of applications, providing efficient and fast results.

However, it should be noted that, being a heuristic optimisation algorithm, there is no general guarantee for convergence. In [76] proof is provided that the algorithm converges globally in the convex settings. This was further refined and proved by [77].

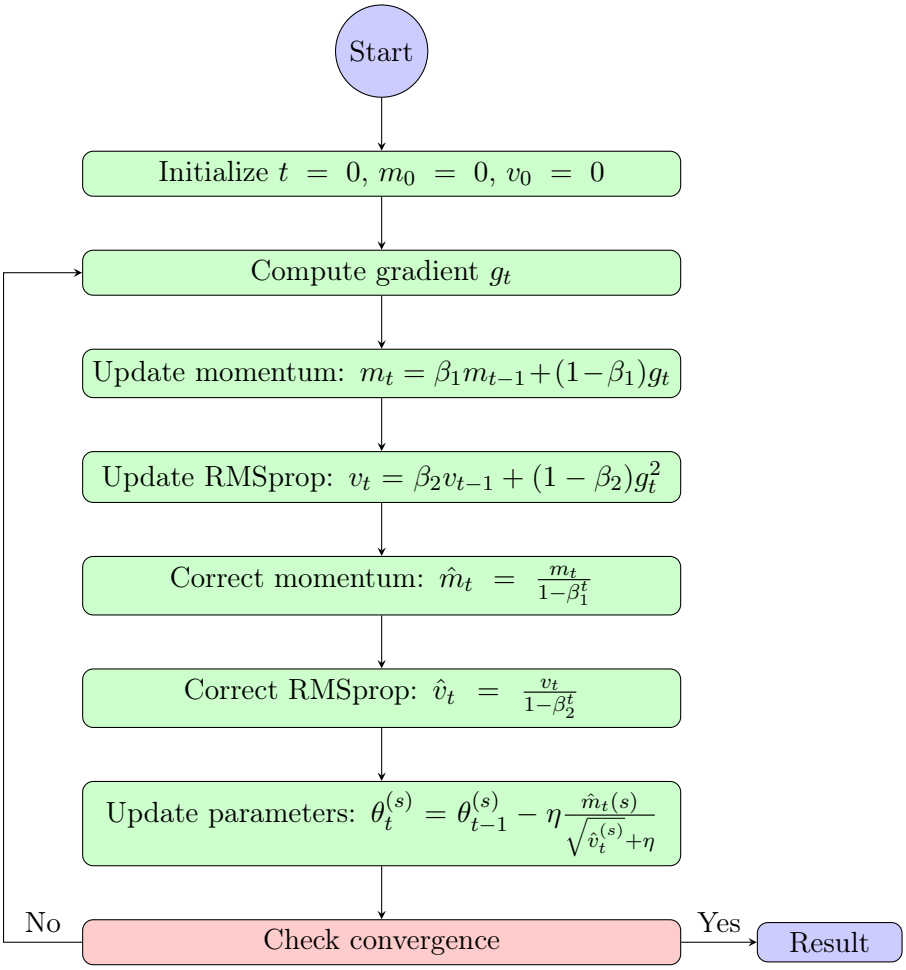


Figure 2.4: Adam Optimization Algorithm flowchart.

Areas of use and main benefits of the ADAM algorithm ADAM is widely used for a reason. It offers fast and accurate results with low numerical complexity. The main advantages over other optimization options are the following:

- Adaptive learning rates: The main difference between the main competitors (e.g., Stochastic Gradient Descent (SGD)) and ADAM is that the latter provides adaptive learning rates for each of the model parameters, which are based on the history of their gradients. This property allows it to converge faster and more accurately towards the global minimum. This is especially noticeable in high-dimensional parameter spaces.
- Momentum: Smoothing out the fluctuations in the optimization using momentum helps ADAM to avoid local minima and saddle points.

- Bias correction: The optimization automatically corrects the bias of first and second-moment estimates to ensure they are, in fact, unbiased estimates of the true values.
- Robustness: The optimization performs well over a wide range of deep learning architectures and is relatively robust to hyper-parameter choices.

The complexity of the ADAM algorithm can be discussed in three parts:

Computational Complexity The computational complexity of ADAM can be evaluated by considering the operations involved in each parameter update. Gradient calculation is generally the most computationally intensive part, with its complexity varying based on the model architecture and the dataset. In neural networks, this involves backpropagation, which requires calculating the gradients of the loss function concerning each parameter. The operations required to update the first and second moments, as well as the bias correction, scale linearly with the number of parameters n.

For each parameter, Adam performs two multiplications and one addition for updating m_t and v_t. Additionally, it requires one division each for computing the bias-corrected estimates $\hat{m}_t$ and $\hat{v}_t$, one square root operation, and one division, one addition, and one subtraction for the final parameter update. Consequently, the per-parameter complexity is $O(1)$, resulting in an overall complexity of $O(n)$ for n parameters. This efficient handling of computational complexity is one reason for Adam's popularity, as it ensures scalability even for large models.

Memory Requirements Adam necessitates the storage of additional parameters for each weight, specifically the first-moment estimate m_t and the second-moment estimate v_t. This requirement effectively triples the memory needed compared to vanilla stochastic gradient descent (SGD), as it must store the weights along with the first and second-moment estimates. While this increased memory usage is a trade-off, it allows Adam to maintain a more sophisticated understanding of the parameter updates, leading to more stable and efficient training processes. In practice, the memory overhead is manageable for most modern hardware configurations, but it can be a consideration for very large models or when running on limited resources.

Design parameters The design parameters of the Adam optimization algorithm introduce a level of complexity due to the necessity of fine-tuning several hyperparameters to achieve optimal performance. These parameters include the learning rate (α), which dictates the step size for each iteration, and the exponential decay rates for the moment estimates, β_1 and β_2, which control the moving averages of the first and second moments of the gradients. Additionally, ϵ is a small constant added to the denominator to improve numerical stability. The interplay between these parameters significantly influences the convergence behaviour and generalization ability of the algorithm, making their selection critical and often dependent on the specific task or dataset at hand.

Toy example for the ADAM algorithm In this example, we aim to minimize the quadratic function $f(x, y) = x^2 + y^2$ using the Adam optimization algorithm in PyTorch with a learning rate of 0.025.

We start by initializing random values for the parameters x and y, marking them as requiring gradients. We then perform a series of optimization steps to minimize the quadratic function. In each step, we evaluate the function value, compute gradients, and update the parameters x and y using the Adam optimizer.

The results of the optimization process are visualized in figures 2.5 and 2.6.

Figure 2.5 shows the value of the loss function $f(x, y)$ as a function of optimization iteration. The x-axis represents the optimization iteration number, while the y-axis represents the loss value. The plot illustrates how the loss decreases over-optimization iterations as the algorithm converges toward the optimal solution. Initially, the loss may be high, but as the optimization progresses, the loss decreases, indicating that the algorithm is approaching the optimal solution. It might seem like the algorithm needs many iterations to reach the optimal solution, but this is highly dependent on the learning rate set initially. In this toy example, the algorithm's hyperparameters were not optimized for the fastest convergence since it was not necessary.

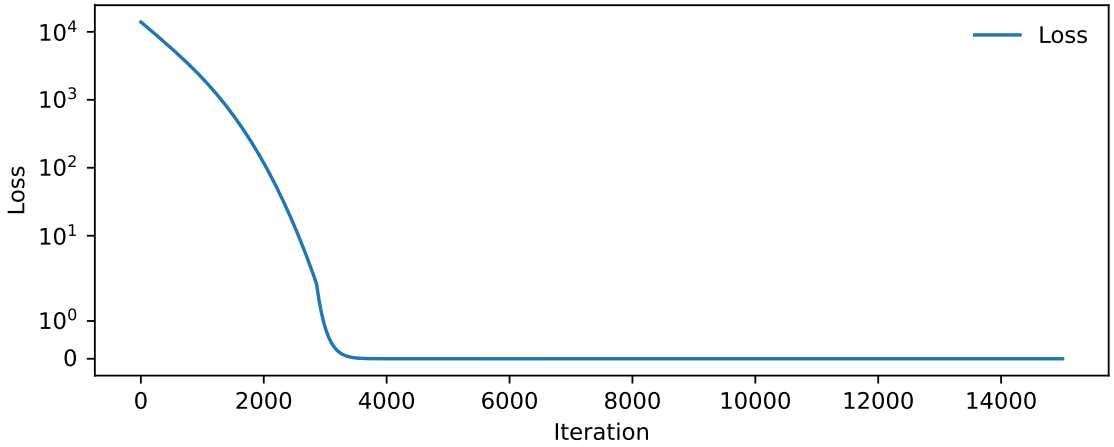


Figure 2.5: The loss function values over iterations for the ADAM optimization on the simulated example. We can see that the optimizer converges.

Figure 2.6 visualizes the quadratic function in a 2D parameter space. Contour lines represent the values of the loss function, with each line indicating a specific loss value. The red cross marks the global minimum of the function, representing the optimal solution. The black circle indicates the starting point of the optimization process, where we initialize the parameters x and y. The dashed blue line represents the optimization path, illustrating how the parameters x and y evolve over the optimization iterations as the algorithm searches for the optimal solution.

2.4.3 The MCMC algorithm

A well-known staple in the Bayesian inference area, the MCMC method is widely used throughout the field. The basic idea behind MCMC is to sample from a probability distribution by constructing a Markov chain with the desired distribution as its equilibrium distribution. Through this process, the algorithm produces very accurate results by allowing for the approximation of complex integrals and posterior distributions that are otherwise analytically intractable. However, MCMC does suffer from the curse of dimensionality, as problems can become computationally untractable with datasets of high dimensions.

Strengths and weaknesses of MCMC MCMC methods are utilized in a wide range of applications. In Bayesian statistics, they are employed to estimate the posterior distributions of model parameters. In machine learning, MCMC plays a crucial role in algorithms such as Bayesian neural networks and probabilistic graphical models. The field of statistical physics uses MCMC for simulating the behavior of systems at thermal equilibrium.

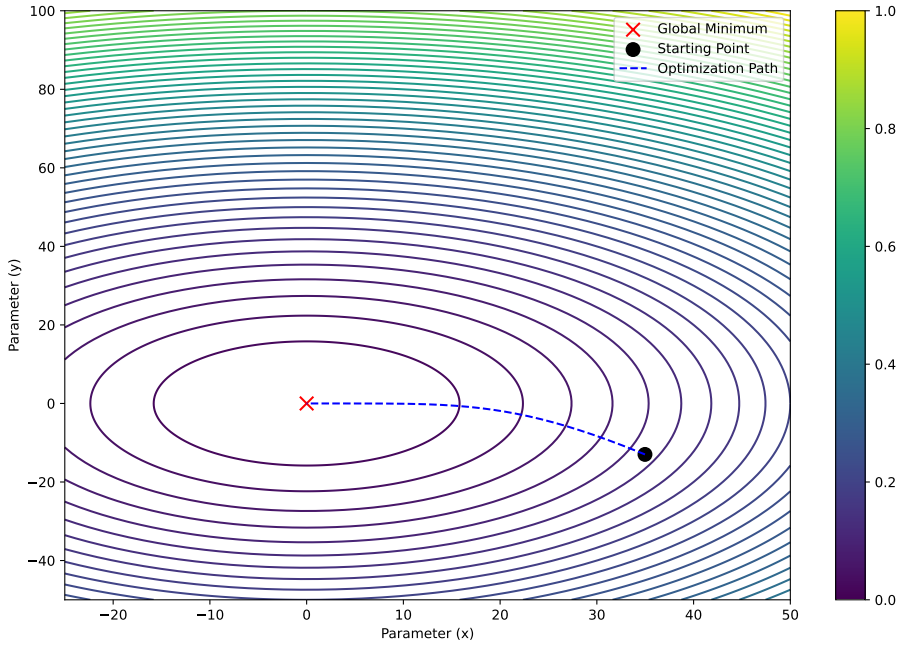


Figure 2.6: Visually presented optimization steps of the ADAM algorithm for the presented toy example. The starting spot is presented with a black dot, and the final optimized values are presented with a red x mark.

Additionally, in genetics, MCMC aids in the analysis of genetic data and evolutionary biology studies. Econometrics benefits from MCMC in Bayesian econometric models, including time-series analysis, while bioinformatics applies MCMC methods to phylogenetic tree construction and other computational biology applications.

The flexibility of MCMC methods allows them to be applied to a wide range of problems with different types of distributions and models. They provide highly accurate approximations for posterior distributions and complex integrals. With appropriate algorithms and computational resources, MCMC can handle moderately high-dimensional problems. Additionally, given sufficient iterations and the right conditions, MCMC algorithms are guaranteed to converge to the target distribution.

However, MCMC methods are not without their drawbacks. As the dimensionality of the data increases, MCMC can become computationally expensive and slow to converge, a phenomenon known as the curse of dimensionality. Proper tuning of parameters, such as step size and number of iterations, is crucial for effective performance, which can be a challenging task. Diagnosing the convergence of the Markov chain to the target distribution can also be difficult. The computational cost in terms of time and resources can be high, especially for large datasets or complex models. Furthermore, the initial phase, known as the burn-in period, where the chain has not yet converged, must be discarded, increasing the required computation.

2.4.4 Numerical example of VB approach on simulated SOFC data

For the numerical simulation, the following model with $N = 3$ RQ elements was selected:

$$Z(\omega) = R_s + \sum_{i=1}^{N} \frac{R_i}{(j\omega)^{\alpha_i} Q_i R_i + 1} + j\omega L \tag{2.17}$$

The model (2.17) was simulated on the frequency interval $f \in [10^{-4}, 10^{4}]$ Hz with parameters listed in Table 2.1. The value of inductance L is taken as low as 100 nH. Zero-mean Gaussian noise with standard deviation $\sigma = 0.17$ was added in order to simulate the measurement noise. The simulated Nyquist curve is shown with black dots in Figure 2.7.

Table 2.1: Model parameters and prior distributions calculated using equations (2.18) and (2.19).

ECM parameter	Prior distribution
$R_s = 3\ \Omega$	$\sim$ Lognormal$(0.84, 0.39)$
$R_1 = 1\ \Omega$, $R_2 = 2\ \Omega$, $R_3 = 3\ \Omega$	$\sim$ Lognormal$(1.59, 0.20)$
$Q_1 = 0.1\ Fs^{\alpha_1}$	$\sim$ Lognormal$(-0.35, 0.83)$
$Q_2 = 5\ Fs^{\alpha_2}$	$\sim$ Lognormal$(1.96, 0.83)$
$Q_3 = 150\ Fs^{\alpha_3}$	$\sim$ Lognormal$(4.99, 0.55)$
$\alpha_1 = 0.88$, $\alpha_2 = 0.82$, $\alpha_3 = 0.99$	$\sim$ Beta$(13.91, 5.68)$

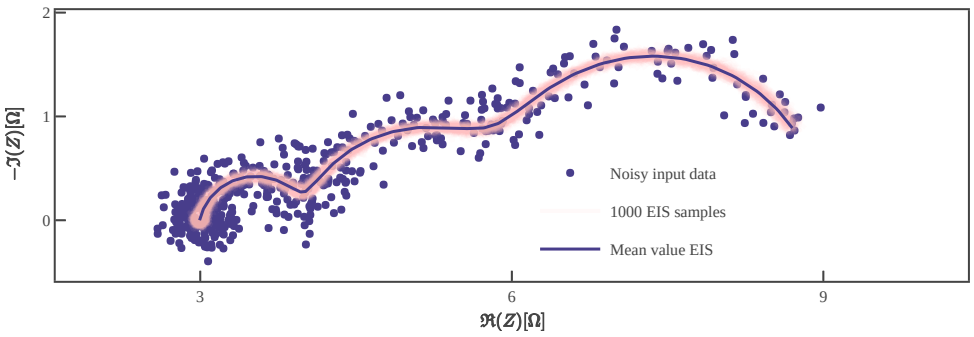


Figure 2.7: The simulated EIS values used for the simulated example problem is presented with purple dots, which represent noisy EIS measurements. The spread of the EIS curves calculated from the posterior distributions of the ECM parameters can be seen in light pink, and the purple line represents the EIS curve derived from the means of the same distributions.

The first step in applying the VB approach is the definition of prior distributions. This can be done in two ways: either by using uninformative priors or by introducing prior knowledge. When lacking manufacturer data, the parameters of the prior distributions can be inferred from the shape of the measured EIS curve.

This line of thought can be demonstrated in the present numerical example. Let us assume that the true parameter values are unknown. Looking at Figure 2.7 one can first notice that the real part of the curve "starts" around 3 Ω. This provides a rough estimate for the parameter R_s thus selecting prior with mean 2.5 Ω and variance 1 seems a reasonable choice. The next step is setting the priors for the parameters $R_i, i \in \{1, 2, 3\}$. Just by looking at the axis values, it is apparent that they should be located within the interval (1,10) Ω. The mean value of their distributions was, therefore, set in the middle, with the scale being sufficiently large to give the algorithm enough space for search. Priors of the parameters $Q_i, i \in \{1, 2, 3\}$ can be assessed by taking into consideration the frequency bandwidth and the location of the arcs. Lastly, the fractional order powers $\alpha_i, i \in \{1, 2, 3\}$ are expected to have values in the interval [0.5,1]. Following this rationale, the rough

estimates of informative priors can be assessed in an efficient manner. Another important step is selecting the type of distribution that fits the parameters best. For the system 2.17, the parameters must all be positive, with an additional condition for α_i $i \in \{1, 2, 3\}$ to be between $[0, 1]$. Beta distribution was therefore chosen as the best fit for the α parameters and Log-normal distribution for the rest of the parameters.

Since we determined our priors as if they would follow a normal distribution, we had to transform them to fit the ones we chose. Obtaining lognormal distribution parameters (σ_{ln}, μ_{ln}) from $\mathcal{N}(\sigma, \mu)$ can be done by using the following equations:

$$\sigma_{ln} = \ln\left(\frac{\sigma^2}{\sqrt{\sigma^2 + \mu^2}}\right) \qquad \mu_{ln} = \ln\left(1 + \frac{\mu^2}{\sigma^2}\right) \tag{2.18}$$

and for beta distribution parameters α and β e.g. Beta(α, β) we use:

$$\alpha = \frac{\sigma^2 - \sigma^3}{\mu} - \sigma \qquad \beta = \frac{\alpha}{\sigma} - \alpha \tag{2.19}$$

Distributions with their "normal" values are presented in Table 2.1.

Optimization was performed using the global learning rate of 0.05. The optimization was completed in 20,000 steps. By sampling the resulting posterior distributions and simulating the model (2.17), it is possible to visualize the obtained results. For that purpose, 1000 samples were drawn for the posterior distributions and are shown in Figure 2.7. The EIS curve simulated from the posteriors' mean values is presented in blue colour. The variance of the posterior distributions is presented with the sampled curves and the noise variance parameter bounds. The mean value EIS curve appears to fit the input data well.

Taking a look at ELBO loss plot reveals some inner performance of the algorithm. In Figure 2.8, we can see that the plot smooths out nicely as the algorithm progresses. Local abrupt "jumps" in the values of the ELBO loss can be associated with the ever-adapting learning rates of the ADAM optimizer and possible shifts in searching space for the parameters.

Decreasing ELBO curve with a relatively low amount of jumps means that the algorithm is steadily converging towards the optimal values for the parameters.

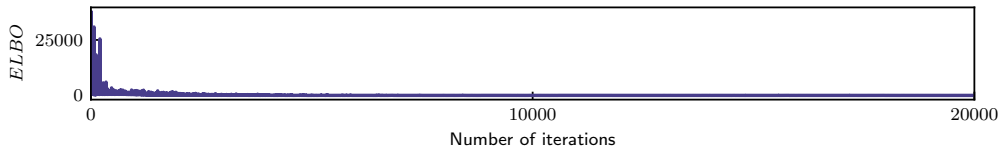


Figure 2.8: The ELBO loss over the iterations of the VB approach for the simulated example. We can observe the obvious convergence of the optimization algorithm.

The evolution of the optimization process for each separate parameter is presented in Figure 2.9. The true parameters we used to simulate our input data are represented by a light-blue dashed line. Progress in the algorithm's confidence can be easily observed with the help of the evolution of scales (variance) for parameter distributions, shown in light pink. As seen in the Figure 2.9, at the beginning, the scales[1] of the posterior distributions are wide. As the optimization progresses and explores a larger search space, the scale steadily decreases. Note here also that the variance of the overall noise parameter was estimated at 0.168, which is very close to the simulated value of 0.17.

[1]The scale parameter of a probability distribution function is closely related to the variance.

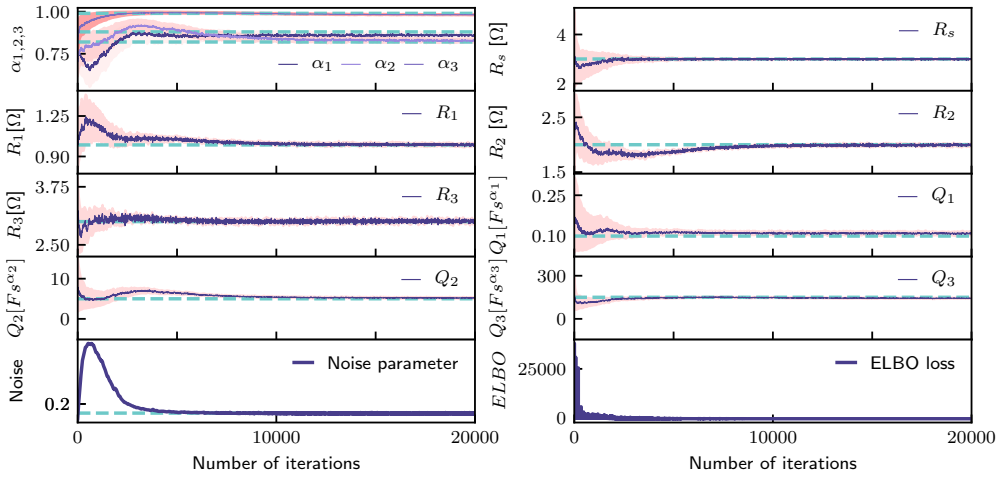


Figure 2.9: Progress of the estimation for each parameter over the iterations. The means of the parameters are presented with dark, hard lines, while their variance is shown in light, soft colors. The actual parameter values used for the simulation are marked with light blue dashed lines. We can see that the parameters converge nicely toward their actual values, meaning that the VB approach correctly estimated the values of the ECM parameters.

Comparison with the MCMC algorithm

The MCMC method is widely used throughout the field. The basic idea behind MCMC is to sample from a probability distribution by constructing a Markov chain with the desired distribution as its equilibrium distribution. Through this process, the algorithm produces very accurate results by allowing for the approximation of complex integrals and posterior distributions that are otherwise analytically intractable. However, MCMC does suffer from the curse of dimensionality, as problems can become computationally untractable with datasets of high dimensions. The first application of the MCMC to the parameter estimation of the ECM of SOFC was proposed in [78].

The resulting posterior distributions of the model parameters from subsection 2.4.4 as well as the posteriors obtained with MCMC are shown in Figure 2.10. Three key observations can be made. First, comparing posterior and prior distributions, the optimization process was capable of reaching the "true" mean values even for cases when the prior means were "far" from the simulated values. Secondly, the scales (variance) of the posterior distributions are low, leading to the conclusion that the obtained model has low uncertainty. Table 2.2 lists all resulting posterior parameter distributions transformed back to normal distributions for easier discussion. Finally, the empirical posteriors obtained through MCMC method are similar to the ones obtained by VB. It should be noted that the VB posteriors are under-dispersed (lower variance) than the ones obtained by MCMC. This is a known effect since the entropy of the VB posterior is always lower or equal than the "true" posterior [79], [80]

Comparison of MCMC and VB VB and MCMC are both methods used for approximating posterior distributions in Bayesian inference, each with distinct strengths and weaknesses suitable for different scenarios. VB is generally more scalable to large datasets and high-dimensional models because it often converges faster than MCMC, which can be computationally intensive and slow, especially for complex models with many parameters. Additionally, VB is usually faster since it converts the inference problem into an

Table 2.2: Posterior distributions of parameters.

$R_s \sim \text{LN}(1.097, 0.002)$		
$R_1 \sim \text{LN}(-9.8e^{-4}, 0.012)$	$R_2 \sim \text{LN}(0.69, 0.009)$	$R_3 \sim \text{LN}(1.1, 0.012)$
$Q_1 \sim \text{LN}(-2.22, 0.05)$	$Q_2 \sim \text{LN}(1.64, 0.025)$	$Q_3 \sim \text{LN}(4.97, 0.017)$
$\alpha_1 \sim \text{Beta}(436.15, 71.59)$	$\alpha_2 \sim \text{Beta}(1117.65, 227.44)$	$\alpha_3 \sim \text{Beta}(2168.12, 40.78)$

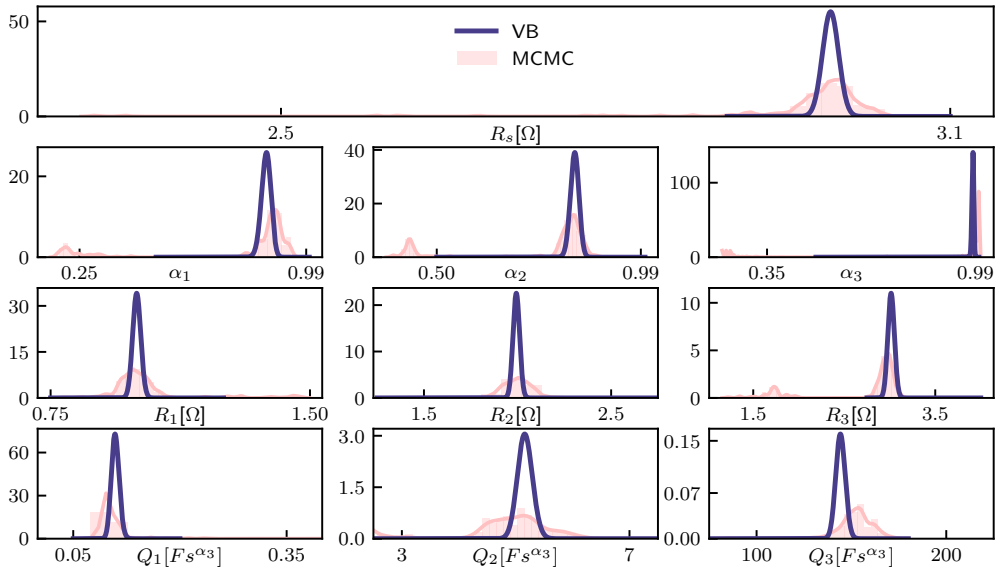


Figure 2.10: The posterior distributions of ECM parameters derived from simulated data are presented with dark purple lines, while the MCMC estimations of the same parameters are shown in pink for comparison. It is evident that the results are comparable, but the overconfidence in the VB approach results is clearly visible.

optimization problem solvable by efficient deterministic optimization algorithms. VB provides deterministic approximations, meaning repeated runs on the same data yield the same results, which is advantageous for reproducibility and debugging. Moreover, VB typically converges faster than MCMC, which might require a large number of samples to approximate the posterior accurately.

For complex models with intractable posteriors, VB can approximate such posteriors by transforming the problem into an optimization framework, making it more efficient and practical than MCMC. Furthermore, VB can be adapted for online learning scenarios where data arrives sequentially, with stochastic variational inference (SVI) allowing VB to update posterior estimates in real-time, a challenging task for MCMC.

Finally, VB can be simpler to implement for certain models, particularly when leveraging variational autoencoders or other existing frameworks. In contrast, MCMC might require more sophisticated techniques to ensure proper mixing and convergence.

In summary, VB is preferable when faster, scalable, and deterministic approximations are needed, particularly with large datasets, high-dimensional models, real-time inference, or complex models with intractable posteriors. MCMC might provide better results when exact sampling is crucial, the computational cost is manageable, and detailed posterior understanding is required.

2.5 The Application of Variational Bayes Approach on Experimental SOFC data

2.5.1 Process description

The VB based approach was validated using EIS curves measured on six planar anode-supported cells, installed in an insulated ceramic housing. The active area of a single cell was 80 cm^2 [63]. The SOFC short stack was fed with a mixture of hydrogen and nitrogen with a flow rate H_2/N_2=0.216/0.144 Nl $h^{-1}cm^{-2}$ whereas the air flow rate was 4 Nl $h^{-1}cm^{-2}$. Stack was operated at nominal current of 32 A (0.4 A cm^{-2}) with FU=77.5 %.

The EIS curves were obtained by inducing current excitation (galvanostatic mode). The excitation current had a DC value of I_{DC}=32 A with a peak-to-peak amplitude of 2 A. The amplitude was chosen to stay in the linearity region but still be big enough to maintain a sufficient signal-to-noise ratio.

The excitation was imposed using an electronic load, Rigol DL3031A, connected in series with the SOFC. The electrical current was measured using a Hall probe sensor LEM HASS 50-S. Such a measurement technique has no influence on the performance of the system under test and offers sufficient accuracy. Furthermore, the selected sensor has a sufficiently wide frequency bandwidth with a cut-off frequency of 240 kHz. The cell voltages were measured independently using a differential 16-bit NI USB-6215 data acquisition system. The analogue signals were firstly low pass filtered at 10.8 kHz and sampled with sampling frequency $f_s = 50$ kHz.

2.5.2 Experimental results

The measured EIS curve is shown in Figure 2.11 with dark blue dots. The noise level in the measured EIS curve is quite low. The first step in the identification process is to select the model structure, i.e., the number of poles. In the context of fractional order systems, the problem of structural identification is still an open one [81]. Since this issue is beyond the scope of this approach, a fallback solution is to employ background knowledge to determine the number of poles in the system.

For the type of the SOFC system used in this analysis, it was determined that an ECM with three poles would suffice. Consequently, the model structure was selected to be the same as in (2.17) with $N = 3$. Prior distributions were selected similarly as for the case of the numerical example by examining the intervals covered by the measured EIS curve. Looking at the Figure 2.11, we can see that the high-frequency part of the real component has values around 3mΩ, which will serve as a mean value for the R_s parameter. Its distribution scale was set quite wide to allow the optimization process to explore wide intervals in the search for the optimum. Next, priors of $R_{1,2,3}$ parameters can be inferred from the complex and real values our EIS curve takes that their values should be between the interval [0.1,8] mΩ. From previous testing, we can assume that with similar R values and visible poles, the $Q_{1,2,3}$ parameters will be at least by an order of 10 apart, so we assumed they are located at $1, 50$ and 500 Fs^{α_1} respectively. The scales for their prior distributions were set respectively large. Lastly, for $\alpha_{1,2,3}$ we can assume from experience that they probably take value between 0.5 and 1.0. All the prior distributions can be seen on Table 2.3.

The learning rate was set to 0.025. This value is lower than in the numerical example. The goal was to allow steady and efficient estimation. The optimization finished in 12500 steps. It should be pointed out that the algorithm seems to be quite robust, and its performance is mostly unaffected by the hyper-parameter values.

Table 2.3: Prior distributions of parameters.

$R_s \sim \text{Lognormal}(-5.86, 0.32)$	$R_{1,2,3} \sim \text{Lognormal}(-8.41, 1.27)$	$\alpha_{1,2,3} \sim \text{Beta}(12, 3)$
$Q_1 \sim \text{Lognormal}(-2.31, 2.14)$	$Q_2 \sim \text{Lognormal}(3.57, 0.83)$	$Q_3 \sim \text{Lognormal}(6.20, 0.20)$

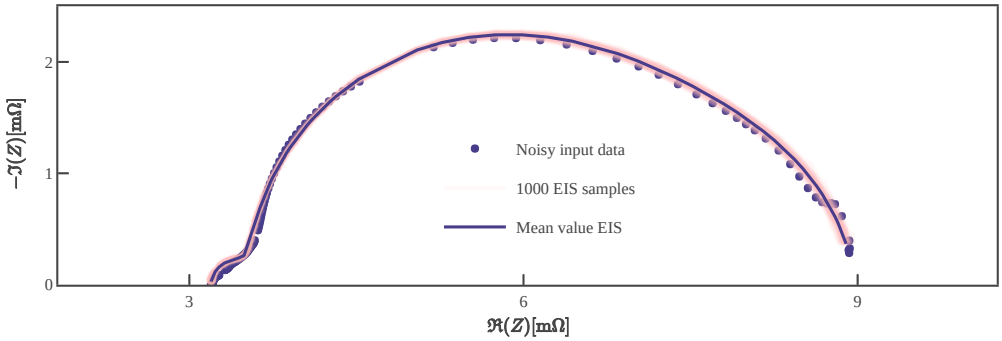


Figure 2.11: The EIS curve obtained from the experiments in a laboratory setting is presented in dark purple dots. We can notice the low noise in the data. The EIS curve spread obtained from the posterior distributions of the ECM parameters from the VB approach can be seen in light pink, and the mean EIS curve calculated from the same distributions is shown as a dark purple line.

Algorithm performance is best observed by looking at EIS curves obtained from the posterior distributions of the estimated parameters. A batch of 1000 samples from the inferred distributions is shown in Figure 2.11 with a light pink colour. The EIS curve obtained from the posterior mean values is shown with a blue line. The results confirm that despite the low number of iterations, the resulting parameters represent an accurate fit for the measured EIS data.

The progression of ELBO values can be seen in Figure 2.12. Evidently, the plot smoothed out nicely as we iterated. Compared to the Figure 2.8 we can notice that the smoothing rate is slower, which can be associated with the much lower learning rate and a smaller number of iterations used for the run on experimental data. As we approach the end of the computation, ELBO value changes minimally, which means we have found our optimum for the estimation.

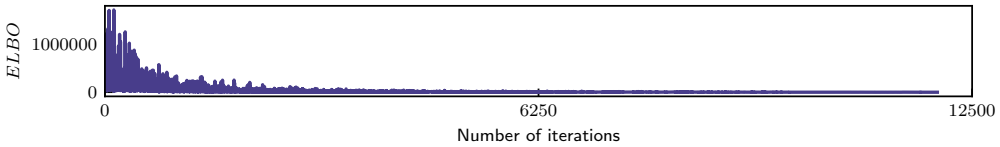


Figure 2.12: Progression of the ELBO loss over iterations of the optimizer for experimental data.

Progress of parameter estimation values is shown in Figure 2.13. The mean values of the distributions can be observed as blue lines, while the variances for each parameter are shown in light pink. The algorithm steadily converges towards the optimum, while the a priori-set distribution scales keep degrading throughout the process, representing

increased confidence in the results. The parameter evolution is very smooth for each parameter. The final shape of the posterior distributions is shown in Figure 2.14. It is apparent that the scales (variance) of the posterior distributions are quite low. Posterior parameter distributions are listed in Table 2.4, where Lognormal is denoted as LN.

Table 2.4: Posterior distributions of ECM parameters.

$R_s \sim \text{Lognormal(LN)}(-5.75, 8.3e^{-4})$		
$R_1 \sim \text{LN}(-7.995, 0.011)$	$R_2 \sim \text{LN}(-6.33, 0.0039)$	$R_3 \sim \text{LN}(-5.61, 0.0075)$
$Q_1 \sim \text{LN}(2.25, 0.0086)$	$Q_2 \sim \text{LN}(6.22, 0.0025)$	$Q_3 \sim \text{LN}(3.96, 0.0028)$
$\alpha_1 \sim \text{Beta}(1081.56, 41.42)$	$\alpha_2 \sim \text{Beta}(2216.36, 187.46)$	$\alpha_3 \sim \text{Beta}(399.28, 0.495)$

2.5.3 Discussion of the results

Comparing prior and posterior parameter distribution values, found in Table 2.3 and Table 2.4 respectively, it is apparent that the algorithm found new optimal values for the parameter means and lowered the variances as it got more confident. This is further emphasized in Figure 2.13. EIS curve simulated from the parameter means appears to be a great fit for the data at hand, while the 1000 sampled curves in Figure 2.11 nicely envelop the data points. Both of those two facts highlight the performance of VB method.

Figure 2.14 shows an interesting trend. First, the variance of the distribution of the R_s parameter is extremely low. This parameter can be assessed quite easily since its value determines the location of the EIS curve. As a result, its prior can be selected particularly close to its true value. Secondly, coupling posterior parameter distributions as pairs $[R_1, Q_1]$, $[R_2, Q_2]$ and $[R_3, Q_3]$, leads to a a similar conjecture. The second pair has a much narrower spread than $[R_1, Q_1]$, while the third couple is somewhere in between.

The pole located at lower frequency seems to be the least important for the complete EIS curve, whereas the other two carry much information about the whole data set, with the second pole at middle frequencies being the most important. It is easy to check that the poles at low, middle, and high frequencies correspond with posterior distributions of parameters $[R_2, Q_2, \alpha_2]$, $[R_3, Q_3, \alpha_3]$ and $[R_1, Q_1, \alpha_1]$, in that particular order. This coincides with the reasoning for parameter R_s.

2.6 Empowering the Baseline Variational Bayes Approach with Mixture Distributions

To tackle the bias of the VB algorithm and to make it more flexible in the sense of approximating as broad a family of potential ECM parameters distributions as possible, we researched the potential of mixture distributions as priors for the ECM parameters. Extending the VB algorithm with mixture distributions provides several benefits.

Firstly, incorporating mixture distributions increases the flexibility of the VB algorithm. This flexibility allows the model to better capture the underlying structure of the data by representing more complex distributions. Secondly, this approach maintains similar computational complexity while achieving the same or better accuracy of results compared to the traditional VB methods. This makes it an efficient alternative without compromising performance.

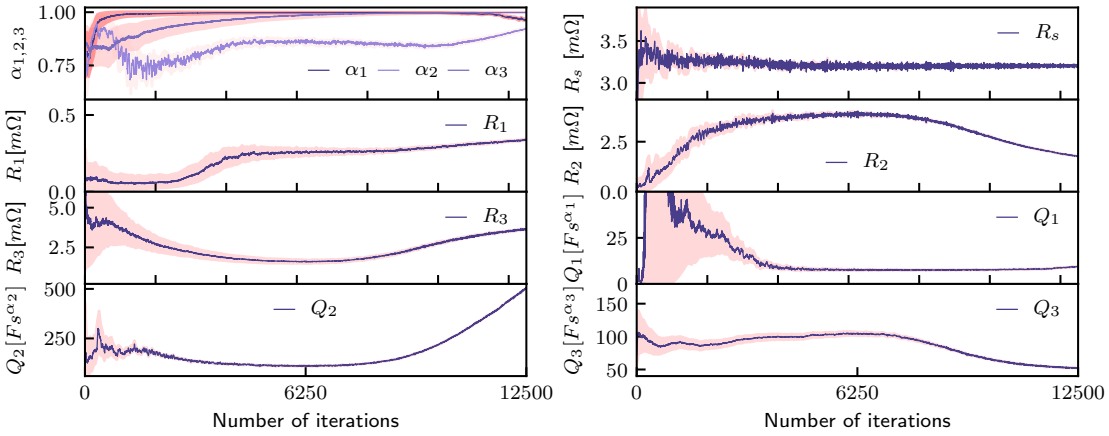


Figure 2.13: Progression of the ECM parameter estimation for experimental data. The means of the parameter distributions are shown in dark purple colors, and their variances are in pink.

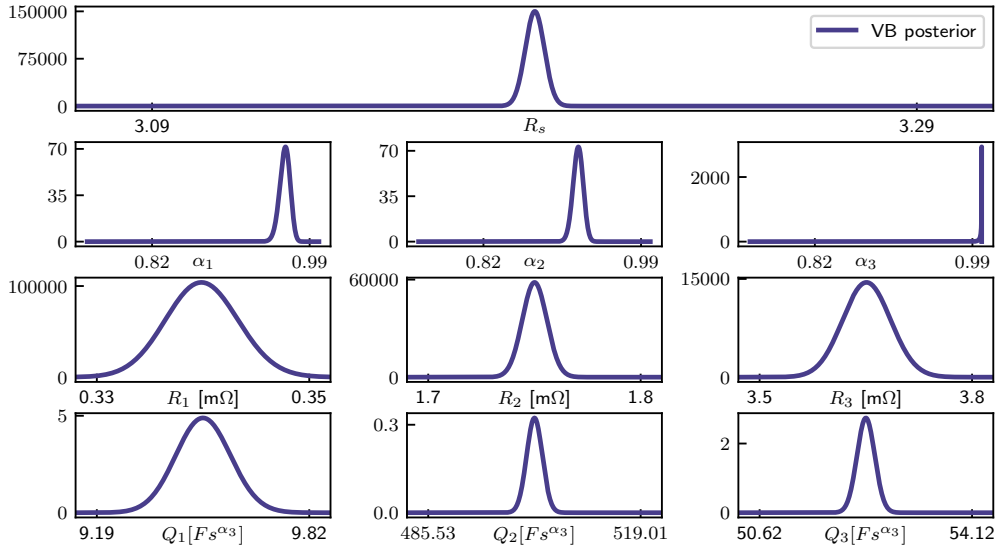


Figure 2.14: Posterior distributions of parameters for the experimental data example. Once again, we can notice that the posterior distributions have quite low variance.

Another significant advantage is the reduction of the inherent bias of the VB algorithm. By using mixture distributions, the algorithm can approximate a wider range of posterior distributions more accurately, leading to less biased inference. Additionally, this method proves beneficial in tackling high noise ratios, which is particularly useful for in-field applications where data is often noisy. The increased robustness to noise enhances the algorithm's applicability in real-world scenarios, making it more reliable for practical use.

In summary, the extension of the VB algorithm with mixture distributions as priors offers increased flexibility, maintains or improves accuracy, reduces bias, and handles noisy data more effectively. These improvements make it a compelling approach for enhancing the performance and applicability of the VB algorithm in various complex inference tasks.

2.6.1 Methodology

The methodology remains the same as for the VB approach but with one important change. The prior distributions for the ECM parameters are now mixture distributions, namely mixtures of LogNormal and Beta distributions. The Beta distributions for the α parameters and the LogNormal for everything else. This keeps us in line with the physical properties of the estimated parameters.

A n-dimensional mixture distribution M_n is defined as a weighted sum of n component distributions $f_{1,\dots,n}$:

$$M_n = \sum_{i=1}^{n} w_i f_i,$$

where $w_{1,\dots,n}$ are mixing weights[2] such, that $\sum_{i=1}^{n} w_i = 1$. Let us note here that the distributions f are not required to be from the same family of distributions. The weights were set as a parameter for the estimation as members of the Dirichlet distributions family. At each iteration, we took a sample from our weight distribution for each parameter and, based on that, assigned the corresponding mixture part and optimized the weight distribution parameters.

2.6.2 A toy example

To present the results and flexibility of the mixture extension, the mixture algorithm is tested on simulated and experimental EIS data and compared with the plain VB approach.

Implementation specifics The problem of estimating the parameters of the SOFC was already extensively researched by our department, so we had quite some prior experience and expectations about the distributions of the system parameters. The parameters we are estimating are physical quantities that are bound to be positive, therefore we wanted to ensure that the posterior distributions follow the same rule. As stated the weights were set to Dirichlet distributions

$$M_{DIR} = \mathrm{Dir}(\alpha_1^d, \alpha_2^d),$$

where $\sum_{i=1}^{2} \alpha_i^d = 1$. Next, we chose mixtures of Lognormal distributions

$$M_{LN} = \sum_{i=1}^{2} \mathrm{Lognormal}_i(\mu_i,\, \sigma_i^2)$$

as the family of candidate distributions for variational distributions for every parameter except α. We know that the α parameters are bounded on the interval $(0, 1]$. Combining this with prior knowledge, we have chosen the Beta distribution mixtures

$$M_B = \sum_{i=1}^{2} \mathrm{Beta}_i(\alpha_i,\, \beta_i)$$

as their variational distributions. Experimenting with different mixture dimensions n leads us to believe that there is no significant improvement in estimation performance with $n > 2$ for the problem at hand.

[2] i. e. how much each individual component distribution contributes to the mixture.

Table 2.5: Parameters used for numerical simulation of EIS.

$R_s[\Omega]$	$R_1[\Omega]$	$Q_1[Fs^{\alpha_3}]$	α_1
3	1	0.1	0.88

Results on simulated data For the first simulation, we used the ECM parameters presented in Table 2.5. This simulation was done without adding any noise to the input and output signal. Signals can be seen on Figure 2.16. Note here that only a fraction of the signal is shown for clarity and presentation. On Figure 2.15 we can see the progress of the cost function ELBO for the first optimization example. By taking multiple samples from our posterior parameter distributions presented on Figure 2.17 we can present the spread in estimated EIS curves and compare them to the original simulated one, this is presented on Figure 2.16.

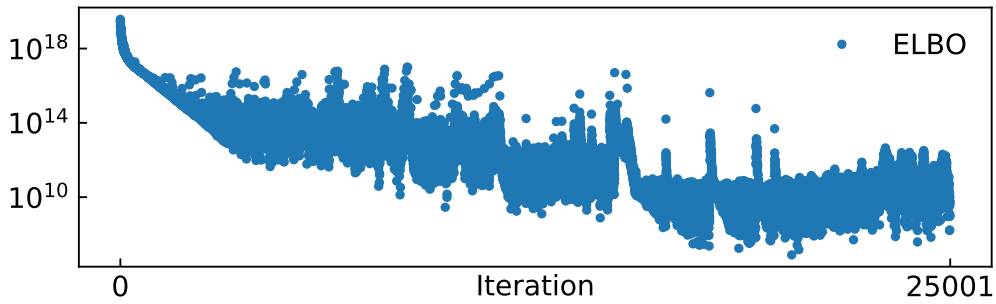


Figure 2.15: The evolution of ELBO loss over iterations during the optimization process for the toy example of the Mixture VB optimization steps.

The second simulated example had a relatively large amount of noise added to the input and output signal before continuous wavelet transform (CWT) was used to obtain the EIS curve. Specifically, the variance of n_1 and n_2 was set to 0.01. Parts of the signals can be seen in Figure 2.19. Parameters used for this simulation were the same as for the first one and can be seen in Table 2.5. Similarly, as before, we can see the progress of the ELBO function on Figure 2.18 and the spread of EIS curves from posterior parameter distributions (Figure 2.20) visible on Figure 2.19.

The simulated EIS curve can be seen in Figure 2.19.

Results on experimental data To further test the Mixture VB algorithm, we used data obtained during the experiments conducted during the REACTT project. The EIS data used for testing can be found in Figure 2.22.

The progression of ELBO fit function can be seen in Figure 2.21. We can conclude that the algorithm reached the desired convergence.

To illustrate the obtained estimated distributions of ECM parameters, we took 1000 samples of their posterior distributions and calculated the corresponding EIS curves to compare them with the experimental data. The results are presented in Figure 2.22.

The resulting posterior distributions of the estimated ECM parameters are presented in Figure 2.23.

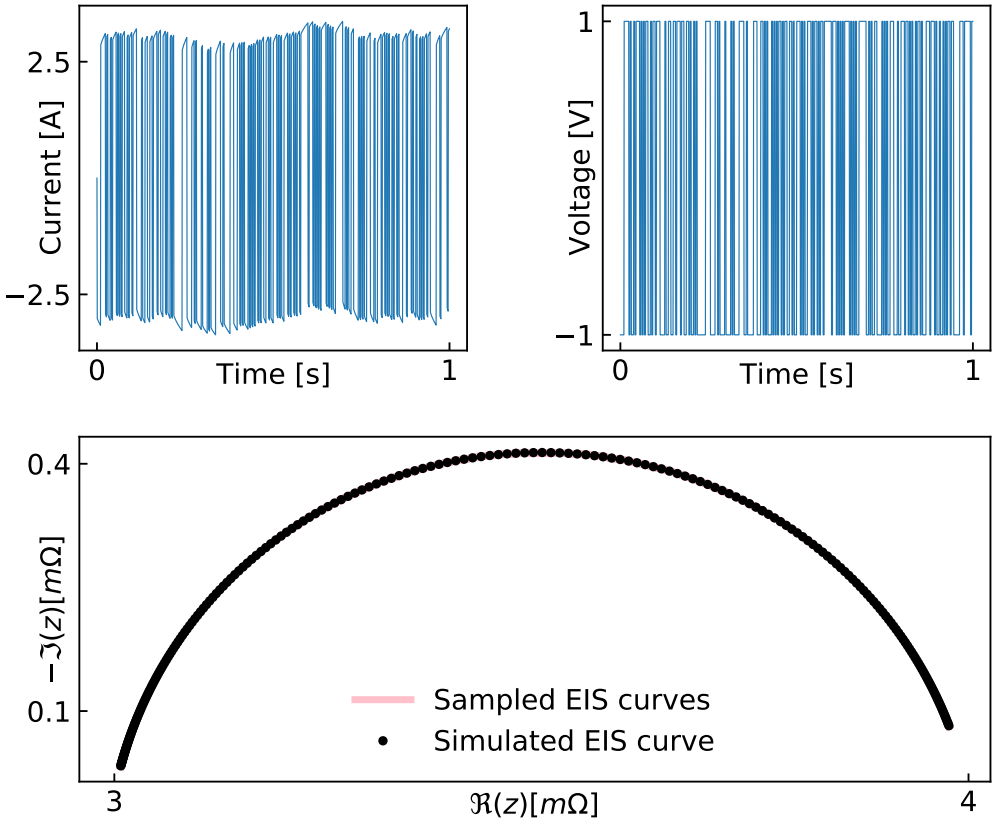


Figure 2.16: Current and voltage used in the simulation of the EIS curve (above). The simulated EIS curve is presented with black dots, and the results of the mixture VB algorithm are shown in pink (below). Because the posterior distributions of the parameters are very narrow and the predicted EIS curves fit the data exactly with a very narrow confidence band, the result is barely visible.

2.6.3 Mixture Extension of VB: Final Remarks

The mixture extension provided accurate results at the expense of the higher computational complexity. This increased flexibility offers benefits primarily on noisy data, while results on data with relatively low noise remain comparable to those obtained from the VB approach. This suggests that the added complexity and flexibility may lead to overfitting when noise levels are high, resulting in "good" fits that inaccurately model the noise as well.

The conclusion is that the benefits of the mixture extension are too marginal to justify the increased complexity at this time. Further testing on in-field application data of SOFC and SOEC might reveal greater advantages of the mixture approach over the plain VB algorithm, but such data were unavailable for this study. Despite its easy implementation and encouraging preliminary results, the mixture extension was not incorporated into the unified framework for SOC systems control due to the reasons mentioned.

Extending this analysis, it's important to note that while the mixture extension shows potential, the current evidence suggests that its practical application is limited by its tendency to overfit noisy data. This overfitting can obscure the true underlying trends and lead to less reliable results in practice. Therefore, while the theoretical benefits of increased flexibility are acknowledged, the real-world applicability remains questionable

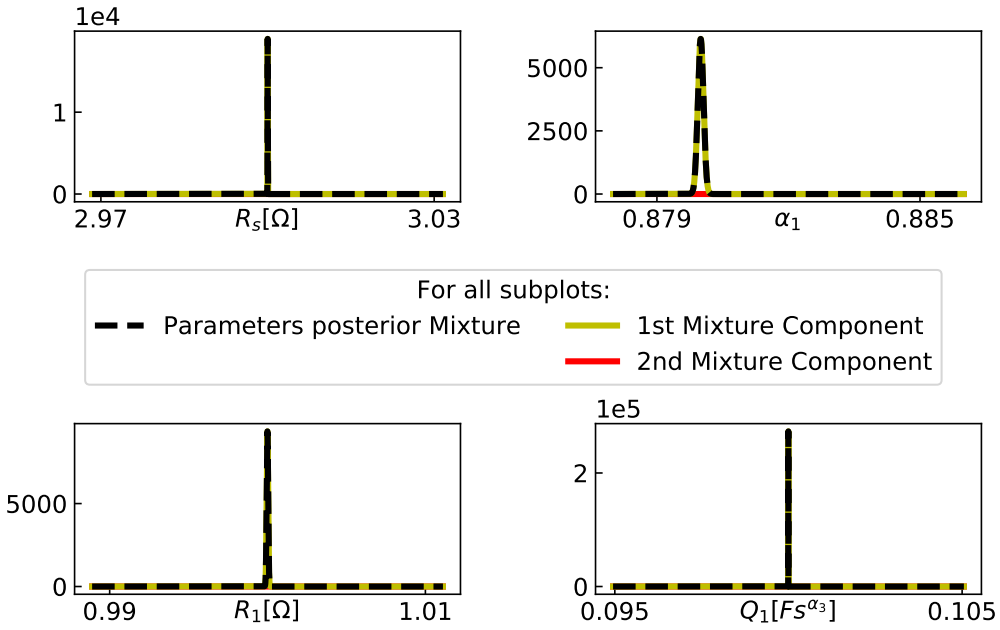


Figure 2.17: Posterior distributions of the ECM parameters for the noiseless simulated example. The distributions are visibly very thin, which can be attributed to the overconfidence of the VB approach and additionally to the non-existing noise in the simulated data.

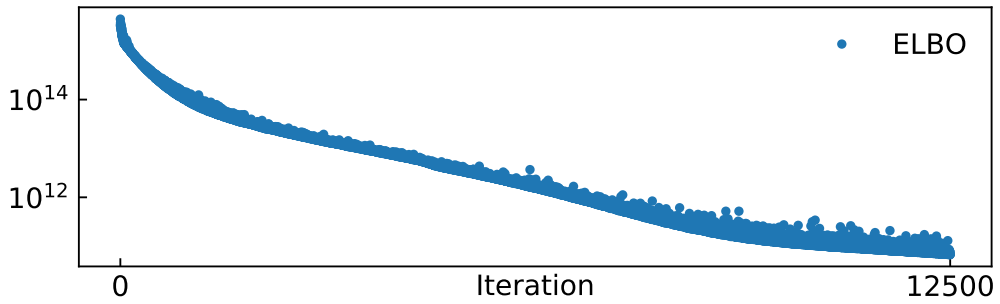


Figure 2.18: The evolution of ELBO loss over the iterations during the optimization process for the simulated toy example with added noise.

without further empirical validation.

The potential for future research lies in obtaining and analyzing extensive in-field data from SOFC and SOEC systems. Such data could provide a more comprehensive understanding of the mixture approach's performance in varied conditions, potentially highlighting scenarios where its flexibility is advantageous. Additionally, refining the algorithm to mitigate overfitting while retaining its adaptability could enhance its practical utility.

In summary, the mixture extension presents an intriguing and insufficiently explored option. The attempt to apply mixture models to describe the distributions of the ECM parameters showed that the benefit is small, if any, compared to the use of specific families of distributions. The focus should remain on improving the robustness to noise and external perturbations of existing methods and conducting thorough empirical testing to evaluate new approaches. As more real-world data becomes available, revisiting the mixture extension could yield new insights, potentially making it a viable component of advanced SOC

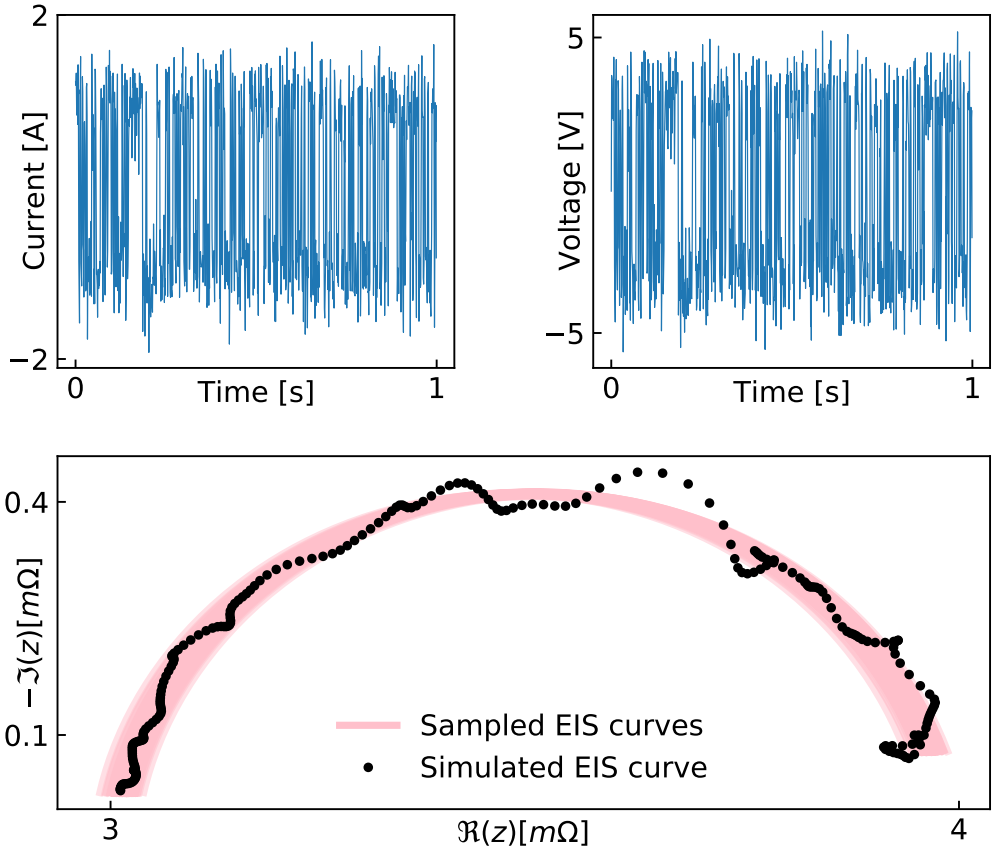


Figure 2.19: Current and voltage used in the simulation of the EIS curve with added noise (above). The simulated EIS curve is presented with black dots, and the results of the mixture VB algorithm are shown in pink (below).

system control frameworks.

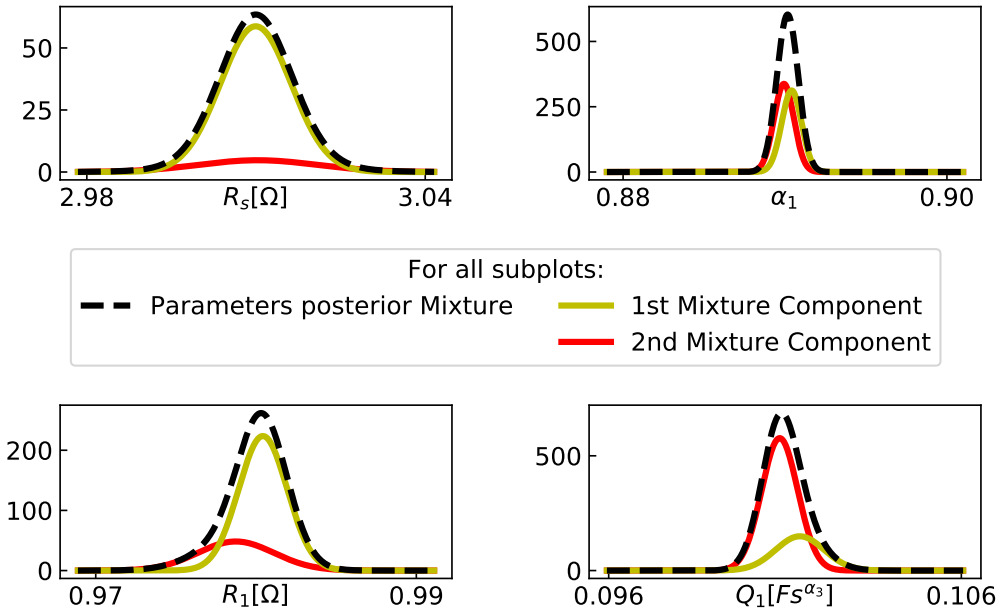


Figure 2.20: Posterior distributions of the ECM parameters for the noisy simulated example. The distributions are much less confident, which can be attributed to the strong added noise in the simulation of the EIS measurements used for this toy example.

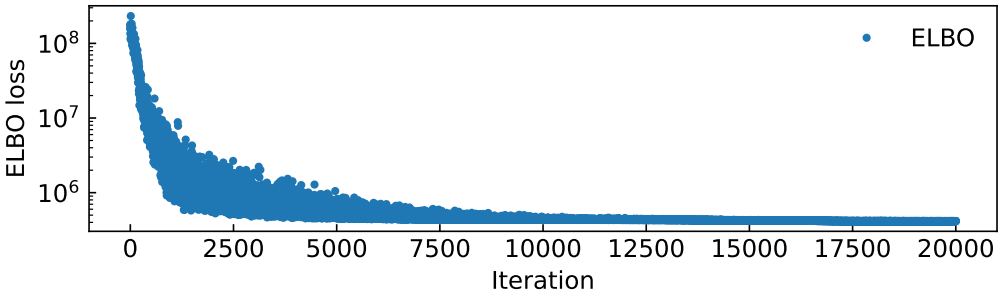


Figure 2.21: The evolution of ELBO loss over the iterations during the optimization process for the experimental example.

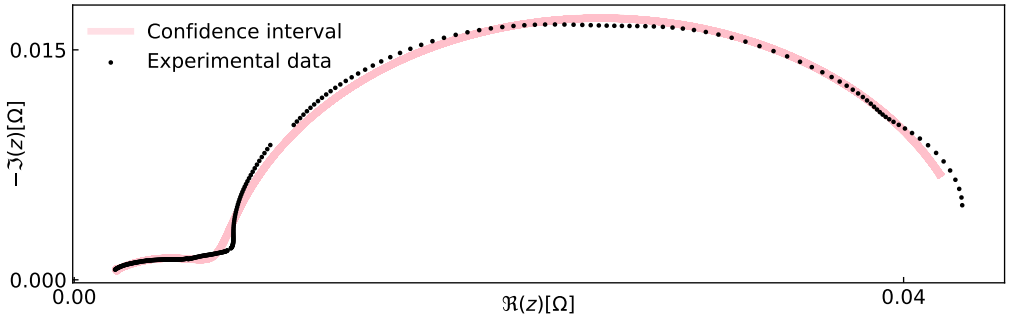


Figure 2.22: The EIS curve measured in the laboratory setting is presented with black dots. The EIS curves calculated from the estimated posterior distributions of the ECM parameters are shown in pink.

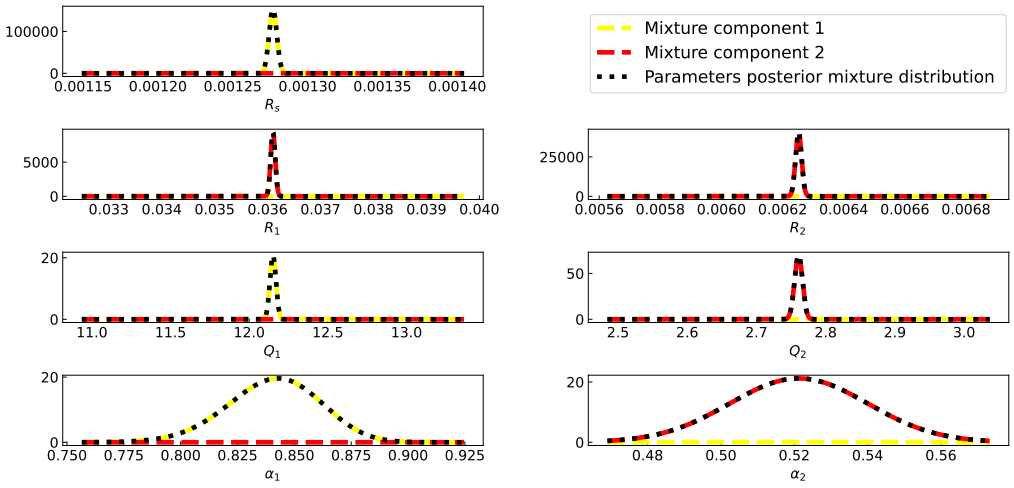


Figure 2.23: Posterior distributions of the ECM parameters for the experimental example. The distributions are somewhat confident, which can be attributed to the low noise in the measured EIS data, as well as the inherent overconfidence of the VB approach. The weights of the posterior mixture distributions are visibly leaning dominantly to one of the two mixture parts, which may mean that the mixture expansion is not needed for the data measured in a laboratory setting.

Chapter 3

EIS-based diagnosis of SOEC systems in non-stationary operating conditions

3.1 Problem Statement

The conventional EIS-based diagnosis procedure tacitly assumes that the process operates at constant operating conditions. Hence, the important stationarity requirement that ensures correct EIS measurement (stated in subsection 2.1) is fulfilled. During the in-field operation, it is not always possible to ensure identical operating conditions for EIS evaluation and, hence, consistent comparison of the obtained spectra. Namely, if the spectra are evaluated under equal operating conditions, one can relatively unambiguously associate changes in spectra with the faults or the degradation modes. If EIS spectra are taken at different operating conditions, then the latter can significantly impact the spectra and the associated ECM parameters. A question that naturally arises is how to consistently compare EIS curves measured at different operating conditions. To better demonstrate the problem, Figure (3.1) shows EIS curves taken at different current, temperature and fuel flow conditions. The differences can be rather substantial, especially in the low-frequency part of the spectra.

3.2 The Solution Outline

The conventional EIS-based diagnosis (Figure 3.3) starts by first perturbing the stack during the EIS session and then evaluating the EIS characteristic from voltage and current recordings. Finally, the equivalent circuit parameters θ from the EIS curve are calculated and compared with the reference parameters $\boldsymbol{\theta}^*$. The difference, referred to as residuals, takes values around zero if the stack is in the healthy (or reference) condition. Non-zero residuals $\boldsymbol{r} = \boldsymbol{\theta} - \boldsymbol{\theta}^* \neq \mathbf{0}$ signify a change in the condition, be it a fault or onset of a degradation. The root cause for the discrepancy can hopefully be inferred from the pattern of zero and non-zero residuals[1] in the vector $\boldsymbol{r}$.

The solution to the problem of non-stationary operating conditions relies on the presumption that ECM parameters depend on the operational conditions in a consistent way (Figure 3.2). Finding a nominal model that relates the ECM parameters $\boldsymbol{\theta}$ and the vector of the operating parameters $\boldsymbol{x}$ in the nominal condition is the key prerequisite for the diagnostic procedure suggested in Figure 3.2). To learn the model from data, fast EIS using

[1] What is "zero" and what "non-zero" is the matter of qualitative resoning

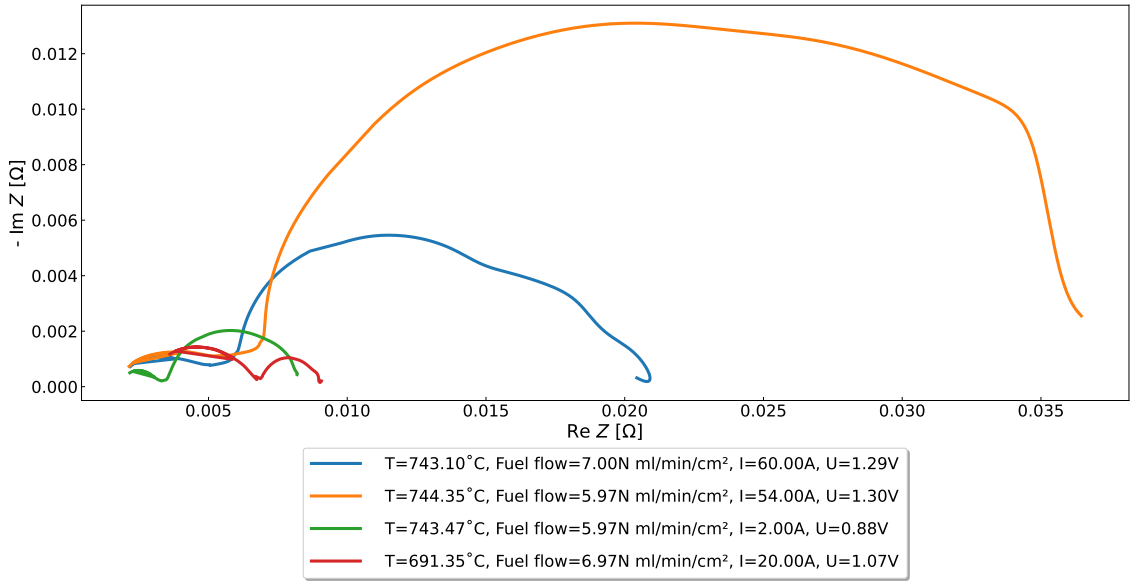


Figure 3.1: EIS characteristics measured at the same (nominal) health condition but at different operating points (temperatures, fuel flows, currents, and voltages).

discrete random binary sequence (DRBS) excitation is performed at different operating points under healthy stack conditions (reference condition). Due to the nonlinear character of the SOEC process and a complicated relationship between $\boldsymbol{\theta}^*$ and $\boldsymbol{x}$, data-driven modelling seems a reasonable option. For that purpose, the probabilistic data-driven GP model is used. The GP model maps the vector of the operational parameters into the probability density functions of the ECM parameters. The idea of GP modelling is highlighted in Appendix A.

The outline of the probabilistic diagnostic procedure, tailored for the non-stationary operating conditions, is shown in Figure 3.3 on the right side. After the perturbation session, the Fast EIS evaluation procedure is applied and is followed by the probabilistic decomposition of the Nyquist curve by means of the VB approach described in Chapter 2. The probabilistic ECM evaluation results in the probability density distribution of the ECM parameters. On the other hand, the nominal GP model results in the distribution of the ECM parameters conditioned on the process operational parameters. In the case of no fault and no degradation, the two probability densities are nearly the same. To quantify the discrepancy between them, the WD is applied. For similar distributions $WD \approx 0$. The greater the dissimilarity of the two distributions, the higher is the WD. The idea of Wasserstein distance is highlighted in Appendix B.

The last step in the diagnostic reasoning concerns the classification of the vector of Wasserstein distances (associated with the ECM parameters), which is performed by means of the SVM to link them with one of the presumed faulty regimes. The essentials of SVM are presented in Appendix C.

Figure 3.2: The idea of modelling the equivalent circuit model parameters $\theta_i, i = 1, ..., m$ as a function of the vector of process variables $\boldsymbol{x}$.

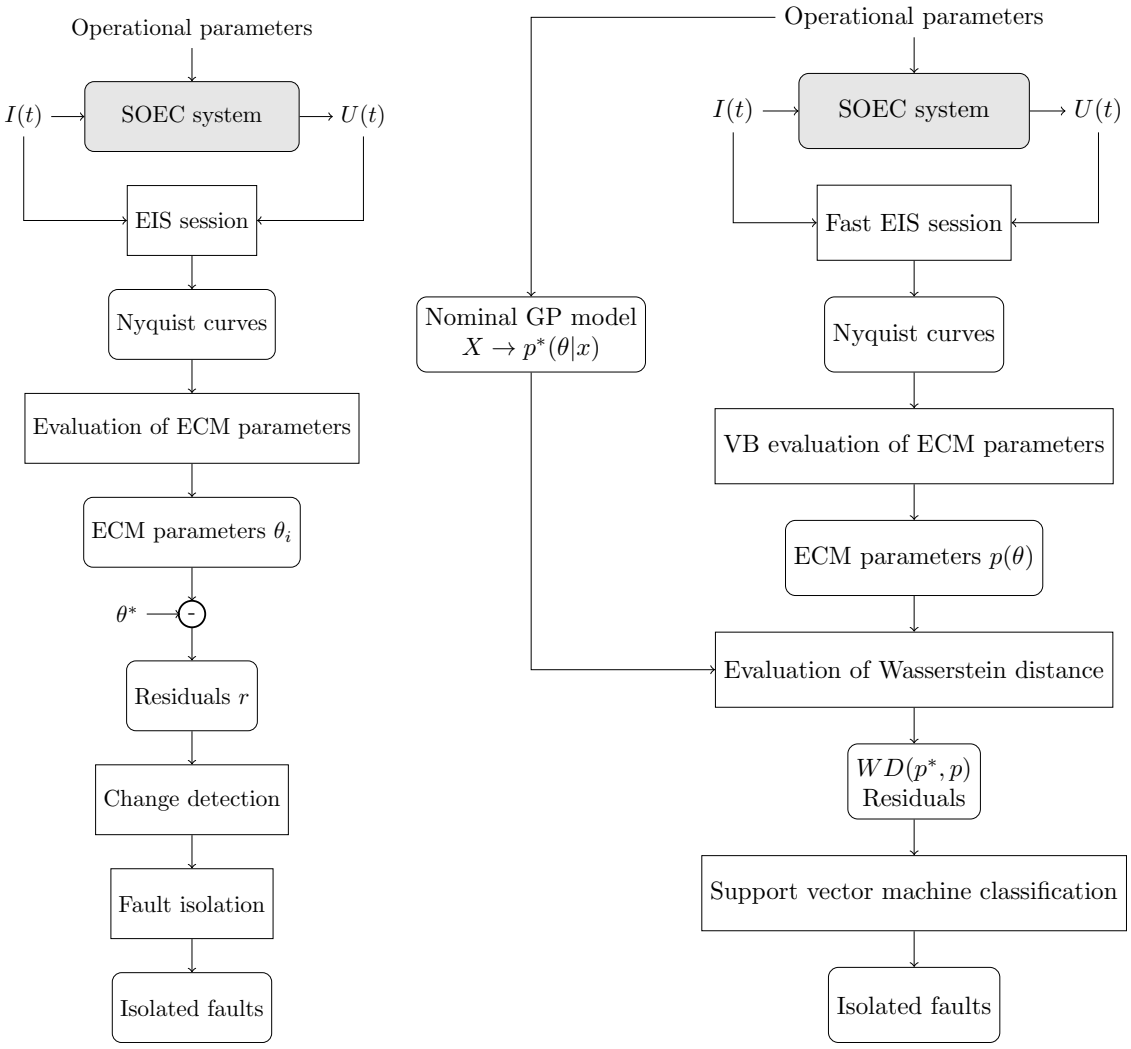


Figure 3.3: Comparison between the conventional EIS procedure applicable in stationary operating conditions (left) and the novel EIS-based diagnostic procedure tailored for the non-stationary operating conditions (right).

From the data obtained during the two experimental campaigns on short stacks at CEA, 24 different nominal working points were investigated. An excerpt is shown in Figure 3.1 from which it is clear that changes in operating conditions implicate significant changes in the shape and size of the EIS spectra. In other words, cancellation of the influence of the operating conditions on EIS spectra is mandatory for stable and reliable diagnosis.

3.3 The Case Study on a Short SOEC Stack System

3.3.1 Process description

The study comprises a 6-cell short stack manufactured by SolydEra and tested at the CEA lab within the H2020 project REACTT. Each cell has an $80cm^2$ area, with the cross-flow configuration of the reactants inlet/outlet. On the cathode inlet, a mixture of steam and hydrogen is delivered, whereas, at the anode inlet, the flow of air is supplied. The short stack is placed inside a furnace whose temperature determines the temperature of the stack.

The short stack and the test bench are shown in Figure 3.4. A simplified process

flowsheet with the indicated essential sensors is presented in Figure 3.5. The list of sensors with their description is presented in Table 3.2.

(a)  (b) 

Figure 3.4: (a) The 6-cells short stack from SolydEra and (b) the test bench at CEA.

3.3.2 Description of the experiment

The stack has been operated for 90 days at different regimes according to the protocol that included performance mapping, stack response in the regime referred to as 'H2 inlet starvation' and 'high steam conversion'. Measurements of the stack current, cell voltages, temperatures at the stack top and bottom, air inlet and outlet temperatures, fuel inlet and outlet temperatures, air inlet and outlet pressures, fuel inlet and outlet pressures, inlet flows of air and fuel (i.e. hydrogen and water steam), and outlet flow of hydrogen were taken continuously with a 1 s sampling interval. EIS measurements were done every 2 h by using the conventional sinusoidal perturbation and the DRBS perturbation signal. All the details of the experiment can be found on the IJS's portal at https://portal.ijs.si/nextcloud/apps/files/?dir=/reactt/Data. In total, the experiment produced over 600 EIS curves over the course of 90 days.

Table 3.1: Event and Regime descriptions over the time of the experiment.

Time Interval	**Type**	**Description**
Day 1 - Day 69	Regime	Nominal operation
Day 32 - Day 48	Regime	Performance mapping (still nominal)
Day 69 - Day 72	Event	Induced fault: Humidification
Day 74 - Day 78	Event	Induced fault: H_2 starvation
Day 78 - Day 79	Event	Spontaneous fault: Steam pump failure
Day 82 - Day 84	Event	Induced fault: Steam conversion

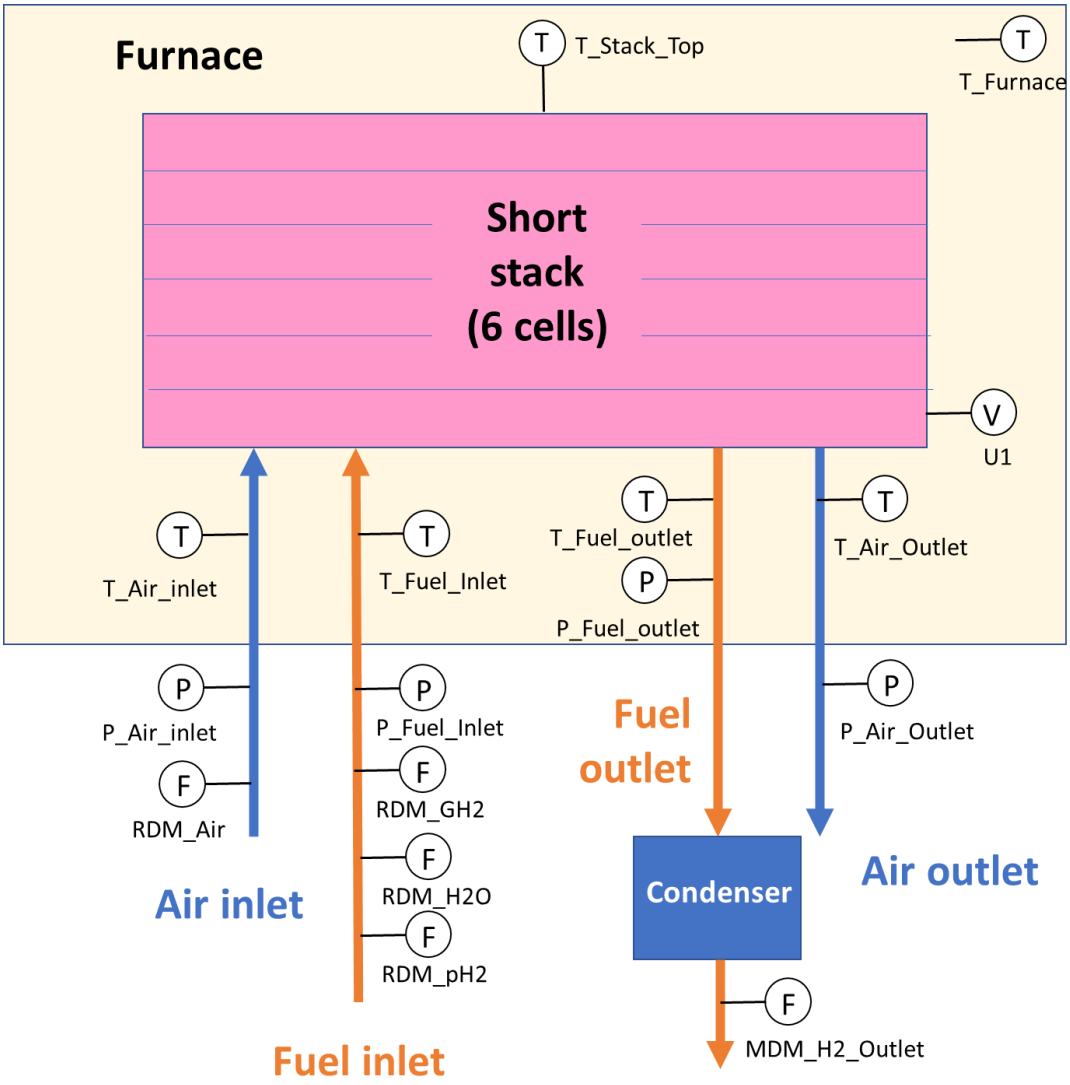


Figure 3.5: The simplified flowsheet of the SOEC system with a 6-cell short stack and indicated main instrumentation.

3.3.3 Modelling the ECM parameters with GP models

For the GP modelling of the ECM parameters, all of the process variables indicated in Table 3.2 are used as regressors. No detailed optimal selection of the regressors is performed at this research stage. An additional reason to operate with all the regressors is that it is impossible to define the informative regressors based on prior physical knowledge.

The entire dataset was split into two sets. One set contained the data gathered while the SOEC was operating in a guaranteed nominal mode, i.e., a nominal set, and the rest was obtained during the normal operation of the stack and formed a test set. The nominal set was further split into training and validation sets. The split was based on the number of measurements performed at each state of the system, which was three. Therefore, for every three consecutive measurements, we randomly chose one to belong to the validation set and the other two to the training set. The division of the dataset is compactly presented in Figure 3.7. The choice of the kernel was decided based on the Bayesian information criterion (BIC).

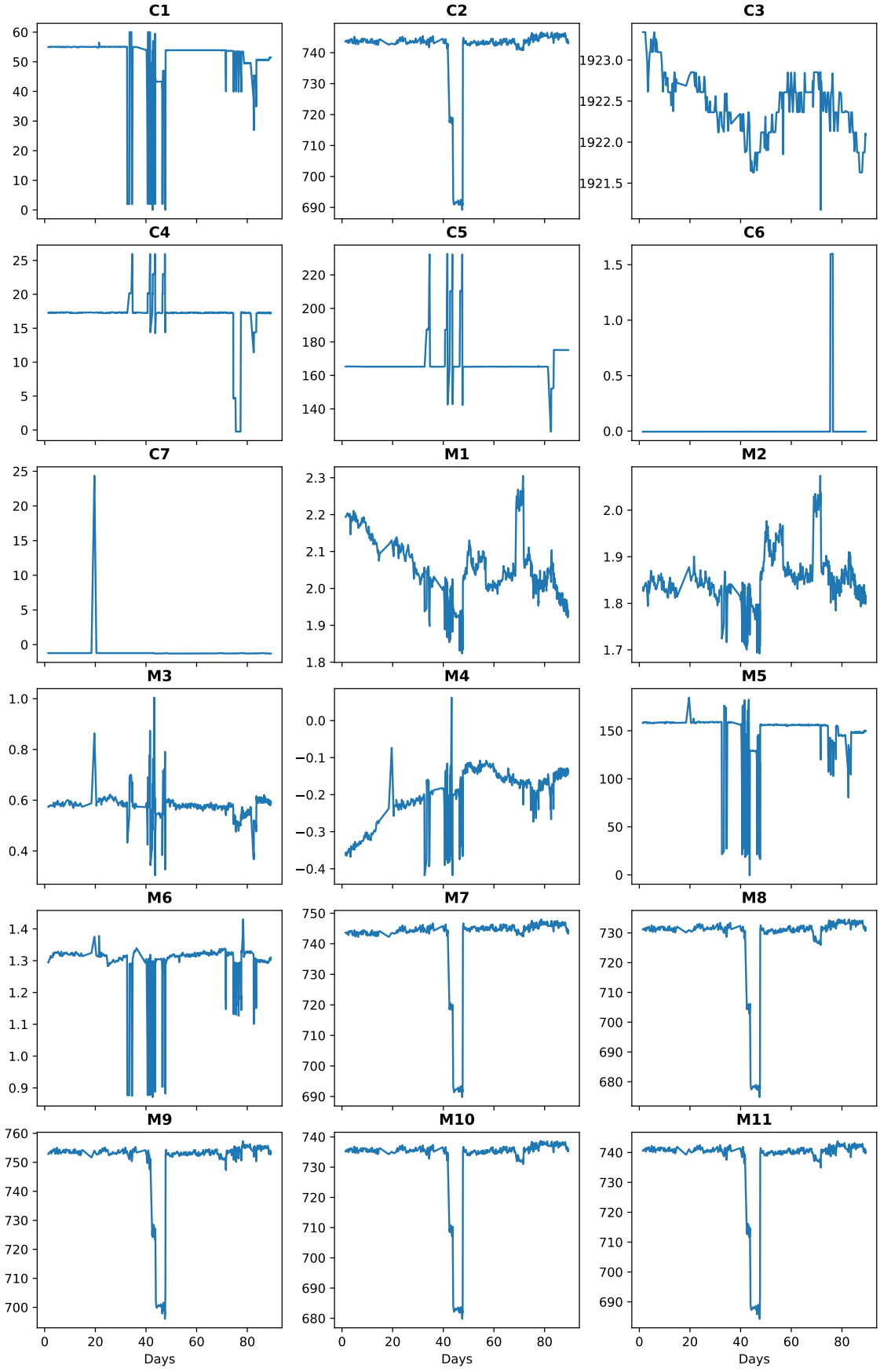


Figure 3.6: Measured signals during the REACTT experiment used in the fault detection and isolation phases with the GP. Descriptions of the signals can be found in Table 3.2.

Table 3.2: Table with the descriptions of the measured signals used in the fault detection and fault isolation phases.

Par.	Name	Description
C1	Current	Stack current [A]
M6	T_Furnace	Furnace temperature [°C]
C3	RDM_Air	Air inlet flow rate [slph]
C5	RDM_GH2	Hydrogen inlet flow rate [slph]
C6	RDM_H2O	Water vapour inlet flow rate [slph]
C4	RDM_pH2	H2 fraction [%]
C7	RDM_pH20	Water vapour fraction [%]
M1	P_Air_Inlet	Air inlet pressure [mbar]
M2	P_Air_Outlet	Air outlet pressure [mbar]
M3	P_Fuel_Inlet	Fuel inlet pressure [mbar]
M4	P_Fuel_Outlet	Fuel outlet pressure [mbar]
M5	MDM_H2_Outlet	Hydrogen outlet volumetric flow [slph]
M7	U1	Cell #1 voltage [V]
M8	T_Stack_Top	Temperature at stack top [°C]
M9	T_Air_Inlet	Temperature at air inlet [°C]
M10	T_Air_Outlet	Temperature at air outlet [°C]
M11	T_Fuel_Inlet	Temperature at fuel inlet [°C]
M12	T_Fuel_Outlet	Temperature at fuel outlet [°C]

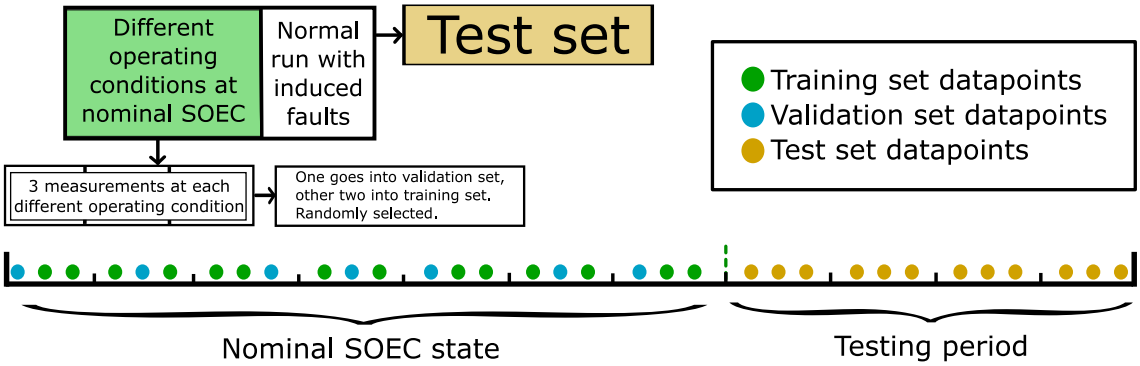


Figure 3.7: The way of selecting learning, validation and test data from the available dataset.

Kernel selection The Bayesian information criterion (BIC) is a model selection criterion used in statistical model evaluation [82]. It is defined as:

$$\mathrm{BIC} = -2\ln(\hat{L}) + c\ln(n)$$

where $\hat{L}$ is the maximized value of the likelihood function of the model (expression (A.1) in Appendix A), c is the number of hyper-parameters, and n is the number of data points. The BIC balances model fit and complexity, penalizing models with more parameters to avoid overfitting.

To select the most appropriate kernel function, we trained 10 different kernel functions on the training set. Then, we calculated their BIC scores from predictions on the validation dataset. Their BIC scores are given in Table 3.3.

The best kernel function, balancing prediction quality and model complexity, was the Matérn kernel with $\nu = \frac{3}{2}$ combined with white noise. For simplicity, we will refer to this combined kernel as Matérn($\frac{3}{2}$) from now on. The Matérn kernel is defined as [83]:

$$k_{\text{Matérn}}(x, x') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\,|x - x'|}{l} \right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu}\,|x - x'|}{l} \right) \tag{3.1}$$

where K_ν is the modified Bessel function of the second kind. For $\nu = \frac{3}{2}$, this simplifies to:

$$k_{\text{Matérn}}(x, x') = \sigma_f^2 \left(1 + \frac{\sqrt{3}\,|x - x'|}{l} \right) \exp \left(-\frac{\sqrt{3}\,|x - x'|}{l} \right) \tag{3.2}$$

By adding the White Noise kernel, the combined kernel function becomes:

$$k_{\text{combined}}(x, x') = \sigma_f^2 \left(1 + \frac{\sqrt{3}\,|x - x'|}{l} \right) \exp \left(-\frac{\sqrt{3}\,|x - x'|}{l} \right) + \sigma_n^2 \delta(x, x') \tag{3.3}$$

Here, x and x' are points in the regressor space, representing locations where function values are compared; $k(x, x')$ is the covariance function (kernel) that defines the covariance between function values at x and x', encoding assumptions about the smoothness and other properties of the function; σ_f^2 is the signal variance controlling the overall variability of function values; l is the length scale parameter determining how quickly correlations decrease with distance; ν controls the smoothness of the function modelled by the Matérn kernel; and K_ν is the modified Bessel function of the second kind, defining the kernel's shape and behaviour over input point distances. For the White Noise kernel, the σ_n^2 represents the noise variance, and $\delta(x, x')$ is the Kronecker delta function.

Performance of the GP model on the training dataset. The results of the GP models of all ECM parameters obtained in the training dataset by using the selected kernel Matérn $\frac{3}{2}$ are presented in Figure 3.8 in the first left column. Let us note that instead of R_1, R_2, Q_1, Q_2 we used their logarithm $log(R_1), log(R_2)$ etc. since their posterior distributions derived with VB are LogNormal. Furthermore, it should be noted that abscissas on the graphs in the first column do not represent absolute time. A point on the abscissa represents the consecutive number of the EIS measurement from the training dataset, which is referred to as the index. The explanation comes from Figure 3.7. Namely, two EIS measurements defined by two consecutive indexes do not necessarily follow one after another if viewed in absolute time.

The GP model demonstrates high accuracy and precision on the learning dataset. The mean values of the GP are very close to the measured values so that the model error is within the confidence interval. The variance of the GP model error is pretty low.

It should be noted that the abrupt jumps in the ECM parameter values are due to the abrupt jumps in the operating points of the system during the nominal operational regime.

Performance of the GP model on the validation dataset. The performance of the GP models with Matérn($\frac{3}{2}$) kernel on the validation dataset is shown in Figure 3.8 in the middle column. Again, note that the results are indexed and not plotted time-wise since the measurements in the validation dataset are not sequential. The quality of prediction still remains very good, although the variance of the prediction error is now higher than in

Table 3.3: 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

the case of the training data. That is within the expectation since the operational regimes of the system backing the validation data slightly differ from those backing the training data. Differences in the operating points mainly manifest in the increased variance of the prediction error.

Performance of the GP model on the test dataset. The performance of the validated GP model (Matérn($\frac{3}{2}$) kernel) for ECM parameters on the test set is presented in Figure 3.8 in the last column. The test dataset contains four major faults that were either induced or emerged spontaneously (c.f. Table 3.1). Please note that the abscissas on the plots in the rightmost column represent absolute time, making it more convenient to demonstrate the performance of the GP predictor model in different faulty operating regions. The main observations can be summarised as follows:

- Generally, it seems that the GP prediction model is sensitive to even slight changes in the internal condition.
- The induced humidification causes no significant changes in the prediction of the ECM parameters except for Q_1, where the prediction error variance clearly increased.
- After the humidification is turned off, the system is assumed to return to nominal operating conditions. However, from the predictions of the GP model, it seems that the system underwent a slight change, which remains poorly understood. The variance of the parameter R_2 and the bias (the difference between the predicted mean

and the "measured" R_2) increased. No significant changes can be observed in the prediction of other ECM parameters.

- The induced H_2 humidification regime visibly affects the predicted ECM parameters both in terms of bias as well as in terms of the variance of the prediction error.
- At the beginning of the 77^{th} day of the experiment, an unexpected pump problem occurred. That significantly affected the prediction error of all ECM parameters, resulting in a substantial increase in the bias of the prediction error.
- After the fault was mitigated around day 79, the precision of the predictions for all ECM parameters considerably degraded. Both bias and variance of the prediction error increased. This is reasonable since, after the pump change, the system is (strictly speaking) not the same as it was before.

3.3.4 Feature extraction by means of the Wasserstein distance

The WD distances between the GP predicted distributions and the distributions of the ECM parameters obtained by VB approach are presented in Figure 3.9 for each ECM parameter. At around day 80, the system mass-flow sensor was subjected to a mechanical modification. In other words, we have a modified system from that day on. That poses a significant problem, which is, unfortunately, rather marginalized by the research community. In principle, in such a case, the diagnostic system should be re-calibrated, which in turn requires repeated learning and validation of the relevant models. Due to the lack of data, we used another remedy.

Note that the Wasserstein distance W_2, from the day 80 on, is composed of the two major parts

$$W_2 = W_{ECM} + W_{\Delta}, \tag{3.4}$$

where W_{ECM} is the WD between the GP output and the VB estimates, while W_{Δ} is the WD value attributed to the manual change in the system operation by the system's operator. One should note that manual intervention performed on the sensor introduces a change in the system, making the new instance of the system different from the original one. However, from the process operator's point of view, the system is considered healthy. Consequently, the WD associated with the updated system instance might mislead the detection and isolation algorithms because changes in WD caused by system modifications might be incorrectly interpreted as caused by the fault. Therefore, the evaluated WD must be properly corrected to account for the change that occurs only due to the system fault.

The correct approach after the incidence would be to re/calibrate the model. Unfortunately, we were not in a position to repeat the learning procedure. Threfore, we mitigated the issue by subtracting the estimated offset W_{Δ}, attributed to the system modification. This offset was calculated as the difference between the average WD values from the most recent nominal operation and the average WD values obtained post-intervention.

The newly prepared data can be seen in Figure 3.10.

Fault detection algorithm. After obtaining the corrected WD values, one can proceed with fault detection and isolation by using the Support Vector Classifier (SVC) method outlined in the Appendix C. The SVC builds on the RBF kernel function.

On the input of the SVC, there is the vector of the WD values $\boldsymbol{wd} = [\boldsymbol{wd_{R_1}}, ..., \boldsymbol{wd_{Q_2}}]$ obtained after an EIS session. The output of the SVC is a categorical variable d with two possible values, i.e. $d \in \{\text{Fault, No-Fault}\}$. That means an element of vector $\boldsymbol{wd}$ is

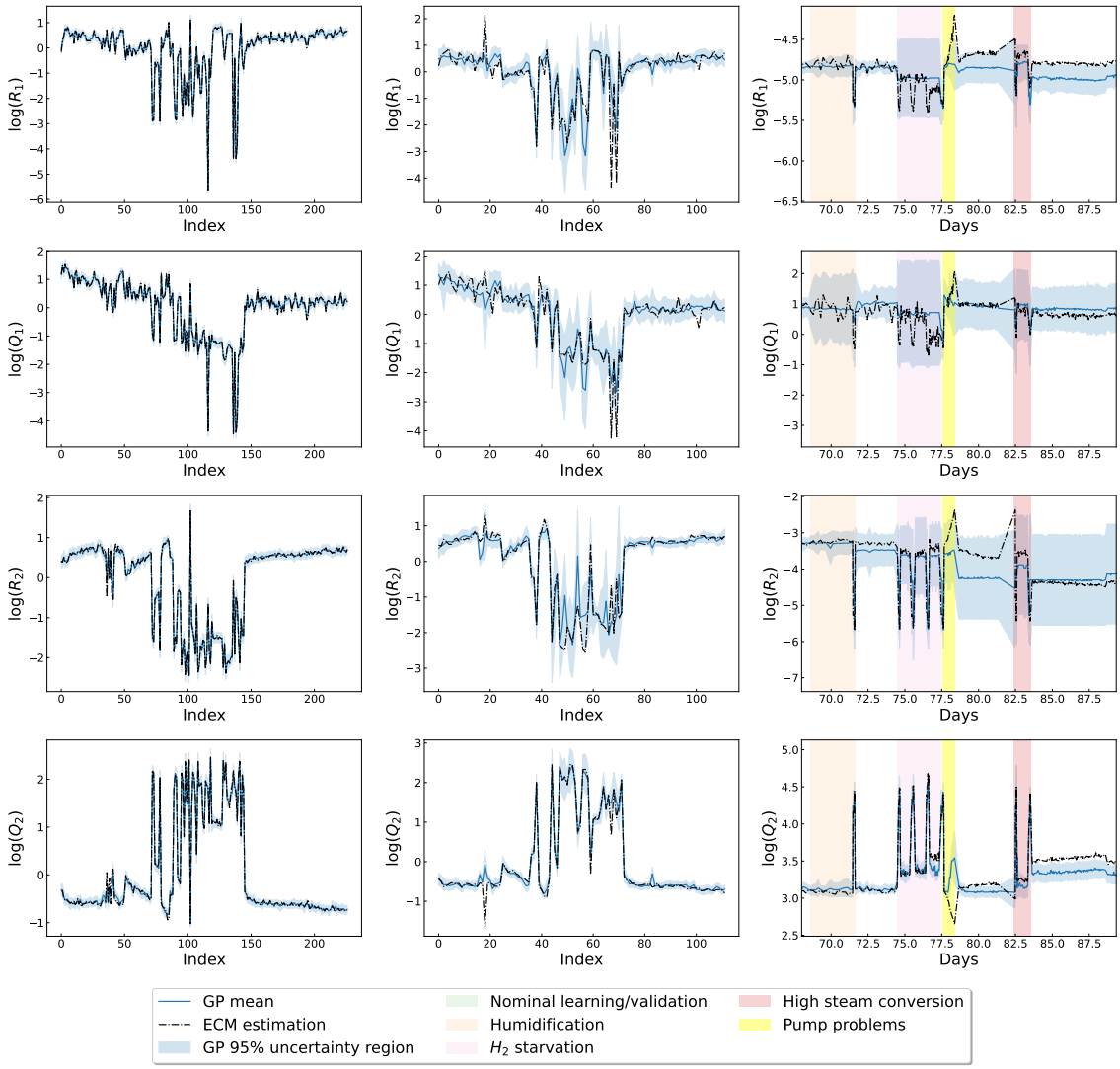


Figure 3.8: Performance of the GP model for the logarithm of ECM parameters across training, validation, and test datasets. The GP model's mean predictions and variances of the ECM parameters $\log(R_1)$, $\log(R_2)$, $\log(Q_1)$, and $\log(Q_2)$ are presented in blue. For the training and validation datasets (first and middle columns), each parameter is plotted against the measurement index since these datasets are not sequential. In the test dataset (last column), parameters are plotted against time in days because the data are sequential. The GP model closely fits the training data with low variance, shows some discrepancies and higher variances in the validation data, and highlights significant deviations in the test data corresponding to system faults.

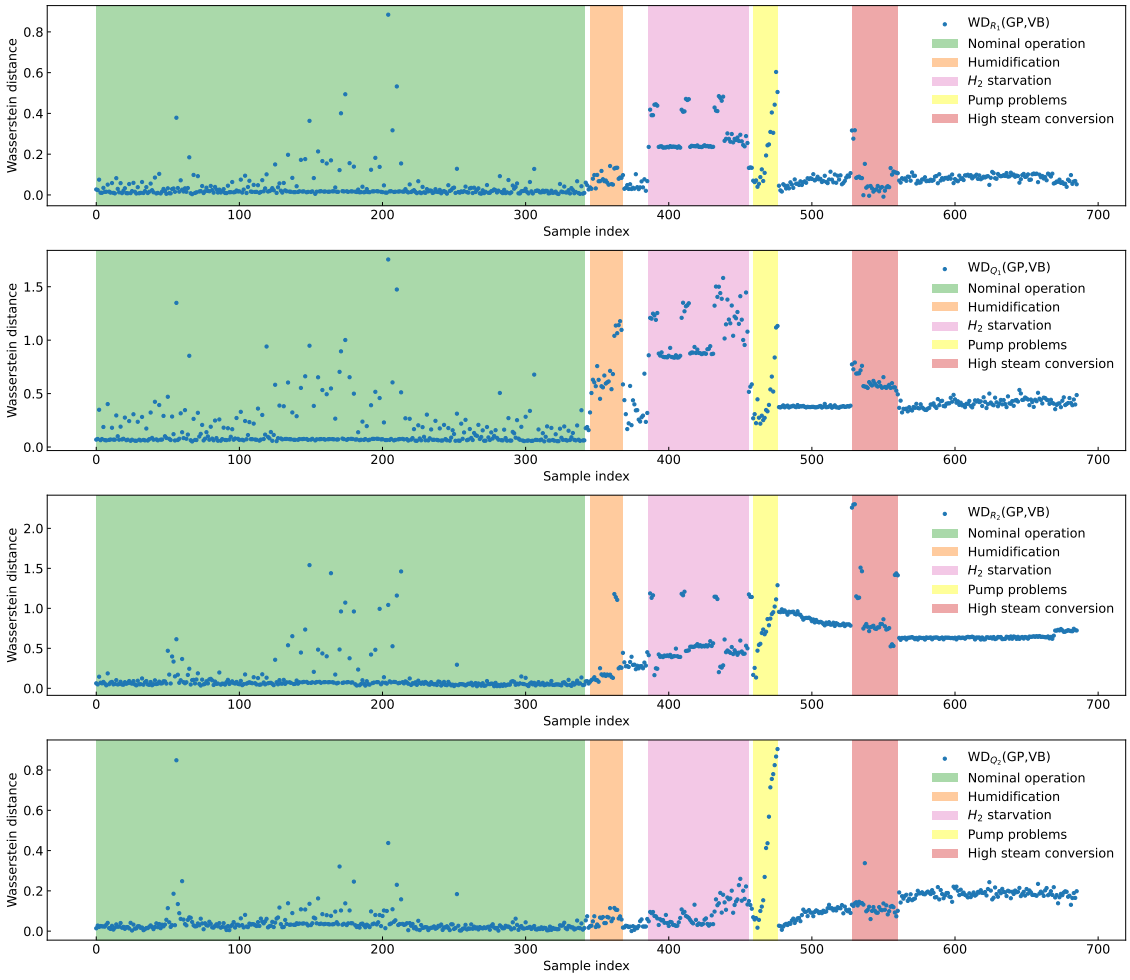


Figure 3.9: The uncorrected Wasserstein distances for the ECM model parameters, represented by blue dots. The intentionally induced faults are appropriately highlighted.

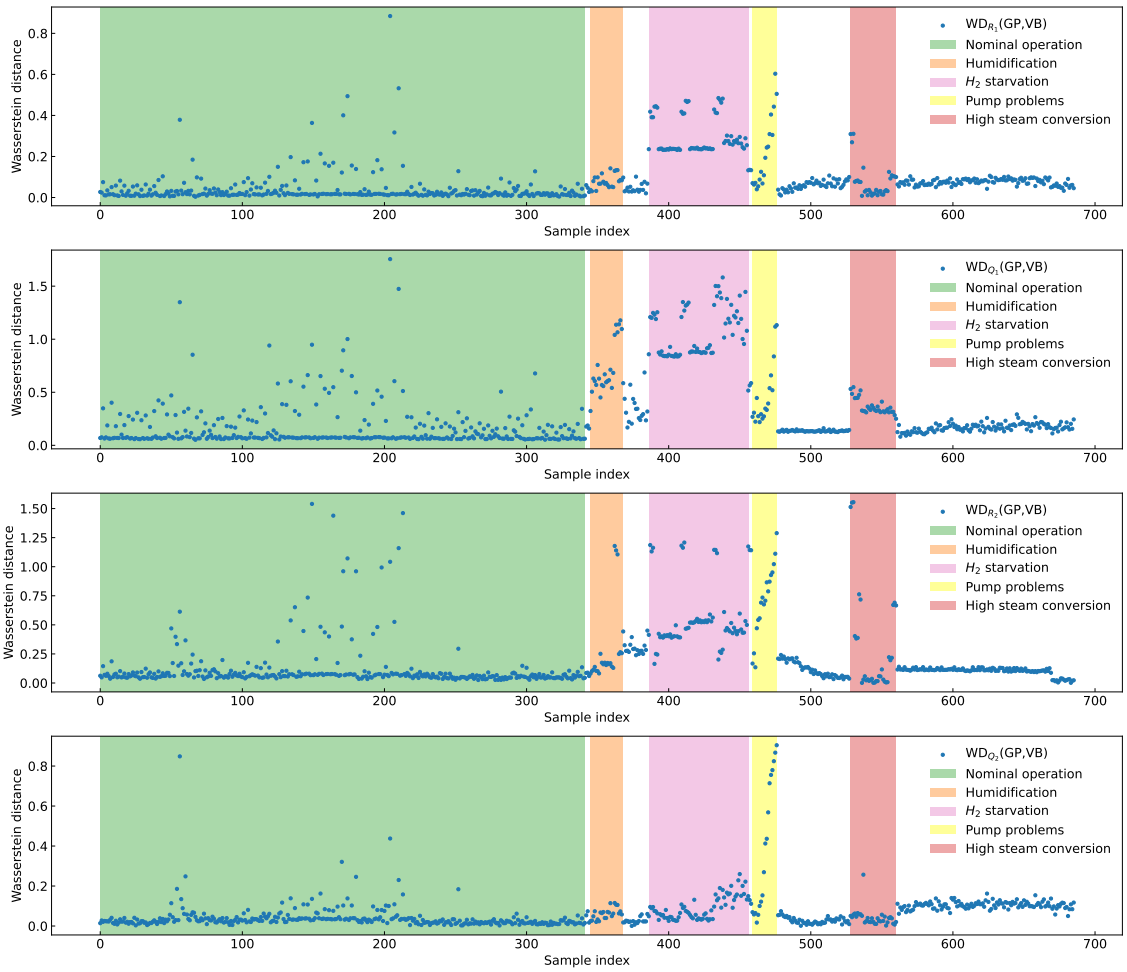


Figure 3.10: The corrected Wasserstein distances, indicated by blue dots, across all estimated parameters, with the induced faults clearly marked.

associated with d = Fault or d = No-Fault health status. By means of Platt calibration (more details are given velow), we can obtain the probability of fault for each element of $\boldsymbol{wd}$.

Platt calibration Platt calibration is a method used to convert the decision function outputs of a Support Vector Classifier (SVC) into calibrated probabilities, making the SVC's predictions interpretable in probabilistic terms. The decision function $f(x)$ from the SVC, which typically outputs uncalibrated scores, is transformed using a sigmoid function to map these scores into the range $[0, 1]$, hence representing probabilities. The sigmoid model is defined as:

$$P(y = 1 \mid f(x)) = \frac{1}{1 + \exp(Af(x) + B)}$$

where A and B are scalar parameters learned by fitting a logistic regression model to the decision scores. These parameters are optimized by minimizing the negative log-likelihood on a calibration dataset.

Selecting the training and validation sets for fault detection Let $\mathcal{S}$ denote the set of all pairs $(\boldsymbol{wd}_i, d_i)$, where i represents the index of each EIS measurement.

The elements of $\mathcal{S}$ are randomly divided into the two subsets: $\mathcal{T}$ and $\bar{\mathcal{T}}$. The subset $\mathcal{T}$ is treated as the test set, comprising approximately 40% of the total data from $\mathcal{S}$.

To perform 5-fold cross-validation[2], the remaining subset $\bar{\mathcal{T}}$ is divided into five equal parts. Four of these parts are combined to form the training set $\mathcal{L}$, while the remaining one is used as the validation set $\mathcal{V}$.

The SVC is trained five times, each time using a different fifth of $\bar{\mathcal{T}}$ as the validation set $\mathcal{V}$. In each run, the model is validated on the selected validation set $\mathcal{V}$, and the set of hyperparameters that performs the best is selected for further classification tasks.

Fault detection results. The classifier achieved an accuracy of 97%, with a dataset comprising 76 nominal samples and 61 faulty samples. The precision for predicting both nominal states and faults was 97%, reflecting a balanced performance across both classes. As shown in Table 3.4, the F1-scores were also 97% for both fault and no-fault classes, further demonstrating the model's balanced effectiveness.

Both classes achieved a recall of 97%, indicating the classifier's robustness in correctly identifying both faulty and nominal conditions. The macro and weighted averages for precision, recall, and F1-score were consistently high at 97%, underscoring the model's strong performance across varying conditions. These results suggest the classifier is effective even with potential noise or external influences. The confusion matrix for the test set is shown in Figure 3.11, visually confirming the classifier's effectiveness in fault detection.

The algorithm correctly predicted 74 nominal values and 59 faulty ones. However, there were some misclassifications: 2 nominal values were incorrectly predicted as faulty (false positives), and 2 faulty values were incorrectly predicted as nominal (false negatives). This misclassification may be attributed to high noise levels in the measurements, especially after the occurrence of faults, which may have obscured the distinctions between faulty and nominal conditions.

Extending this analysis, the results indicate robust diagnostic performance with a high accuracy rate for both nominal and faulty values. The false positive rate, where nominal values were misclassified as faulty, highlights a potential area for improvement, particularly in handling noisy data. Addressing this issue may involve refining the noise filtration

[2]It is generally recommended to use k-fold cross-validation to prevent overfitting in classification models.

Table 3.4: 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
Macro average	0.97	0.97	0.97	137
Weighted average	0.97	0.97	0.97	137

techniques or enhancing the model's sensitivity to differentiate between actual faults and noise-induced anomalies.

The false negative rate, where faulty values were misclassified as nominal, suggests there is still room for improvement in detecting all faulty conditions, although the model remains generally effective. This is crucial for the reliability and safety of SOC systems, where undetected faults could lead to significant operational issues.

Further research should focus on improving the model's robustness to noise and its ability to identify faults accurately. Future improvements could involve integrating advanced noise reduction techniques but mainly focusing on expanding the training dataset to include more varied examples of nominal and faulty conditions, enhancing the model's generalization capabilities.

In conclusion, while the predictive model demonstrates strong performance, particularly in identifying nominal and faulty values, addressing the challenges posed by noisy data and false negatives will be essential for optimizing its accuracy and reliability. Continuous refinement and testing, especially with more diverse and extensive datasets, will achieve a more reliable fault detection framework for SOC systems. The predicted fault probabilities were calibrated using Platt scaling [84]. They can be seen on the test dataset in Figure 3.12.

To provide further insight, we can examine the probability of fault detection at each measurement across the entire dataset used for classification, as shown in Figure 3.13. It is important to note that this analysis does not fully reflect the algorithm's performance, as some of the data was used during the training phase.

Next, we can revisit Figure 3.10, which displays the predicted faults over the entire timeline. The instances where faults were detected are highlighted in orange for all relevant ECM parameters, as illustrated in Figure 3.14.

3.4 Fault Isolation

Fault isolation algorithm. Similarly to fault detection, the SVC algorithm was used for fault isolation, with the selected kernel function being the RBF kernel.

The input to the SVC is a vector of WD values, denoted as $\boldsymbol{wd} = [wd_{R_1}, \ldots, wd_{Q_2}]$, obtained after an EIS session, just as in fault detection. The only difference lies in the output. For fault isolation, the SVC outputs a categorical variable d_{iso} with five possible values, corresponding to the four known faults and the "No Fault" condition, i.e., $d_{\text{iso}} \in$ {No Fault, Fault 1, Fault 2, Fault 3, Fault 4}. Thus, each element i of the vector $\boldsymbol{wd}$ is associated with a specific system health status, such as $d^i_{\text{iso}} =$ Fault $\{1, 2, 3, 4\}$ or $d^i_{\text{iso}} =$ No Fault.

Using Platt calibration, the probability of each health state can be obtained for each element of $\boldsymbol{wd}$, providing probabilistic insights into the system's condition.

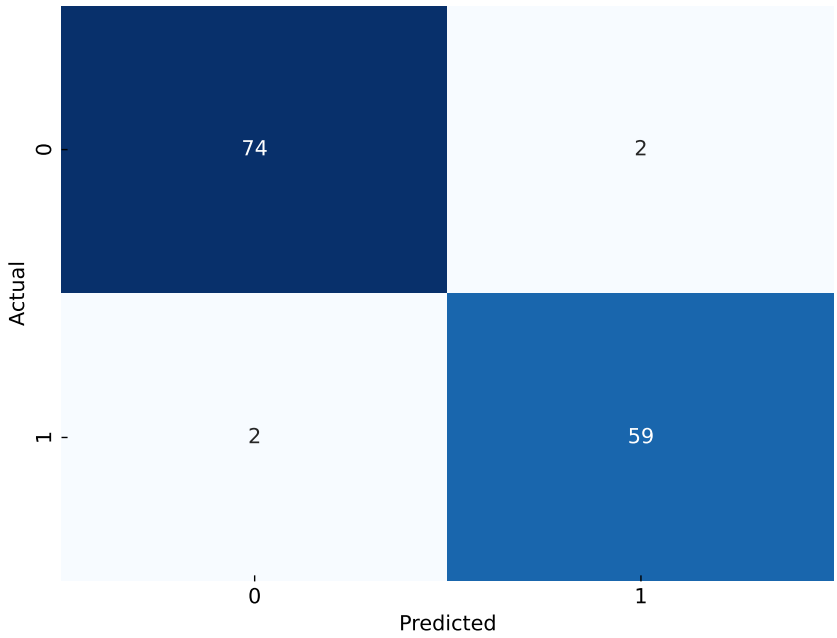


Figure 3.11: Resulting confusion matrix for fault detection done with SVC method on the test dataset.

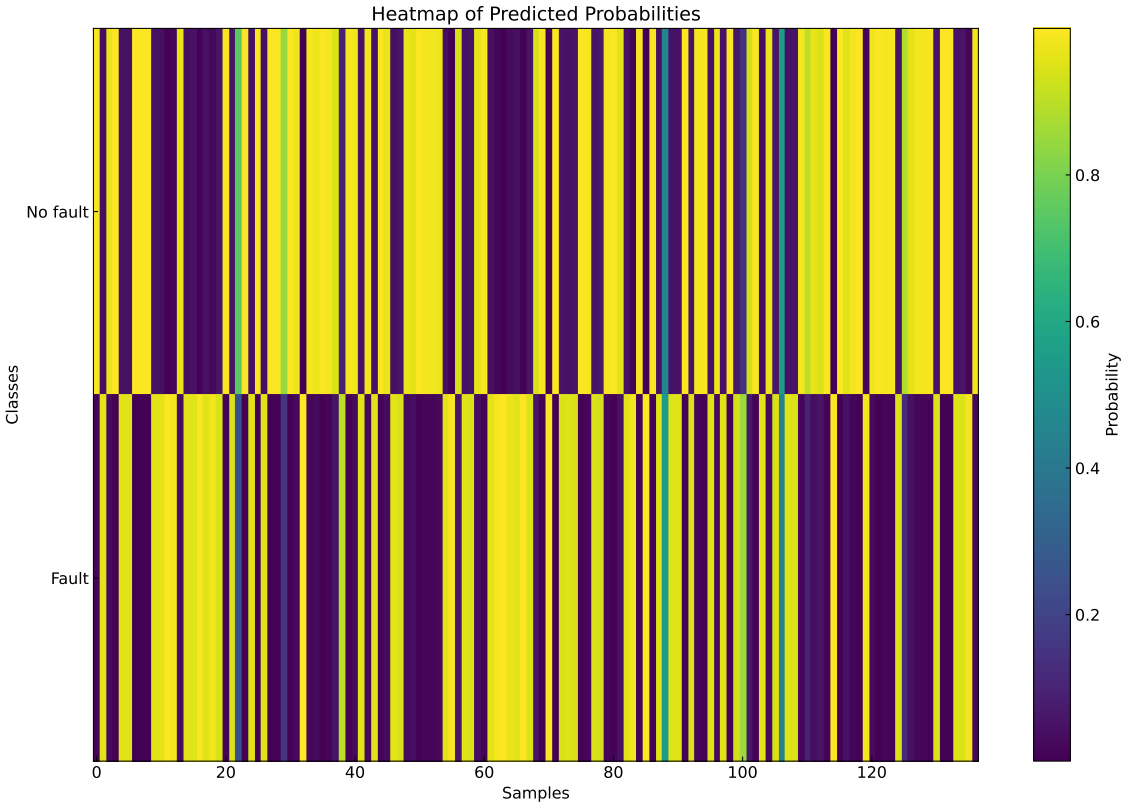


Figure 3.12: Probabilities of a fault in terms of the index on the test dataset.

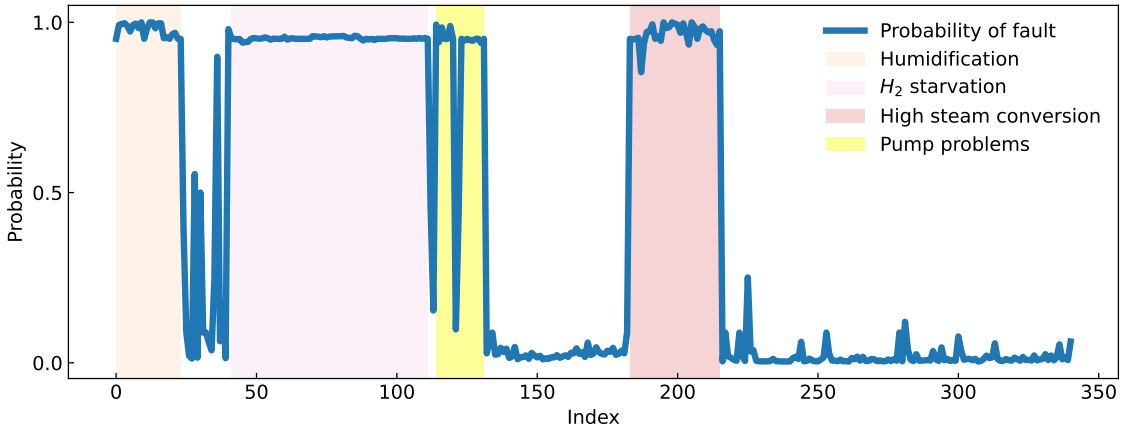


Figure 3.13: Probability of fault on the fault detection dataset.

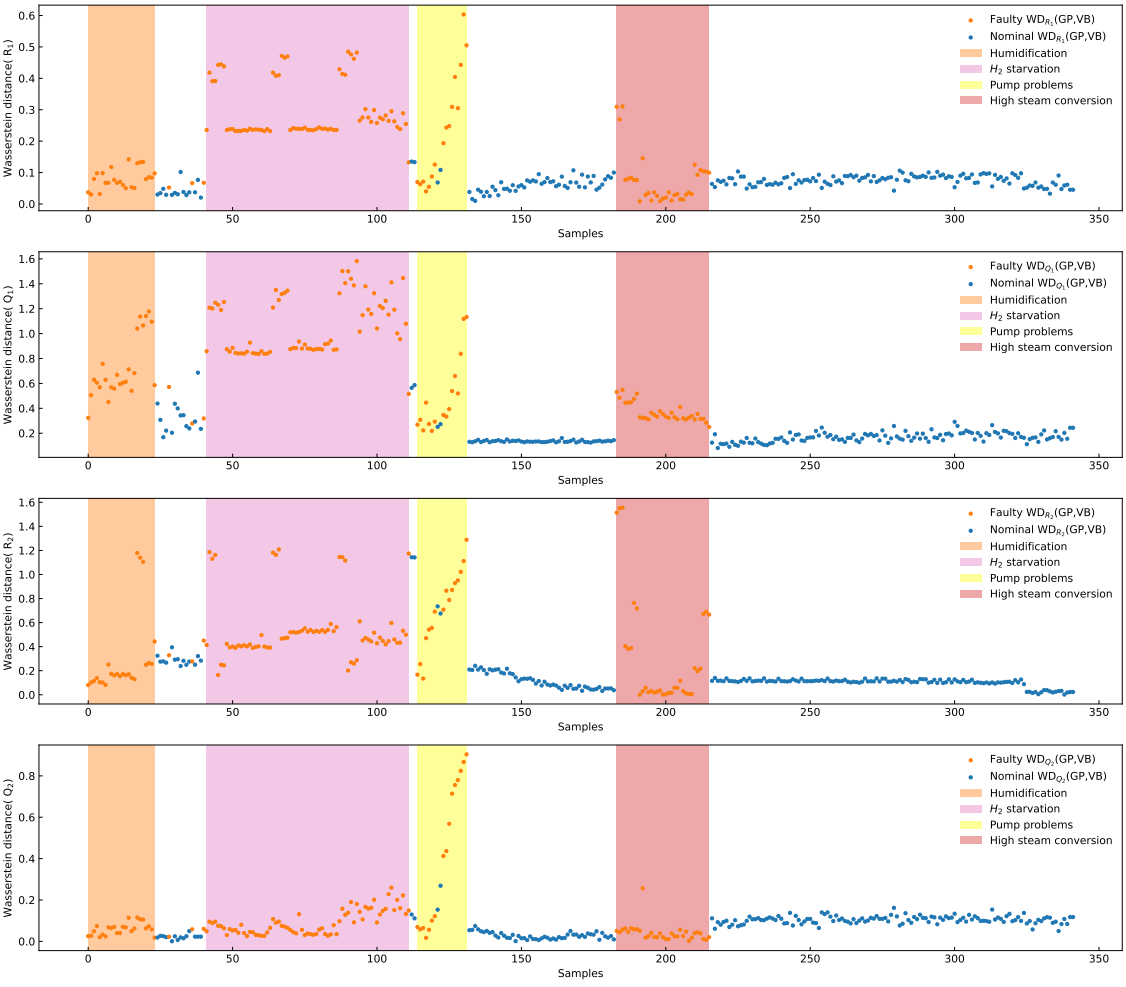


Figure 3.14: Detected faults on the WD values for each parameter with marked faults in orange and detected nominal operation mode in blue dots. The induced errors are also marked with coloured rectangles.

Training and validation Set Construction for the SVC fault isolation. The training set $\mathcal{L}$, validation set $\mathcal{V}$, and test set $\mathcal{T}$ were constructed in a nearly identical manner

as for fault detection. The only difference is that the construction of pairs $(\boldsymbol{wd}_i, d_i^{\text{iso}})$ now takes the values from the d^{iso} categorical array of 5 possible categories instead of d, which only had "Fault", "No Fault" values.

To avoid overfitting, we again performed 5-fold cross-validation on the generated datasets and obtained the predictions using the best set of hyperparameters.

Fault isolation results. To summarise the results of the SVC for fault isolation, it is useful to examine the corresponding confusion matrix in Figure 3.15. The classifier achieved an overall accuracy of 96% on the test set, with a balanced performance across different fault types.

For nominal values, the classifier demonstrated strong precision (99%) and recall (95%), correctly predicting 72 out of 76 instances. However, two nominal instances were misclassified as Fault 3, and one nominal instance was misclassified as Fault 1. For Fault 1, the classifier achieved a perfect recall (100%), correctly identifying all eight faulty instances, but there was one false positive, where a nominal instance was classified as Fault 1.

Fault 2 had excellent precision (97%) and perfect recall (100%), with all 31 instances classified correctly. Similarly, Fault 4 showed perfect performance with a precision and recall of 100%, correctly identifying all 15 instances. Fault 3, however, presented some challenges. While the classifier had good recall (86%) and precision (75%), one instance of Fault 3 was misclassified as nominal.

The classification report, as shown in Table 3.5, further highlights these performance metrics. Despite a few misclassifications, the model performed robustly across most classes, especially for Faults 2 and 4. The macro and weighted averages for precision, recall, and F1-score were 0.92, 0.96, and 0.94, respectively, indicating solid overall performance.

These results suggest that while the SVC is generally effective at isolating faults, it shows some limitations, particularly in distinguishing between nominal values and certain fault types like Fault 1 and Fault 3. Further refinement, especially in noise handling and feature extraction, and increasing the available data could improve the fault isolation capability for these edge cases.

Table 3.5: 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
Macro average	0.92	0.96	0.94	137
Weighted average	0.97	0.96	0.96	137

The predicted probability of faults during the whole test dataset is presented in Figure 3.16.

To better understand the capabilities of the SVC and put the predicted fault probabilities into a time-wise order, we also calculated the probability of each fault for the whole dataset. Let us note here that this biases the algorithm, artificially increasing the performance of this dataset since the data used for training will naturally have a higher chance

of being predicted correctly. The predicted probability of faults for the whole dataset is shown on Figure 3.17.

Finally, the WD values over the whole dataset are presented in Figure 3.18. To illustrate the predicted faults of the algorithm, the plot was color-coded based on the fault that the SVC predicted. The falsely diagnosed faults can also be seen at the start of the timeline, usually during one of the algorithm WD values raising unexpectedly. We can assume that the correctly predicted no-fault points that unexpectedly increased residuals come from the training dataset.

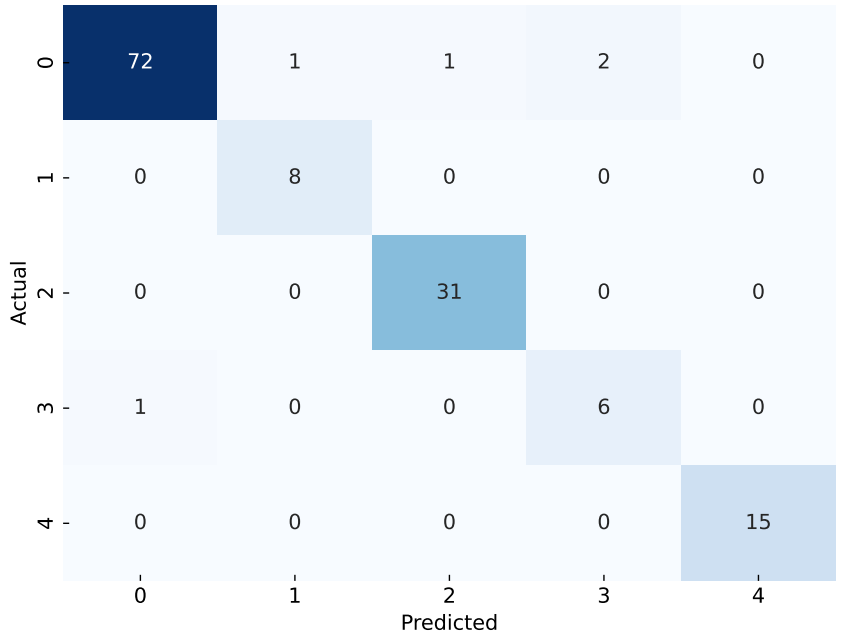


Figure 3.15: Resulting confusion matrix of the fault isolation on the test dataset.

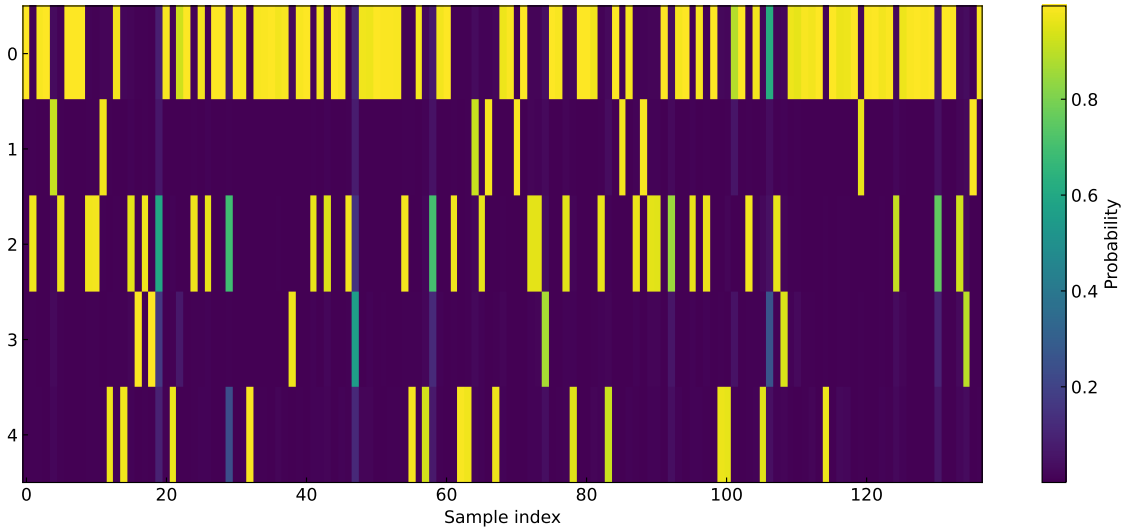


Figure 3.16: Probabilities for each fault over all of the test sample indexes.

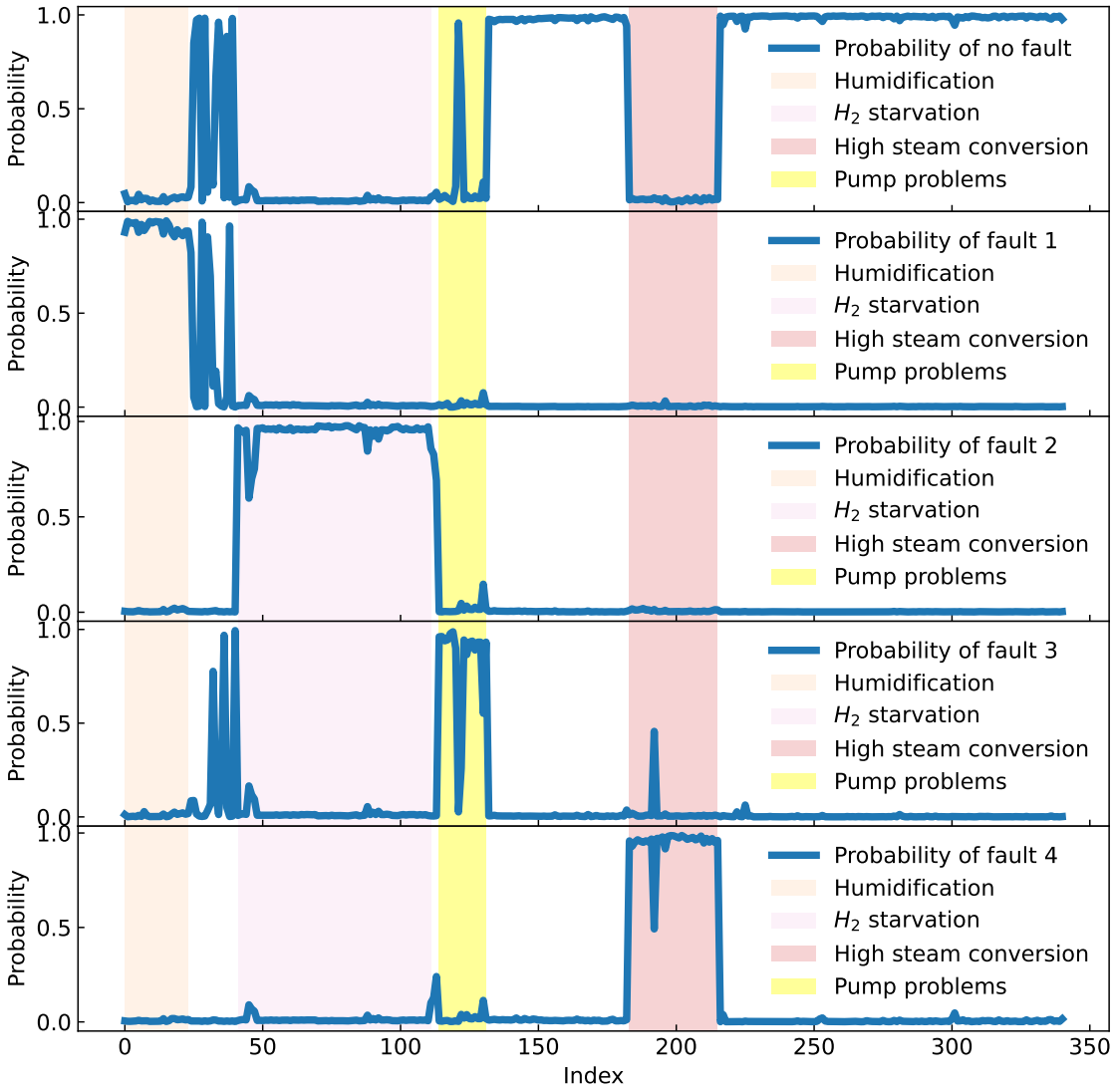


Figure 3.17: Probabilities for each fault and the nominal operation over the whole fault isolation dataset.

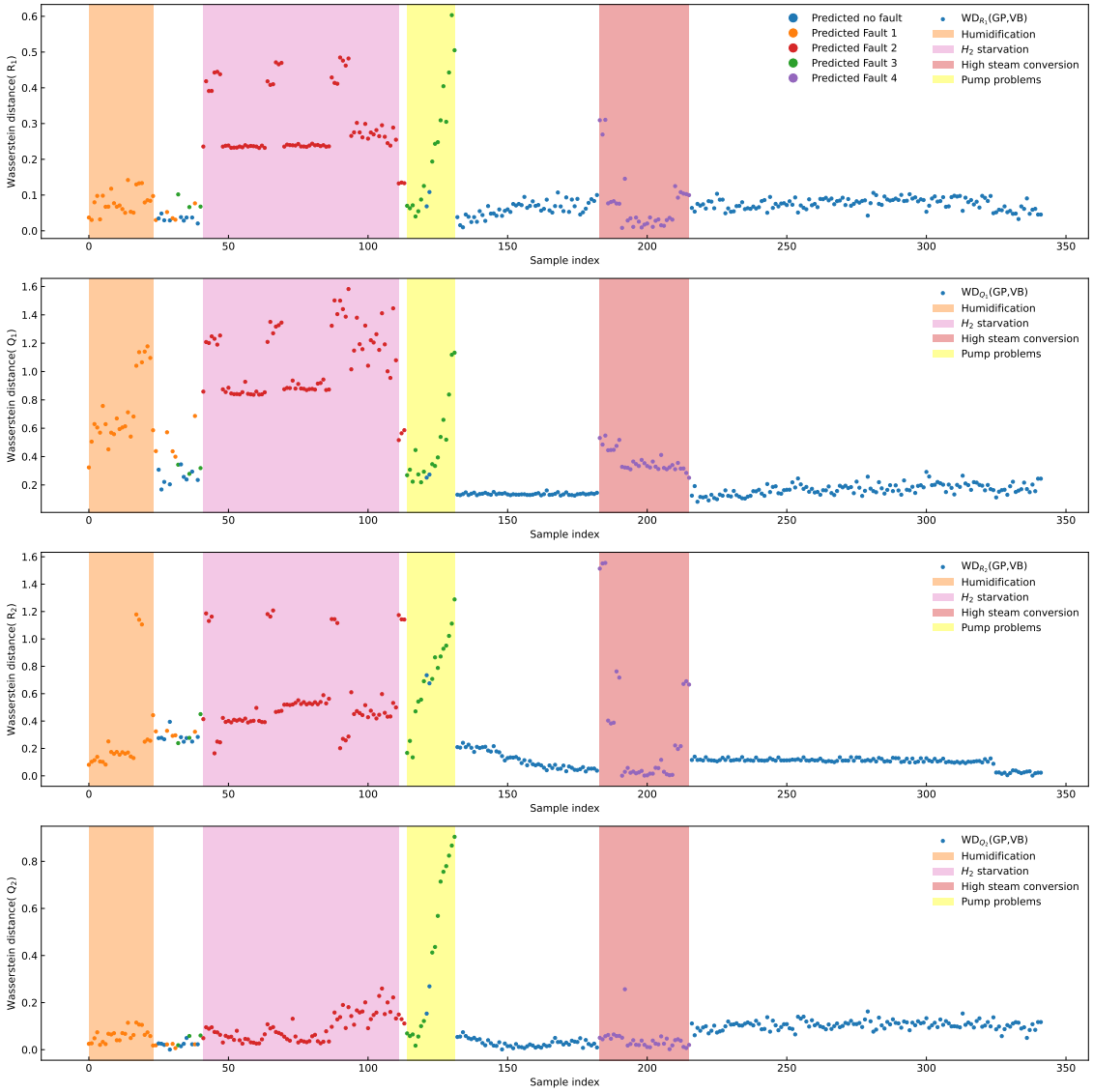


Figure 3.18: Isolated faults over each separate parameter's whole fault isolation dataset. Each fault is marked with a separate colour, and the nominal operation values are marked with blue.

Chapter 4

Health Prognosis

In the context of predictive maintenance and asset management at large prognosis focuses on predicting the health state of a system and determining the Remaining Useful Life (RUL) of its components or the system as a whole. Prognosis offers valuable insights into the degradation processes and failure mechanisms of SOC systems. By comprehending the factors and conditions that cause cells to deteriorate over time, advancements can be made in the design, materials, and operational practices to improve performance and extend longevity [85]. By adopting effective prognosis methods one can significantly boost the safety of SOC systems. Preventive actions can be taken by anticipating potential failures before they happen, thus reducing the risks associated with abrupt and unforeseen cell failures.

As pointed in the subsection 1.3, degradation is a complex process, which depends on a range of internal and external factors. Definitely, there are many uncertainties that have to be taken into account during the prediction process, such as sensor noise, environmental conditions, and variations in usage patterns. That is why the result of prognosis should be expressed in the probabilistic terms.

Much of the research endeavour in prognostics has been focused on deriving detailed multiphysics and multiscale models from the micro-structural level to the system level to understand the conditions that cause the degradation and its rate as well as the relationship with the reliability and the life span [86], [87]. Due to the complexity of the underlying models and the required computational load, their use in online diagnosis and prognosis is not feasible.

Thanks to the increasing volume of historical data acquired from the rising number of SOC systems under test in the laboratory and in situ operation, room for using different machine learning approaches to either predict the degradation trend or estimate the RUL has become widely open.

Many prognostic approaches rely on the idea that the degradation can be viewed as a hidden random walk process. To deconvolute the degradation dynamics from the available process data, the authors developed a nonlinear lumped model of the process and the black-box model of the degradation dynamics. The temporal evolution of the ASR is obtained using filtering techniques. Based on that, the RUL predictions are evaluated using MC simulation. Similar efforts are evident in [88], employing a particle filter for ASR estimation. This method demonstrates superior accuracy in contrast to the Kalman filtering approach. In an attempt to include more prior knowledge in the degradation model, in [55], the authors showed that under proper conditions, the quality of prognostics can be slightly improved compared to the pure black box model. Similarly, in [89], the authors propose a model-based fault prognosis method for SOFC that combines degradation modeling and particle filtering to estimate the state of health (SOH), EOL, and RUL.

The hidden random walk models of the degradation might face two issues. First, it should be noted that a range of phenomenological models has been proposed in [90] based on a thorough investigation of the degradation phenomena under accelerated tests. The decision of which model to apply depends on correctly identifying the degradation mode. Incorrect diagnosis implicates the use of the incorrect degradation model, which might, in turn, lead to misleading results. Second, it can happen in practice that one degradation mode eventually unfolds other degradation modes so that such cascades are complicated to identify.

From own experience based on past experimental runs, it seems the degradation progresses mainly as a monotone process [91] or even a piece-wise linear one [92]. Once the health indicator is selected, the prediction of its future evaluation can be anticipated using time series analysis tools. In this dissertation, we apply a probabilistic linear trend analysis to find the optimum trade-off between the simplicity of the prognostic algorithm and the quality of the prognosis. The idea is to estimate the local trend using Ordinary Least Squares (OLS) and then estimate the RUL when the health indicator is anticipated to hit a limit value that marks the beginning of the failure state. A novelty of this approach presented in [93] is that the distribution of the FHT can be expressed in closed form, meaning that the computational load is almost nil. Thus, the algorithm can be easily applied to industrial controllers with limited computational power.

4.1 Fault Prognosis Methodology

The idea of estimating the RUL relies on predicting the future evolution of a health indicator. Then RUL equals to the time needed for the health indicator to hit a limit value (value that marks the end of life). To do so we use the idea of FHT of a random process. To provide as much information about the RUL we aimed to provide a distribution of FHT, which would be much more telling than a single value.

The idea behind the presented prognosis method is basic and trivial since it assumes that the trend is linear and monotone. Therefore, we can imagine the health metric as a noisy line or a random walk with a trend. A model for such a behaviour can be a simple line with noise, which makes it rather straightforward to accurately represent the system. The FHT value of a line model $x(t) = n + kt, \quad n, k \in \mathbb{R}$, can be calculated from the following equation:

$$t = \frac{\alpha - n}{k} = \frac{\zeta}{k}, \tag{4.1}$$

where t is the remaining time to failure or FHT, α is the predetermined threshold value, $\zeta = \alpha - n$ and n and k are the intercept and slope of the line model, respectively. Considering the presented values as distributions, we can derive the distribution of FHT in closed form. Since the FHT distribution is a ratio of two normal distributions, we denote it *ratio distribution*. This is presented below in more detail.

4.1.1 Local linear trend model

The observed degradation metric is taken as a time series. A rectangular window function of length M is used to select a part of the data - usually the latest M measurements. For each instance of windowed data, we propose the latent state of the health index, denoted as x, to behave as a linear function

$$x(t) = n + kt, \quad n, k \in \mathbb{R}, \tag{4.2}$$

while the observed state, denoted as y, is just the latent state with superimposed normally distributed white noise

$$y(t) = x(t) + \eta(t), \quad \eta(t) \sim \mathcal{N}(0, \sigma_\eta^2), \ \sigma_\eta \in \mathbb{R}^+. \tag{4.3}$$

Since we only get our measurements at discrete times t_j, $j \in (1, \ldots, N)$, we can rewrite the upper equation to matrix form

$$\boldsymbol{Y} = \begin{bmatrix} y(t_{N-M+1}) \\ y(t_{N-M}) \\ \vdots \\ y(t_N) \end{bmatrix} = \boldsymbol{F} \begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} \eta(t_{N-M+1}) \\ \eta(t_{N-M}) \\ \vdots \\ \eta(t_N) \end{bmatrix}, \text{ where } \boldsymbol{F} = \begin{bmatrix} 1, t_{N-M+1} \\ 1, t_{N-M} \\ \vdots \\ 1, t_N \end{bmatrix}. \tag{4.4}$$

From well-known OLS theory [94] follows that the most likely value of $[n, k]^\top$, denoted as $\boldsymbol{Z}$, is

$$\mathrm{E} \begin{bmatrix} n \\ k \end{bmatrix} = \boldsymbol{Z} = (\boldsymbol{F}^\top \boldsymbol{F})^{-1} \boldsymbol{F}^\top \boldsymbol{Y}, \tag{4.5}$$

the uncertainty of $[n, k]^\top$, denoted as $\boldsymbol{R}$, is

$$\mathrm{Cov} \begin{bmatrix} n \\ k \end{bmatrix} = \boldsymbol{R} = (\boldsymbol{F}^\top \boldsymbol{F})^{-1} \sigma_\eta^2, \tag{4.6}$$

and the noise level is

$$\sigma_\eta^2 = \frac{1}{M-2} (\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{Z})^\top (\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{Z}). \tag{4.7}$$

The idea of the method is showcased in Figure 4.1, highlighting using a data window to predict the distribution of FHT.

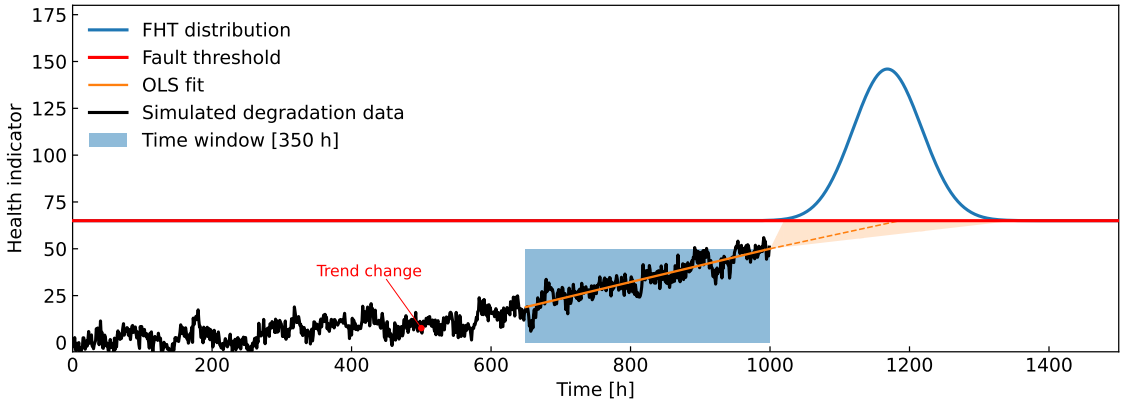


Figure 4.1: Example of the general inner workings of the proposed method. Since a change in the trend of the observed health metric is present, not all of the available data should be taken into account. A (blue rectangle) data window chooses the most recent data, from which a prediction of the FHT distribution (blue curve) is calculated.

4.1.2 Distribution of the ratio of two jointly normal variables

Again, the degradation with time, denoted as $x(t)$, is modeled as a linear function with slope k and intercept n. To account for the uncertainty of the said two parameters, we consider them probabilistic. Their joint distribution is taken to be multivariate-normal

$$\begin{bmatrix} n \\ k \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_n \\ \mu_k \end{bmatrix}, \begin{bmatrix} \sigma_n^2, & \rho_{n,k}\sigma_n\sigma_k \\ \rho_{n,k}\sigma_k\sigma_n, & \sigma_k^2 \end{bmatrix} \right), \tag{4.8}$$

where

$$\begin{bmatrix} \mu_n \\ \mu_k \end{bmatrix} = \boldsymbol{Z}, \quad \mu_n, \mu_k \in \mathbb{R}, \qquad \text{and} \tag{4.9}$$

$$\begin{bmatrix} \sigma_n^2, & \rho_{n,k}\sigma_n\sigma_k \\ \rho_{n,k}\sigma_k\sigma_n, & \sigma_k^2 \end{bmatrix} = \boldsymbol{R}, \quad \sigma_n, \sigma_k \in \mathbb{R}^+,\ \rho_{n,k} \in [-1, 1]. \tag{4.10}$$

The threshold, denoted as α, is also considered uncertain, meaning

$$\alpha \sim \mathcal{N}\left(\mu_\alpha, \sigma_\alpha^2\right), \quad \mu_\alpha \in \mathbb{R},\ \sigma_\alpha \in \mathbb{R}^+. \tag{4.11}$$

So if $\alpha = kt + n$, then t, interpreted as the FHT, is

$$t = \frac{\alpha - n}{k} = \frac{\zeta}{k}, \tag{4.12}$$

where $\zeta = \alpha - n$. The joint distribution of ζ and k is then

$$\begin{bmatrix} \zeta \\ k \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_\zeta \\ \mu_k \end{bmatrix}, \begin{bmatrix} \sigma_\zeta^2, & \rho_{\zeta,k}\sigma_\zeta\sigma_k \\ \rho_{\zeta,k}\sigma_k\sigma_\zeta, & \sigma_k^2 \end{bmatrix} \right), \tag{4.13}$$

where

$$\begin{aligned} \mu_\zeta &= \mu_\alpha - \mu_n, \\ \sigma_\zeta^2 &= \sigma_\alpha^2 + \sigma_n^2, \\ \rho_{\zeta,k} &= \tfrac{-\rho_{n,k}\sigma_n}{\sigma_\zeta}. \end{aligned} \tag{4.14}$$

It is clear that $t = \zeta/k$ is a ratio distribution of two (possibly correlated) normal variables. Fortunately, the probability density function for such distributions is available in closed form [95]

$$\begin{aligned} p_{\mathrm{FHT}}(t) = {} & \frac{\sigma_k\sigma_\zeta\left(1-\rho_{\zeta,k}^2\right)^{1/2}}{\pi\left(\sigma_k^2t^2 - 2\rho_{\zeta,k}\sigma_k\sigma_\zeta t + \sigma_\zeta^2\right)} \exp\left[-\frac{1}{2\left(1-\rho_{\zeta,k}^2\right)}\left(\frac{\mu_k^2}{\sigma_k^2} - 2\rho_{\zeta,k}\frac{\mu_k}{\sigma_k}\frac{\mu_\zeta}{\sigma_\zeta} + \frac{\mu_\zeta^2}{\sigma_\zeta^2}\right)\right] \\ & + \frac{\mu_k\sigma_\zeta^2 - \mu_\zeta\rho_{\zeta,k}\sigma_k\sigma_\zeta + \left(\mu_\zeta\sigma_k^2 - \mu_k\rho_{\zeta,k}\sigma_k\sigma_\zeta\right)t}{\sqrt{2\pi}\left(\sigma_k^2t^2 - 2\rho_{\zeta,k}\sigma_k\sigma_\zeta t + \sigma_\zeta^2\right)^{3/2}} \\ & \times \exp\left(-\frac{(\mu_\zeta - \mu_k t)^2}{2\left(\sigma_k^2t^2 - 2\rho_{\zeta,k}\sigma_k\sigma_\zeta t + \sigma_\zeta^2\right)}\right) \\ & \times \left[1 - 2Q\left(\frac{\mu_k\sigma_\zeta^2 - \mu_\zeta\rho_{\zeta,k}\sigma_k\sigma_\zeta + \left(\mu_\zeta\sigma_k^2 - \mu_k\rho_{\zeta,k}\sigma_k\sigma_\zeta\right)t}{\sigma_k\sigma_\zeta\left(1-\rho_{\zeta,k}^2\right)^{1/2}\left(\sigma_k^2t^2 - 2\rho_{\zeta,k}\sigma_k\sigma_\zeta t + \sigma_\zeta^2\right)^{1/2}}\right)\right], \end{aligned} \tag{4.15}$$

where

$$Q(x) = \frac{1}{2}\left(\text{erfc}\left(\frac{x}{\sqrt{2}}\right)\right). \tag{4.16}$$

The erfc is the complementary error function, defined as $\text{erfc}(z) = 1 - \text{erf}(z)$, where $\text{erf}(z)$ stands for the error function

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt. \tag{4.17}$$

Equation (4.15) can then be used to calculate the probability density function of the FHT distribution to obtain the estimates for RUL in a closed-form manner.

It is important to stress that it is tacitly assumed that the indicator increases together with the degradation level, i.e., it should have a positive trend. In cases of a high noise level combined with a short time window, the slope μ_k of the estimated local trend model at time t_N might turn negative. In that case, the intersection (4.12) occurs at $t < t_N$, which is not a feasible solution and is therefore discarded.

Rationale behind the estimation of the model We considered different methods for estimating the model parameters from the health indicator trend before settling on the LLS method.

We considered Gauss-Wishart (G-W) and ARMA algorithms, but the improvements were minimal, while the complexity, albeit still low, did increase. We have therefore selected the most simple and non-complex solution that did produce accurate results.

4.2 Results on Simulated Health Indicator Data

Four different simulation scenarios are analyzed to demonstrate the performance of the method above.

First, a noisy time series with a fixed linear trend in the latent variable $x(t)$ is presented.

In the second case, the latent variable undergoes an abrupt change in trend. This case highlights the impact of the window length on the adaptivity of the local linear trend model.

In the third case, the latent variable suffers from oscillations, modeled by an ARMA [96] process. This example aims to demonstrate the proposed method's ability to account for the additional uncertainty in the FHT caused by the oscillations in the latent variable.

The fourth case is similar to the third one with one difference: the parameters of the latent model are subjected to an abrupt change, once again illuminating the importance of window length.

4.2.1 Case 1: Time series with a fixed linear trend

In this simplest case, the latent variable $x(t)$ exhibits a fixed linear trend, while the measured health indicator $y(t)$ is corrupted with noise

$$x(t) = n + kt, \quad n, k \in \mathbb{R} \quad \text{(Latent health metric)}, \tag{4.18}$$

$$y(t) = x(t) + \xi(t), \quad \xi \sim \mathcal{N}(0, \sigma_\xi^2), \quad \sigma_\xi \in \mathbb{R}^+ \quad \text{(Observed health metric)}. \tag{4.19}$$

The parameters used in the simulation are given in Table 4.1.

Table 4.1: Parameters of the linear trend model with fixed parameters (case 1).

Simulated scenario	n	k	σ_η^2	α
Case 1	0	1	30	600

The simulated time series with a fixed linear trend is plotted in Figure 4.2. The time between consecutive samples is $t_{i+1} - t_i = 1\text{h}$. At time $t_N = 400\text{h}$, several local linear trend models are estimated over data windows $M \in \{400, 200, 60\}$ and then used to evaluate the distribution of FHT. Not surprisingly, utilizing all the available data provides the most accurate prediction of FHT.

4.2.2 Case 2: Time series with changing linear trend

Like case 1, the latent variable follows a linear function, but its parameters undergo a change during the simulated process

$$x(t) = (1 - \phi(t))\,(k_1 t + n_1) + \phi(t)(k_2 t + n_2), \quad k_1, k_2, n_1, n_2 \in \mathbb{R} \quad \text{(Latent health metric)}, \tag{4.20}$$

$$\phi(t) = \frac{1}{1 + \exp\left(-\kappa(t - \tau)\right)}, \quad \kappa \in \mathbb{R}^+,\ \tau \in \mathbb{R}, \tag{4.21}$$

$$n_2 = (k_1 - k_2)\tau + n_1, \tag{4.22}$$

$$y(t) = x(t) + \xi(t), \quad \xi \sim \mathcal{N}(0, \sigma_\xi^2), \quad \sigma_\xi \in \mathbb{R}^+ \quad \text{(Observed health metric)}, \tag{4.23}$$

where $\phi(t)$ is the sigmoid function which changes the system from $x(t) = k_1 t + n_2$ to $x(t) = k_2 t + n_2$. With the parameters τ and κ, one defines the transition time and the

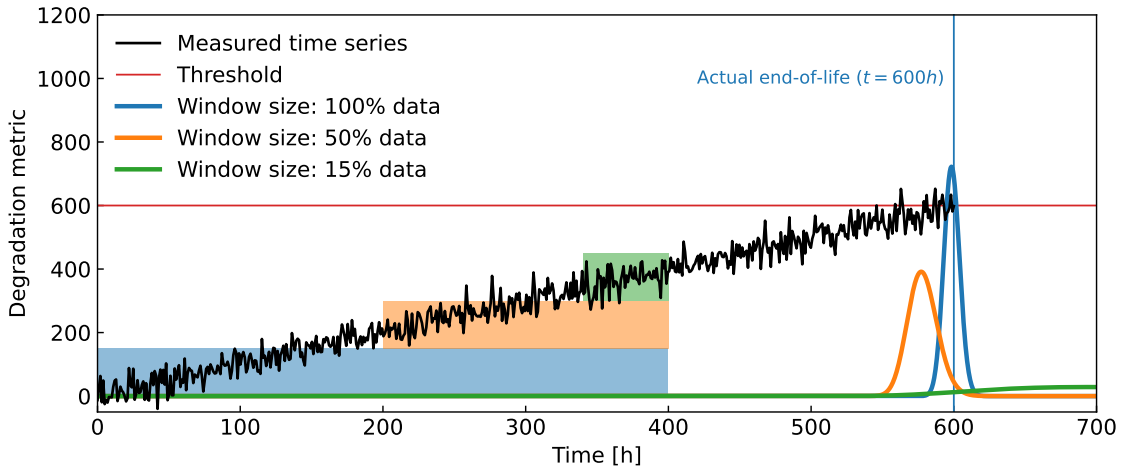


Figure 4.2: Simulated time series with a fixed linear trend. The measured $y(t)$ are collected for 0h $\leq t \leq$ 400h, and the distributions of FHT are evaluated at t=400h for three different windows. The shortest window is [340h,400h], which means the last 15% of the available data, while the longest window [0h,400h] utilizes all available data.

Table 4.2: Parameters of the linear trend model with switching parameters (case 2).

Simulated scenario	n	k	σ_η^2	τ [h]	α	κ [h^{-1}]
Case 2	$n_1 = 0$	$k_1 = 1$	30	250	800	0.03
	n_2 (eq. 4.22) $= -250$	$k_2 = 3$	30			

transition rate, respectively. Note that n_2 is not a free parameter to ensure a more natural-looking transition. The parameters used to simulate the second example are provided in Table 4.2.

In Figure 4.3, the simulated time series and the predicted distributions of the FHT obtained from local models estimated on three different data windows are presented. Due to shifts in the trend, the forecasted FHT distribution from shorter windows proves more accurate than predictions derived from longer time windows. However, excessively brief windows are not advisable. The predictive quality diminishes as windows become too short, primarily due to heightened noise interference, particularly when the noise-to-signal ratio is high. In this case, we opted to show the predictions from window lengths that effectively highlight both the advantages and drawbacks of varying window durations.

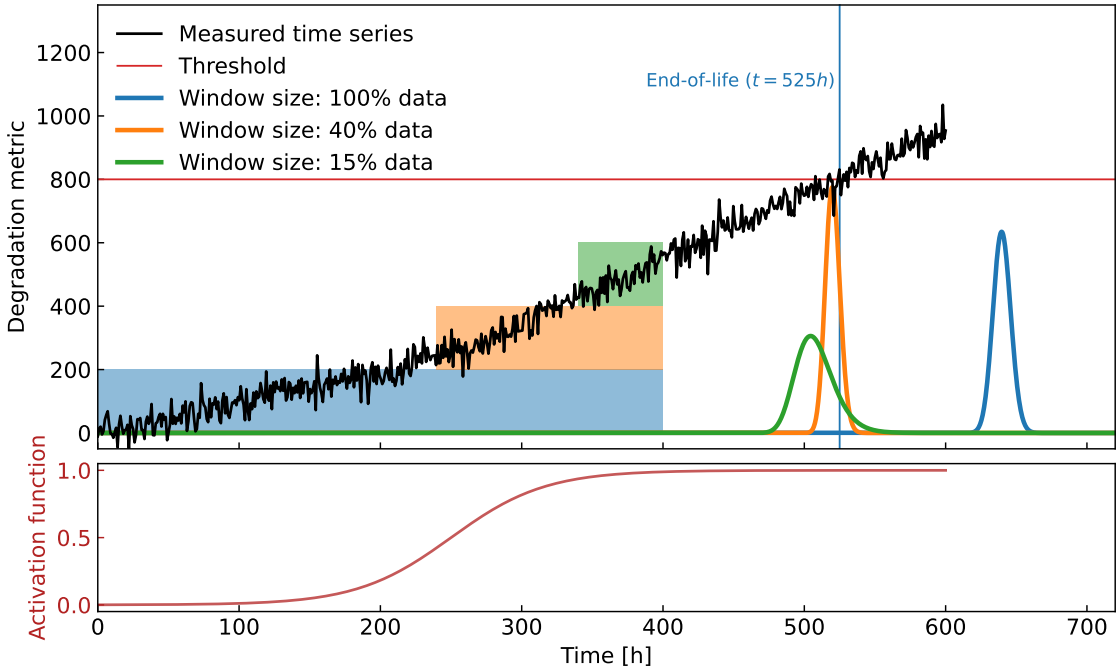


Figure 4.3: (Top) Time series with an abrupt change in the linear trend and the distributions of the FHT obtained from local models evaluated over different data windows. Note that the trend model evaluated from the most recent data provides the most accurate prediction of the FHT. In this case, utilizing only 40% of the latest data yields the best result. (Bottom) The sigmoid function switches between two linear trend models with different parameters.

4.2.3 Case 3: Time series resulting from ARMA process with fixed parameters

A more general case of a time series containing a linear trend and an oscillatory component is analyzed $x(t) = x_{\text{trend}}(t_j) + x_{\text{osc.}}(t_j)$. For that purpose ARMA$(p, q)$ is utilized

$$x(t_j) = x_{\text{trend}}(t_j) + x_{\text{osc.}}(t_j) \quad \text{(Latent health metric)}, \tag{4.24}$$

$$x_{\text{trend}}(t_j) = lt_j + m, \quad l, m \in \mathbb{R}, \tag{4.25}$$

$$x_{\text{osc.}}(t_j) = \sum_{i=1}^{p} \phi_i x_{\text{osc.}}(t_{j-i}) + \sum_{i=1}^{q} \theta_i \epsilon(t_{j-i}) + \epsilon(t_j), \quad \epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2), \quad \phi_i, \theta_i \in \mathbb{R},\ \sigma_\epsilon \in \mathbb{R}^+, \tag{4.26}$$

$$y(t_j) = x(t_j) + \xi(t_j), \quad \xi(t_j) \sim \mathcal{N}(0, \sigma_\xi^2), \quad \sigma_\xi \in \mathbb{R}^+ \quad \text{(Observed health metric)}. \tag{4.27}$$

Following the same steps as in section 4.2.1 yields FHT distributions that are too narrow since the oscillations of the latent variable are falsely regarded as white measurement noise. To circumvent this problem, we note that the scale of oscillations $\mathrm{E}[x_{\text{osc.}}^2]$ is approximately the same as the scale of the perceived white noise σ_η. We can now transfer this perceived uncertainty, caused by oscillations, to uncertainty of the threshold, so $\sigma_\alpha = \sigma_\eta$ (see Eq. 4.11). Such adaptation of the problem statement rescales the distribution of FHT to a much better width, as shown in Figure 4.4. Note that this approximation is meaningful only under relatively high signal-to-noise ratios. Otherwise, the uncertainty of the resulting

Table 4.3: Parameters of ARMA models used to generate the time series.

Simulated scenario	n	k	σ_η^2	σ_ξ^2	ϕ_1	θ_1
ARMA $(1,0)$	0	1	5	5	0.9	0
ARMA $(0,1)$	0	1	5	5	0	0.3
ARMA $(1,1)$	0	1	5	5	0.9	0.3

FHT distributions may be overestimated. Alternatively, if the actual noise level σ_ξ is known, one can subtract the actual noise level from the perceived white noise level, so $\sigma_\alpha = \sqrt{\sigma_\eta^2 - \sigma_\xi^2}$.

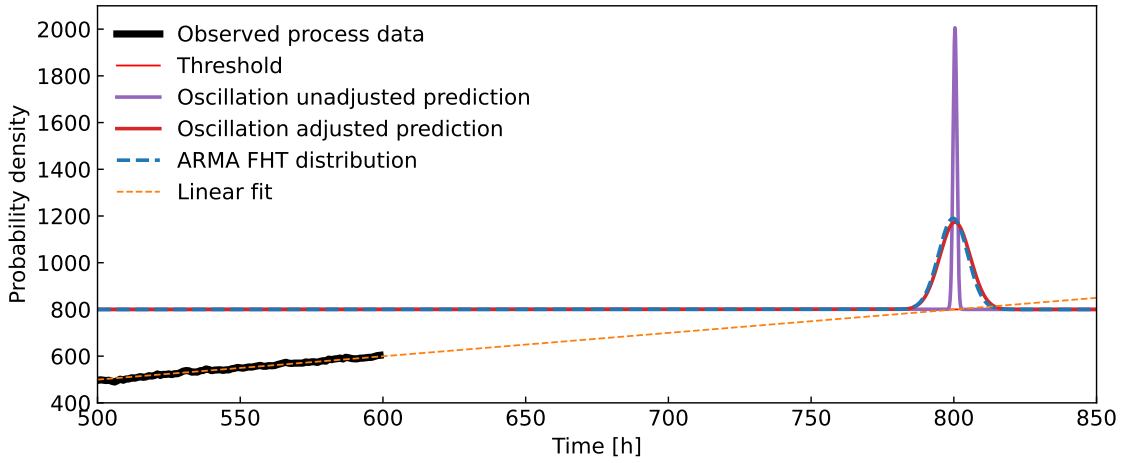


Figure 4.4: Accounting for the variability in the internal variable by using the distribution of α. The predictions rely on time series resulting from an ARMA process. The blue curve shows the ground truth, i.e., the prediction obtained with MC sampling of the latent model composed of a deterministic linear trend with ARMA$(1,0)$ oscillations, see Eq. 4.24.

To adequately test the algorithm, three different ARMA models are employed to simulate oscillating latent variables. Their parameters are presented in Table 4.3. Their typical behavior is showcased in Figure 4.5.

In Figure 4.6, the algorithm was tested using three data windows of different sizes. For a meaningful comparison to the resulting distributions, the ground-truth FHT distribution was approximated by MC sampling of the known ARMA process and compared to the results.

Examining the same situation differently, Figure 4.7 features the same latent processes as Figure 4.6. Still, instead, the quantiles of the resulting FHT distributions from various sizes of the data windows are presented.

It turns out that with enough data and far enough away from the threshold, the algorithm stands as a viable alternative to a full regression of an ARMA model for FHT prediction. It is also worth noting that retaining as much information as possible leads to more accurate predictions in this case, where the time series has a constant trend.

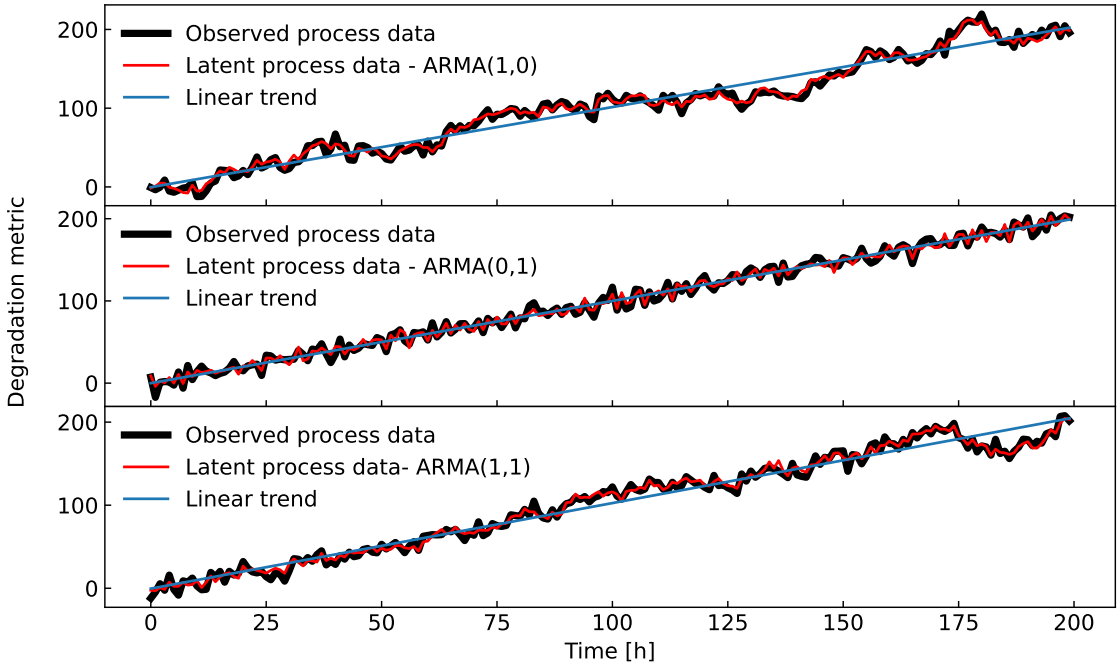


Figure 4.5: Time series generated by three different ARMA processes with the parameters given in Table 4.3.

Table 4.4: Parameters of the ARMA model with switching parameters.

Simulated scenario	n	k	σ_η^2	Time of change [h]	α
Case 4	$n_1 = 0$	$k_1 = 0.02$	0.04	500	60
	$n_2 = -5$	$k_2 = 0.03$	0.04		

4.2.4 Case 4: Time series resulting from ARMA process with abruptly changing parameters

An important feature of the prediction algorithm in dynamic situations is its flexibility to adapt quickly to changes in trends. For that purpose, we again utilize limited-size time windows. Similarly to section 4.2.2, the algorithm is tested on data generated by ARMA process featuring a change in the trend. The model parameters used for simulation are shown in Table 4.4.

In Figure 4.8 and Figure 4.9 the simulated ARMA process with a change in trend is depicted for two examples. The first example has the data from $t = [0, 1000]$ h available and for the second one, we use $t = [0, 1800]$ h. The change in trend is initiated around $t = 500$ h for both examples. For the data interval up to about $t = 500$ h, the slope of the trend is 0.02, which changes to 0.03 afterward. Resulting FHT distributions derived from different data windows are illustrated in Figure 4.8 and Figure 4.9. Incorporating only the data after the trend change can clearly improve the quality of the prognosis.

In Figure 4.10 and Figure 4.11, we show quantiles of FHT distributions from local trend models obtained from differently sized windows on the data. For example, a window $[800, 1000]$h includes 20% of the data up to 1000h. The consequence of such flexibility can be seen in frequent changes in the mean of the FHT distribution acquired from the most recent data, where the trend is not yet apparent. Ground-truth FHT distribution (dark

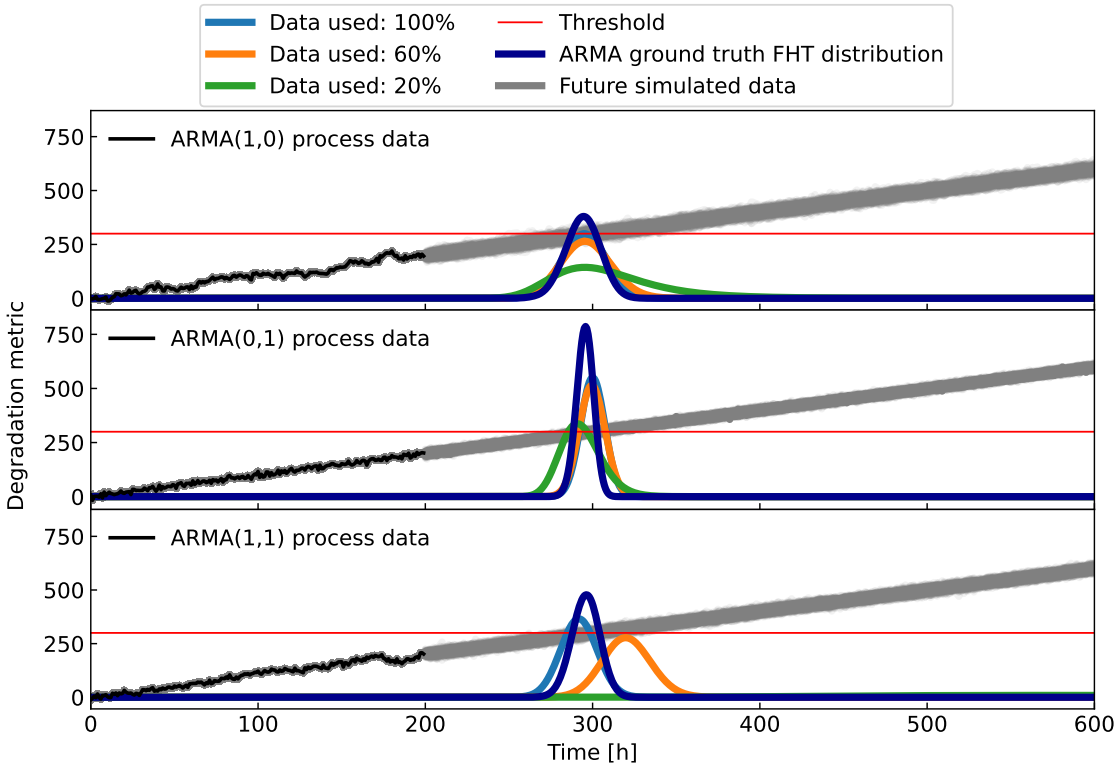


Figure 4.6: The distributions of FHT evaluated at time $t_N = 200$ for three different ARMA models. Additionally, three different windows for estimation of the local trends are used: $[160, 200]$ or the last 20%, $[120, 200]$ or the last 60% and $[0, 200]$ or 100% of the available data.

blue curve) is also provided for comparison.

We notice that adding data improves precision and accuracy up to a point. If we also include data before the trend change, the precision of the predictions does not suffer as much as the accuracy. This highlights a potential limitation in assessing the reliability of the prediction. Without the variance of the prediction ballooning, this could lead to a false sense of confidence in the operator's prediction.

We also notice a slight bias in predictions when the prediction time is close to the EOL, see Figure 4.9. Due to the limited time available, any large enough oscillation significantly impacts the threshold crossing point. Therefore, the latent variable is unable to regress to the mean before the crossing, which significantly alters the ground truth distribution.

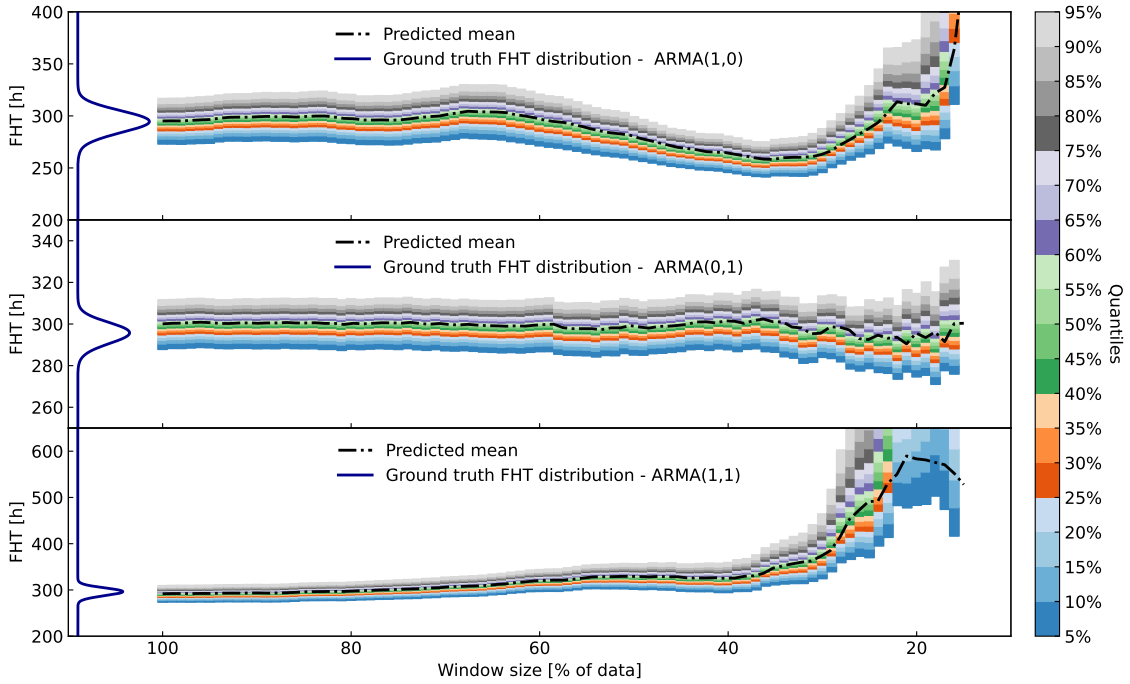


Figure 4.7: Quantile plots over time for three different simulated processes (see Fig. 4.6 for their time-signals). The predicted FHT distributions calculated by the proposed algorithm stay in line with the ground truth when at least 2/3 (from the start) of the data are considered. When data are taken at the beginning of the process (far from the FHT, e.g., the first 1/3 of the dataset), the variance of the predicted distributions grows significantly, thus signaling an unreliable prognosis.

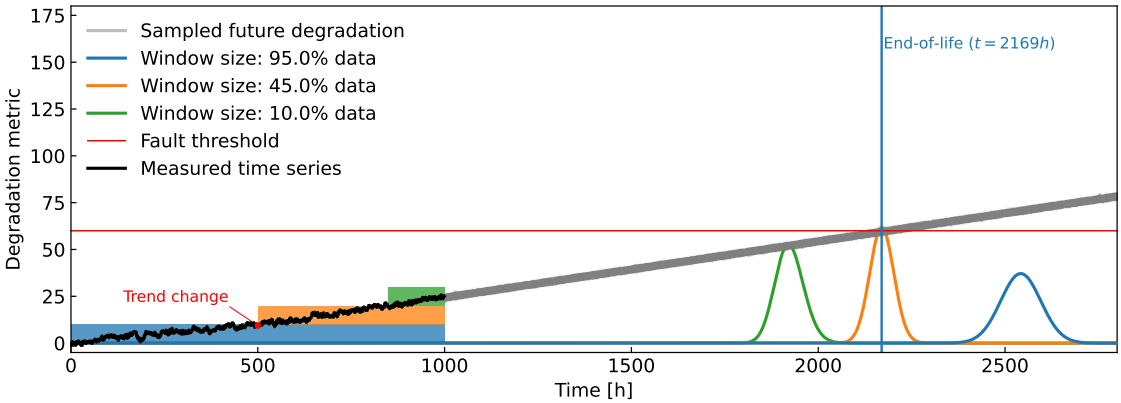


Figure 4.8: Synthetic experiment on ARMA time signal with an abrupt change in the trend (around $t = 500$h). Data available for the prognosis was limited up to $t = 1000$h. Actual EOL is compared to results from the developed algorithm with various amounts of data considered.

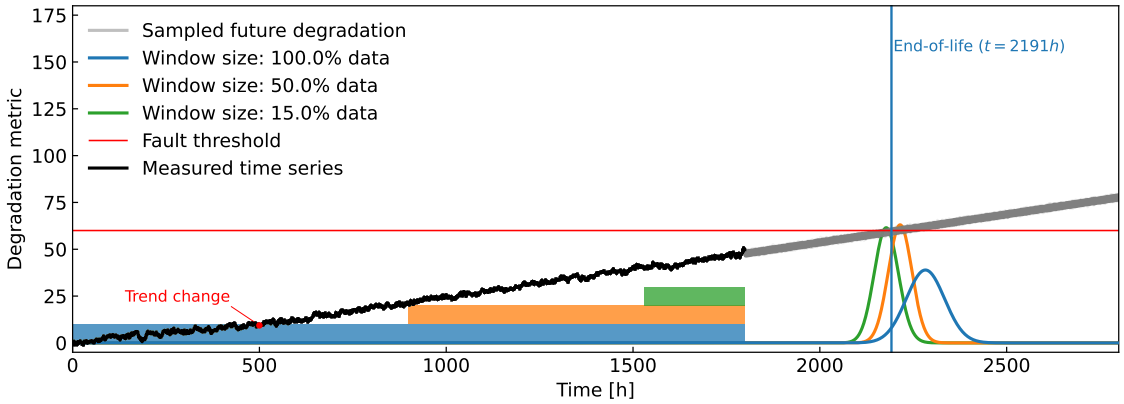


Figure 4.9: Second example of synthetic time signals of the ARMA process with an abrupt change in the trend (around $t = 500$h). The FHT distributions are derived with various amounts of data taken into account. Relating to Figure 4.8, there were more available data for the prognostic algorithm here, up to $t = 1800$h. But on the other hand, EOL is now much closer, and any large oscillations in time signal will importantly affect the actual breach of the threshold.

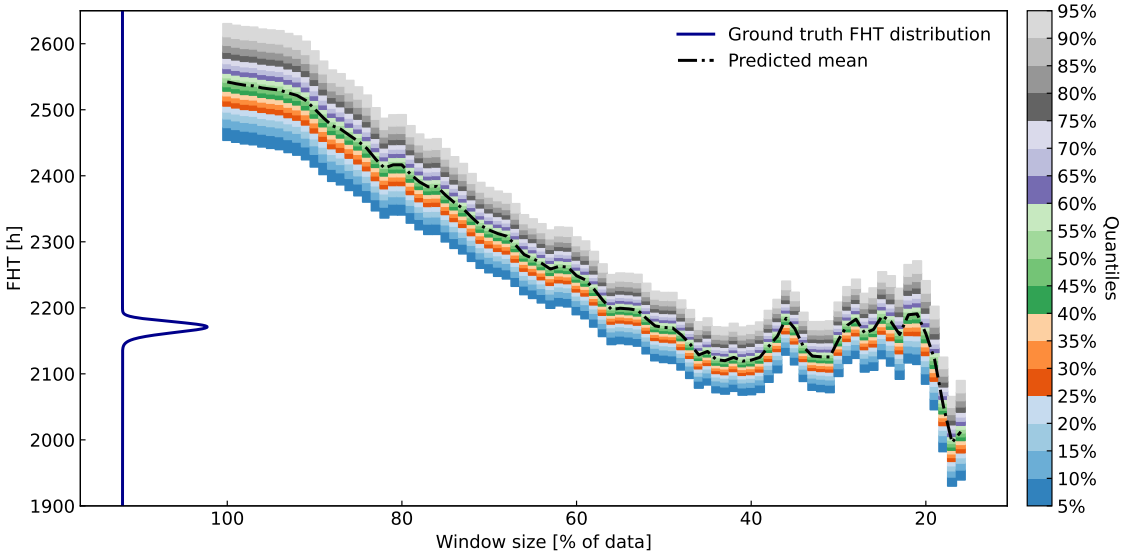


Figure 4.10: Quantiles of FHT distributions over various percentages of available data at $t = 1000$h taken into account. The artificial data is from an ARMA process with a change in the trend. On the left, ground-truth FHT is presented (dark blue curve). The prediction is best when taking at most 50% of the recent data since that is when the trend change has occurred.

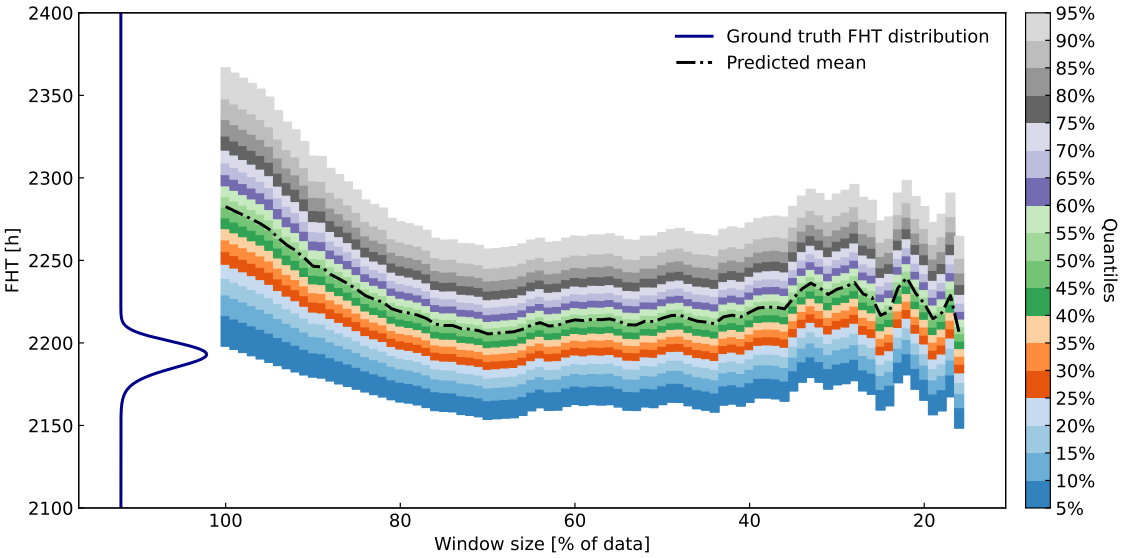


Figure 4.11: Quantiles of FHT distributions over various percentages of available data at $t = 1800$h taken into account. The artificial data is from an ARMA process with a change in the trend. On the left, ground-truth FHT is presented (dark blue curve). The prediction improves when taking up at most 60% of the recent data. Since the prediction time is relatively close to the EOL, there is a noticeable difference between the bias and the variance in the ground truth and the predictive FHT distribution.

4.3 Results on Experimental Health Indicator Data

In Figures 4.12 and 4.13, operational data of almost 1850h is presented. The system was operated in a non-stationary regime during which there were frequent changes in the load. Therefore, the stack voltage is not the best candidate for the health indicator. Instead, ASR was chosen as the health metric since it is invariant to changes in process variables (technically, we have chosen ASR_0 as the metric. For validation of the algorithm, the threshold of ASR was set to 0.11 Ω mm^2 (a red horizontal line in Figure 4.12).

Validation of the proposed method by comparison to an alternative prognostic algorithm

To assess the new algorithm on experimental data, where the latent model is unknown, we compared it to an ARMA process whose parameters were obtained by the maximum-likelihood method. Specifically, we have used statsmodels v0.14.1 [97], a Python package for inference for statistical models. After comparing several different variations of the model via the Akaike information criterion, we have landed on $\mathrm{AR}(2, 0)$ with a deterministic linear function as an exogenous variable (same framework as in (4.24)).

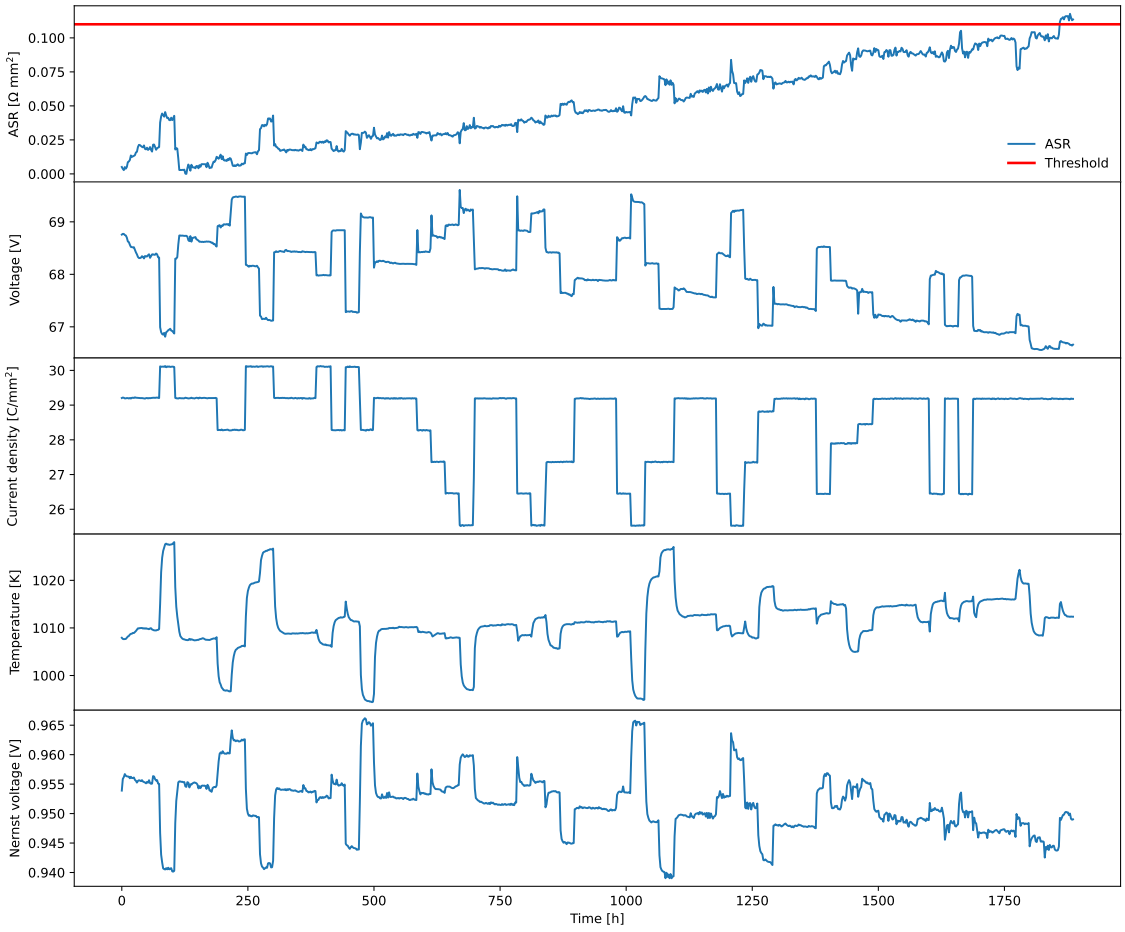


Figure 4.12: Data from SOFC stack experiment in [98]. From top to bottom: ASR, stack voltage, current density, outlet temperature and Nernst voltage.

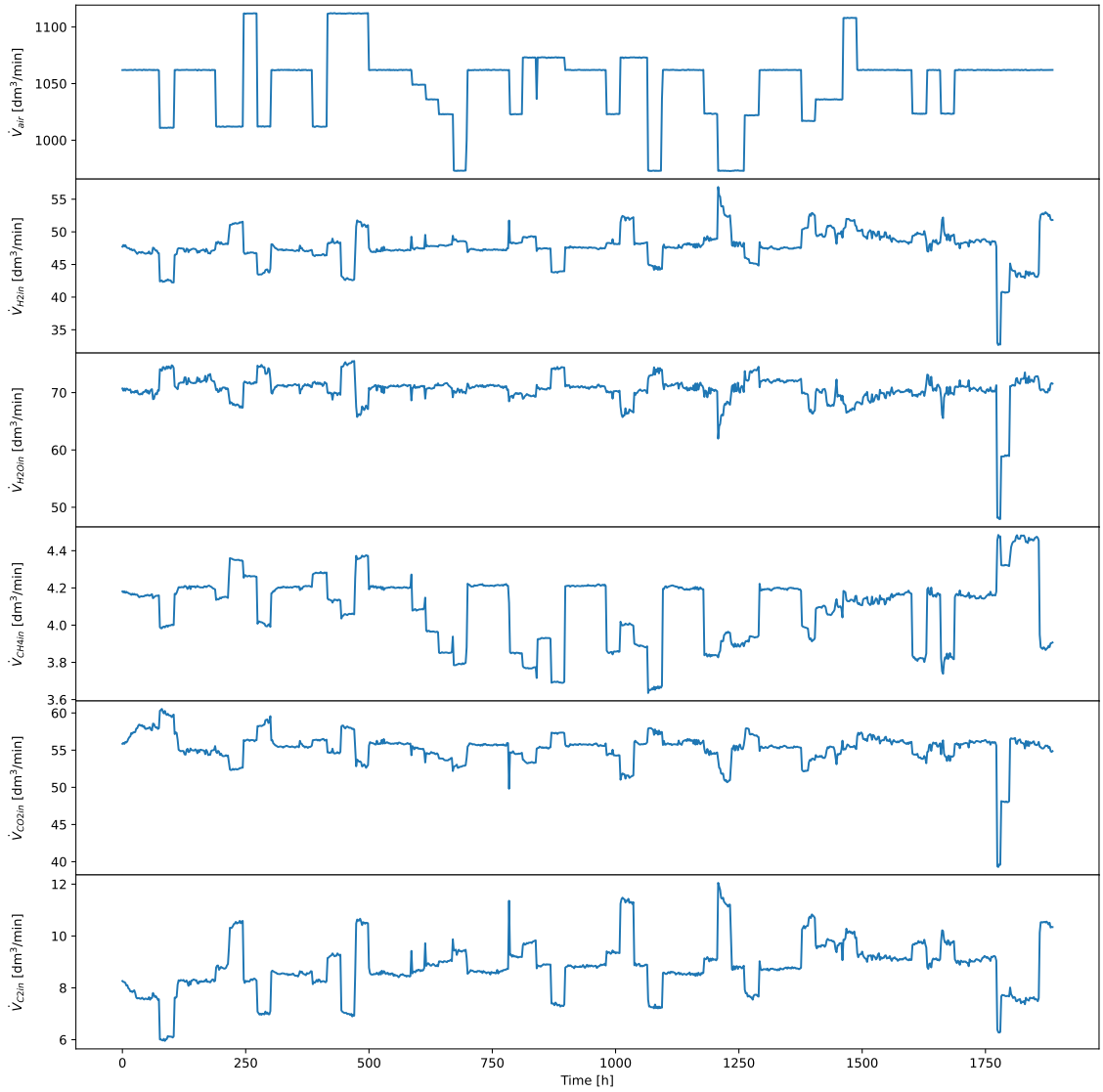


Figure 4.13: Inflow data from SOFC stack experiment in [98]. From top to bottom: Air, H_2,H_2O,CH_4,CO_2,C_2 flows.

4.3.1 Results

Results for the experimental data can be seen in Figure 4.14. The improvement in the quality of resulting FHT distribution is proportional to the number of data points included in the prognosis, which is expected since the ASR time signal includes no obvious trend changes. The alternative FHT distribution, sampled from fitted ARMA model, can be seen in Figure 4.14 as well. It was smoothed using kernel density estimation (KDE) [99] for easier comparison.

To further compare the quality of the derived algorithm to an ARMA model, Figure 4.15 shows the evolution of variance and mean of both predictive distributions over the amount of data taken into account. The perceived EOL time is shown with a horizontal red line.

Again, the presented algorithm shows results comparable to those of the alternative. The larger variance at the start can be attributed to insufficient data. Such uncertainty in the prediction can be valuable for the operator, as it suggests a need for caution when interpreting the prognosis. Over time, we can observe a slow regression of the mean to the

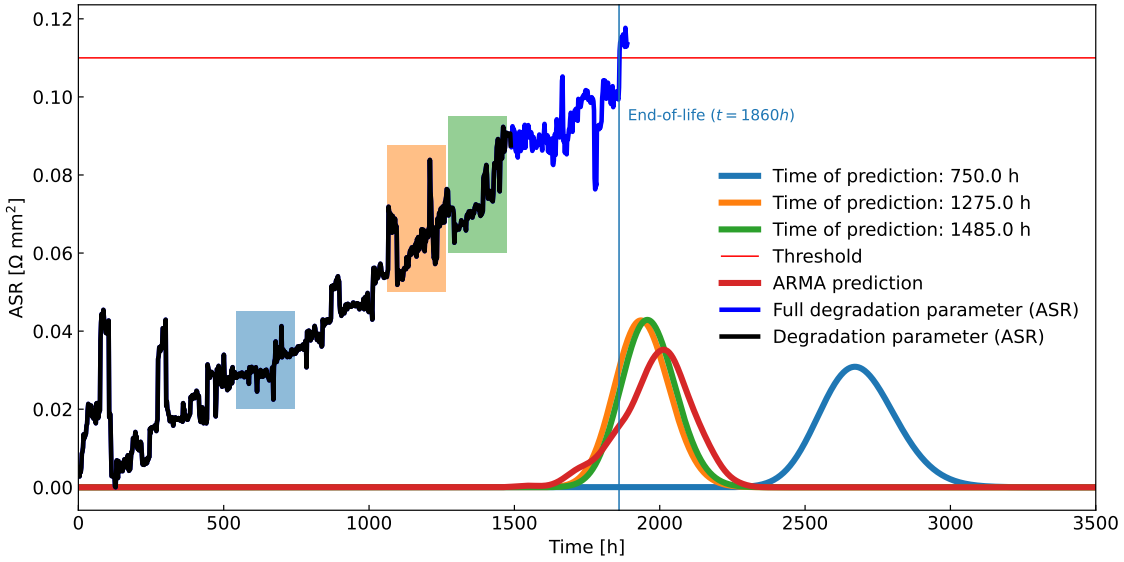


Figure 4.14: A demonstration of the application of the new algorithm on real-life ASR measurements. A major chunk of the data is used for learning (black curve), while the minor chunk is used for validation (dark blue curve). Few data windows, shown as rectangular coloured windows, were used on the learning part of the data to derive FHT predictions. The figure includes a comparison to an alternative prognostic algorithm, an ARMA process, fitted on the whole learning part of the data, which obtains similar results.

ground truth FHT, together with the variance shrinking to a useful level.

On the other hand, the alternative prediction from an ARMA model suffers from overconfidence issues in the beginning since only a point-wise fit is considered for the parameters. Besides that, it also slowly converges to the vicinity of the actual FHT.

Similarly to Figure 4.15, Figure 4.16 shows quantiles of FHT distributions derived throughout the experiment. Abrupt changes in the mean of the FHT distributions can be observed, which indicate volatile oscillations of ASR during the measurement process. In the magnified inset of the Figure, we again compare the actual FHT to the prediction. Upon closer examination, we conclude that the predicted FHT distributions successfully encompass the eventual EOL as early as $t = 800$h.

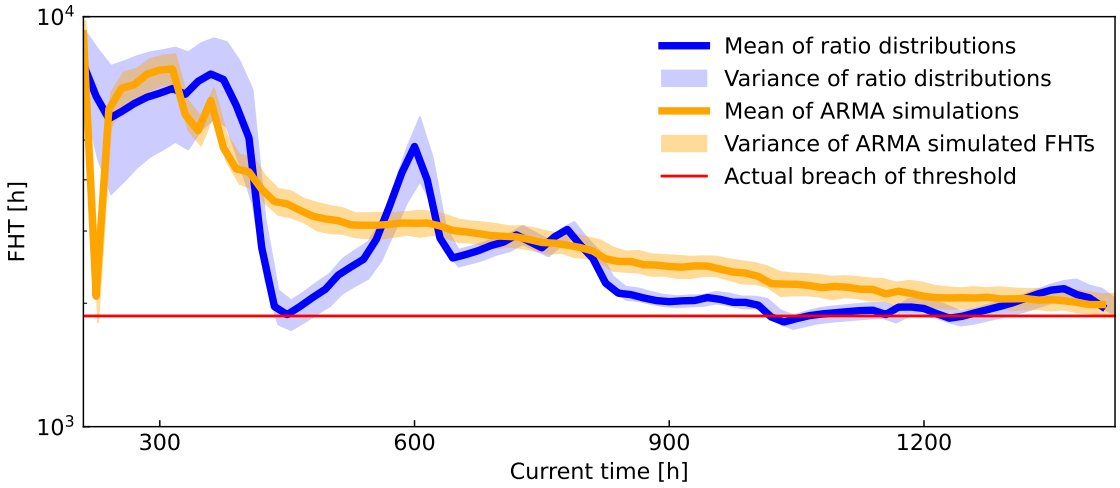


Figure 4.15: Comparison of ARMA and the derived prognostic algorithm on the experimental data. The x-axis represents the percentage of the full set of data used, simulating the real-life scenario of data accumulating over time. The horizontal red line denotes EOL for the stack. While both models converge to a similar solution that aligns with the actual ASR's breach of threshold, when at least 2/3 of data is available, they both experience significant deviations from the correct result when the data is insufficient. The significant difference between the methods is that the new prognostic algorithm undergoes variance explosion when data is scarce, while ARMA stays overconfident.

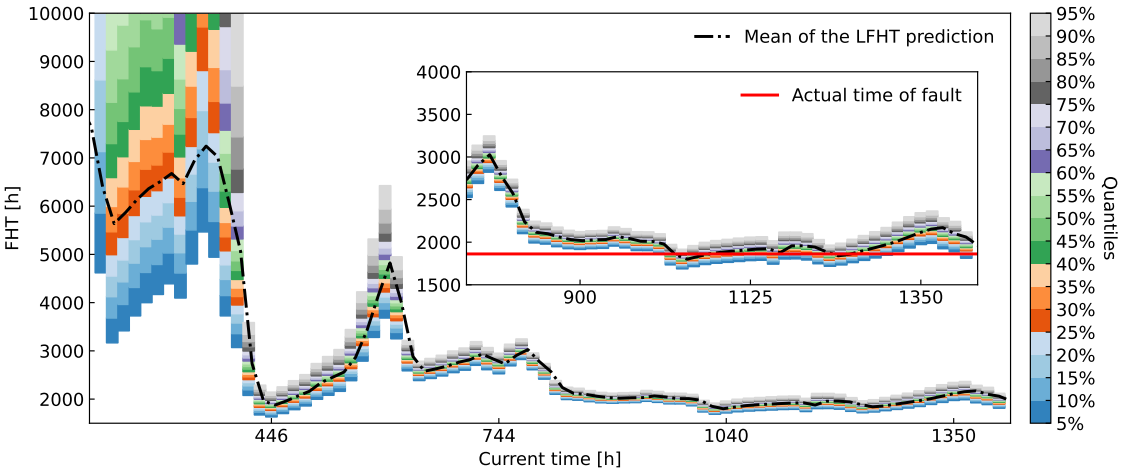


Figure 4.16: Quantile plot over time for ASR. The figure depicts the evolution of the FHT distribution over time, starting at 15% of the data and progressing to the full dataset. It simulates a real-life application where data comes in over time and the prognostic capabilities improve. The inset highlights the latter part of the measurement process, where the predicted FHT distribution can be trusted to assess the probable EOL.

Chapter 5

Conclusions

Let us summarise the accomplishments in the context of the main hypotheses:

- *Hypothesis 1: The VB approach significantly improves the efficiency of EIS deconvolution compared to traditional MCMC, thereby making the approach usable for online applications.*

  The primary conclusion is that the VB approach enhances the estimation of variability in EIS data, particularly when compared to point-wise estimation methods. This improvement is crucial in evaluating the results of ECM parameter estimates. While the VB method achieves nearly the same accuracy as MCMC methods, it exhibits a slight overconfidence bias, which is well-documented but has minimal impact in this case. Moreover, the reduced complexity of the VB algorithm makes it highly suitable for online applications.

- *Hypothesis 2: Probabilistic diagnostic inference, which treats all variables as random variables, provides a consistent framework allowing for systematic calibration of the diagnostic system, in particular, easy tuning of the detection thresholds.*

  Probabilistic fault detection represents a logical advancement when model parameters are treated as distributions. The fault detection methods presented in this thesis incorporate complete information about the estimated ECM parameters, leading to highly accurate results. This research demonstrates that probabilistic inference significantly enhances fault detection frameworks by allowing automatic threshold derivation from sufficient nominal operation data, streamlining system calibration.

- *Hypothesis 3: The proposed statistical approach for estimating the remaining useful life of SOC systems offers significant benefits over existing methods.*

  This hypothesis focuses on the prognosis of EECD systems. Though straightforward, the proposed method proves to be highly effective. The results are comparable to those of current state-of-the-art methods in system prognosis and may even outperform some. Its simplicity also makes it ideal for online implementation. A key advantage of this prognosis framework is its probabilistic output for FHT, estimated as a distribution, which provides a much more reliable assessment of the system's current state, offering valuable additional information for the system operator.

In conclusion, this thesis demonstrates that the VB approach for EIS deconvolution, probabilistic diagnostic inference for fault detection, and the proposed statistical method for system prognosis represent a novel contribution to the advancement of SOC systems. These methodologies not only enhance the accuracy and reliability of the results but also

ensure that the solutions are practical for online applications, addressing some of the critical challenges in the field.

Further research should explore the performance of the proposed diagnostic and prognostic systems across additional datasets, incorporating a wider variety of faults. Addressing the challenge of recalibrating the diagnostic system after maintenance while minimising system intrusion will also require careful investigation. Finally, distinguishing between system degradation and faults remains an unresolved issue which future work should aim to address.

Appendix A

Gaussian Process Model

A random vector $\boldsymbol{y} = (y_1, \ldots, y_d)^\top$ is said to follow a multivariate Gaussian distribution if for all $\boldsymbol{a} \in \mathbb{R}^d$, the linear combination $\boldsymbol{a}^\top \boldsymbol{y}$ have a univariate Gaussian distribution [100]. The corresponding probability density function (pdf) models the joint stochastic relationship between the variables. As in the univariate case, the pdfs are defined through the first two moments, namely the mean $\boldsymbol{\mu} := \mathbb{E}(\boldsymbol{y})$ and the covariance $\boldsymbol{\Sigma} := \mathrm{cov}(\boldsymbol{y})$. The equation for the joint Gaussian pdf is:

$$p(\boldsymbol{y}) = \mathcal{N}(\boldsymbol{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{\mu})\right),$$

where $|\boldsymbol{\Sigma}|$ denotes the determinant of $\boldsymbol{\Sigma}$.

A GP is a collection of random variables, any finite number with a joint Gaussian distribution. A GP is fully specified by its mean function $m(x)$ and covariance function $k(x, x')$. Since this is not a primary focus of this thesis, the definition and description of the GP will be kept quite plain. More detailed information can be found in [83].

Definition Given a set of inputs $X = \{x_1, x_2, \ldots, x_n\}$, a GP is defined as:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

where

$$m(x) = \mathbb{E}[f(x)]$$

and

$$k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))]$$

Mean and Covariance Functions - The mean function $m(x)$ typically assumed to be zero for simplicity:

$$m(x) = 0$$

- The covariance function $k(x, x')$, also known as the kernel, defines the structure of the function space:

$$k(x, x') = \mathrm{Cov}(f(x), f(x'))$$

Commonly used kernels include:

- **Squared Exponential (RBF) Kernel**:

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

- **Matérn Kernel**:

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

- **Periodic Kernel**:

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

Training the GP Model In a GP setting, training involves learning the hyperparameters of the covariance function, which control the properties of the functions that the GP can model. Common hyperparameters include the length scale l, signal variance σ_f^2, and noise variance σ_n^2.

The hyperparameters are typically learned by maximizing the marginal likelihood of the observed data, which is given by:

$$L = \log p(y \mid X, \theta) = -\frac{1}{2} y^\top (K(X, X) + \sigma_n^2 I)^{-1} y - \frac{1}{2} \log |K(X, X) + \sigma_n^2 I| - \frac{n}{2} \log 2\pi \quad \text{(A.1)}$$

where θ represents the set of hyperparameters. This optimization is usually performed using gradient-based methods, such as conjugate gradient or L-BFGS, due to the non-convex nature of the problem.

In addition to marginal likelihood maximization, one can also use alternative approaches such as Variational Inference or Expectation Propagation for hyperparameter learning in more complex GP models, where the assumptions of standard GP might not hold.

Inference Given a set of observations $y = \{y_1, y_2, \ldots, y_n\}$ at inputs X, and taking into account the variance matrix M, derived from the VB approach ECM parameters estimation, the joint distribution of the observed values and the values at test points X_* is:

$$\begin{pmatrix} y \\ f_* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m(X) \\ m(X_*) \end{pmatrix}, \begin{pmatrix} K(X, X) + MI & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{pmatrix} \right)$$

The predictive distribution for the test points X_* is:

$$f_* | X, y, X_* \sim \mathcal{N}(\bar{f}_*, \mathrm{Cov}(f_*))$$

where

$$\bar{f}_* = K(X_*, X)[K(X, X) + MI]^{-1} y$$

and

$$\mathrm{Cov}(f_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + MI]^{-1} K(X, X_*)$$

Appendix B

Wasserstein W_2 Distance

The W_2 distance is a metric used to measure the dissimilarity between two probability distributions [101]. It offers several key advantages in applications involving probability distributions.

The W_2 distance is robust to variations in both support and density of distributions. This robustness makes it suitable for comparing distributions with different shapes and sizes.

Geometrically, the W_2 distance can be interpreted as the minimum cost of transforming one distribution into another, where the squared Euclidean distance between corresponding points defines the cost. This geometric interpretation provides insights into the structural differences between distributions.

The W_2 distance is closely connected to optimal transport theory, which concerns finding the most efficient way to transport mass from one distribution to another. This connection enables the use of powerful optimization techniques for computing the distance.

In machine learning, the W_2 distance has applications in domain adaptation, generative modelling, and clustering tasks. Quantifying the dissimilarity between probability distributions facilitates the comparison and alignment of data distributions, leading to improved model performance.

Normal distributions:

Consider two non-degenerate Gaussian measures $\mu_1 = \mathcal{N}(m_1, C_1)$ and $\mu_2 = \mathcal{N}(m_2, C_2)$ on $\mathbb{R}^n$, with expected values m_1 and $m_2 \in \mathbb{R}^n$ and symmetric positive semi-definite covariance matrices C_1 and $C_2 \in \mathbb{R}^{n \times n}$. Then, with respect to the usual Euclidean norm on $\mathbb{R}^n$, the W_2 distance between μ_1 and μ_2 is given by:

$$W_2(\mu_1, \mu_2)^2 = \|m_1 - m_2\|_2^2 + \operatorname{trace}\left(C_1 + C_2 - 2\left(C_2^{1/2} C_1 C_2^{1/2}\right)^{1/2}\right).$$

Here, $C^{1/2}$ denotes the principal square root of C. Note that the second term involving the trace is precisely the unnormalized Bures metric between C_1 and C_2.

This result generalizes the earlier example of the Wasserstein distance between two point masses (at least in the case $p = 2$). A point mass can be regarded as a normal distribution with a covariance matrix equal to zero, in which case the trace term disappears, and only the term involving the Euclidean distance between the means remains.

Wasserstein metric The Wasserstein distance quantifies the dissimilarity between probability distributions by measuring the optimal transport cost required to transform one distribution into another. It is a fundamental metric in probability theory, statistics, machine learning, and image processing and is widely used for comparing and analyzing distributions.

The choice of the distance function $d(x, y)$ and the order p of the Wasserstein distance influence its sensitivity to different characteristics of the distributions, such as their shapes and spatial arrangements. For instance, when $p = 1$, the Wasserstein distance corresponds to the total variation distance, capturing the total variation between distributions.

An advantage of the Wasserstein distance is its ability to handle distributions with different supports or shapes and distributions of arbitrary dimensions. Even in such cases, it provides a meaningful measure of dissimilarity, making it a versatile tool in various fields.

Definition: Given two probability distributions μ and ν defined on a metric space (X, d), the Wasserstein distance $W_p(\mu, \nu)$ of order p is defined as the p-th root of the minimum cost of transporting mass from μ to ν. A suitable ground cost function and the transport plan determine this cost. Mathematically, it is expressed as:

$$W_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times X} d(x, y)^p \, d\gamma(x, y) \right)^{\frac{1}{p}}$$

where $\Gamma(\mu, \nu)$ is the set of all joint distributions with marginals μ and ν, and $d(x, y)$ represents the distance between points x and y in the metric space.

Appendix C

Support Vector Machine

Support Vector Machines (SVM) are powerful supervised learning algorithms primarily used for classification and regression tasks [102]. The main objective of SVM is to find the optimal hyperplane that best separates the data points of different classes in the feature space. This hyperplane is chosen to maximize the margin, defined as the distance between the hyperplane and the closest data points from each class, which are known as support vectors.

Given a training dataset $\{(x_i, y_i)\}_{i=1}^n$, where $x_i \in \mathbb{R}^d$ represents the feature vector and $y_i \in \{-1, +1\}$ denotes the class label, the goal is to find a hyperplane defined by $w \cdot x + b = 0$, where $w \in \mathbb{R}^d$ is the weight vector and $b \in \mathbb{R}$ is the bias. The decision function can be expressed as:

$$f(x) = \text{sign}(w \cdot x + b)$$

The optimal hyperplane is found by solving the following optimization problem, which maximizes the margin $\frac{2}{\|w\|}$:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to } y_i(w \cdot x_i + b) \geq 1, \forall i$$

This is a convex quadratic optimization problem and can be solved using Lagrange multipliers. The solution involves finding the weight vector w that minimizes the objective function while satisfying the constraints.

For cases where the data is not linearly separable, SVM introduces slack variables $\xi_i \geq 0$ to allow for some misclassifications. The optimization problem is then modified as:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i \quad \text{subject to } y_i(w \cdot x_i + b) \geq 1 - \xi_i, \forall i$$

Here, $C > 0$ is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error. A higher value of C leads to a smaller margin if it results in fewer misclassifications.

In many real-world scenarios, data is not linearly separable in the original feature space. SVM handles this by using a kernel function $K(x_i, x_j)$ that maps the data into a higher-dimensional space where a linear separation is possible. Commonly used kernel functions include:

- The linear kernel: $K(x_i, x_j) = x_i \cdot x_j$
- The polynomial kernel: $K(x_i, x_j) = (\gamma x_i \cdot x_j + r)^d$

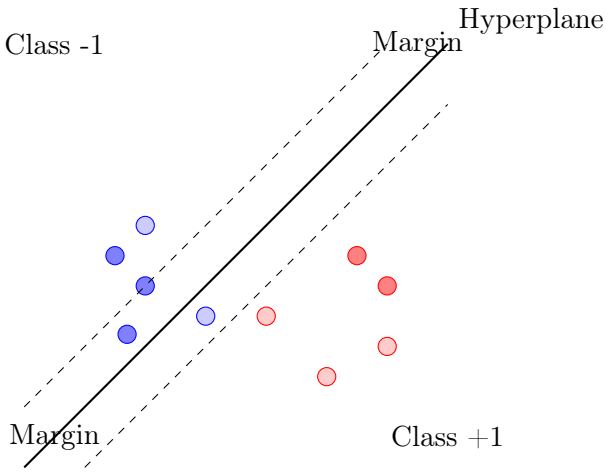


Figure C.1: Illustration of Support Vector Machine: Optimal hyperplane separating two classes with maximum margin. The circled points are the support vectors.

- The radial basis function (RBF) kernel: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$
- The sigmoid kernel: $K(x_i, x_j) = \tanh(\gamma x_i \cdot x_j + r)$

The decision function in the kernelized SVM is given by:

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right)$$

where α_i are the Lagrange multipliers obtained from the dual form of the optimization problem.

SVMs are particularly effective in high-dimensional spaces and are well-suited for problems where the number of dimensions exceeds the number of samples. Additionally, SVMs are versatile, allowing for the use of different kernel functions to adapt to various types of data distributions. This versatility, combined with the ability to manage non-linear decision boundaries and control overfitting through the parameter C, makes SVM a robust and widely-used machine learning algorithm.

Appendix D

Confusion Matrix for SOC Diagnosis

A confusion matrix in the context of SOFC diagnosis is a tabular representation used to evaluate the performance of diagnostic models by comparing predicted diagnoses with actual diagnoses for a set of SOFC health or fault conditions [103]. It comprehensively summarizes the model's diagnostic accuracy, highlighting its ability to correctly identify different types of SOFC health states or fault conditions.

The confusion matrix typically consists of four categories:

- **True Positive (TP):** SOFC fault conditions correctly diagnosed by the diagnostic model.
- **True Negative (TN):** Non-fault conditions correctly diagnosed by the diagnostic model.
- **False Positive (FP):** Non-fault conditions incorrectly diagnosed as fault conditions by the diagnostic model.
- **False Negative (FN):** SOFC fault conditions incorrectly diagnosed as non-fault conditions by the diagnostic model.

Using these categories, the confusion matrix enables the calculation of various performance metrics such as accuracy, precision, recall (sensitivity), specificity, and F1 score, which provide insights into the diagnostic model's effectiveness in correctly identifying and diagnosing SOFC fault conditions.

For instance, in the confusion matrix shown in Figure D.1:

- The value **74** represents **True Negatives (TN)**, indicating that 74 non-fault conditions were correctly diagnosed.
- The value **59** represents **True Positives (TP)**, indicating that 59 fault conditions were correctly diagnosed.
- The value **2** represents **False Positives (FP)**, indicating that 2 non-fault conditions were incorrectly diagnosed as faults.
- The value **2** represents **False Negatives (FN)**, indicating that 2 fault conditions were incorrectly classified as non-fault conditions.

Using these values, the performance metrics can be calculated as follows:

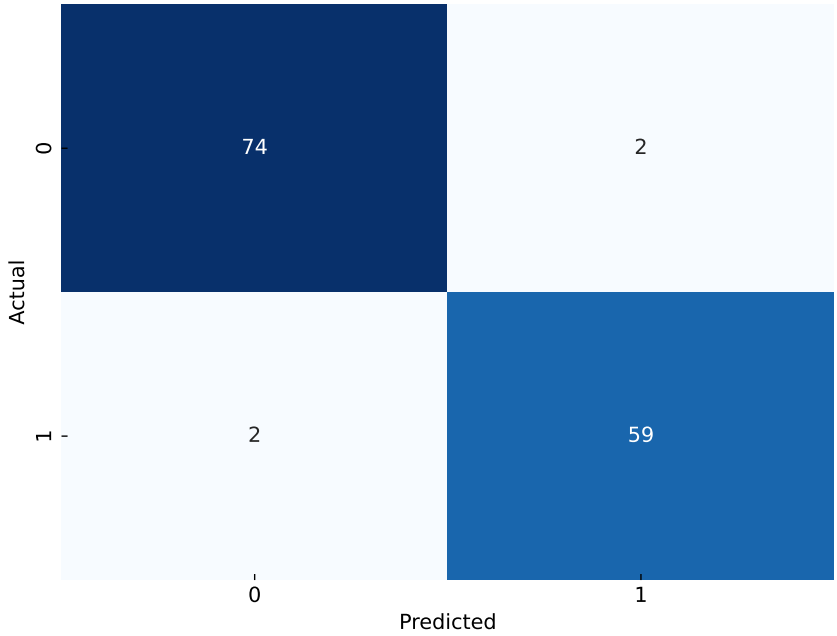


Figure D.1: Resulting confusion matrix for fault detection done with SVC method on the test dataset.

- **Accuracy:** Measures the overall correctness of the diagnostic model:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} = \frac{59 + 74}{59 + 74 + 2 + 2} \approx 0.970.$$

- **Precision:** Indicates the proportion of correctly diagnosed fault conditions out of all predicted faults:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{59}{59 + 2} \approx 0.967.$$

- **Recall (Sensitivity):** Reflects the model's ability to detect actual fault conditions:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{59}{59 + 2} \approx 0.967.$$

- **Specificity:** Measures how well the model identifies non-fault conditions:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} = \frac{74}{74 + 2} \approx 0.974.$$

- **F1 Score:** A harmonic mean of precision and recall, providing a balanced measure of the model's accuracy in diagnosing fault conditions:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0.967 \times 0.967}{0.967 + 0.967} \approx 0.967.$$

The confusion matrix is a valuable tool for evaluating the performance of diagnostic models in SOFC systems, enabling stakeholders to assess model reliability, identify areas for improvement, and make informed decisions regarding maintenance, repair, or replacement actions to ensure the reliability and safety of SOFC systems.

References

[1] O. Schmidt, A. Gambhir, I. Staffell, A. Hawkes, J. Nelson, and S. Few, "Future cost and performance of water electrolysis: An expert elicitation study," *International Journal of Hydrogen Energy*, vol. 42, no. 52, pp. 30 470–30 492, 2017, ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2017.10.045. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360319917339435.

[2] A. Pandiyan, A. Uthayakumar, R. Subrayan, S. W. Cha, and S. B. Krishna Moorthy, "Review of solid oxide electrolysis cells: A clean energy strategy for hydrogen generation," *Nanomaterials and Energy*, vol. 8, no. 1, pp. 2–22, 2019.

[3] Y. Luo, S. Yixiang, W. Li, and N. Cai, "Dynamic electro-thermal modeling of co-electrolysis of steam and carbon dioxide in a tubular solid oxide electrolysis cell," *Energy*, vol. 89, Jul. 2015. DOI: 10.1016/j.energy.2015.05.150.

[4] M. Laguna-Bercero, "Recent advances in high temperature electrolysis using solid oxide fuel cells: A review," *Journal of Power sources*, vol. 203, pp. 4–16, 2012.

[5] J. Mougin, "Hydrogen production by high-temperature steam electrolysis," in Dec. 2015, pp. 225–253, ISBN: 9781782423614. DOI: 10.1016/B978-1-78242-361-4.00008-X.

[6] O. Posdziech, T. Geißler, K. Schwarze, and R. Blumentritt, "System development and demonstration of large-scale high-temperature electrolysis," *ECS Transactions*, vol. 91, no. 1, p. 2537, Jul. 2019. DOI: 10.1149/09101.2537ecst. [Online]. Available: https://dx.doi.org/10.1149/09101.2537ecst.

[7] S. Mahboob and R. M M, "Revisiting the mechanism of solid oxide fuel cells (sofc's); its difficulties and probable new mechanism," vol. 12, pp. 126–128, Apr. 2023.

[8] H. Ghezel-Ayagh and B. P. Borglum, "Review of progress in solid oxide fuel cells at fuelcell energy," *ECS Transactions*, vol. 78, no. 1, p. 77, May 2017. DOI: 10.1149/07801.0077ecst. [Online]. Available: https://dx.doi.org/10.1149/07801.0077ecst.

[9] S. Roy and M. Ethakota, "Solid oxide electrolysis cell (soec): Potential technology for low-cost green hydrogen," *H2 Tech*, vol. Q4, 2022, Accessed: 2024-08-30. [Online]. Available: https://h2-tech.com/articles/2022/q4-2022/special-focus-future-of-hydrogen-energy/solid-oxide-electrolysis-cell-soec-potential-technology-for-low-cost-green-h-sub-2-sub/.

[10] A. Hauch, R. Küngas, P. Blennow, *et al.*, "Recent advances in solid oxide cell technology for electrolysis," *Science*, vol. 370, no. 6513, eaba6118, 2020. DOI: 10.1126/science.aba6118. eprint: https://www.science.org/doi/pdf/10.1126/science.aba6118. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.aba6118.

[11] T. Skafte, J. Hjelm, P. Blennow, and C. Graves, "Quantitative review of degradation and lifetime of solid oxide cells and stacks," Jul. 2016.

[12] T. A. N. Heirung and A. Mesbah, "Input design for active fault diagnosis," *Annual Reviews in Control*, vol. 47, pp. 35–50, 2019, ISSN: 1367-5788. DOI: https://doi.org/10.1016/j.arcontrol.2019.03.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578819300070.

[13] M. Gallo, P. Polverino, J. Mougin, B. Morel, and C. Pianese, "Coupling electrochemical impedance spectroscopy and model-based aging estimation for solid oxide fuel cell stacks lifetime prediction," *Applied Energy*, vol. 279, p. 115 718, 2020, ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2020.115718. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261920312113.

[14] C. Wendel, P. Kazempoor, and R. Braun, "A thermodynamic approach for selecting operating conditions in the design of reversible solid oxide cell energy systems," *Journal of Power Sources*, vol. 301, pp. 93–104, Jan. 2016. DOI: 10.1016/j.jpowsour.2015.09.093.

[15] L. Barelli, E. Barluzzi, and G. Bidini, "Diagnosis methodology and technique for solid oxide fuel cells: A review," *International Journal of Hydrogen Energy*, vol. 38, no. 12, pp. 5060–5074, 2013, ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2013.02.024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S036031991300390X.

[16] Q.-A. Huang, S. R. Hui, B. Wang, and J. Zhang, "A review of ac impedance modeling and validation in sofc diagnosis," *Electrochimica Acta*, vol. 52, pp. 8144–8164, Nov. 2007. DOI: 10.1016/j.electacta.2007.05.071.

[17] Y. Wang, S. Li, and Y. Yang, "Data-driven fault diagnosis-tolerant control integrated technique for solid oxide fuel cell," *International Journal of Smart Grid and Clean Energy*, pp. 357–365, Jan. 2020. DOI: 10.12720/sgce.9.2.357-365.

[18] A. Nechache, M. Cassir, and A. Ringuedé, "Solid oxide electrolysis cell analysis by means of electrochemical impedance spectroscopy: A review," *Journal of Power Sources*, vol. 258, pp. 164–181, 2014, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2014.01.110. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775314001578.

[19] A. Debenjak, P. Boškoski, B. Musizza, J. Petrovcic, and Đ. Juričić, "Fast measurement of proton exchange membrane fuel cell impedance based on pseudo-random binary sequence perturbation signals and continuous wavelet transform," *Journal of Power Sources*, vol. 254, pp. 112–118, May 2014. DOI: 10.1016/j.jpowsour.2013.12.094.

[20] R. Du, X. Wang, H. Dai, X. Wei, and P. Ming, "Online impedance spectrum measurement of fuel cells based on morlet wavelet transform," *International Journal of Hydrogen Energy*, vol. 46, May 2021. DOI: 10.1016/j.ijhydene.2021.05.012.

[21] P. Boškoski, A. Debenjak, and B. M. Boshkoska, *Fast Electrochemical Impedance Spectroscopy: As a Statistical Condition Monitoring Tool*. Springer, 2017.

[22] G. Nusev, B. Morel, J. Mougin, Đ. Juričić, and P. Boškoski, "Condition monitoring of solid oxide fuel cells by fast electrochemical impedance spectroscopy: A case example of detecting deficiencies in fuel supply," *Journal of Power Sources*, vol. 489, p. 229 491, Mar. 2021, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2021.229491.

[23] B. Königshofer, M. Höber, P. Boškoski, *et al.*, "Performance investigation and optimization of an soec stack operated under industrially relevant conditions," English, in *17th International Symposium on Solid Oxide Fuel Cells, SOFC 2021*, ser. ECS Transactions, United Kingdom: IOP Publishing Ltd., 2021, pp. 519–528. DOI: 10.1149/10301.0519ecst.

[24] C. Xie and Q. shuhai, "Drawing impedance spectroscopy for fuel cell by eis," *Procedia Environmental Sciences*, vol. 11, pp. 589–596, Dec. 2011. DOI: 10.1016/j.proenv.2011.12.092.

[25] E. Adinolfi, M. Gallo, P. Polverino, D. Beretta, S. Araya, and C. Pianese, "Ecm-based algorithm for on-board pemfcs diagnosis," English, in *ELECTRIMACS 2019 - Selected Papers - Volume 2*, W. Zamboni and G. Petrone, Eds., ser. Lecture Notes in Electrical Engineering, Germany: Springer, 2020, pp. 103–115, ISBN: 9783030569693. DOI: 10.1007/978-3-030-56970-9_9.

[26] P. Raffaele, C. Pianese, P. Polverino, and M. Sorrentino, "Method and apparatus for monitoring and diagnosing electrochemical devices based on automatic electrochemical impedance identification," May 2016.

[27] A. Nechache, A. Mansuy, M. Petitjean, *et al.*, "Diagnosis of a cathode-supported solid oxide electrolysis cell by electrochemical impedance spectroscopy," English, *Electrochimica acta*, vol. 210, pp. 596–605, 2016, ISSN: 0013-4686. DOI: 10.1016/j.electacta.2016.05.014.

[28] A. Baldinelli, A. Staffolani, G. Bidini, L. Barelli, and F. Nobili, "An extensive model for renewable energy electrochemical storage with solid oxide cells based on a comprehensive analysis of impedance deconvolution," *Journal of Energy Storage*, vol. 33, p. 102 052, 2021, ISSN: 2352-152X. DOI: https://doi.org/10.1016/j.est.2020.102052. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X20318867.

[29] A. Leonide, "Sofc modelling and parameter identification by means of impedance spectroscopy," Ph.D. dissertation, 2010, 147 pp., ISBN: 978-3-86644-538-3. DOI: 10.5445/KSP/1000019173.

[30] A. Hauch, A. Ploner, S. Pylypko, G. Cubizolles, and J. Mougin, "Test and characterization of reversible solid oxide cells and stacks for innovative renewable energy storage," *Fuel Cells*, vol. 21, no. 5, pp. 467–476, 2021. DOI: https://doi.org/10.1002/fuce.202100046. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/fuce.202100046. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/fuce.202100046.

[31] M. Hörlein, "Degradation study on solid oxide steam electrolysis," Ph.D. dissertation, Jul. 2020.

[32] Y. Yan, "Degradation study of soc stacks with impedance spectroscopy," Published on the publication server of RWTH Aachen University; Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2018, PhD Thesis, Rheinisch-Westfälische Technische Hochschule Aachen, Aachen, 2018. DOI: 10.18154/RWTH-2018-230369. [Online]. Available: https://publications.rwth-aachen.de/record/749842.

[33] A. Leonide, "Sofc modelling and parameter identification by means of impedance spectroscopy," Ph.D. dissertation, Karlsruher Institut für Technologie, 2010, ISBN: 978-3-86644-538-3.

[34] A. Staffolani, A. Baldinelli, L. Barelli, G. Bidini, and F. Nobili, "Early-stage detection of solid oxide cells anode degradation by operando impedance analysis," *Processes*, vol. 9, p. 848, May 2021. DOI: 10.3390/pr9050848.

[35] V. Subotić, B. Königshofer, D. Juricic, *et al.*, "Detailed insight into processes of reversible solid oxide cells and stacks using drt analysis," *Energy Conversion and Management*, vol. 226, p. 113 509, Dec. 2020. DOI: 10.1016/j.enconman.2020.113509.

[36] V. Subotić, S. Pofahl, V. Lawlor, N. Menzler, T. Thaller, and C. Hochenauer, "Online monitoring tools for soh diagnostic and prognostic of remaining lifetime of reversible solid oxide cell (rsoc) systems," *Energy Procedia*, vol. 158, pp. 2329–2334, Feb. 2019. DOI: 10.1016/j.egypro.2019.01.271.

[37] R. Isermann, "Model-based fault-detection and diagnosis - status and applications," *Annual Reviews in Control*, vol. 29, pp. 71–85, Dec. 2005. DOI: 10.1016/j.arcontrol.2004.12.002.

[38] D. Marra, C. Pianese, P. Polverino, and M. Sorrentino, *Models for Solid Oxide Fuel Cell Systems*. Mar. 2016, ISBN: 978-1-4471-5657-4. DOI: 10.1007/978-1-4471-5658-1.

[39] P. Polverino, C. Pianese, M. Sorrentino, and D. Marra, "Model-based development of a fault signature matrix to improve solid oxide fuel cell systems on-site diagnosis," *Journal of Power Sources*, vol. 280, pp. 320–338, Apr. 2015. DOI: 10.1016/j.jpowsour.2015.01.037.

[40] P. Polverino, M. Sorrentino, and C. Pianese, "A model-based diagnostic technique to enhance faults isolability in solid oxide fuel cell systems," *Applied Energy*, vol. 204, pp. 1198–1214, 2017, ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2017.05.069. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261917305627.

[41] A. Sorce, A. Greco, L. Magistri, and P. Costamagna, "Fdi oriented modeling of an experimental sofc system, model validation and simulation of faulty states," *Applied Energy*, vol. 136, pp. 894–908, 2014, ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2014.03.074. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261914003146.

[42] D. Marra, M. Sorrentino, C. Pianese, and B. Iwanschitz, "A neural network estimator of solid oxide fuel cell performance for on-field diagnostics and prognostics applications," *J. Power Sources*, vol. 241, pp. 320–329, 2013.

[43] M. Sorrentino, D. Marra, C. Pianese, *et al.*, "On the use of neural networks and statistical tools for nonlinear modeling and on-field diagnosis of solid oxide fuel cell stacks," *Energy Procedia*, vol. 45, pp. 298–307, 2014, ATI 2013 - 68th Conference of the Italian Thermal Machines Engineering Association, ISSN: 1876-6102. DOI: https://doi.org/10.1016/j.egypro.2014.01.032. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1876610214000332.

[44] D. Marra, C. Pianese, P. Polverino, and M. Sorrentino, *Models hierarchy*, 2016. DOI: 10.1007/978-1-4471-5658-1_2.

[45] M. Basseville and I. Nikiforov, *Detection of Abrupt Change Theory and Application*. Apr. 1993, vol. 15, ISBN: 0-13-126780-9.

[46] P. Costamagna, A. De Giorgi, L. Magistri, G. Moser, L. Pellaco, and A. Trucco, "A classification approach for model-based fault diagnosis in power generation systems based on solid oxide fuel cells," *IEEE Transactions on Energy Conversion*, vol. 30, pp. 1–12, Nov. 2015. DOI: 10.1109/TEC.2015.2492938.

[47] K. Wang, M. Péra, D. Hissel, N. Yousfi Steiner, A. Pohjoranta, and S. Pofahl, "Sofc modelling based on discrete bayesian network for system diagnosis use," *IFAC Proceedings Volumes*, vol. 45, no. 21, pp. 675–680, 2012, 8th Power Plant and Power System Control Symposium, ISSN: 1474-6670. DOI: https://doi.org/10.3182/20120902-4-FR-2032.00118. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S147466701632050X.

[48] B. Yang, Z. Guo, J. Wang, *et al.*, "Solid oxide fuel cell systems fault diagnosis: Critical summarization, classification, and perspectives," *Journal of Energy Storage*, vol. 34, p. 102 153, 2021, ISSN: 2352-152X. DOI: https://doi.org/10.1016/j.est.2020.102153. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X20319794.

[49] M. Gasperin, D. Juricic, P. Boškoski, and J. Vizintin, "Model-based prognostics of gear health using stochastic dynamical models," *Mechanical Systems and Signal Processing*, vol. 25, pp. 537–548, Feb. 2011. DOI: 10.1016/j.ymssp.2010.07.003.

[50] D. Hissel and M. Pera, "Diagnostic and health management of fuel cell systems: Issues and solutions," *Annual Reviews in Control*, vol. 42, pp. 201–211, 2016, ISSN: 1367-5788. DOI: https://doi.org/10.1016/j.arcontrol.2016.09.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578816300918.

[51] H. Liu, J. Chen, Q. Ouyang, and H. Su, "A review on prognostics of proton exchange membrane fuel cells," Oct. 2016, pp. 1–6. DOI: 10.1109/VPPC.2016.7791594.

[52] M. Gallo, D. Marra, M. Sorrentino, C. Pianese, and S. Au, "A versatile computational tool for model-based design, control and diagnosis of a generic solid oxide fuel cell integrated stack module," *Energy Conversion and Management*, vol. 171, Jun. 2018. DOI: 10.1016/j.enconman.2018.06.062.

[53] X. Wu and Q. Ye, "Fault diagnosis and prognostic of solid oxide fuel cells," *Journal of Power Sources*, vol. 321, pp. 47–56, 2016, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2016.04.080. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775316304232.

[54] B. Dolenc, D. Vrečko, Đ. Juričić, A. Pohjoranta, and C. Pianese, "Online gas composition estimation in solid oxide fuel cell systems with anode off-gas recycle configuration," *Journal of Power Sources*, vol. 343, pp. 246–253, 2017, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2017.01.038. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775317300381.

[55] B. Dolenc, D. Juričić, and P. Boškoski, "Identification of the coupling functions between the process and the degradation dynamics by means of the variational bayesian inference: An application to the solid-oxide fuel cells," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 377, no. 2160, p. 20 190 086, 2019. DOI: 10.1098/rsta.2019.0086. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2019.0086. [Online]. Available: https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2019.0086.

[56] S.-J. Cheng, W.-K. Li, T.-J. Chang, and C.-H. Hsu, "Data-driven prognostics of the sofc system based on dynamic neural network models," *Energies*, vol. 14, no. 18, 2021, ISSN: 1996-1073. DOI: 10.3390/en14185841. [Online]. Available: https://www.mdpi.com/1996-1073/14/18/5841.

[57] M. Li, J. Wu, Z. Chen, *et al.*, "Data-driven voltage prognostic for solid oxide fuel cell system based on deep learning," *Energies*, vol. 15, no. 17, 2022, ISSN: 1996-1073. DOI: 10.3390/en15176294. [Online]. Available: https://www.mdpi.com/1996-1073/15/17/6294.

[58] B. Dolenc, P. Boškoski, M. Stepančič, A. Pohjoranta, and Đ. Juričić, "State of health estimation and remaining useful life prediction of solid oxide fuel cell stack," *Energy Conversion and Management*, vol. 148, pp. 993–1002, 2017, ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2017.06.041. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0196890417305861.

[59] A. Barai, G. Chouchelamane, Y. Guo, A. McGordon, and P. Jennings, "A study on the impact of lithium-ion cell relaxation on electrochemical impedance spectroscopy," English, *Journal of Power Sources*, vol. 280, pp. 74–80, Apr. 2015, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2015.01.097.

[60] F. M. Kindermann, A. Noel, S. V. Erhard, and A. Jossen, "Long-term equalization effects in li-ion batteries due to local state of charge inhomogeneities and their impact on impedance measurements," English, *Electrochimica Acta*, vol. 185, no. C, pp. 107–116, 2015. DOI: 10.1016/j.electacta.2015.10.108.

[61] U. Westerhoff, T. Kroker, K. Kurbach, and M. Kurrat, "Electrochemical impedance spectroscopy based estimation of the state of charge of lithium-ion batteries," *Journal of Energy Storage*, vol. 8, pp. 244–256, 2016, ISSN: 2352-152X. DOI: https://doi.org/10.1016/j.est.2016.09.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X16301359.

[62] P. Boškoski, A. Debenjak, and B. M. Boshkoska, *Fast electrochemical impedance spectroscopy : as a statistical condition monitoring tool* (SpringerBriefs in applied sciences and technology), eng. Cham, Switzerland: Springer, 2017, ISBN: 9783319533902.

[63] G. Nusev, *Time-efficient method for monitoring electrochemical energy conversion systems : doctoral dissertation = Časovno učinkovita metoda za spremljanje elektrokemijskih sistemov za pretvorbo energije : doktorska disertacija*. Ljubljana: [G. Nusev], 2021.

[64] G. Nusev, Đ. Juričić, M. Gaberšček, J. Moškon, and P. Boškoski, "Fast impedance measurement of li-ion battery using discrete random binary excitation and wavelet transform," *IEEE Access*, vol. 9, pp. 46 152–46 165, 2021. DOI: 10.1109/ACCESS.2021.3058368.

[65] S. Mallat, *A Wavelet Tour of Signal Processing*. Jan. 1999, ISBN: 0-12-466606-X.

[66] P. Boškoski, A. Debenjak, and B. Mileva Boshkoska, *Fast Electrochemical Impedance Spectroscopy*. Springer International Publishing, 2017, ISBN: 9783319533902. DOI: 10.1007/978-3-319-53390-2.

[67] M. Orazem and B. Tribollet, *Electrochemical Impedance Spectroscopy* (The ECS Series of Texts and Monographs). Wiley, 2011, ISBN: 9781118209943.

[68] A. Lasia, *A. Lasia, Electrochemical Impedance Spectroscopy and its Applications, book, Springer, 2014*. Jul. 2014, ISBN: 978-1-4614-8932-0. DOI: 10.1007/978-1-4614-8933-7.

[69] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American Statistical Association, Vol. 112 , Iss. 518, 2017*, 2016. DOI: 10.1080/01621459.2017.1285773. arXiv: 1601.00670v9 [stat.CO].

[70] P. Boškoski, L. Žnidarič, Ž. Gradišar, and V. Subotić, "Probabilistic deconvolution for electrochemical impedance through variational bayesian inference," *Journal of Power Sources*, vol. 622, p. 235 359, 2024, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2024.235359. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775324013119.

[71] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American Statistical Association*, vol. 112, no. 518, pp. 859–877, Apr. 2017, ISSN: 1537-274X. DOI: 10.1080/01621459.2017.1285773. [Online]. Available: http://dx.doi.org/10.1080/01621459.2017.1285773.

[72] B. Wang and D. M. Titterington, "Inadequacy of interval estimates corresponding to variational bayesian approximations," in *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics*, R. G. Cowell and Z. Ghahramani, Eds., ser. Proceedings of Machine Learning Research, Reissued by PMLR on 30 March 2021., vol. R5, PMLR, Jan. 2005, pp. 373–380. [Online]. Available: https://proceedings.mlr.press/r5/wang05a.html.

[73] H. Attias, "A variational baysian framework for graphical models," in *Advances in Neural Information Processing Systems*, S. Solla, T. Leen, and K. Müller, Eds., vol. 12, MIT Press, 1999. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/1999/file/74563ba21a90da13dacf2a73e3ddefa7-Paper.pdf.

[74] L. Žnidarič, G. Nusev, B. Morel, J. Mougin, Đ. Juričić, and P. Boškoski, "Evaluating uncertainties in electrochemical impedance spectra of solid oxide fuel cells," *Applied Energy*, vol. 298, p. 117 101, 2021.

[75] K. P. Murphy, *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023. [Online]. Available: http://probml.github.io/book2.

[76] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2014. arXiv: 1412.6980 [cs.LG].

[77] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of adam and beyond," 2019. arXiv: 1904.09237v1 [cs.LG].

[78] B. Dolenc, G. Nusev, P. Boškoski, B. Morel, J. Mougin, and Đ. Juričić, "Probabilistic deconvolution of solid oxide fuel cell impedance spectra," *Elecetrimacs 2019 Salerno*, pp. 53–54, 2019, Nasl. z nasl. zaslona. [Online]. Available: http://www.electrimacs2019.unisa.it/files/ELECTRIMACS2019_Booklet_20190514.pdf.

[79] Y. Wang and D. M. Blei, *Variational bayes under model misspecification*, 2019. arXiv: 1905.10859 [stat.ML].

[80] Y. Wang and D. M. Blei, "Frequentist consistency of variational bayes," *Journal of the American Statistical Association*, vol. 114, no. 527, pp. 1147–1161, Aug. 2018, ISSN: 1537-274X. DOI: 10.1080/01621459.2018.1473776. [Online]. Available: http://dx.doi.org/10.1080/01621459.2018.1473776.

[81] E. Ivanova, R. Malti, and X. Moreau, "Pseudo-state-space fractional system identification," in *Proceedings of International Conference on Fractional Differentiation and its Applications*, 2018.

[82] A. Neath and J. Cavanaugh, "The bayesian information criterion: Background, derivation, and applications," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, Mar. 2012. DOI: 10.1002/wics.199.

[83] J. Kocijan, *Modelling and Control of Dynamic Systems Using Gaussian Process Models*. Jan. 2016, ISBN: 978-3-319-21020-9. DOI: 10.1007/978-3-319-21021-6.

[84] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," 1999.

[85] P. Tanasini, M. Cannarozzo, P. Costamagna, *et al.*, "Experimental and theoretical investigation of degradation mechanisms by particle coarsening in sofc electrodes," *Fuel Cells*, vol. 9, no. 5, pp. 740–752, 2009. DOI: https://doi.org/10.1002/fuce.200800192. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/fuce.200800192. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/fuce.200800192.

[86] E. D. R. Silva, M. Hubert, B. Morel, H. Moussaoui, J. Debayle, and J. Laurencin, "A dynamic multi-scale model for solid oxide cells validated on local current measurements: Impact of global cell operation on the electrodes reaction mechanisms," *ECS Transactions*, vol. 103, no. 1, p. 893, Jul. 2021. DOI: 10.1149/10301.0893ecst. [Online]. Available: https://dx.doi.org/10.1149/10301.0893ecst.

[87] E. D. R. Silva, G. Sassone, M. Prioux, M. Hubert, B. Morel, and J. Laurencin, "A multiscale model validated on local current measurements for understanding the solid oxide cells performances," *Journal of Power Sources*, vol. 556, p. 232 499, 2023, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2022.232499. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775322014768.

[88] L. Cui, H. Huo, G. Xie, J. Xu, X. Kuang, and Z. Dong, "Long-term degradation trend prediction and remaining useful life estimation for solid oxide fuel cells," *Sustainability*, vol. 14, no. 15, 2022, ISSN: 2071-1050. DOI: 10.3390/su14159069. [Online]. Available: https://www.mdpi.com/2071-1050/14/15/9069.

[89] Y. Xu, C. Jiang, J. Peng, X.-L. Wu, L. Xiao, and X. Li, "Fault prognosis method for solid oxide fuel cells based on mechanism degradation process model and particle filtering," *IEEE Transactions on Power Electronics*, vol. 38, no. 6, pp. 6831–6840, 2023. DOI: 10.1109/TPEL.2023.3246068.

[90] P. Polverino, M. Gallo, and C. Pianese, "Development of mathematical transfer functions correlating solid oxide fuel cell degradation to operating conditions for accelerated stress test protocols design," *Journal of Power Sources*, vol. 491, p. 229 521, 2021, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2021.229521. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775321000707.

[91] M. Guida, F. Postiglione, and G. Pulcini, "A random-effects model for long-term degradation analysis of solid oxide fuel cells," *Reliability Engineering & System Safety*, vol. 140, pp. 88–98, 2015, ISSN: 0951-8320. DOI: https://doi.org/10.1016/j.ress.2015.03.036. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0951832015001027.

[92] M. Naeini, J. S. Cotton, and T. A. Adams, "An eco-technoeconomic analysis of hydrogen production using solid oxide electrolysis cells that accounts for long-term degradation," *Frontiers in Energy Research*, vol. 10, 2022, ISSN: 2296-598X. DOI: 10.3389/fenrg.2022.1015465. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fenrg.2022.1015465.

[93] L. Žnidarič, Ž. Gradišar, and Đ. Juričić, "Predicting the remaining useful life of solid oxide fuel cell systems using adaptive trend models of health indicators," *Energies*, vol. 17, no. 11, p. 2729, Jun. 2024, ISSN: 1996-1073. DOI: 10.3390/en17112729.

[94] A. Goldberger, W. Shenhart, and S. Wilks, *Econometric Theory* (WILEY SERIES in PROBABILITY and STATISTICS: APPLIED PROBABILITY and STATIST ICS SECTION Series). J. Wiley, 1964, ISBN: 9780471311010. [Online]. Available: https://books.google.si/books?id=KZq5AAAAIAAJ.

[95] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists* (International series in engineering and computer science). Springer US, 2007, ISBN: 9780387476940. [Online]. Available: https://books.google.si/books?id=zjJdP0CJUAYC.

[96] R. H. Shumway and D. S. Stoffer, "Arima models," in *Time Series Analysis and Its Applications: With R Examples*. Cham: Springer International Publishing, 2017, pp. 75–163, ISBN: 978-3-319-52452-8. DOI: 10.1007/978-3-319-52452-8_3. [Online]. Available: https://doi.org/10.1007/978-3-319-52452-8_3.

[97] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with python," in *9th Python in Science Conference*, 2010.

[98] B. Dolenc, P. Boškoski, A. Pohjoranta, M. Noponen, and Đ. Juričić, "Hybrid approach to remaining useful life prediction of solid oxide fuel cell stack," *ECS Transactions*, vol. 78, no. 1, p. 2251, May 2017. DOI: 10.1149/07801.2251ecst. [Online]. Available: https://dx.doi.org/10.1149/07801.2251ecst.

[99] W. K. Härdle, M. Müller, S. Sperlich, and A. Werwatz, *Nonparametric and Semiparametric Models*. Springer Berlin, Heidelberg, Jan. 2006, pp. 68–69, ISBN: 978-3-642-62076-8. DOI: 10.1007/978-3-642-17146-8.

[100] A. Gut, *An Intermediate Course in Probability*. New York: Springer, 1995. DOI: 10.1007/978-1-4757-2431-8. [Online]. Available: https://link.springer.com/book/10.1007/978-1-4757-2431-8.

[101] C. Villani, *Optimal Transport: Old and New* (Grundlehren der mathematischen Wissenschaften). Berlin, Heidelberg: Springer-Verlag, 2009, vol. 338, ISBN: 978-3-540-71049-3. DOI: 10.1007/978-3-540-71050-9.

[102] C. Bishop, *Pattern Recognition and Machine Learning* (Information Science and Statistics). Springer New York, 2016, ISBN: 9781493938438. [Online]. Available: https://books.google.si/books?id=kOXDtAEACAAJ.

[103] R. Kohavi and F. Provost, "Glossary of terms," *Machine Learning*, vol. 30, no. 2-3, pp. 271–274, 1998. DOI: 10.1023/A:1017181826899.

Bibliography

Publications Related to the Thesis

L. Žnidarič, G. Nusev, B. Morel, J. Mougin, Đ. Juričić, and P. Boškoski, "Evaluating uncertainties in electrochemical impedance spectra of solid oxide fuel cells," *Applied Energy*, vol. 298, p. 117 101, 2021.

L. Žnidarič, Ž. Gradišar, and Đ. Juričić, "Predicting the remaining useful life of solid oxide fuel cell systems using adaptive trend models of health indicators," *Energies*, vol. 17, no. 11, p. 2729, Jun. 2024, ISSN: 1996-1073. DOI: 10.3390/en17112729.

P. Boškoski, L. Žnidarič, Ž. Gradišar, and V. Subotić, "Probabilistic deconvolution for electrochemical impedance through variational bayesian inference," *Journal of Power Sources*, vol. 622, p. 235 359, 2024, ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2024.235359. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378775324013119.

Biography

Luka Žnidarič was born in Ptuj, Slovenia. In 2011, he graduated from II. Gimnazija Maribor and went on to study mathematics at the Faculty of Natural Sciences and Mathematics in Maribor. He completed his Bachelor's degree in 2016 and continued with a Master's program in theoretical mathematics, finishing in 2018.

That same year, he joined the Young Researchers Programme under the supervision of Prof. Dr. Đani Juričić at the Jožef Stefan Institute. His research focused on Systems and Control, contributing to projects involving the analysis of Solid Oxide Fuel Cells (SOFC) and Solid Oxide Electrolysis Cells (SOEC), emphasising probabilistic approaches to various problems.