



School II: School of Computing Science, Business Administration, Economics and Law
Department of Computing Science

*Data-Driven State-of-Health Quantification for
Industrial Proton Exchange Membrane (PEM) Water Electrolyzer Fleets*

The School of Computing Science, Business Administration, Economics and Law of
the University of Oldenburg has accepted the dissertation for the degree and title of

Doctor of Engineering (Dr.-Ing.)

by *Ms. Xuqian Yan*
born on December 8, 1994, in Shandong, China.

June 30, 2025

Supervisor:

Prof. Dr.-Ing. Astrid Nieße

Other examiners:

Prof. Dr.-Ing. Richard Hanke-Rauschenbach

Prof. Dr. Wolfram Wingerath

Date of disputation:

June 25, 2025

Abstract

Producing green hydrogen through water electrolysis with electricity from renewable sources offers a promising solution to decarbonize various industrial sectors. As a result, the global installed capacity of water electrolyzers is expected to grow rapidly in the near future. One of the mature water electrolysis technologies is the proton exchange membrane (PEM) water electrolysis. It has several advantages such as allowing for quick dynamic operation and having a smaller size. However, its market expansion is hindered by its high cost and short lifetime. To mitigate these disadvantages, PEM electrolyzer operators must maximize its utilization by carefully planning maintenance activities and optimizing the operation schemes. This requires close monitoring of the state of health (SOH) of the industrial PEM electrolyzers. Existing SOH quantification methods for PEM water electrolyzers are primarily based on direct measurements in laboratory environments under strictly controlled operating conditions, which are not suitable for industrial applications. To bridge this gap, this research focuses on the easily measurable degradation indicator – voltage, and uses data-driven methods to correct voltage measured under arbitrary operating conditions to the reference conditions, denoted as U_{rc} – the SOH indicator for PEM water electrolyzers.

Two methods are developed in this research. They share the same working principle: Voltage models are fitted over time with the help of fleet knowledge; the fitted models are used to calculate voltage under the reference condition and obtain U_{rc} . The first method is based on the transfer linear regression algorithm, which adjusts a pre-existing fitted voltage model based on new data in a sequential manner. The second method is based on Bayesian inference, using fleet knowledge as the prior probability distribution. These methods are validated with both industrial and synthetic data. In addition, this research suggests a framework to guide industrial practitioners in developing fleet-based algorithms based on their business requirements.

The proposed methods will enable automatic and continuous degradation monitoring for industrial PEM water electrolyzer. It provides a strong foundation to enhance the operation safety and profitability. It also facilitates knowledge discovery from the large amount of industrial data to gain more insights into the degradation behavior.

Kurzzusammenfassung

Die Erzeugung von grünem Wasserstoff durch Wasserelektrolyse mit Strom aus erneuerbaren Quellen bietet eine vielversprechende Lösung, um verschiedene Industriesektoren zu dekarbonisieren. Dementsprechend wird erwartet, dass die weltweit installierte Kapazität von Wasserelektrolyseuren in naher Zukunft schnell wachsen wird. Eine der ausgereiften Technologien für Wasserelektrolyse ist die Protonenaustauschmembran (PEM) Wasserelektrolyse. Sie hat mehrere Vorteile, wie zum Beispiel die Ermöglichung eines schnellen, dynamischen Betriebs und eine kleinere Größe. Allerdings wird ihre Marktausweitung durch hohe Kosten und eine kurze Lebensdauer gehemmt. Um diese Nachteile zu mildern, müssen Betreiber von PEM-Elektrolyseuren deren Nutzung maximieren, indem sie Wartungsaktivitäten sorgfältig planen und die Betriebsschemata optimieren. Dies erfordert eine genaue Überwachung des Gesundheitszustands (engl. State of Health, SOH) der industriellen PEM-Elektrolyseure.

Bestehende Methoden zur Quantifizierung des SOH für PEM-Wasserelektrolyseure basieren hauptsächlich auf direkten Messungen in Laborumgebungen unter streng kontrollierten Betriebsbedingungen, die für industrielle Anwendungen nicht geeignet sind. Um diese Lücke zu schließen, konzentriert sich diese Forschung auf den leicht messbaren Degradationsindikator – die Spannung – und verwendet datengetriebene Methoden, um die unter beliebigen Betriebsbedingungen gemessene Spannung auf die Referenzbedingungen zu korrigieren, die als U_{rc} – der SOH-Indikator für PEM-Wasserelektrolyseure – bezeichnet wird.

In dieser Forschung werden zwei Methoden entwickelt. Sie teilen das gleiche Arbeitsprinzip: Spannungsmodelle werden über die Zeit mit Hilfe von Flottenwissen angepasst; die angepassten Modelle werden verwendet, um die Spannung unter der Referenzbedingung zu berechnen und U_{rc} zu erhalten. Die erste Methode basiert auf dem Transfer-Linear-Regression-Algorithmus, der ein bereits vorhandenes angepasstes Spannungsmodell basierend auf neuen Daten sequenziell anpasst. Die zweite Methode basiert auf der Bayesschen Inferenz, wobei das Flottenwissen als die a-priori-Wahrscheinlichkeitsverteilung verwendet wird. Diese Methoden werden sowohl mit industriellen als auch mit synthetischen Daten validiert. Darüber hinaus schlägt diese Forschung einen Rahmen vor, um industrielle Praktiker bei der Entwicklung von flottenbasierten Algorithmen, basierend auf ihren Geschäftsanforderungen, zu leiten.

Die vorgeschlagenen Methoden ermöglichen eine automatische und kontinuierliche Degradationsüberwachung für industrielle PEM-Wasserelektrolyseure. Sie bieten eine starke Grundlage, um die Betriebssicherheit und Rentabilität zu erhöhen. Zudem erleichtern sie die Wissensentdeckung aus der großen Menge an industriellen Daten, um tiefere Einblicke in das Degradationsverhalten zu gewinnen.

Acknowledgments

First, I would like to thank the sponsor of this research, the Federal Ministry of Education and Research of Germany (Bundesministerium für Bildung und Forschung, BMBF), for their support within the scope of the SEGIWA (grant number: 03HY121A) and DERIEL (grant number: 03HY122A) projects.

Besides, this research would not have been possible without the unwavering support of my family, friends, and colleagues. First and foremost, to my fiancé, Jakob Fittler: Your 24/7 support – whether through countless cups of coffee, delicious meals, or simply your cheering smile during challenging times – meant the world to me. Also, I owe my deepest gratitude to my parents: Thank you for giving me the freedom to move abroad for my studies and career. Even from afar, I feel calm and secure knowing that you are always there, ready to support me in any way I need.

My heartfelt thanks go to my supervisor, Prof. Dr.-Ing. Astrid Nieße: I am deeply grateful for your commitment to our monthly meetings, despite a fully packed schedule. Your continued supervision, even after handing over the projects at an early stage, provided security that is so vital for a doctoral student.

To my colleagues at Siemens Energy across various departments: I am thankful for your technical inputs and commercial insights, which helped bring real-world impact to my research. A special thank you goes to Dr. Kunyuan Zhou, who made the journey from Berlin to attend my disputation.

To my project partners at OFFIS – Nils Huxoll, Sharaf Aldin Alsharif, and Jelke Wibbeke: It has been a pleasure navigating this journey together with you. Our collaborations have always been enriching, and every visit to Oldenburg felt like a reunion.

To my project partners at the Leibniz University of Hannover – Janis Woelke, Dr.-Ing. Christoph Eckert, and Dr.-Ing. Boris Bensmann: Thank you for your openness in sharing your expertise in PEM water electrolysis. I greatly appreciate our collaboration, whether during conferences, workshops, or our joint publication.

Finally, I would like to thank the members of my examination committee – Prof. Dr.-Ing. Astrid Nieße, Prof. Dr.-Ing. Richard Hanke-Rauschenbach, Prof. Dr. Wolfram Wingerath, Prof. Dr.-Ing. Andreas Rauh, and Dr. Michael Brand: Thank you for taking the time to examine my work. Many of you traveled to Oldenburg to attend the disputation in person – your presence made the day truly special.

Contents

Abbreviations	xi
Nomenclature	xiii
Indices	xv
1 Introduction	1
2 Foundation and state of the art	9
2.1 PEM water electrolyzer and its degradation	9
2.2 PEM water electrolyzer models	12
2.2.1 Variety of models	12
2.2.2 Voltage models	13
2.3 SOH quantification methods	17
2.3.1 SOH applications and types	17
2.3.2 In-situ measurements for PEM water electrolyzers	19
2.3.3 Statistical methods to mitigate the impact of operation modes	21
2.3.4 Challenge of limited data coverage	22
2.4 Fleet-based algorithms	22
2.4.1 Basics of fleet	23
2.4.2 Categories of fleet-based algorithms	23
2.5 Research gap	27
3 Method and data overview	29
3.1 Method overview	29
3.2 Method evaluation	31
3.2.1 Data for evaluation	31
3.2.2 Evaluation metrics	34
3.2.3 Benchmark method	36

4	Method 1: SOH quantification with transfer linear regression	37
4.1	Algorithm	37
4.1.1	Overview	37
4.1.2	Voltage model	37
4.1.3	Model fitting	40
4.2	Method evaluation with synthetic data	45
4.3	Method evaluation with industrial data	51
4.4	Limitations	55
4.4.1	Elevated error for degraded electrolyzers	55
4.4.2	Difficulty in determining hyperparameters	55
4.4.3	Rely on the last fitted model	56
5	Method 2: SOH quantification with Bayesian inference	59
5.1	Algorithm	59
5.1.1	Overview	59
5.1.2	Voltage model	59
5.1.3	Model fitting	63
5.2	Method evaluation with synthetic data	66
5.2.1	Evaluation with a synthetic fleet	66
5.2.2	Sensitivity analysis 1: impact of data coverage	72
5.2.3	Sensitivity analysis 2: impact of fleet-target discrepancy	73
5.2.4	Sensitivity analysis 3: impact of fleet size	74
5.3	Method evaluation with industrial data	76
5.4	Limitations	79
5.4.1	Similarity between the fleet and the target electrolyzer	79
5.4.2	Influence of operation modes	80
5.4.3	Nonlinear voltage models	80
6	Discussions on the methods	81
6.1	Impact of using different temperature data	81
6.2	Impact of model fitting interval length	81
6.3	Other impact factors on voltage	83
6.4	Transferability of the method to other health indicators	84
7	Practical guide for fleet-based algorithm development	85
7.1	Core steps of fleet-based algorithms	85
7.2	Fleet-based algorithm development process	87

7.3	Example of using the guide	90
8	Conclusion	91
8.1	Answer to the research question	91
8.2	Methods reflection	93
8.3	Outlook	94
A	Appendix	97
A.1	Appendix A.1	97
	List of Figures	99
	List of Tables	103
	Bibliography	105

Abbreviations

CCM	catalyst coated membrane
GDL	gas diffusion layer
GPR	Gaussian process regression
MEA	membrane electrode assembly
PEM	proton exchange membrane
PFSA	perfluorosulfonic acid
PTL	porous transport layer
PV	photovoltaic
RMSE	root mean squared error
SLR	simple linear regression
SOH	state of health
TLR	transfer linear regression

Nomenclature

a	Tafel slope [V dec^{-1}]
A	Cell area [cm^2]
α	Charge transfer coefficient [-]
b	Tafel constant [V]
β	A vector of unknown model coefficients
c	Coefficients to be fitted
C	Concentration [mol m^{-3}]
γ	Model coefficients shift
ϵ	Errors
η	Overpotential [V]
F	Faraday's constant, $96\,485\text{ C mol}^{-1}$
I	Electric current [A]
i	Electric current density [A cm^{-2}]
i_0	Exchange current density [A cm^{-2}]
λ	A hyperparameter used in transfer linear regression to control the parameter shift
m	Mean
n	Number of data [-]
$\mathcal{N}$	Normal distribution
OH	Operating hours since last start [h]

p	Probability distribution
P	Pressure [Pa]
R	Ideal gas constant, $8.314 \text{ J K}^{-1} \text{ mol}^{-1}$
R_{ohm}	Ohmic resistance [Ω]
S	Covariance
σ	Standard deviation
T	Temperature [$^{\circ}\text{C}$]
U	Voltage [V]
U_0	Reversible cell voltage [V]
X	Independent variables
X_T	A hyperparameter used in transfer linear regression to control the parameter shift
Y	Dependent variables
z	Number of electrons exchanged in the chemical reaction, $z = 2$ for water electrolysis reaction

Indices

0	Reference working condition used in Nernst equation
act	Activation overpotential
an	Anode
bub	Bubbles overpotential
ca	Cathode
dif	Diffusion overpotential
f	Fleet
future	Data in the future
H ₂	Hydrogen
O ₂	Oxygen
ohm	Ohmic overpotential
pol	Data from polarization curves
rc	Reference condition
trans	Mass transfer overpotential
true	Ground truth

1

Introduction

Hydrogen is a versatile resource with a wide range of uses across industry, transport, and power sectors [1]. Global hydrogen demand reached a new high of 97 Mt in 2023, with a 2.5% year-on-year increase [2]. As consuming hydrogen does not generate CO₂ emission, it is recognized as an essential component for decarbonization, especially for hard-to-abate sectors such as steel and chemical production [1, 3]. In the Net Zero Emission by 2050 scenario, the International Energy Agency anticipates that the global hydrogen demand will reach 400 Mt in 2050 (Fig. 1.1). The three largest demand sources are projected to be aviation and marine fuel (e.g., synthetic kerosene, ammonia, methanol), power generation (e.g., fuel cells, co-firing hydrogen in gas turbines), and chemical production (e.g., ammonia for fertilizers) [4, 5].

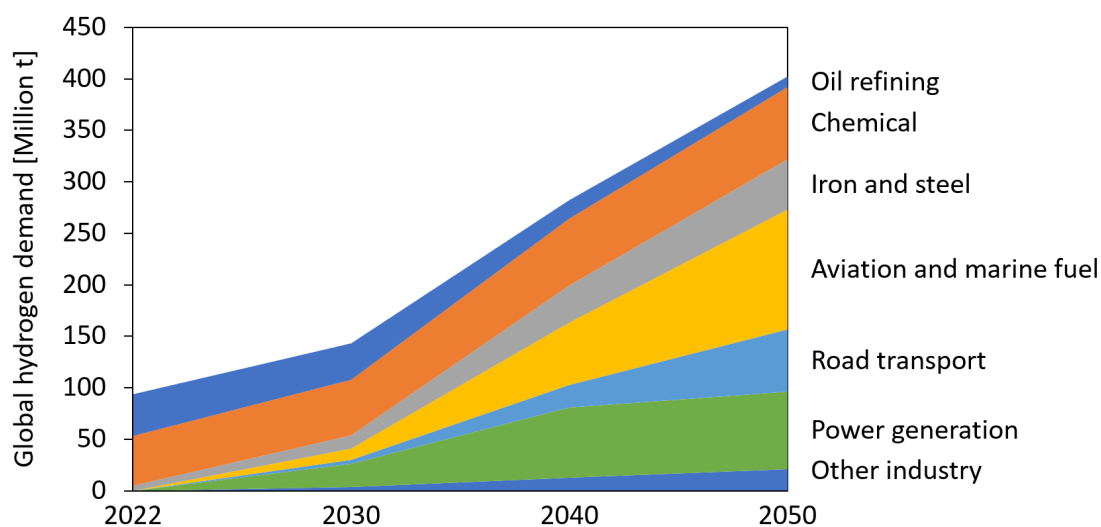


Figure 1.1: Global hydrogen demand by sector in the Net Zero Emission scenario [6].

Nevertheless, to achieve decarbonization over the lifecycle, hydrogen production must become cleaner. As of now, the production of hydrogen is almost entirely fossil fuel-based, dominated by natural gas reforming and coal gasification [7]. These carbon-intensive production processes make hydrogen production responsible for more than 900 Mt of CO₂

emission, equivalent to 2.7% of global total CO₂ emission (33572.1 Mt in 2021 [8]). As a comparison, the total CO₂ emission of Germany in 2021 is 624.1 Mt [8].

One low-emission hydrogen production technology is water electrolysis using electricity from renewable sources (such as solar and wind) or nuclear power [4]. In 2023, water electrolysis constitutes only 0.1% of global hydrogen production [4]. However, the installed capacity is expected to grow rapidly, from 2.2 GW in 2023 to over 400 GW by 2030 (Fig. 1.2), corresponding to 28 Mt hydrogen production per year [4].

Among the announced projects, alkaline and proton exchange membrane (PEM) water electrolysis are the most selected technologies (Fig. 1.2). Other technologies such as solid oxide electrolysis and anion exchange membrane electrolysis only account for a marginal share. The focus of this research lies on the PEM water electrolysis technology.

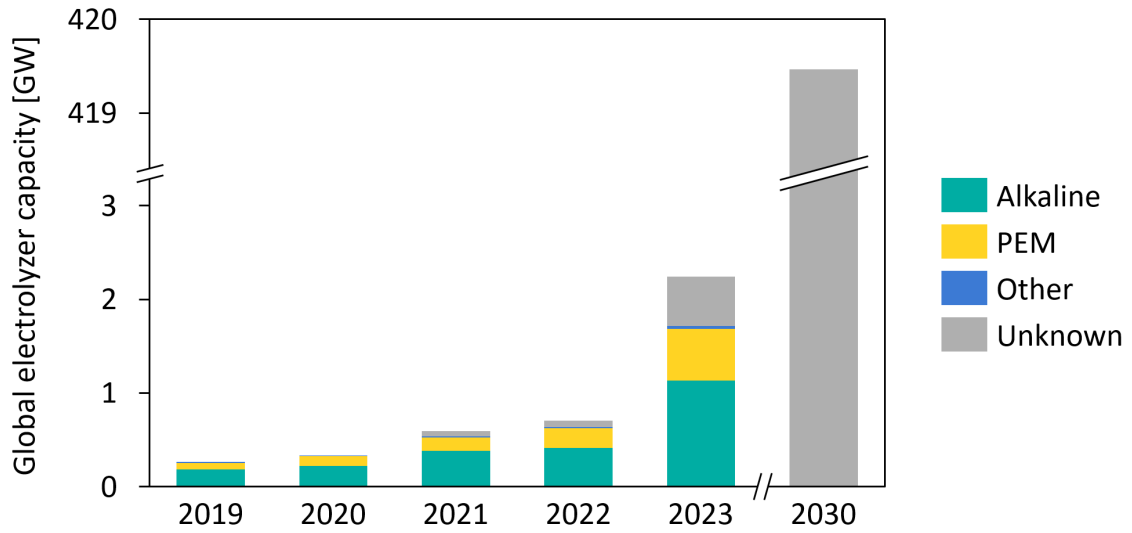


Figure 1.2: Global electrolyzer capacity by technology based on announced projects [4]. The expected capacity in 2030 includes early-stage projects (e.g., projects with only a stakeholder cooperation agreement).

In comparison to alkaline water electrolysis (Fig. 1.3), PEM electrolyzer has several advantages, including the capability to operate at higher current densities (resulting in higher hydrogen production rate for a given size), higher gas purity, less corrosion, smaller size, and quicker start. However, its main disadvantages are the high manufacturing cost (mainly due to the precious metal for catalysts) and short lifetime, hindering the market expansion. To mitigate these disadvantages, operators of PEM electrolyzers must maximize their utilization by carefully planning maintenance activities and optimizing operation schemes, which require close monitoring of the state of health (SOH) of the PEM electrolyzers.

Tracking the SOH of PEM electrolyzers not only provides financial benefits for the

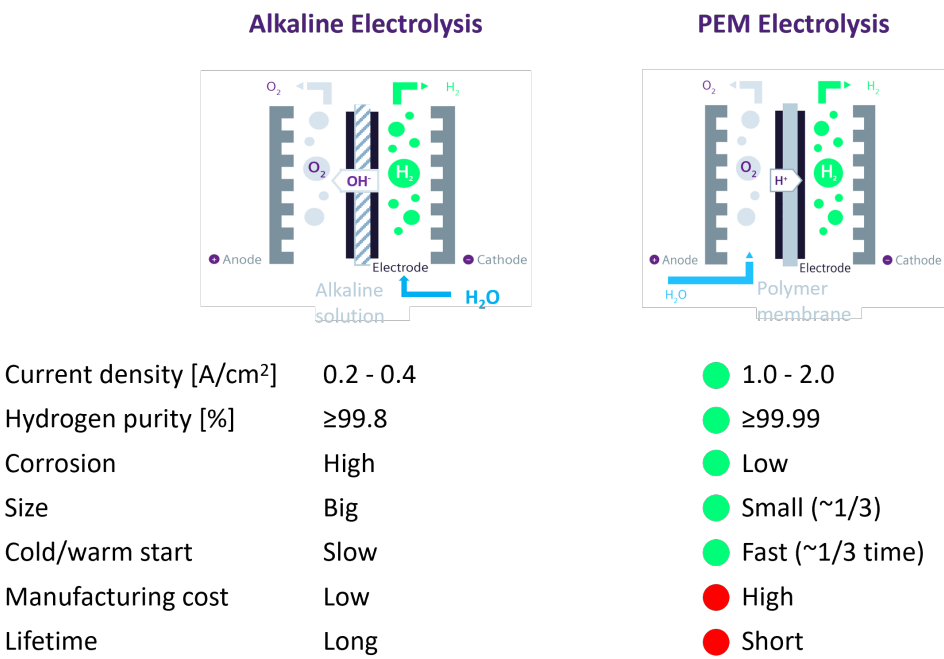


Figure 1.3: Comparison between alkaline and [PEM](#) electrolysis technologies [9–11].

operators, but also facilitates knowledge discovery to further understand the degradation behavior, supports risk assessment, and enables early detection of safety hazards.

The [SOH](#) of [PEM](#) electrolyzers can be observed from several aspects, such as increase in voltage, decrease in gas purity, and fluoride release into the effluent water. Among the various indicators, voltage degradation is the focus of this research. First of all, it is easily measurable in industrial environments without the need for costly equipment. Moreover, it has a significant economic relevance: An increase in voltage is directly associated with higher energy demand (i.e., lower efficiency) for hydrogen production. Specifically, the goal of this research is to track the temporal evolution of the overall voltage, without delving into the degradation mechanisms.

Existing studies on [PEM](#) water electrolyzer degradation focus on quantifying the [SOH](#) through direct measurements in laboratory environments, such as measuring voltage under predefined current levels. However, quantifying degradation for industrial electrolyzers is not as straightforward and faces various challenges:

Few testing opportunity. The operation modes of industrial electrolyzers are business-driven, making it impractical to frequently run predefined operational procedures to obtain [SOH](#) values.

Various influencing factors. Multiple operating variables (e.g., electric current and tem-

perature) influence degradation-related measurements such as voltage. Therefore, a simple 1-dimensional characteristic curve (e.g., a voltage-current curve) is insufficient to capture the impact of these variables.

Limited data coverage. The coverage of industrial operation data can be limited. For example, when an electrolyzer is operated under a constant electric current, the coverage of the current data is very narrow. Such limited data coverage poses challenges in modeling the impact of operating variables on the degradation-related measurements.

Heterogeneous fleet. As PEM water electrolyzer just recently gained industrial interest, it is experiencing rapid design updates. This makes data collected from one machine not representative for another. Based on such diverse datasets, it is hard to construct a data-driven model that is consistently applicable for all electrolyzers.

Lack of degradation model. Given the complex degradation mechanism and limited operation experience, comprehensive degradation models for industrial PEM water electrolyzers are not yet available. Therefore, filter-based methods (e.g., Kalman filter) to identify physical parameters are not applicable.

To address these challenges, this research seeks to answer the research question:

How to quantify the state-of-health of industrial PEM water electrolyzers?

Through investigating the research question, this research has made the following contributions.

Empirical voltage models. This research presents two empirical models that capture the impacts of operating parameters on the electrolyzer voltage. These models not only include the commonly assessed parameters – electric current and temperature, but also address the impact of restarting an electrolyzer after it has been shut down. The models feature simple and transparent linear structures, yet they have proven to be effective in describing the voltage behaviour of the industrial electrolyzers.

Transfer linear regression (TLR) algorithm. To tackle the model fitting difficulty under limited data coverage, the TLR algorithm is developed. This algorithm adjusts the parameters of an existing fitted model based on new data. The parameter adjustment is controlled using fleet experience (e.g., how much the parameters are expected to change over time). When applied sequentially, the TLR algorithm effectively captures the temporal drift in the voltage model caused by degradation.

Incorporating fleet knowledge with Bayesian inference. In addition to TLR, this research presents a second approach to assist model fitting under limited data coverage. This approach involves incorporating aggregated information from a fleet of electrolyzers (fleet knowledge) with Bayesian inference. The fleet knowledge is formed by aggregat-

ing model parameters from multiple electrolyzers, and used as the prior probability distribution in the Bayesian inference scheme.

Synthetic data generation. This research introduces a framework to generate synthetic data based on industrial data. The central component of the framework is a model that describes the relationships between the input variables (e.g., electric current) and the output variable (voltage). The model's coefficients, input variables, and noise levels are set based on the industrial data and can also be customized to simulate various scenarios. This framework ensures that the generated synthetic data exhibit a similar behaviour as the industrial data and provides flexibility of varying the setup to support sensitivity analysis.

Practical guide for industrial applications. This research provides a practical guide to assist industrial practitioners in developing data-driven models using fleet data. This guide consists of three core steps of fleet-based algorithms: data selection, model development, and model adjustment. It guides the practitioners to integrate these steps according to their use cases and accuracy requirements.

These contributions are presented in the following publications.

- Yan, X., Locci, C., Hiss, F., Nieße, A.: State-of-health estimation for industrial H₂ electrolyzers with transfer linear regression. *Energies* **17**(6), 1374 (2024). doi:[10.3390/en17061374](https://doi.org/10.3390/en17061374) [12]
- Yan, X., Helmers, L., Zhou, K., Nieße, A.: Fleet-based degradation state quantification for industrial water electrolyzers. *Electrochemical Science Advances* 5(3), 70002 (2025). doi:[10.1002/elsa.70002](https://doi.org/10.1002/elsa.70002) [13]
- Yan, X.⁼, Woelke, J.⁼, Bensmann, B., Eckert, C., Hanke-Rauschenbach, R., Nieße, A.: Cross-method overview of fleet-based machine health estimation and prediction: a practical guide for industrial applications. *IEEE Access* 13, 60131–60147 (2025). doi:[10.1109/access.2025.3556251](https://doi.org/10.1109/access.2025.3556251) (= equal contribution) [14]

It is also discussed with the scientific community through a poster presentation.

- Yan, X., Helmers, L., Lettenmeier, P.: Fleet-based performance model and degradation estimation. In: Book of Abstracts of the 4th International Conference on Electrolysis, p. 160 (2023). <https://engineering.nwu.ac.za/hysa/book-abstracts> [15]

Additionally, an international patent is filed.

- Yan, X., Helmers, L.: Method of Determining a Degradation of an Energy System and Computer Program Product. International patent, application number: PC-T/EP2024/074685. <https://register.epo.org/application?number=EP24765160>

This research is conducted as part of the hydrogen flagship project H2Giga, specifically within the scope of SEGIWA (serial production of gigawatt-scale PEM electrolyzers, grant number: 03HY121A) and DERIEL (de-risking for PEM electrolyzers, grant number: 03HY122A) projects, sponsored by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF). During the course of this research, the author is employed as a data analyst at Siemens Energy Global GmbH & Co. KG, Munich.

This dissertation is structured as shown in Fig. 1.4. Following this introduction chapter, it reviews the state of the art of SOH quantification methods and fleet-based algorithms (Chapter 2), then presents two SOH quantification methods for industrial electrolyzers (Chapter 3 to Chapter 6). Next, it generalizes the algorithm development process and provides a guide for industrial practitioners (Chapter 7). Finally, it concludes the research by answering the research question, reflecting on the methods, and suggesting future research directions (Chapter 8).

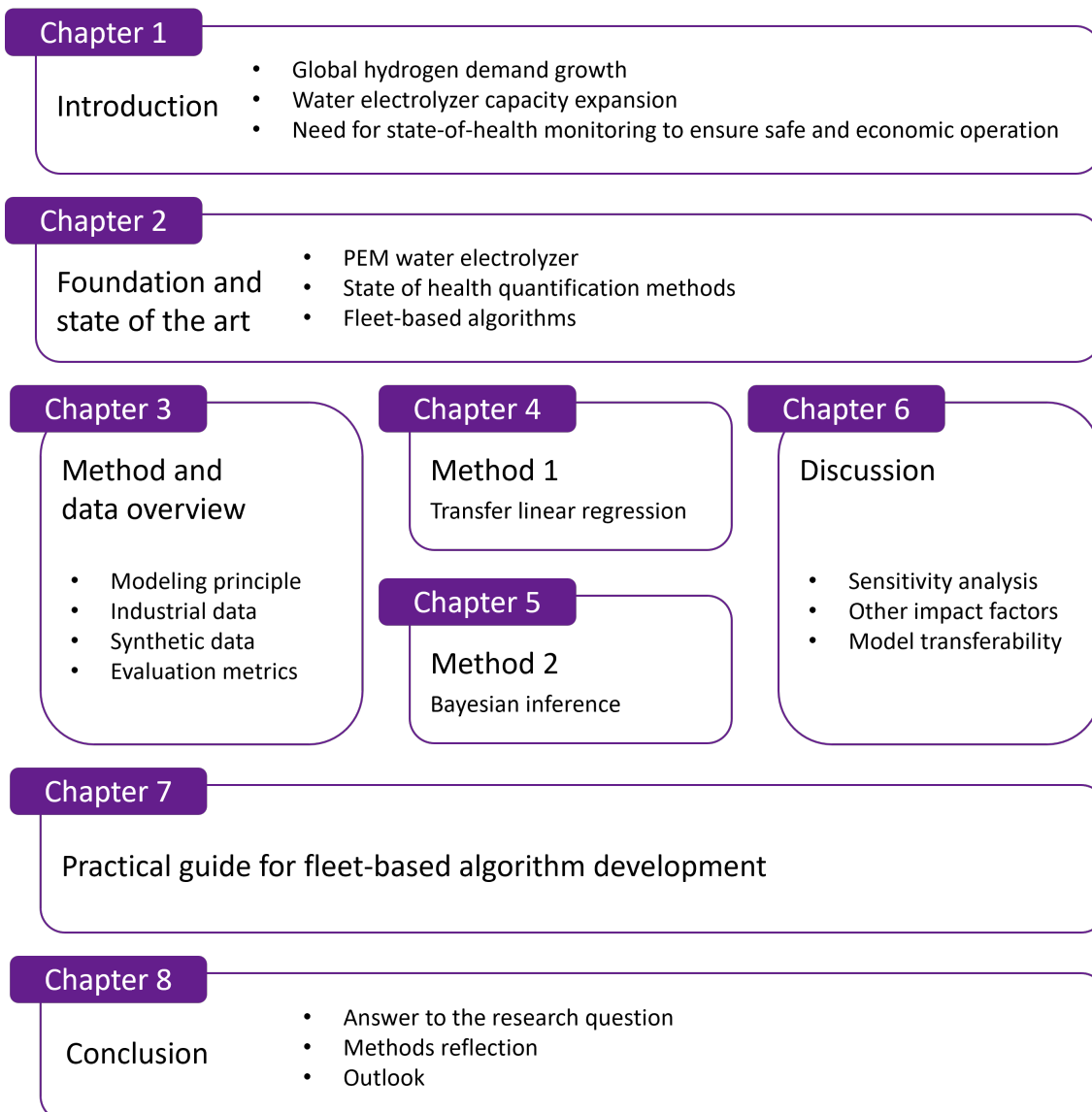


Figure 1.4: Dissertation structure.

2

Foundation and state of the art

This chapter first introduces the fundamentals of **PEM** water electrolysis, including its components and degradation mechanisms (Section 2.1). Next, it provides a deep dive into **PEM** water electrolyzer models, with a particular focus on the voltage models (Section 2.2). Furthermore, it reviews the state of the art of **SOH** quantification methods (Section 2.3) and fleet-based algorithms (Section 2.4). Finally, the identified research gaps are summarized (Section 2.5).

2.1 PEM WATER ELECTROLYZER AND ITS DEGRADATION

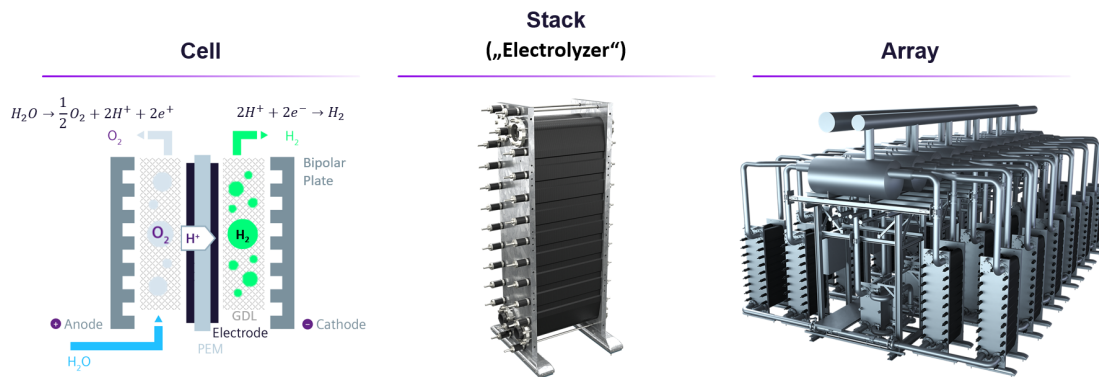


Figure 2.1: **PEM** water electrolyzer cell, stack, and array (source: Siemens Energy).

Water electrolysis is a technology that uses electricity to split water into oxygen and hydrogen. **PEM** electrolysis uses a proton exchange membrane as the central element. As shown in Fig. 2.1 (left), the membrane separates the anode and cathode electrodes and allows protons (H^+) to pass through. Water is fed into the anode side and split into oxygen and protons. The protons flow to the cathode side and get converted to hydrogen. To yield a higher hydrogen generation rate, multiple electrolyzer cells are assembled together as an electrolyzer stack (Fig. 2.1, middle). Multiple electrolyzer stacks can be further aggregated as an electrolyzer array to simplify auxiliary systems such as gas/water treatment equipment (Fig. 2.1, right).

An PEM electrolyzer cell consists of a membrane, electrodes, gas diffusion layers (GDLs), and bipolar plates (Fig. 2.1, left). Their functions, materials, and degradation mechanisms are introduced below. Unless specified otherwise, the content is primarily summarized from [16–18].

1. Membrane. The membrane is typically a thin perfluorosulfonic acid (PFSA) polymer membrane with a thickness of 100–400 μm . It serves to separate and support the electrodes and to conduct the protons. Degradation mechanisms related to the membrane include:
 - Mechanical. The membrane might suffer from mechanical failures (such as puncture and cracking) caused by impurities introduced during the manufacturing process, inhomogeneous compression, local heating, hydraulic pressure, etc.
 - Chemical. The membrane can get attacked by the oxidizing species generated by the crossover gases and metal impurities (such as Fe^{x+}), leading to the loss of functional groups and thinning. Additionally, metallic cations can occupy the ion exchange sites of the membrane, causing the membrane proton conductivity to decrease.
2. Electrodes. Anode and cathode electrodes are coated on each side of the membrane to form a membrane electrode assembly (MEA) or catalyst coated membrane (CCM). Due to the harsh electrochemical environment (e.g., high potential, low PH) and the inactive chemical reaction, precious metals are typically used as catalysts. Platinum is a common catalyst on the cathode side, and Iridium or its oxides is the state-of-the-art catalyst on the anode side. The catalysts are often dispersed on the supporting materials (such as TiO_2 or carbon-based materials) to maximize the catalyst utilization and improve their stability. Degradation mechanisms related to the electrodes include:
 - Catalyst dissolution. Iridium can convert to unstable forms (such as $\text{Ir}(\text{OH})_2$ and Ir_2O_3) with high solubility at low PH. Platinum can be oxidized to Pt^{2+} and get dissolved.
 - Catalyst agglomeration. Catalyst agglomeration is the increase in the size of catalyst particles, causing the active surface area to decrease. Although a reaction-specific particle agglomeration mechanism has not yet been identified, studies have shown that platinum particle agglomeration is related to the hydrogen generation rate [19] and iridium agglomeration is related to the ionomer loss of the membrane [20].
 - Metallic poisoning. Cationic impurities (such as Na^+ , Ca^{2+} , Fe^{3+}) do not only occupy the ion exchange sites of the membrane (mentioned under point 1), but also block the catalyst surface area, decreasing the catalyst reactivity.

- Support passivation. Metallic support material can form an oxide layer that reduces the electrical conductivity. This is influenced by current density and impurities in the cell.
3. **GDL**. **GDL** is also named as the porous transport layer (**PTL**) or the current collector. Its porous structure allows the generated gas to evacuate away from the electrodes and water to flow towards the electrodes. It also provides good electrical conductivity as well as mechanical support for the **MEA**. Carbon material is commonly used on the cathode side, and metallic material (such as titanium and stainless steel) is preferred on the anode side to sustain the corrosive environment. Degradation mechanisms related to the **GDL** include:
 - Mechanical degradation. During the electrolyzer manufacturing process, **MEA**, **GDL**, and bipolar plates are mechanically pressed together to form a stack. High compress pressure improves the electrical contact but also decreases the gas/water permeability.
 - Chemical degradation. Different materials face different degradation mechanisms: Titanium **GDL** suffers from hydrogen embrittlement, passivation, and corrosion, whereas stainless steel's main concern is corrosion. (Similar to the degradation mechanisms of the bipolar plate under point 4.)
 4. **Bipolar plate**. Bipolar plates are placed between each two adjacent cells in an electrolyzer stack. It conducts electrical current from one cell to the next, serving as the anode of one cell and the cathode of the next. It also has flow channels to distribute the water input and remove the produced gas. Titanium is the commonly used material for bipolar plates because of its corrosion resistance, high conductivity, strong mechanical support, and lightweight. Degradation mechanisms related to the bipolar plate include:
 - Hydrogen embrittlement. Titanium absorbs hydrogen and forms TiH_2 at high temperature and pressure, which makes the material very brittle and might lead to mechanical failure.
 - Passivation. An oxide film can form on the surface of titanium, which increases the electrical resistance between the bipolar plate and the **GDL**.
 - Corrosion. Titanium can be corroded in the presence of certain chemicals, such as fluoride ions. Fluoride ions attack the TiO_2 film on the surface of the bipolar plates, resulting in Ti-F^- compounds. An alternative material for bipolar plates is stainless steel, which reduces the embrittlement concern but has a larger corrosion problem: Stainless steel corrosion releases multivalent cations such as Fe^{2+} and Ni^{2+} which will further damage the catalyst and membrane.

The above-mentioned degradation mechanisms occur at a microscopic scale. On a macro-

scopic level, degradation can be observed from several aspects, such as increase in voltage, decrease in gas purity, and fluoride release into the effluent water. Commonly used degradation indicators in the literature will be summarized in Section 2.3.2.

2.2 PEM WATER ELECTROLYZER MODELS

2.2.1 Variety of models

In the literature, a variety of PEM water electrolyzer models are available, encompassing diverse physical domains, targeting different aggregation levels, and employing various modeling approaches.

One or more physical domains can be included in PEM water electrolyzer models, such as the electrochemical, thermal, mass transport, and fluidic domains [21]. The electrochemical domain is the core in PEM water electrolyzer models, as it captures the key function of electrolyzers: using electrical energy to trigger chemical reactions. For example, [22] models the electrolyzer voltage with respect to current density and proposes an efficient way to identify the model parameters; [23] uses an electrochemical model as the basis to optimize a photovoltaic (PV)-PEM energy system. In addition, to capture the impact of temperature on the electrochemical reaction, thermal models are sometimes coupled with the electrochemical model. For example, [24] considers the thermal effects and captures the temperature inhomogeneity in industrial-scale PEM water electrolyzers; [25] includes a thermal model to facilitate fault identification. Next, mass transport models can also be considered to quantify the resulting overpotential, gas production, and gas crossover. For example, [26] models the mass flow to calculate the resulting overpotential, thereby investigating the impacts of temperature and pressure on the electrolyzer voltage; [27] summarizes the mechanisms and models for the mass transport of H₂, and evaluated their contribution to H₂ crossover from the cathode to the anode side (a safety concern due to the formation of explosive gas mixture with O₂ produced on the anode side). Finally, some models also include the fluidic domain. For example, [28] models the water flow in the bipolar plate to examine the distribution of pressure and velocity; [29] models the biphasic flow in the GDL to analyze the impact of geometrical and operating parameters on the electrolyzer performance.

PEM water electrolyzer models can target various aggregation levels, ranging from the cell or stack level, to the component level, and even down to the molecular level. At a high level of aggregation, a cell or a stack is modeled as a whole, also named as a single-box model [30]. For example, [31] models the energy demand of an electrolyzer to optimize the design of a power-to-gas plant; [12] models the voltage of an electrolyzer to monitor

its performance degradation over time. One level deeper, the model can consider the individual components within an electrolyzer [32]. For example, [33] breaks down the voltage into different overpotentials to investigate the impact of catalyst loading and GDL properties; [34] extends the voltage breakdown with more details on the ohmic resistance of each component and the mass transport loss, thereby offering insights into the performance at high current densities. Delving further, the models can be at the molecular level. For example, [35] models the membrane degradation process by considering the formation of H_2O_2 and 23 related chemical reactions; [36] models the dissolution of iridium in the anode catalyst layer, to investigate its impact on the performance degradation.

Regarding the modeling approach, PEM water electrolyzer models can be physics-based or empirical models [21, 37]. In physics-based models, all parameters have a physical meaning, although some of them can be fitted with data or obtained from empirical equations [21]. For example, [38] models different sources of overpotentials as well as the flow of water and gas based on physical laws; [39] models the electrolyzer voltage as an equivalent circuit and uses the current interrupt technique to identify the resistance and capacity. In empirical models, the parameters do not have any physical meanings [21]. Empirical models can have explicit functional forms to describe the relationship of variables. For example in [40, 41], the electrolysis voltage is modeled as a linear function of current. Empirical models can also have complex structures (such as neural networks) to learn complex patterns based on a large amount of data. For example, [42] models the electrolysis voltage based on a neural network using electric current and temperature as inputs; [43] uses five different machine learning algorithms to model the hydrogen production rate and current density as a function of design parameters (e.g., electrode and catalyst types). Physical knowledge can also be incorporated into empirical models to enhance its performance, such as [44], which models the activation overpotential with a neural network and includes a physical formula of the overpotential in the loss function.

2.2.2 Voltage models

As already mentioned in Chapter 1, this research focuses on tracking voltage degradation, as voltage is easily measurable and has direct economic relevance. Therefore, this section deep dives into the existing PEM water electrolyzer voltage models.

2.2.2.1 Physics-based voltage model

The electrolysis voltage (U) is the sum of the reversible voltage (U_0), activation overpotential (η_{act}), ohmic overpotential (η_{ohm}), and mass transport overpotential (η_{trans}) [17]:

$$U = U_0 + \eta_{\text{act}} + \eta_{\text{ohm}} + \eta_{\text{trans}}. \quad (2.1)$$

Reversible voltage is sometimes also referred to as the open circuit voltage [45]. It is the minimum voltage required to split 1 mole of water, determined by the Gibbs free energy change (ΔG) which is a function of temperature (T) and pressure (P) [46]:

$$U_0 = \frac{\Delta G(T, P)}{zF}, \quad (2.2)$$

where z is the number of electrons exchanged in the reaction ($z = 2$) and F is Faraday's constant ($F = 96\,485 \text{ C mol}^{-1}$).

Activation overpotential is the extra voltage required to overcome the molecular bonds and begin the reaction [45]. It can be further split into the activation overpotential of the anode ($\eta_{\text{act,an}}$) and cathode ($\eta_{\text{act,ca}}$) side, where the contribution from the cathode side is often neglected because it is a small contribution compared to the anode side [47]:

$$\eta_{\text{act}} = \eta_{\text{act,an}} + \eta_{\text{act,ca}} \approx \eta_{\text{act,an}}. \quad (2.3)$$

There are different formulations to model the activation overpotentials, mostly based on the Butler-Volmer equation [21]. One popular formulation is the Tafel equation, adopted in many publications such as [22–25, 33, 47–50]. It is a simplification of the Butler-Volmer equation where only the forward reaction is considered and the backwards reaction is neglected:

$$\eta_{\text{act,an}} = \frac{RT}{z\alpha F} \ln \left(\frac{i}{i_0} \right), \quad (2.4)$$

where R is the ideal gas constant ($R = 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$), α is the charge transfer coefficient, i is the electric current density, and i_0 is the exchange current density. Both α and i_0 are temperature dependent [24]. The equation above can also be simplified to be

$$\eta_{\text{act,an}} = a \ln i - b, \quad (2.5)$$

where a is named as the Tafel slope and b as the Tafel constant [17].

Ohmic overpotential arises from the resistance to the electron flow through the cell components as well as to the proton flow through the membrane [17]. It can be modeled

with the standard Ohm's law:

$$\eta_{\text{ohm}} = IR_{\text{ohm}}, \quad (2.6)$$

where R_{ohm} is the sum of all electronic and protonic resistances because they are connected in series [17, 51, 52]. The resistance is temperature dependent [21].

Mass transport overpotential comprises two components: diffusion overpotential (η_{dif} , also named as concentration overpotential [26, 52]) and bubbles overpotential (η_{bub}) [17]:

$$\eta_{\text{trans}} = \eta_{\text{dif}} + \eta_{\text{bub}}. \quad (2.7)$$

Diffusion overpotential is caused by gas bubbles hindering the water supply through the porous media, while bubbles overpotential arises when gas bubbles shield the electrochemical active area, thereby reducing the catalyst utilization [17]. There are different models to describe the diffusion overpotential [21]. One widely used model is based on the Nernst equation, adopted in [17, 22, 24, 26, 51–54]:

$$\eta_{\text{dif}} = \frac{RT}{zF} \ln \left(\frac{C_{O_2}}{C_{O_2,0}} \right)^{\frac{1}{2}} + \frac{RT}{zF} \ln \left(\frac{C_{H_2}}{C_{H_2,0}} \right). \quad (2.8)$$

C_{O_2} and C_{H_2} stand for the gas concentrations at the membrane-electrode interface. $C_{O_2,0}$ and $C_{H_2,0}$ are the concentrations at a reference working condition (saturation concentrations are used in [24, 55]). Bubbles overpotential also has various formulations [21]. However, it is rarely explicitly modeled in literature, because it only becomes significant at high current density [17]. Overall, mass transport overpotential receives less attention compared to other overpotentials (ohmic and activation overpotentials) [21], because of its relatively minor influence at low to medium current densities.

The voltage model is often graphically presented with a polarization curve: the relationship between voltage and current density under constant operating temperature and pressure [46, 56]. An example is shown in Fig. 2.2. Polarization curves and their drift over time are used in many studies to investigate the degradation of PEM water electrolyzers. For example, it is used in [57] to analyze the degradation caused by iron ions and in [58] to investigate the impact of intermittent operation.

While physics-based voltage models are solidly grounded in theoretical principles, they have certain disadvantages. First, they require numerous parameters that are sometimes challenging to determine. For example, the Tafel parameters to describe the activation overpotential (Eq. 2.5) can only be determined with data at very low current densities (typically between 0.01 and 0.1 A cm⁻² [58–60]). However, in industrial applications, operation

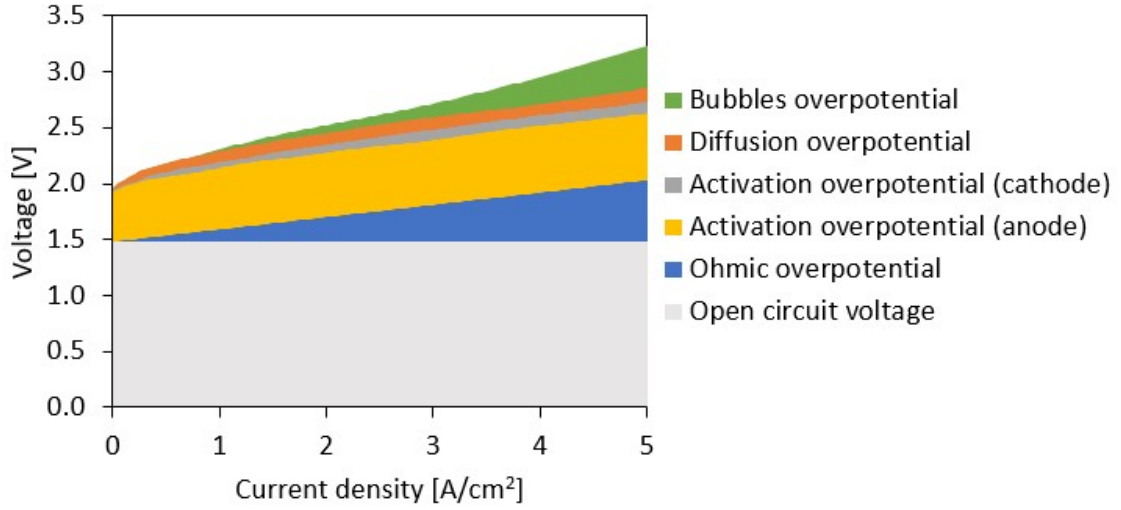


Figure 2.2: An exemplary polarization curve with the individual voltage components (adapted from [17]).

at low current density is very rare due to the risk of high gas crossover [39] and the high leveled cost of H_2 [61]. In addition, the more impact factors to include in the model, the more parameters are needed. For example, to capture the impact of temperature on the membrane's resistance, parameters such as the membrane's water content and thickness are required [24]. While these parameters can be obtained from the membrane's supplier or lab measurements for a pristine membrane, getting these values for a degraded membrane in use is very difficult.

Second, to reduce model complexity, simplified assumptions are sometimes necessary, which can lead to poor accuracy [62]. For example, the mass transport overpotential is often omitted in the voltage model [21], but [33] shows that this can lead to a biased Tafel fit. Another example is regarding the inhomogeneity within large-scale electrolyzers. [24] demonstrates the inhomogeneous temperature and current density distribution in an industrial scale electrolyzer stack (40 cells, each with 1 m^2 active cell area). If such inhomogeneities are neglected in a model, degradation or even failure caused by local effects might be overlooked.

2.2.2.2 Empirical voltage model

To avoid the above-mentioned difficulties in fully capturing the physics, some studies employ empirical models to describe the electrolysis voltage. While empirical models have the limitation of being accurate only for a specific operating range and electrolyzer design [37], they can be sufficient for industrial applications, which are typically engineered

for a predefined operating range using the same design.

In [40], voltage is modeled as a linear function of current:

$$U = 0.326I + 1.476. \quad (2.9)$$

Although the models' simplicity is appealing, it might not adequately capture the complexities of industrial electrolyzers. In an industrial setting, operating parameters besides current (such as temperature) might not be constant. Therefore, to ensure that the model accurately reflects the real-world conditions of industrial operations, these additional parameters also need to be included in the model.

Study [63] uses a more complex formula, which includes non-linear terms and temperature dependency based on the model proposed in [64]:

$$U = U_0 + \frac{r_1 + r_2 T}{A} I + (s_1 + s_2 T + s_3 T^2) \ln \left(\frac{t_1 + \frac{t_2}{T} + \frac{t_3}{T^2}}{A} I + 1 \right), \quad (2.10)$$

where U_0 represents the reversible cell voltage, A is the cell area, r_i , s_i , and t_i are all fitting parameters. As the model complexity increases, the number of unknown parameters to be fitted also increases. Consequently, more data are needed to accurately determine these parameters.

Studies [42] and [65] further increase the model's flexibility by using machine learning models, which do not require any predefined relationship between voltage and its impact factors. [42] uses a neural network model with electric current and temperature as inputs; [65] uses both neural network and support vector regression models with various operating and material parameters as inputs. Although these machine learning models offer great flexibility and the potential for high predictive accuracy, they sacrifice the model's interpretability, which hinders their adoption for industrial applications.

2.3 SOH QUANTIFICATION METHODS

2.3.1 SOH applications and types

The **SOH** of an industrial machine refers to its performance, efficiency, and reliability. For fault detection or diagnosis purposes, **SOH** can be represented by categorical labels such as healthy, faulty, or the exact fault types. Whereas for tracking the gradual degradation processes, **SOH** is represented with low-dimensional (typically 1-dimensional [66]) numerical values. The latter is the focus of this research.

Tracking the **SOH** of industrial machines can support various business needs across the

entire product lifecycle, from the research and development (R&D) phase to the end of life.

- Knowledge discovery. To support R&D, the SOH data can be analyzed to identify degradation patterns, trends, and root causes, thereby providing a basis for improving the system design.
- Performance warranty calculation. To support sales, the SOH data can be used to assess the performance and risk of the product. This information can be utilized to determine warranty terms and risk mitigation strategies.
- Operation optimization. During the operation phase, continuously monitoring the machines' condition and performance allows for data-driven operation or dispatch decisions to optimize overall efficiency and resource utilization.
- Safety monitoring. During the operation phase, continuous SOH monitoring enables early detection of potential safety hazards to prevent unexpected equipment failures or accidents.
- Electrolyzer replacement planning. When electrolyzers approach the end of life, SOH monitoring helps to schedule replacement based on the degradation trend, thereby maintaining the operation efficiency and minimizing downtime.

There are two types of SOH indicators: physical and synthetic indicators [66]. Examples of physical indicators include

- for batteries: capacity [67], characteristics of the discharge capacity curve [68], charging duration for a predefined voltage range [69], normalized voltage [70];
- for fuel cells: normalized voltage [71], internal resistance [72];
- for gas turbines: normalized power output [73];
- for rolling bearings: root mean square of vibration signals [74], maximum power spectral density [75].

In contrast to physical indicators, synthetic indicators do not have clear physical meanings. They are obtained through statistical techniques such as multiple linear regression - linearly combining multiple sensor measurements [76, 77], principal components analysis - reducing the original multi-dimensional input data to a small number of variables that capture patterns in the data [78, 79], and autoencoders - using a neural network-based model to learn features in an unsupervised manner [80, 81].

For PEM water electrolyzers, existing studies focus on physical health indicators directly measured in laboratory experiments. This is introduced next.

2.3.2 In-situ measurements for PEM water electrolyzers

The upper part of Table 2.1 provides an overview of the commonly used SOH quantification methods for PEM water electrolyzers. Only in-situ methods are considered because of their high relevance to the goal of this research - continuous condition monitoring, while ex-situ or postmortem methods are excluded. As shown in the table, existing studies focus on directly measuring the SOH: Typically, degradation triggers are applied in laboratory environments (e.g., contaminated water [57], extreme operation schemes [58, 82], harsh manufacturing procedures [83], modified design [82]), followed by the measurement of degradation responses. Among the various SOH measurements, voltage-related quantities are the most widely used.

These SOH quantification methods are evaluated against the following requirements of continuous condition monitoring for industrial electrolyzers (Table 2.1).

- Easy to measure in industry. Many laboratory measurement techniques are impractical to implement in the industry. For example, the current interrupt technique requires data with nanosecond [84] or millisecond [24] resolution, leading to expensive sensors and data acquisition systems; the electrochemical impedance spectroscopy technique is constrained by its maximum deliverable current, which is insufficient to gain insights for industrial size electrolyzers [24].
- Not require specific operation mode. Most measurements require a predefined operational protocol. For example, voltage under constant current requires that an electrolyzer is constantly operated under the same current; polarization curve measurement requires the current to vary over a wide range (e.g., 0.001 - 4 A/cm² in [59]) under constant temperature and pressure. Such measurements are inconvenient for the industry as the operation scheme should be aligned with the operator's business requirements. Fluoride release is the only exception, which does not necessarily require a specific operation mode. Here, a distinction between total fluoride release and fluoride release rate needs to be made. Total fluoride release as an SOH indicator does not require a specific operation mode, because it is an accumulated value to reflect the membrane thinning over time. A higher total fluoride release is always more concerning than a lower one, regardless of how the electrolyzer is operated. On the other hand, fluoride release rate faces instantaneous impacts from operation modes such as current density, water flow, and pressure. Therefore, to ensure the comparability of fluoride release rates, they must be measured under the same operation modes.
- Not reliant on degradation model. As PEM water electrolyzer is an emerging technology with limited long-term operation experience and rapid design updates, it is risky to assume a degradation model and rely on it for SOH quantification. This is

not an issue for the direct measurement-based method, as no model or assumption on the degradation behavior is needed.

Table 2.1: Two types of SOH quantification methods evaluated against requirements: (a) in-situ measurements for PEM water electrolyzers that are commonly used in the literature and (b) statistical methods that mitigate the impact of operation modes.

			Easy to measure in industry	Not require specific operation mode	Not reliant on degradation model
(a) In-situ measurements for PEM water electrolyzers	Voltage-related	Voltage under constant current [57, 83]	X		X
		Polarization curve [57, 58, 82, 85–88]	X		X
		Current interrupt [39, 84, 89]			X
		Electrochemical impedance spectroscopy [39, 57, 58, 82, 83, 85–88, 90]			X
		Cyclic voltammetry [58, 82, 86]			X
	Foreign gas concentration [27, 91]				X
	Fluoride release [35, 85–88]			(X)	X
(b) Statistical methods	Filter-based method to identify physical parameters [92, 93]		X	X	
	Measurement correction based on characteristic curves [71, 73, 94–96]		(X)	(X)	X

As mentioned above, most direct measurement-based methods require a predefined operation procedure, which is impractical for the industry. The following section reviews statistical methods that mitigate the impact of operation modes. While none of these statistical methods have been applied to PEM water electrolyzers, their underlying principles are relevant.

2.3.3 Statistical methods to mitigate the impact of operation modes

As shown in the lower part of Table 2.1, two types of statistical methods applicable under various operation modes have been identified. The first type is filter-based methods to identify physical parameters that are invariant to operation modes. The disadvantage of these methods is that they rely on degradation models that explicitly describe the degradation behavior. For example, [92] uses an extended Kalman filter to estimate the SOH indicator α for PEM fuel cells. It is defined as the degradation of resistance (R) and limiting current density (i_L): $\alpha_t = \frac{R_t}{R_0} - 1 = 1 - \frac{i_{L,t}}{i_{L,0}}$, where the subscript 0 and t denote the initial value and the value at time t , respectively. The definition of α reveals the first assumption: R and i_L have the same degradation state. Moreover, it is assumed that α varies linearly overtime: $\alpha_t = \beta t$, where β is the degradation speed. Another example is [93]. It also applies an extended Kalman filter for SOH quantification for PEM fuel cells, with the assumption that the resistance, limiting current, and exchange current all change linearly over time. As mentioned in Section 2.3.2, due to the limited operation experience and rapid design updates, such degradation models for industrial PEM water electrolyzers are not yet established, making filter-based methods unsuitable.

The second type of statistical method is to correct the measured values based on characteristic curves. This approach requires measuring characteristic curves to derive a correction model. For example, [71, 94] use voltage-based SOH indicators for fuel cells, and the impact of current on cell voltage is corrected using a fitted voltage-current model from the polarization curve measured at the beginning of life (and additionally at the end of life in [71]). [95] uses a power-based SOH indicator and accounts for the impact of current with a fitted power-current curve. [96] uses a resistance-based SOH indicator for lithium-ion batteries and compensates for the temperature influence with a resistance-temperature curve. Although obtaining 1-dimensional characteristic curves at the beginning or end of life is relatively straightforward, capturing multi-dimensional curves over time is impractical but essential to ensure comparability across complex operating conditions and various degradation states (hence the parentheses in the last row of Table 2.1). For instance, to capture the voltage degradation for PEM water electrolyzers, re-measuring characterization curves over time is necessary due to the changing voltage-current relationship depending on the degradation mechanism. Additionally, the voltage measurement is impacted by many factors such as current, temperature, and start, thus a multi-dimensional correction for the voltage is required.

The second type of statistical method is the closest to fulfilling all requirements and therefore is the basis for this research. To overcome the difficulty in measuring multi-dimensional characteristic curves over time, this research develops data-driven methods

to regularly generate a multi-dimensional performance model regardless of the operation mode, which is then used to correct the measured electrolyzer voltage to a predefined operating condition. By doing so, the impact of operating modes is eliminated, so the corrected voltage can be compared across time as a SOH indicator.

2.3.4 Challenge of limited data coverage

The challenge of regularly generating a multi-dimensional performance model is to gather sufficient data to fit the model coefficients. As mentioned in Chapter 1, the coverage of industrial operation data can be limited. For example, when an electrolyzer is operated under a constant electric current, the coverage of the current data is very narrow. This scarcity of data complicates the quantification of the voltage-current relationship.

The approaches to address the limited data coverage can be broadly classified into data- and model-level approaches. Data-level approaches involve adding new data. This can be achieved by identifying the data gap and then collecting additional data in a targeted manner for the underrepresented group [97]. However, such data collection is only possible if one has control over the data source. Another way to add new data is to generate synthetic data with data augmentation techniques. This has been widely used for image data [98] but is less researched for time series data [99], especially for multivariate time series data [100]. Demir et al. [100] proposed using autoencoders, variational autoencoders, and generative adversarial networks to augment multivariate time series data. However, training such neural network-based algorithms on a daily basis—a typical frequency for SOH monitoring—is computationally expensive. Therefore, data augmentation is not a suitable solution for this research.

Model-level approaches use the limited data to train a model without altering the dataset, which is also named as few-shot learning [101]. This is the focus of this research. Specifically, this research utilizes fleet knowledge to enhance the model fitting under limited data coverage. In the next section, existing fleet-based algorithms in the literature are reviewed.

2.4 FLEET-BASED ALGORITHMS

This section is a recap of a more detailed publication: Yan, X.⁼, Woelke, J.⁼, Bensmann, B., Eckert, C., Hanke-Rauschenbach, R., Nieße, A.: Cross-method overview of fleet-based machine health estimation and prediction: a practical guide for industrial applications. IEEE Access 13, 60131–60147 (2025). doi:[10.1109/access.2025.3556251](https://doi.org/10.1109/access.2025.3556251) (⁼ equal contribution) [14].

2.4.1 Basics of fleet

Data-driven methods require a large amount of data to train a reliable model. This is often challenging in the industry due to design variations and diverse operating conditions. This makes data collected from one machine not directly applicable to another. Especially for costly machines with limited testing opportunities, generating a large amount of data for each design and operating condition is unrealistic. To address the data scarcity problem, fleet data can be leveraged to enrich the database, allowing for improved data-driven models.

A fleet is a group of machines that share some technical and/or operational characteristics [102, 103]. It can be described by the following dimensions [103].

- Purpose (e.g., transportation, power generation).
- Context, i.e. general information of the fleet without details (e.g., unit, size, lifetime).
- Composition, i.e. technical and functional characteristics of the units in the fleet (e.g., design, manufacturer).
- Operating condition, i.e. the environment the fleet is exposed to (e.g., usage, load, stress).
- Data characteristics (e.g., structure, type, source).

From the manufacturer's perspective, a fleet consists of machines that are technically similar but may be operated differently; while from the operator's perspective, machines in a fleet perform the same function but have different technical characteristics [104]. Depending on the level of similarity, a fleet can also be categorized as identical, homogeneous, or heterogeneous fleets [103]. To address differences within a fleet, it can be helpful to divide it into subfleets - groups of machines exhibiting higher similarities, e.g., regarding technical characteristics or degradation patterns, depending on the modeling objective [103].

Fleet data can be used in modeling through various approaches. This research reviews the existing fleet-based algorithms in the literature and the results are presented in the next section.

2.4.2 Categories of fleet-based algorithms

As this research focuses on tracking the gradual degradation process of industrial electrolyzers, literature within the following scope is deemed the most relevant.

- Goal: SOH estimation and/or prediction.
- Application: A fleet of industrial machines, sub-systems, components, and equipment.
- Method: Regression algorithm.
- Data: Time series data.
- Type: Peer-reviewed publication.

- Language: English.

By reviewing the collected literature, six categories of fleet-based methods have been identified and illustrated in Fig. 2.3. They are differentiated in the data used for modeling (input) and the resulting model (output). On the input side, the data basis consists of data from all machines in the fleet. They can be experimental data from the lab, operational data from the field, or simulation data from a suitable model. Some methods - fleet modeling, parameter-based, and feature-based transfer learning, use data from all machines as input; while other methods use only data from a part of the machines. On the output side, fleet and subfleet modeling produce a generic model for multiple machines; while similarity-based modeling and transfer learning produce a dedicated model for the target machine.

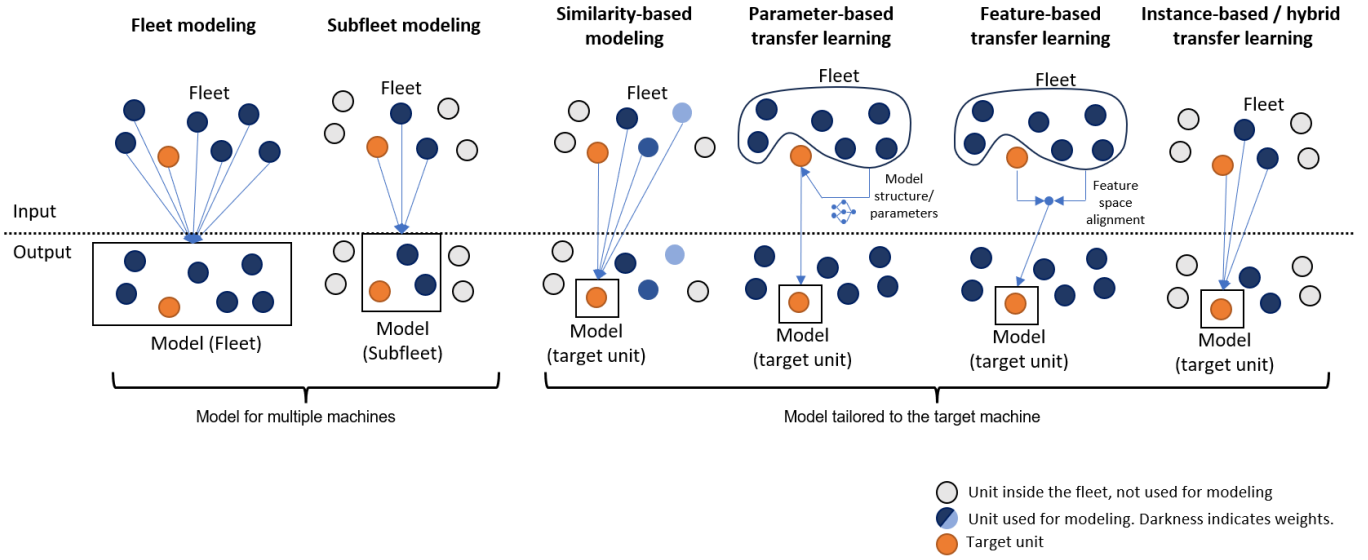


Figure 2.3: Categories of fleet-based algorithms.

- Fleet modeling is the most straightforward method of utilizing fleet data. It simply uses data from all units to construct a model that applies to the entire fleet. For example, [68] pools degradation data of 124 lithium-ion batteries and builds a linear model to predict the batteries' lifetime.
- Subfleet modeling first identifies a group of similar units (the subfleet) within the fleet, then builds a model that applies to all the machines within this subfleet. For example, [105] measures the Euclidean distance of the degradation trajectories between the target unit and other units. Units with a distance lower than a predefined threshold are grouped together as a subfleet to train a degradation forecast model.
- Similarity-based modeling measures the statistical similarity between the target unit and other units in the fleet, selects the most similar ones as the referential instances,

and then makes predictions for the target machine based on the referential instances. In some studies, an additional weighting is applied: Higher weights are assigned to units with higher similarities in the modeling process. Similarity-based modeling is primarily used for remaining useful life prediction, which has been reviewed by Xue et al. [66]. For example, [106] quantifies the similarity among fleet units with the maximum mean discrepancy metric, which measures the distance between probability distributions. Then data from similar units are aggregated with Weibull distribution fitting, where the mean of the fitted distribution is the predicted lifetime.

- Parameter-based transfer learning starts with developing an initial model with data from non-target units (named as "source domain" in transfer learning), and then its parameters are fine-tuned with data of the target unit ("target domain"). For example, in [69], a neural network-based SOH prediction model is first trained using data from one battery. Then, the parameters of one layer in the neural network are fine-tuned using data from another battery with a different design.
- Feature-based transfer learning typically starts from multi-dimensional sensor data and aims to find features that are effective for both the target and non-target units (i.e., creating domain-invariant features). For example, in [107], a neural network-based model is constructed to predict the SOH of lithium-ion batteries, with features extracted using convolutional neural networks. The difference between the features from the source and target domains is measured using the maximum mean discrepancy, and this metric is incorporated into the loss function to be minimized during the model training process. Consequently, the final model minimizes the feature differences between the domains.
- Instance-based transfer learning selects the most relevant data from the fleet to assist model training for the target unit. It is usually combined with parameter-based or feature-based TL to form a hybrid approach [104]. Hence this category is named instance-based/hybrid transfer learning. For instance, in [108], a neural network-based algorithm is developed for predicting the lifetime of lithium-ion batteries. The algorithm selects the most relevant batteries based on factors such as degradation rate and degradation curve shape. Then, a pre-trained model is fine-tuned using the data from the selected batteries.

In Table 2.2, these fleet-based algorithms are evaluated against the following requirements.

- Machine-specific model. As PEM water electrolyzer is an emerging technology with quick technology iterations, a machine-specific model is advantageous for accommodating the differences among the electrolyzers. For this requirement, fleet and subfleet modeling fall behind as they build an average model for multiple machines instead

- of emphasizing individual differences.
- **Interpretable model.** For industrial applications, model interpretability is crucial because it enhances the credibility of decision-making, reduces safety risks, and facilitates knowledge discovery [109,110]. However, most transfer learning algorithms are based on neural network structures, which are hard to interpret. The other method categories do not have a predominant model structure, for which both interpretable and non-interpretable models are available in the literature.
 - **Quick retraining.** To accurately capture the unknown long-term degradation behavior and adapt to new technologies, regularly and quickly updating the model is beneficial. However, most algorithms start with the entire fleet database, resulting in time-consuming model training, especially when dealing with complex model structures. The only exception is parameter-based transfer learning, which allows for fine-tuning an existing model without the need for fully retraining the entire model.

Table 2.2: Fleet-based algorithms evaluated against requirements. Only one exemplary reference is given for each category. More references can be found in [14].

		Machine-specific model	Interpretable model	Quick retraining
Fleet modeling [68]			(X)	
Subfleet modeling [105]			(X)	
Similarity-based modeling [106]		X	(X)	
Transfer learning	Parameter-based transfer learning [69]	X		(X)
	Feature-based transfer learning [107]	X		
	Instance-based / hybrid transfer learning [108]	X		

To fulfill all the requirements, this research proposes two SOH quantification methods with the following highlights: (1) Model is trained for each electrolyzer and is regularly updated over time. (2) Simple linear model structures are used that are easy to interpret. (3) Fleet data are aggregated into a prior probability distribution that can be directly used for model fitting.

2.5 RESEARCH GAP

Reviewing the existing literature reveals the following research gaps.

- Existing in-situ **SOH** measurement methods for **PEM** water electrolyzers are not suitable for industrial electrolyzers. The main gap is the requirement for a predefined operation modes, which is not practical for industrial electrolyzer due to the non-standardized but business-driven operation modes. (Section 2.3.2)
- Existing statistical approaches to mitigate the impact of operation modes are also not suitable, because they either require established degradation models or rely on measuring characteristic curves. (Section 2.3.3)
- An alternative to measuring characteristic curves is fitting performance models with operation data. However, this approach faces challenges due to limited data coverage. Existing algorithms that leverage fleet data to address this issue fall short in providing tailored, machine-specific models that are interpretable and can be updated quickly with new data. (Section 2.4)

To address these research gaps, two **SOH** quantification methods designed for industrial electrolyzers are developed in this research. Their fundamental principles are presented in the next section.

3

Method and data overview

This research develops two methods for **SOH** quantification for industrial water electrolyzers. The commonalities between these two methods are presented in this chapter, including the basic principle (Section 3.1), the data used for method evaluation (Section 3.2.1), the evaluation metrics (Section 3.2.2), and the benchmark method for comparison (Section 3.2.3).

3.1 METHOD OVERVIEW

As mentioned in Chapter 1, the target of this research is to track the voltage degradation of **PEM** water electrolyzers. A degraded electrolyzer requires a higher voltage to maintain the electrochemical reaction at a given rate [17]. Nonetheless, the measured voltage cannot be directly used as a degradation indicator, because it depends not only on degradation, but also on operating conditions such as current and temperature. In laboratory environments, dedicated tests under strictly controlled operating conditions can be performed to measure the voltage degradation. But as stated in Section 2.5, this is impractical for industrial electrolyzers due to the arbitrary, business-driven operation modes.

To address the impact of arbitrary operating conditions in the industry, this research develops data-driven **SOH** quantification methods that adjust the measured voltage under arbitrary operating conditions to the reference condition (U_{rc} in Fig. 3.1). The basis for this adjustment is mapping the relationship between voltage and operating conditions by fitting a voltage model (Case A in Fig. 3.1). However, in cases of limited data coverage (e.g., when an electrolyzer is operated under a constant electric current), the model cannot be accurately determined (Case B in Fig. 3.1). To address this difficulty in model fitting, fleet knowledge is leveraged. This research presents two methods that incorporate fleet knowledge in two different manners. The general working principle of the two methods is presented below, and details will be introduced in Section 4 and 5.

Both proposed methods follow a similar principle, as illustrated in Fig. 3.2. The input is operation data, segmented into intervals. For each interval, a voltage model is fitted with the data, and the fitted model can be used for the following purposes.

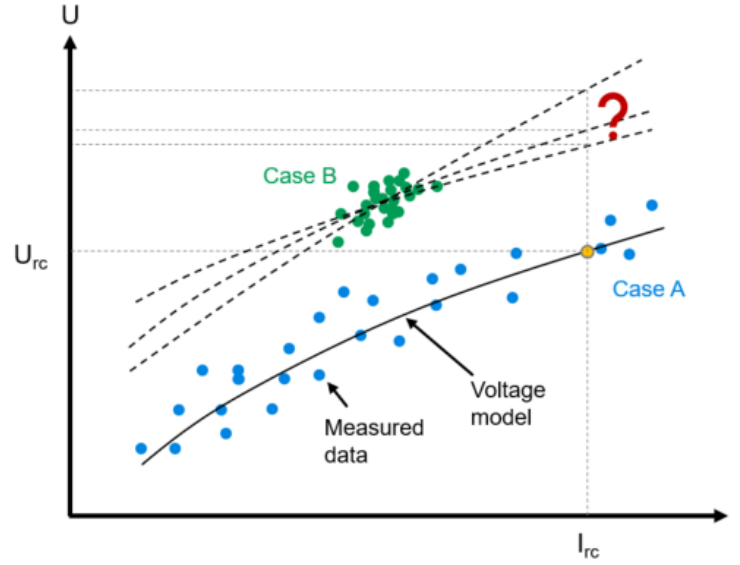


Figure 3.1: Case A: Voltage model fitting and U_{rc} calculation based on arbitrary operation data. Case B: Model fitting ambiguity under limited data coverage. (Note: Voltage is a function of several variables. To simplify the visualization, only the electric current is plotted on the x-axis here.)

- It can be used to calculate voltage under a predefined reference condition (U_{rc}), which serves as the [SOH](#) indicator for electrolyzers.
- It can be used to calculate voltages at different electric current levels to obtain polarization curves (Section 2.2.2.1), providing a basis for understanding the degradation mechanism.
- It can be used to predict the voltage performance in the near future. This forecast capability enables operators to optimize the power input to the electrolyzers, ensuring efficient hydrogen production.

As shown in Fig. 3.2, this process involves two main considerations: what voltage model to use, and how to fit the model. On an abstract level, the two proposed voltage models in this research both have a linear structure. They can be written in the standard linear regression matrix form:

$$Y = X\beta + \epsilon, \quad (3.1)$$

where Y is a vector of the dependent variable (e.g., voltage), and X is a matrix of the independent variables (e.g., operating conditions that impact voltage). Both X and Y are directly obtained from the operation data. The parameter β is a vector of unknown

coefficients that need to be fitted, and ϵ is a vector of errors that follow a normal distribution with an expectation of zero. There are different model fitting methods to estimate the unknown parameter β : a basic benchmark method (will be introduced in Section 3.2.3) and two fleet-based methods developed in this research (will be introduced in Section 4.1.3 and 5.1.3). With the fitted coefficient β and the predefined reference condition X_{rc} , the corresponding Y_{rc} can be calculated with

$$Y_{rc} = X_{rc}\beta. \quad (3.2)$$

The subscript "rc" denotes the reference condition, which can be tailored based on the use case. For example, for maintenance planning, where the focus is the performance under the nominal condition, X_{rc} should correspond to the nominal values; while for method validation, where more measured data are desired as the ground truth, X_{rc} should be set at the most common operating conditions.

Based on this common working principle, the two SOH quantification methods proposed in this research differ in (1) the voltage model and (2) the model fitting approach to address limited data coverage. Details will be introduced in Section 4 and 5.

3.2 METHOD EVALUATION

The proposed methods are evaluated using two types of data: industrial data and synthetic data. Industrial data allow to develop realistic voltage models and to validate the methods' effectiveness for actual industrial applications (Section 3.2.1.1). Synthetic data offer more detailed ground truth to validate the methods' accuracy and support sensitivity analyses with varying data setups (Section 3.2.1.2). These two types of data complement each other, ensuring a comprehensive evaluation of the proposed methods. The evaluation metrics encompass three perspectives: the accuracy of U_{rc} , polarization curves, and voltage forecast (Section 3.2.2). Furthermore, the proposed methods are compared with a benchmark method: simple linear regression (SLR) (Section 3.2.3).

3.2.1 Data for evaluation

3.2.1.1 Industrial data

The data are from 12 multi-cell PEM water electrolyzer stacks used for industrial purposes. Each stack has a nominal power above 0.5 MW. They are operated under atmospheric pressure, at around 60°C. The temperature variation during normal operation is within several degrees. The input current densities are at medium levels, resulting in a nearly

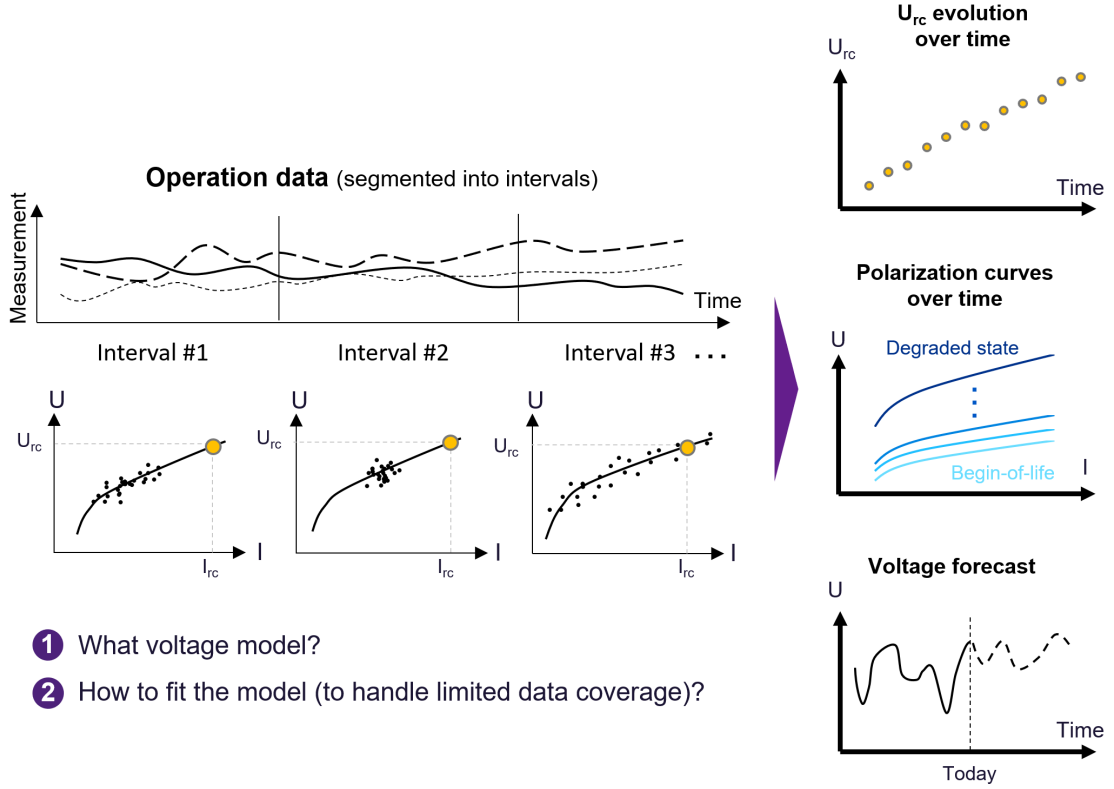


Figure 3.2: The proposed methods involve fitting voltage models to convert measured voltage under arbitrary operating conditions to a reference condition (denoted by the subscript "rc"). The fitted models can be used for calculating the SOH indicator U_{rc} , polarization curves, and voltage forecast.

linear voltage-current relationship. The operation profiles are not standardized but business-driven, including periods of volatile and constant operation. The analyzed data cover a duration of 1 to 3 years with a 1-minute resolution. The following measurements are used in this research. Due to confidentiality reasons, more detailed technical specifications cannot be provided.

- Voltage: the average cell voltage of all electrolyzer cells in a stack, measured with sensors attached to the bipolar plate of each cell.
- Current: the direct current output of the rectifier supplying power to the electrolyzer stack.
- Temperature: the average stack temperature derived from the mean of the process water inlet and gas/water outlet temperature measurements.

The measurement error of industrial data is set as the benchmark to evaluate the accuracy of the proposed methods. Specifically, the error of the modeled voltage or U_{rc} should not

exceed the error of the actual voltage measurement, which is below 10 mV. Hence, the target is to achieve a single-digit-mV level error.

3.2.1.2 Synthetic data

The second type of data used in this research is synthetic data, which is generated based on industrial data analysis. The synthetic data generation scheme is illustrated in Fig. 3.3. The synthetic data also consists of the three variables present in the industrial data: voltage, current, and temperature. Voltage is simulated using an empirical model, with current and temperature as inputs into the model. The empirical model has a linear form with 4 or 5 coefficients (details can be found in Section 4.1.2 and 5.1.2). The coefficients vary over time, representing degradation. Noises are added to the variables to represent measurement error. The operation data (current and temperature), model, model coefficients, and noise are all generated based on industrial experience. However, for confidentiality reasons, they do not necessarily reflect the actual industrial values.

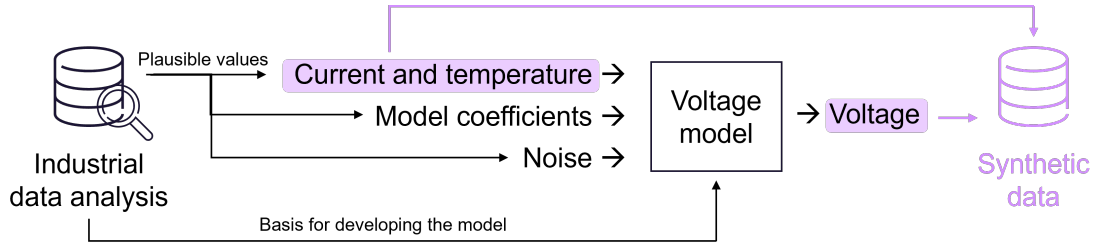


Figure 3.3: Synthetic data generation scheme.

The known coefficients can be used to calculate U_{rc} at each timestamp and under any reference condition, serving as the ground truth for validating the accuracy of the proposed methods. In addition to providing a known ground truth, synthetic data offers the possibility of simulating various operation profiles, degradation behaviors, fleet sizes, etc. This allows for a more comprehensive evaluation of the proposed methods' applicability and sensitivity.

Before using synthetic data for method validation, the fidelity of the synthetic data generation scheme is first verified. This is done by using the generation scheme to reproduce the industrial data. The quality of the reproduction is presented with a histogram of the error distribution. As mentioned in Section 3.2.1.1, error at a single-digit-mV level is considered satisfactory.

3.2.2 Evaluation metrics

3.2.2.1 Accuracy of U_{rc}

The accuracy of U_{rc} is the key metric for evaluating the proposed methods, as U_{rc} is the target quantity of this research - [SOH](#) for water electrolyzers.

With industrial data, the voltage is filtered at reference conditions within a narrow margin and the filtered voltages are regarded as the ground truth $U_{rc,true}$. Three reference conditions covering high, medium, and low current densities are considered. These conditions are selected in a way that abundant measurement data is available. The margins are set as small as possible to obtain precise $U_{rc,true}$ values, yet large enough to yield sufficient data. The resulting $U_{rc,true}$ have the same time resolution as the measurement data (i.e., minute resolution) and are sparsely available over time. Then, to match with the time resolution of the calculated U_{rc} from the proposed methods, $U_{rc,true}$ are resampled and gaps are filled with linear interpolation. Finally, the error is quantified using the root mean squared error ([RMSE](#)):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (U_{rc,true,i} - U_{rc,i})^2}, \quad (3.3)$$

where n denotes the number of U_{rc} data. Using [RMSE](#) as the error metric has the advantage that it has the same units as the target variable. For voltage, it enables easy comparison with the measurement error of the voltage sensor. As mentioned in Section [3.2.1.1](#), given that the voltage measurement error is below 10 mV, the [RMSE](#) of U_{rc} should ideally also be at the single-digit-mV level.

With synthetic data, the ground truth $U_{rc,true}$ is calculated with the model coefficients used to generate the data. The resulting $U_{rc,true}$ have the same time resolution as the synthetic data (i.e., minute resolution) and are continuously available over time. Then, same as above, $U_{rc,true}$ are resampled to match the time resolution of the calculated U_{rc} and the error is quantified with Eq. [3.3](#).

3.2.2.2 Accuracy of polarization curves

As mentioned in Section [2.2.2.1](#), polarization curve is a widely used metric to describe the performance of an electrolyzer and it offers additional insights into degradation mechanisms. Therefore, polarization curves derived from the proposed methods are also examined.

The industrial data contains very few polarization curve test results. These tests are conducted by stepwise varying the electric current from the minimum operating level to beyond the nominal level. For confidentiality reasons, details on the test protocol cannot

be disclosed. For each polarization curve test, the most recently fitted voltage model within 3 days using the proposed method is retrieved. The operating data during the polarization curve test (e.g., current and temperature) is provided as input into the fitted model to calculate the voltage U_{pol} . The measured voltage during the polarization test is used as the ground truth $U_{\text{pol,true}}$ and compared with U_{pol} derived from the proposed method:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (U_{\text{pol,true},i} - U_{\text{pol},i})^2}, \quad (3.4)$$

where n denotes the number of data during the polarization test.

With synthetic data, polarization curves can be readily calculated for any time step with the model coefficients used for data generation. The model is supplied with different current densities, while other operating parameters (e.g., temperature) are kept constant. The voltage obtained is the ground truth $U_{\text{pol,true}}$. Providing the same operating data to the fitted voltage models from the proposed methods yields U_{pol} , which is then compared with $U_{\text{pol,true}}$ using Eq. 3.4.

Same as for U_{rc} , the **RMSE** of the voltage during polarization curve tests should also be comparable with the voltage measurement error, namely at the single-digit-mV level.

3.2.2.3 Accuracy of voltage forecast

The voltage model obtained from the proposed methods not only describes the as-is voltage behavior, but also can be used to predict the voltage performance in the near future. This forecast capability enables operators to optimize the power input to the electrolyzers, ensuring efficient hydrogen production.

With industrial or synthetic data, the measured or synthetic operating voltage in the next 1, 2, and 7 days is the ground truth $U_{\text{future,true}}$. The future operating current and temperature are provided as inputs to the fitted voltage models from the proposed methods to obtain voltage forecast U_{future} . The error is quantified with:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (U_{\text{future,true},i} - U_{\text{future},i})^2}, \quad (3.5)$$

where n is the number of data in the forecast horizon.

The **RMSE** of voltage forecast is calculated for all fitted voltage models, i.e. if voltage models are fitted for each day over a year, there will be 365 **RMSE** results. Box plots are used to depict the distribution of the **RMSE** results (for example in Fig. 4.7c). The box extends from the first quartile (Q1 or 25th percentile) to the third quartile (Q3 or

75th percentile) values of the data, with a line at the median (Q2 or 50th percentile). The whiskers extend from the edges of the box to no more than 1.5 times the interquartile range ($IQR = Q3 - Q1$), ending at the farthest data point within the interval. Outliers are plotted as separate dots [111].

The target **RMSE** for voltage forecast is also at the single-digit-mV level, aligning with the voltage measurement error.

3.2.3 Benchmark method

Besides error quantification by comparing with the ground truth, the proposed methods are also evaluated against a benchmark method. Given that both proposed methods are extensions of the **SLR** algorithm, it is used as the benchmark. **SLR** aims to solve a linear equation (Eq. 3.2) by minimizing the error [112]:

$$\hat{\beta} = \arg \min_{\beta} \|Y - X\beta\|^2, \quad (3.6)$$

where $\|\cdot\|$ denotes the Euclidean norm and the $\hat{\cdot}$ denotes that this is an estimation, not the real value. The closed-form solution is [112]:

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \quad (3.7)$$

In practice, the parameter β is fitted with the `curve_fit` function from the Python package SciPy [113]. With the fitted model, the evaluation metrics introduced in Section 3.2.2 can be applied. The resulting **RMSE** from this benchmark method is then compared with the **RMSE** from the proposed methods.

4

Method 1: SOH quantification with transfer linear regression

This chapter presents the [SOH](#) quantification method based on the [TLR](#) algorithm. The method is detailed in Section [4.1](#), then evaluated with both synthetic data (Section [4.2](#)) and industrial data (Section [4.3](#)). Limitations are discussed in Section [4.4](#). This chapter is largely based on the publication: Yan, X., Locci, C., Hiss, F., Nieße, A.: State-of-health estimation for industrial H₂ electrolyzers with transfer linear regression. *Energies* **17**(6), 1374 (2024). doi:[10.3390/en17061374](https://doi.org/10.3390/en17061374) [[12](#)].

4.1 ALGORITHM

4.1.1 Overview

As described in Section [3.1](#) and illustrated in Fig. [4.1](#), the method involves fitting a voltage model for each interval and using the fitted model to calculate U_{tc} , serving as the [SOH](#) indicator. Specifically, this method adopts an empirical voltage model, featuring a linear voltage-current relationship. The model is fitted with the [TLR](#) algorithm, which adjusts a pre-existing fitted voltage model based on new data in a sequential manner. Fleet knowledge is incorporated into the initial voltage model (e.g., a generic begin-of-life model) and considered in the [TLR](#) algorithm.

Next, the empirical voltage model will be detailed in Section [4.1.2](#) and the [TLR](#) algorithm in Section [4.1.3](#).

4.1.2 Voltage model

Since this research focuses on industrial electrolyzers, which are typically designed for a specific operation range, it is not necessary to use an intricate physics-based model. Instead, a model that effectively describes the voltage within the typical operation range is sufficient. Therefore, a simple linear empirical voltage model is adopted:

$$U = c_1 i + c_2 T + c_3 \ln(OH) + c_4 \quad (4.1)$$

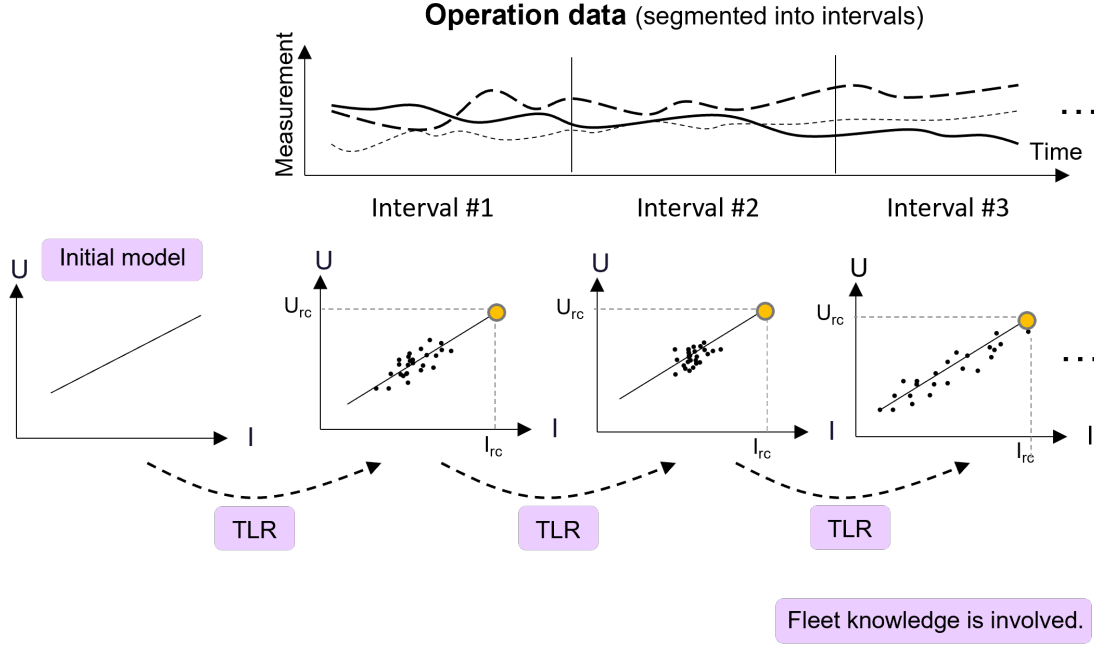


Figure 4.1: Overview of the SOH quantification method using TLR.

where voltage (U) is a linear function of current density (i), temperature (T), and the natural logarithm of operating hours since the last start (OH). Coefficients c_1 – c_4 are unknown and need to be fitted for each interval.

This model is developed by analyzing the operation data of an industrial electrolyzer described in Section 3.2.1.1. Its accuracy is demonstrated with the three exemplary intervals in Fig. 4.2, each with a different operating scheme: (a) constant operation with several starts, (b) volatile operation, and (c) staircase operation. These three intervals all have a duration of 1 day. As the data preprocessing step, data beyond the normal operation range (e.g., during standstill, data with extremely low or high current) are excluded. The proposed empirical model Eq. 4.1 is trained on 75% of the preprocessed data and RMSE is calculated on the remaining 25% of the data. The results show that the model can accurately describe the voltage with RMSE at single-digit-mV level, fulfilling the accuracy requirement defined in Section 3.2.

The model Eq. 4.1 can be written in the matrix form as in Section 3.1:

$$Y = X\beta + \epsilon. \quad (4.2)$$

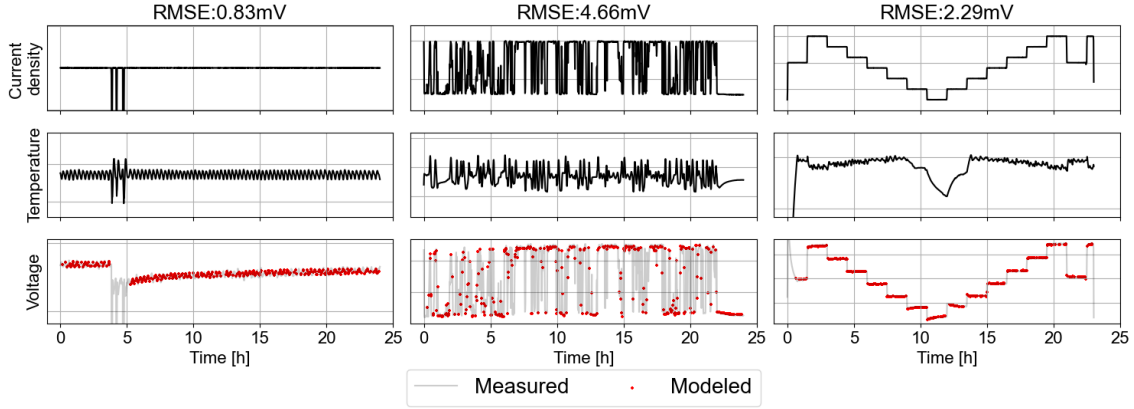


Figure 4.2: Validation of the empirical model Eq. 4.1 with operation data from three different operation schemes. For confidentiality reasons, the y-scales are hidden.

The dependent variable Y is a vector of the measured voltages

$$Y = [U_1 \ U_2 \ \dots \ U_n]^T. \quad (4.3)$$

The independent variable X is a matrix of the impact factors, where the column of ones corresponds to the constant term c_4

$$X = \begin{bmatrix} i_1 & T_1 & \ln(OH)_1 & 1 \\ i_2 & T_2 & \ln(OH)_2 & 1 \\ \dots & \dots & \dots & \dots \\ i_n & T_n & \ln(OH)_n & 1 \end{bmatrix}. \quad (4.4)$$

The parameter β is a vector of the unknown coefficients

$$\beta = [c_1 \ c_2 \ c_3 \ c_4]^T, \quad (4.5)$$

and ϵ is a vector of errors

$$\epsilon = [\epsilon_1 \ \epsilon_2 \ \dots \ \epsilon_n]^T. \quad (4.6)$$

The subscript n denotes the number of data points used for model fitting.

The parameter β is unknown and needs to be fitted. As the benchmark method, [SLR](#) is used to fit β as described in Section 3.2.3, where no fleet knowledge is incorporated in the model fitting process. Whereas to address model fitting difficulty under limited data coverage (as illustrated in Fig. 3.1), the [TLR](#) algorithm is developed to incorporate fleet

knowledge, which will be introduced in the next section. With the fitted coefficient β , the SOH indicator U_{rc} under the predefined reference condition X_{rc} can be calculated:

$$U_{rc} = X_{rc} \beta, \quad (4.7)$$

where

$$X_{rc} = \begin{bmatrix} i_{rc} & T_{rc} & \ln(OH_{rc}) & 1 \end{bmatrix}. \quad (4.8)$$

Additionally, the polarization curve can be obtained by calculating U_{rc} for various i_{rc} values while fixing T_{rc} and OH_{rc} .

4.1.3 Model fitting

The TLR algorithm is inspired by the data-enriched linear regression algorithm proposed in [114]. This is introduced first as the foundation for the TLR algorithm.

4.1.3.1 Background: data-enriched linear regression

The data-enriched linear regression algorithm was proposed by Chen et al. in 2015 [114]. It aims to build a linear regression model for a small dataset with the help of a second dataset that is large but possibly biased. This aim closely aligns with the purpose of this research: building a linear voltage model (Eq. 4.1) for intervals with a limited operation range (the small dataset) with the help of historical data that cover a larger operation range but have another degradation state (the second large but biased dataset).

As illustrated in Fig. 4.3a, the data-enriched linear regression algorithm assumes that the small dataset follows a linear regression model

$$Y_1 = X_1 \beta + \epsilon_1 \quad (4.9)$$

where X_1 and Y_1 are the independent and dependent variables, β is the linear coefficient, and ϵ_1 is the error. The second dataset follows the model

$$Y_2 = X_2(\beta + \gamma) + \epsilon_2 \quad (4.10)$$

where X_2 and Y_2 are the independent and dependent variables, ϵ_2 is the error, and the linear coefficients are shifted by γ . The parameter γ corresponds to the drift and rotation of the model.

The parameters β and γ are estimated by minimizing the loss function

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \arg \min_{\beta, \gamma} (\underbrace{\|Y_1 - X_1\beta\|^2}_{\epsilon_1} + \underbrace{\|Y_2 - X_2(\beta + \gamma)\|^2}_{\epsilon_2} + \underbrace{\lambda\|X_T\gamma\|^2}_{\text{control } \gamma}). \quad (4.11)$$

The first two terms in the loss function are the squared errors of Eq. 4.9 and Eq. 4.10. This is similar to the ordinary least squares regression technique, which estimates model coefficients by minimizing the fitting error. The third term controls the coefficient drift γ . The reason for controlling γ is to restrict the shape of the model.

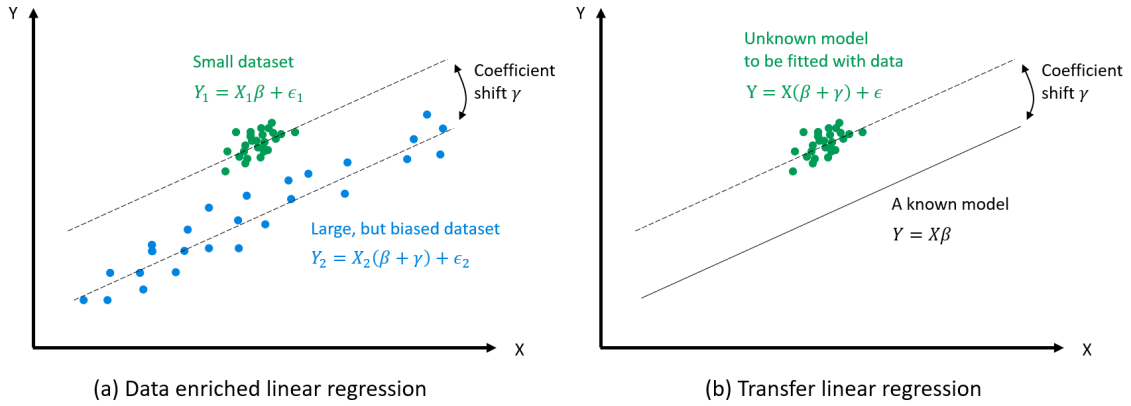


Figure 4.3: Illustration for (a) data-enriched linear regression algorithm and (b) transfer linear regression algorithm.

Two hyperparameters λ and X_T are required to control γ . $\lambda \in [0, \infty]$ controls the overall size of γ . Its impact can be illustrated with two extreme cases: Setting $\lambda = 0$ means no constraint on γ ; therefore, β and $\beta + \gamma$ are fitted *separately* with the two datasets; in contrast, setting $\lambda = \infty$ forces $\gamma = 0$, which is equivalent to fitting β with two datasets *jointly*. The second hyperparameter X_T can be used to control each value in the vector γ in detail, but the paper [114] does not provide concrete instructions on how to construct the X_T matrix.

As a result, the estimated parameters $\hat{\beta}$ and $\hat{\gamma}$ obtained by solving Eq. 4.11 are

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = (\chi^T \chi)^{-1} \chi^T \Upsilon \quad (4.12)$$

where $\chi = \begin{pmatrix} X_1 & 0 \\ X_2 & X_2 \\ 0 & \lambda^{\frac{1}{2}} X_T \end{pmatrix}$ and $\Upsilon = \begin{pmatrix} Y_1 \\ Y_2 \\ 0 \end{pmatrix}$ [114].

4.1.3.2 Transfer linear regression

Mathematical formulation The TLR algorithm is a modification of the data-enriched linear regression algorithm (Fig. 4.3b). The difference is that the data-enriched linear regression algorithm uses two datasets to estimate two parameters β and γ , whereas the TLR algorithm uses only one dataset to estimate γ , assuming that the parameter β is already known (e.g., fitted with historical data or derived from expert knowledge). This setup enables the inclusion of an existing model without requiring a second database and is computationally cheaper by handling smaller data matrices.

The TLR algorithm aims to estimate the parameter γ in the equation

$$Y = X(\beta + \gamma) + \epsilon, \quad (4.13)$$

where X and Y are the data from the target dataset, and β is already known. Analogous to the data-enriched linear regression algorithm, the unknown parameter λ can be estimated by minimizing the loss function

$$\hat{\gamma} = \arg \min_{\gamma} (\underbrace{\|Y - X(\beta + \gamma)\|^2}_{\epsilon} + \underbrace{\lambda \|X_T \gamma\|^2}_{\text{control } \gamma}). \quad (4.14)$$

The hyperparameters λ and X_T control the size of γ in the same manner as the data-enriched linear regression. Their impacts will be detailed later. Solving Eq. 4.14 gives

$$\hat{\gamma} = (X^T X + \lambda X_T^T X_T)^{-1} X^T (Y - X\beta). \quad (4.15)$$

The derivation process to obtain Eq. 4.15 is provided in Appendix A.1.

In summary, the **TLR** algorithm is proposed for **SOH** quantification due to the following reasons.

- It can capture the model drift caused by degradation.
- It tackles the problem of limited data coverage.
- It is suitable for a linear model.
- It is computationally efficient by not including a second dataset.
- Its model transfer mechanism is easy to interpret.

Application on time series data To realize continuous **SOH** estimation along the lifetime of electrolyzers, the **TLR** algorithm is applied iteratively over a time series to continuously update the fitted empirical voltage model. Algorithm 1 shows a minimum example of a time series application. In addition to this minimal example, several modi-

fications can be implemented in practice. For instance, one can set λ and/or X_T flexibly for each interval instead of fixing a global value for all intervals; one can also add a fitting quality check for each interval—if the fitting accuracy is not satisfactory, one can use $\beta_i = \beta_{i-1}$ instead of $\beta_i = \beta_{i-1} + \gamma_i$ (i.e., not updating the coefficients).

Algorithm 1 Apply TLR on time series data

Input: Initial coefficients β_0 , hyperparameters λ and X_T , time series data X and Y segmented into n intervals $(X_1, Y_1), \dots (X_n, Y_n)$

Output: Coefficients for all n intervals $\beta_1, \dots \beta_n$

for $i = 1$ to n **do**

$$\gamma_i = \arg \min_{\gamma} (\|Y_i - X_i(\beta_{i-1} + \gamma)\|^2 + \lambda \|X_T \gamma\|^2) \quad (4.16)$$

$$= (X_i^T X_i + \lambda X_T^T X_T)^{-1} X_i^T (Y_i - X_i \beta_{i-1}) \quad (4.17)$$

$$\beta_i = \beta_{i-1} + \gamma_i \quad (4.18)$$

end for

Setting initial coefficients β_0 Algorithm 1 requires an initial coefficient β_0 . There are different methods to obtain β_0 : It can be derived from the initial time series data, data from other similar systems, or expert knowledge. For the method evaluation with industrial data (Section 4.3), β_0 is derived from a SLR fit using the initial 30 days of data from all electrolyzers stacks described in Section 3.2.1.1, to represent fleet knowledge of the initial model.

Impacts of hyperparameters λ and X_T Hyperparameters λ and X_T need to be set before model fitting. Same as in the data-enriched linear regression algorithm [114], λ is a constant that controls the overall size of γ . Its impact can be seen from the following examples of setting extreme values for λ (also summarized in Table 4.1):

- Setting $\lambda = \infty$ forces $\gamma = 0$. This leads to $\beta_i = \beta_{i-1}$; that is, the linear coefficients do not change over time. In this case, high fitting errors along the time series are expected.
- In contrast, setting $\lambda = 0$ implies no constraint on the size of γ . Equation 4.16 becomes

$$\gamma_i = \arg \min_{\gamma} (\|Y_i - X_i(\beta_{i-1} + \gamma)\|^2);$$

that is, γ_i is estimated freely by minimizing the fitting error. The freedom of γ_i leads to the freedom of β_i (Eq. 4.18), meaning that β_i is not influenced by the previous

coefficient β_{i-1} . This is equivalent to fitting the model independently for each interval.

In this case, fluctuating β and low fitting errors along the time series are expected.

Regarding X_T , the original paper on data-enriched linear regression [114] does not provide concrete instructions on how to construct X_T . For TLR, X_T is designed as a diagonal matrix with dimensions $p \times p$ (p is the dimension of β and γ):

$$\begin{pmatrix} X_{T1} & & \\ & \ddots & \\ & & X_{Tp} \end{pmatrix}.$$

The term $X_T\gamma$ in Eq. 4.14 and Eq. 4.16 is then

$$\begin{pmatrix} X_{T1}\gamma_1 \\ \vdots \\ X_{Tp}\gamma_p \end{pmatrix}.$$

This shows that the diagonal elements $X_{T1} - X_{Tp}$ can control the individual values $\gamma_1 - \gamma_p$ in the vector γ separately. The impact of the diagonal elements X_{Tp} is similar to that of λ (Table 4.1).

Table 4.1: Impacts of setting extreme values for hyperparameters λ and X_T .

Impact on	Values of λ or X_{Tp}	0	∞
Coefficient shift γ		Not constrained	Shrink to 0
Coefficient β_i		β_i is flexibly fitted with data in interval i , not influenced by β_{i-1} .	$\beta_i = \beta_{i-1}$
Trend of β overtime		Fluctuating	Constant
Fitting error		Low	High

Table 4.1 summarizes the impacts of setting extreme values for λ and X_T . It shows that setting small values for λ and X_T results in a low fitting error but fluctuating coefficients, whereas setting high values has the opposite impact. Balancing the trade-off between the fitting accuracy and coefficient stability is key to setting proper values for λ and X_T . This is introduced next.

Setting λ As explained above, setting λ involves a trade-off between the fitting accuracy and coefficient stability. In practice, this trade-off can be observed by testing different λ s

on multiple intervals. This is demonstrated in Section 4.2 and 4.3.

Setting diagonal elements in X_T Recall that X_T is a $p \times p$ diagonal matrix, and p is the dimension of β and γ , namely, the number of variables. Table 4.1 shows that for any variable p , setting $X_{Tp} = \infty$ leads to $\beta_{i,p} = \beta_{i-1,p}$, i.e., the linear coefficient stays constant along the time series; whereas setting $X_{Tp} = 0$ leads to flexible β_p . That means, if the coefficient for variable p is allowed to have a large variation, the corresponding X_{Tp} should be small. Large variations should be allowed when

- (a) the coefficient β_p varies largely by nature over time, and
- (b) the variable p covers a wide range during an interval. (In this case, β_p can be easily identified because of the wide data coverage, so it is allowed to be fitted flexibly with the data. On the contrary, if the variable p covers only a small range in the interval i , $\beta_{i,p}$ is constrained toward the previously fitted coefficient $\beta_{i-1,p}$).

Applying these principles in the electrolyzer context, the X_T matrix is designed as in Algorithm 2.

Algorithm 2 Setting diagonal elements in X_T for electrolyzers

1. Following (a): Fit the model (Eq. 4.1) for all n intervals, to obtain n coefficients $c_{p,1} \dots c_{p,n}$ for each variable p , where $p = 1, 2, 3$ for the voltage model Eq. 4.1. The variations of $c_{p,1} \dots c_{p,n}$ represent the expected variation of β_p through time. Calculate the 1st term for X_{Tp} , which decreases with the expected variation of β_p over time:

$$\text{1st term} = \frac{1}{75\text{th-25th percentile of } c_{p,n}}$$

2. Following (b): For each interval i and each variable p , calculate the 2nd term for X_{Tp} , which decreases with the data coverage of X_p :

$$\text{2nd term} = \frac{\text{Normal operation range of variable } p}{75\text{th-25th percentile of } X_p}$$

3. For each interval i and each variable p ,

$$X_{Tp,i} = \text{1st term} \times \text{2nd term}$$

4.2 METHOD EVALUATION WITH SYNTHETIC DATA

As described in Section 3.2, the TLR algorithm is evaluated with both synthetic data and actual operation data from industrial electrolyzers. In both cases, the TLR algorithm is

compared with the benchmark method [SLR](#), to highlight the necessity and effectiveness of transfer learning.

In this section, synthetic data is employed to validate the proposed method, as the ground truth is known for every timestamp, enabling the quantification of the method's accuracy.

The synthetic data is generated following the scheme described in Section [3.2.1.2](#), with the empirical voltage model Eq. [4.1](#). To assess the fidelity of the synthetic data generation scheme, it is first used to reproduce the industrial data (used in Section [4.3](#)). Due to the model simplification and unknown model coefficients, reproduction error cannot be avoided. The error distribution is shown in Fig. [4.4](#). The average error is -0.1 mV, and 95.9% of the data has an error within ± 10 mV - the same magnitude as the measurement noise (Section [3.2.1.1](#)). Thus, the synthetic data generation scheme is considered satisfactory.

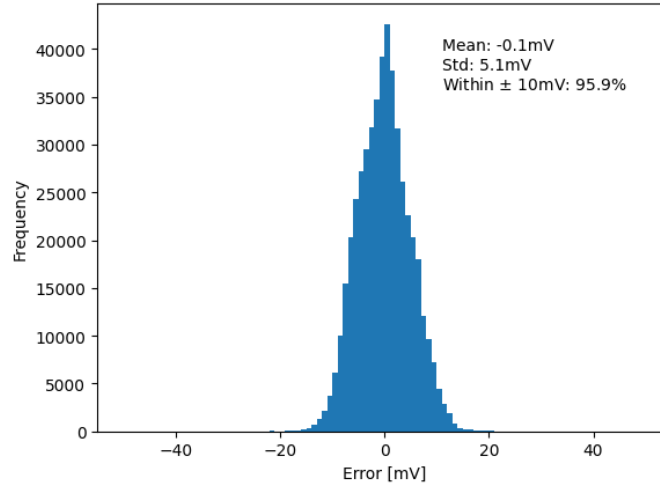


Figure 4.4: Error of using the synthetic data generation scheme to reproduce industrial data.

Then, the following settings are incorporated into the synthetic data generation scheme to produce synthetic data. These settings are inspired by industrial operation data but do not necessarily represent actual values.

- Resolution: 1 min.
- Length: 1 year.
- Current density range: $0.4 - 2$ A/cm².
- Temperature: Sine wave between 58 and 60°C. Decrease to 20°C during standstill.
- Noise: Current density $\pm 1\%$, temperature $\pm 0.1^\circ\text{C}$, voltage ± 1 mV.
- Operation modes: The entire time frame is randomly segmented into periods with

one of the following operation modes.

- Low variation: A random current density $\pm 1\%$ variation.
- Medium variation: A random current density $\pm 10\%$ variation.
- High variation: Random current densities within the full operating range.
- Standstill: Zero current density.
- Degradation behavior is described with model coefficients $c_1 - c_4$ evolving over time:
 - c_1 : Linearly increases from 0.2 to 0.3.
 - c_2 : Linearly decreases -0.003 to -0.0036.
 - c_3 : Stays constant at 0.003.
 - c_4 : Linearly increases from 1.6 to 1.65.

The resulting synthetic operation data is shown in Fig. 4.5.

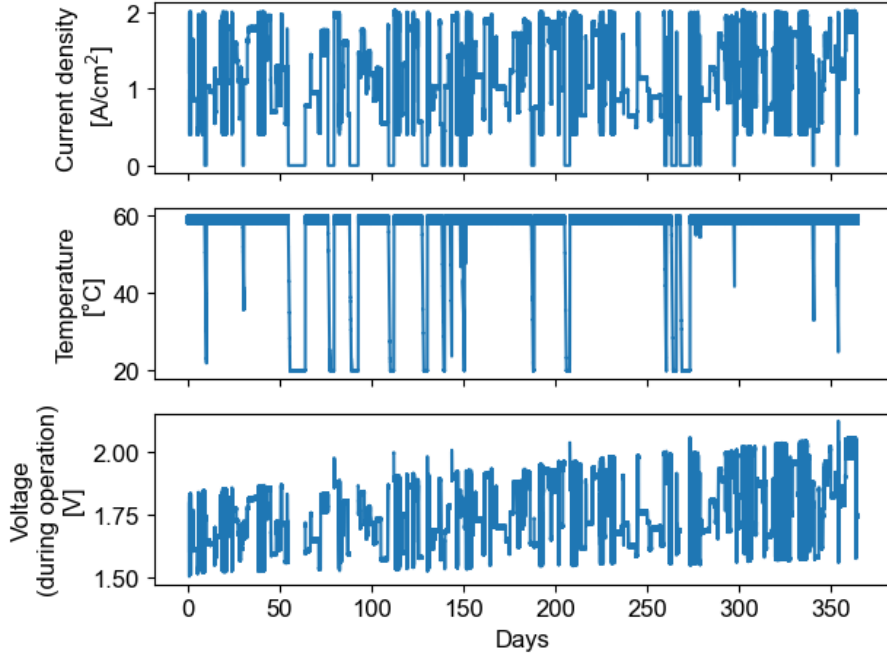


Figure 4.5: Synthetic data.

The [TLR](#) algorithm is executed with the following settings.

- Model fitting interval: 1 day.
- Reference condition to calculate U_{rc} : $i_{rc} = 2\text{A/cm}^2$, $T_{rc} = 60^\circ\text{C}$, and $OH_{rc} = 72\text{h}$.
- Parameters required by the [TLR](#) algorithm (Algorithm 1):
 - Initial model coefficient β_0 . Derived from plain linear regression with the initial 10% of the synthetic data.
 - Hyperparameter X_T . Following Algorithm 2.

- Hyperparameter λ . Selected by testing various λ s to find a trade-off between fitting accuracy and coefficient stability. Eleven different λ values ranging from 10^{-8} to 10^2 (Fig. 4.6) are tested on the first 50 days of synthetic data. For each day, a model fitting with TLR is performed, yielding 50 sets of coefficients c_1 to c_4 and 50 fitting errors RMSE. By calculating the standard deviations of each coefficient and the average RMSE, the impact of the various λ s can be observed: As λ increases, the coefficients get more stable (the standard deviations decrease) and the fitting error increases, which aligns with the expectations in Table 4.1. Ultimately, $\lambda = 10^{-2}$ is selected as a trade-off between coefficient stability and model accuracy.

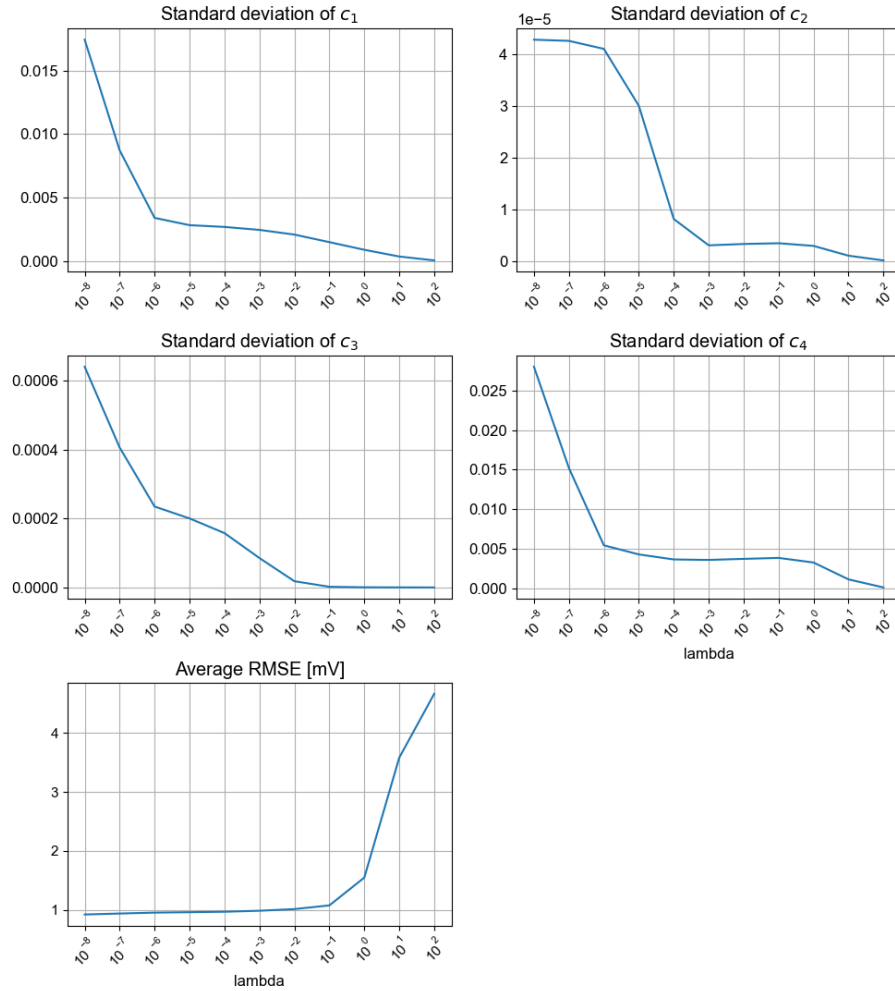


Figure 4.6: Different λ s ranging from 10^{-8} to 10^2 are tested for their impact on the coefficient stability (represented with the standard deviations of c_1 to c_4) and on the fitting error.

Using the selected λ , the **TLR** process (Algorithm 1) is performed. The results are compared with the benchmark method **SLR** (Section 3.2) that does not incorporate fleet knowledge in the model fitting process. Fig. 4.7a shows that the **TLR** algorithm captures the true U_{rc} with a significantly smaller **RMSE** than **SLR**. Regarding the accuracy of polarization curves (generated by applying the fitted voltage model Eq. 4.1 on various current densities between 0.3 A/cm^2 and 2.0 A/cm^2 , while keeping $T = 60^\circ\text{C}$ and $OH = 72\text{h}$), Fig. 4.7b shows that incorporating fleet knowledge with the **TLR** algorithm significantly stabilizes the polarization curves, showing a consistent rise over time. Quantitatively, the **RMSE** is reduced from 20.9 mV to 0.9 mV . Last, Fig. 4.7c shows that voltage models fitted with **TLR** can forecast voltage more accurately than with **SLR**. The superiority is particularly substantial for multi-days ahead forecast: For 7-days ahead forecast, **SLR** produces errors of up to 10^3 mV , while **SLR** produces errors consistently below 10 mV . These comparisons strongly underscore the necessity and effectiveness of the **TLR** algorithm.

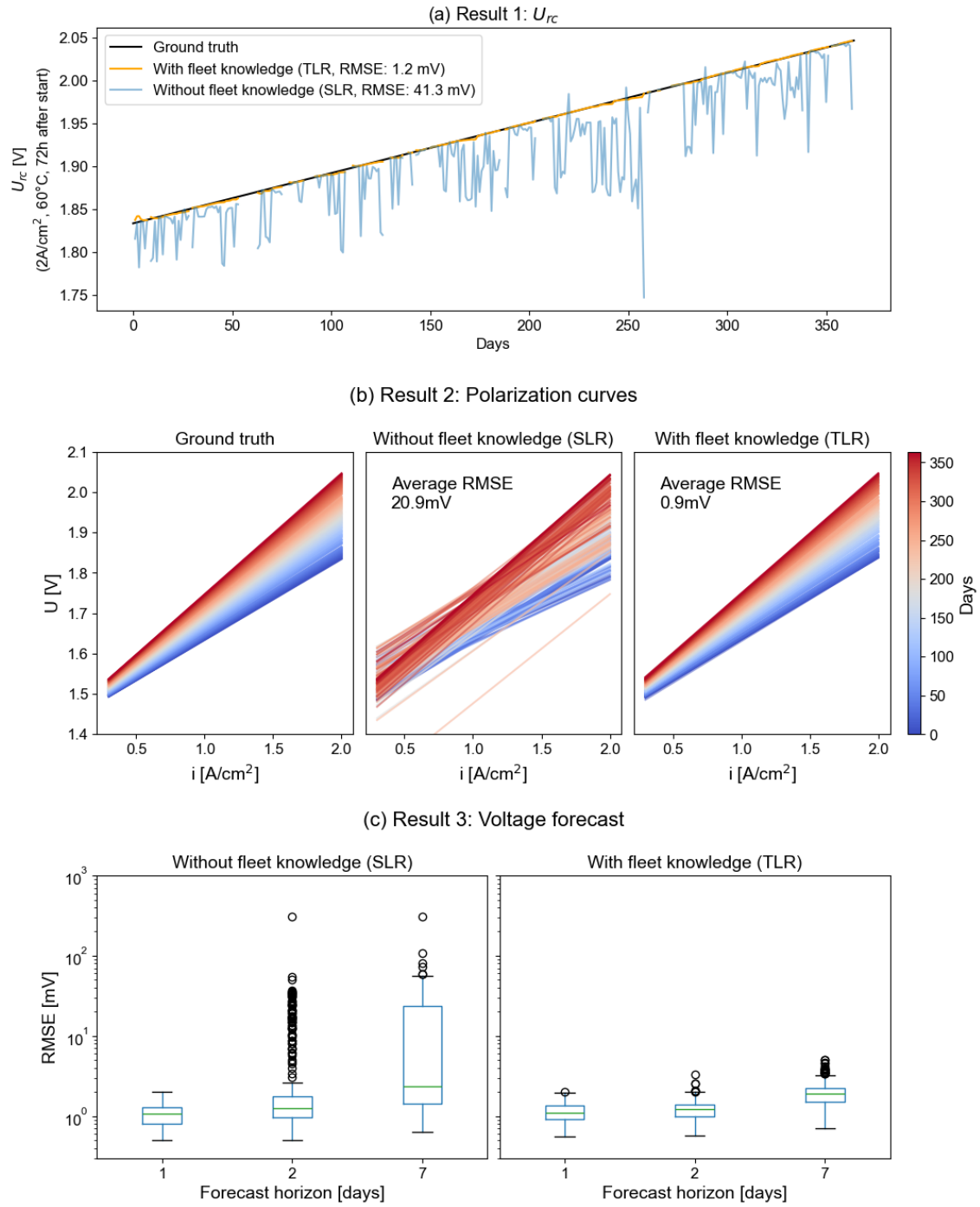


Figure 4.7: **TLR** captures ground truth with significantly smaller errors than the benchmark method **SLR**.

4.3 METHOD EVALUATION WITH INDUSTRIAL DATA

Next, the [TLR](#) algorithm is applied to the operation data of one industrial [PEM](#) water electrolyzer stack described in Section [3.2.1.1](#). It is executed with the following settings.

- Model fitting interval: 1 day.
- Reference condition to calculate U_{rc} : Three different reference conditions. One is the nominal operating condition, and two are operation conditions that occur relatively often with abundant data to facilitate model validation. For confidentiality reasons, the exact conditions cannot be provided.
- Parameters required by the [TLR](#) algorithm (Algorithm [1](#)):
 - Initial model coefficient β_0 : Obtained by plain linear regression fitting using the initial 30 days of data from all electrolyzers stacks described in Section [3.2.1.1](#).
 - Hyperparameter X_T : Following Algorithm [2](#).
 - Hyperparameter λ : Selected by testing various λ s to find a trade-off between fitting accuracy and coefficient stability. Eleven different λ values ranging from 10^{-8} to 10^2 (Fig. [4.8](#)) are tested on the first 50 days of synthetic data. For each day, a model fitting with [TLR](#) is performed, yielding 50 sets of coefficients c_1 to c_4 and 50 fitting errors [RMSE](#). By calculating the standard deviations of each coefficient and the average [RMSE](#), the impact of the various λ s can be observed: As λ increases, the coefficients get more stable (the standard deviations decrease), and the fitting error increases, which aligns with the expectations in Table [4.1](#). Ultimately, $\lambda = 10^{-2}$ is selected as a trade-off between coefficient stability and model accuracy. (A deep dive into the challenge of λ selection can be found in Section [4.4.2](#).)

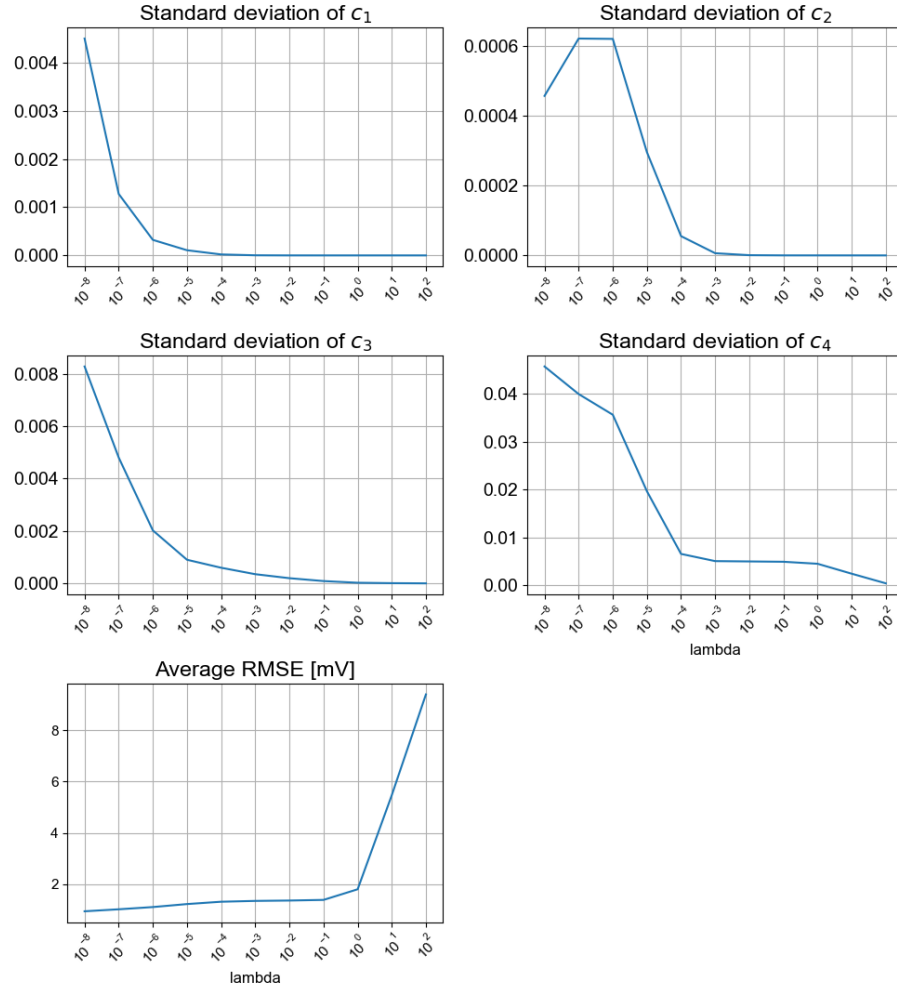


Figure 4.8: Different λ s ranging from 10^{-8} to 10^2 are tested for their impact on the coefficient stability (represented with the standard deviations of c_1 to c_4) and on the fitting error.

Using the selected λ , the TLR process (Algorithm 1) is performed on approximately 2 years of operation data (Fig. 4.9, right). U_{rc} under 3 different reference conditions are calculated and compared with the measured voltage at roughly these reference conditions. As the benchmark method, SLR is also applied on the industrial data, where no fleet knowledge is involved (Fig. 4.9, left). Looking at the measured voltages at the reference conditions, they are only sparsely available (especially at reference condition 1) because the operating conditions are arbitrary and rarely at these reference conditions. This underscores the necessity of the SOH indicator U_{rc} for continuous condition monitoring. Furthermore, upon examining the calculated U_{rc} , it can be observed that SLR produces noisy and

unstable U_{rc} results, which contradicts the gradual degradation process; in contrast, **TLR** produces smoother results with a consistent increasing trend over time. The accuracy of the calculated U_{rc} is quantified with **RMSE** (Section 3.2.2.1). The **RMSE** of U_{rc} computed with **TLR** is at the same magnitude as the voltage measurement error (Section 3.2.1.1) and they are much smaller than with **SLR** at all three reference conditions. This signifies the effectiveness of incorporating fleet knowledge using **TLR**.

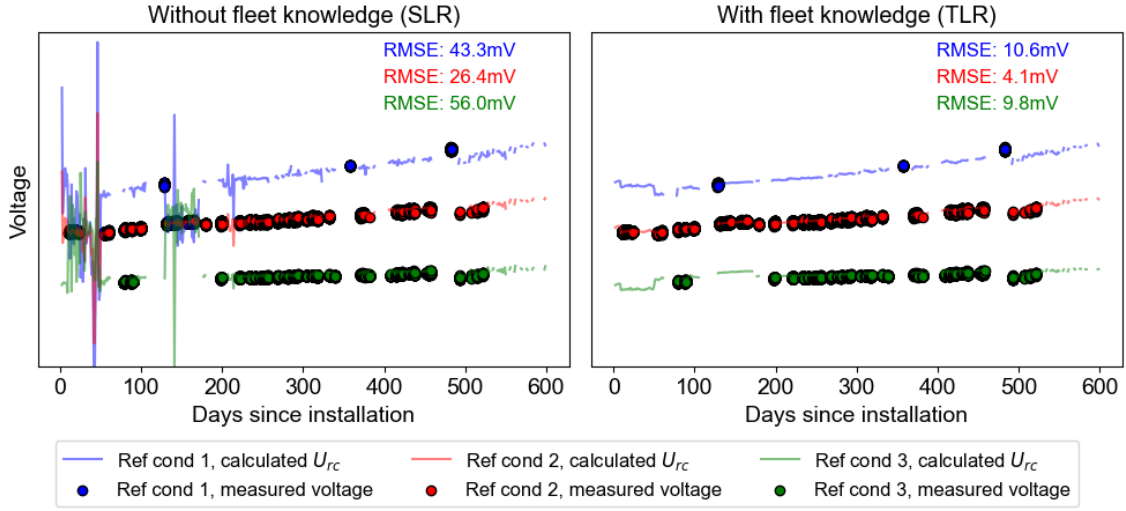


Figure 4.9: U_{rc} calculated with **SLR** (left) and **TLR** (right) under 3 different reference conditions, compared with measured voltage at roughly these conditions. For confidentiality reasons, the y-scales are hidden.

The superiority of the **TLR** algorithm can also be examined with the fitted polarization curves (Fig. 4.10). The polarization curves are plotted by applying the fitted voltage model Eq. 4.1 on various current densities while keeping constant T and OH . Fig. 4.10 shows that **SLR** produces polarization curves with inconsistent shapes, whereas incorporating fleet knowledge with the **TLR** algorithm notably stabilizes the curves, showing a steady rise over time, which aligns with the physical expectation of voltage degradation.

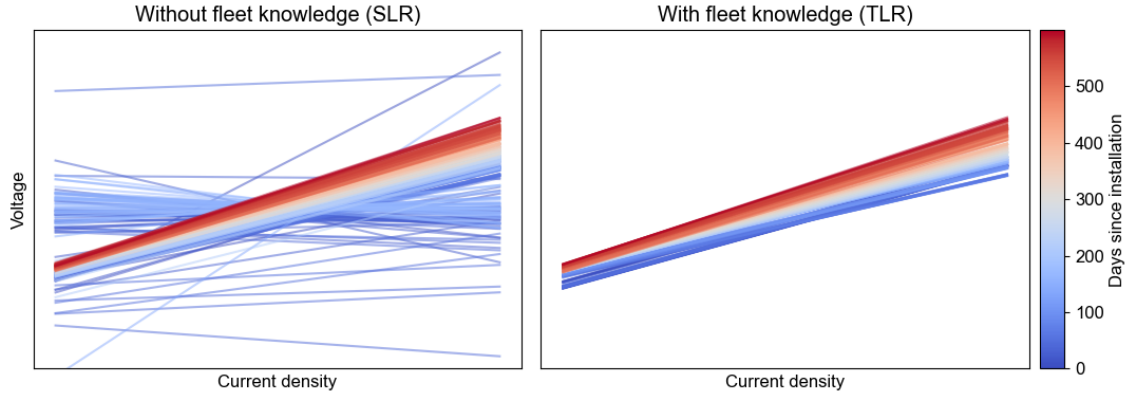


Figure 4.10: [TLR](#) produces more stable polarization curves than the [SLR](#) algorithm. For confidentiality reasons, the x- and y-scales are hidden.

The [TLR](#) algorithm is also evaluated regarding the forecast capability of the fitted voltage models and is compared with the benchmark method [SLR](#) (Fig. 4.11). [RMSE](#) is quantified as introduced in Section 3.2.2.3. For both [SLR](#) and [TLR](#), the forecast error increases as the forecast horizon extends from 1 day to 7 days, due to the changing degradation state over time. The voltage models fitted with [SLR](#) produce errors of up to 10^3 mV, whereas with [TLR](#), the errors consistently remain below 100 mV, with the majority below 10 mV. The error reduction effect of [TLR](#) is particularly pronounced for the outliers, showing the robustness of the algorithm.

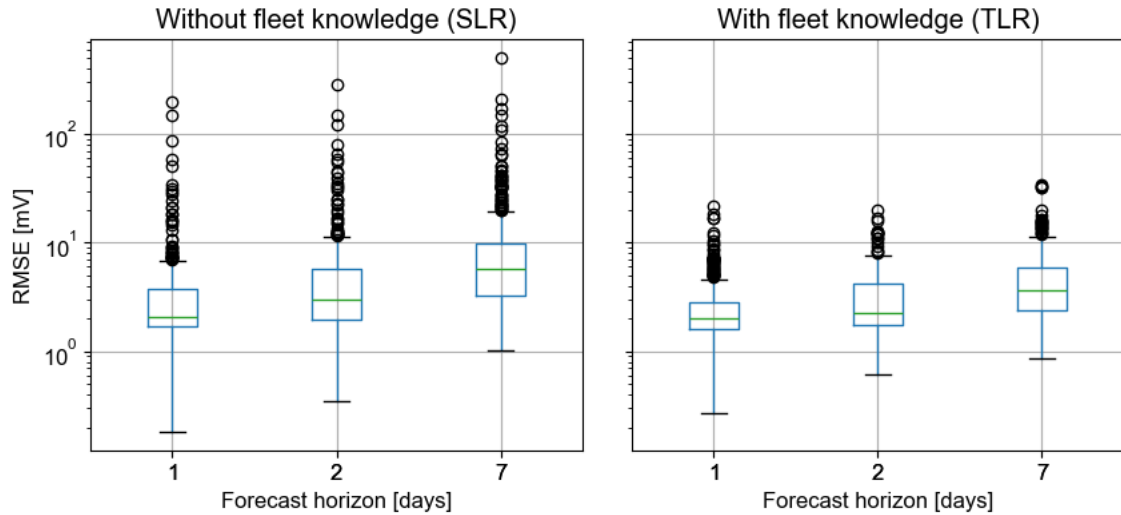


Figure 4.11: Voltage forecast accuracy of the voltage models fitted with [SLR](#) and [TLR](#).

4.4 LIMITATIONS

4.4.1 Elevated error for degraded electrolyzers

The empirical voltage model Eq. 4.1 characterizes a linear voltage-current relationship and is proven to be valid for the industrial operation data with Fig. 4.2. The calculated U_{rc} is also accurate for 600 days (Fig. 4.9). However, as more operation data is gathered during the research, it is observed that the model is no longer accurate for degraded electrolyzers, particularly exhibiting bias at high current densities.

As shown in Fig. 4.12, by extending the evaluation period from 600 days to 1200 days, it becomes evident that the calculated U_{rc} consistently underestimates the measured voltage at reference condition 1, with an RMSE of 64.7 mV. To investigate the cause, the measured polarization curves are compared with the modeled results (Fig. 4.13). RMSE is quantified as introduced in Section 3.2.2.2. The comparison reveals that the voltage-current relationship of the polarization curve measurement at 723 days and 884 days after installation is not anymore linear, where the voltages at higher current levels show an upward bend. To address this, the second proposed method (Section 5) updated the empirical voltage model to capture this non-linear voltage-current behavior.

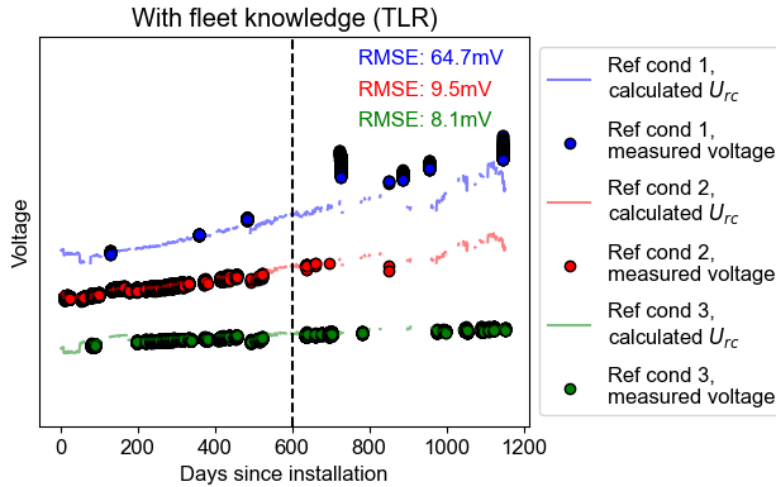


Figure 4.12: U_{rc} calculated with TLR under 3 different reference conditions, compared with measured voltage at roughly these conditions, over roughly 1200 days.

4.4.2 Difficulty in determining hyperparameters

Fig. 4.6 and 4.8 demonstrate the rationale behind the selection of the hyperparameter λ , which involves assessing coefficient stability and fitting accuracy. Nevertheless, this

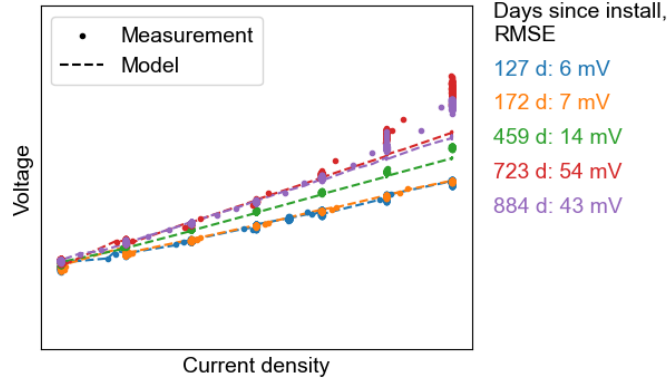


Figure 4.13: Fitted voltage model compared with measured polarization curves.

empirical approach does not guarantee the accuracy of the resulting coefficients. Due to the lack of knowledge of the true coefficients in practice, it is impossible to ensure the correctness of the selected hyperparameter.

Fig. 4.14 showcases the difficulty in determining the hyperparameter λ . Three different λ s 10^{-5} , 10^{-4} , and 10^{-3} are tested on 50 randomly selected days of industrial data. The model coefficient c_2 shows very different trends over time: from almost constant with $\lambda = 10^{-3}$, to varying between 80% to 150% with $\lambda = 10^{-5}$. Despite the different trends, the average fitting errors across the 50 days are only minimally affected, with a narrow range between 2.4 mV and 2.7 mV. Without knowing the true coefficients, it is hard to determine which λ is more appropriate. To simplify the parameter setting, the second proposed method (Section 5) is based on an established algorithm with standard parameters that do not require customization.

4.4.3 Rely on the last fitted model

The essence of the TLR algorithm is to adjust an existing fitted voltage model with newly available data. Simply speaking, if the new data covers a wide range, the model coefficients can be re-fitted with little constraint; on the contrary, if the new data have a limited coverage, the model coefficients are fixed toward the existing coefficients. This behavior is controlled with the hyperparameter X_T (Section 4.1.3.2). As a result, if limited data coverage persists over an extended period, such as during constant operation, the coefficients fitted with the TLR algorithm may eventually drift from the true value.

Fig. 4.15 demonstrates the effect of a long constant operation period using synthetic data. The data is generated similarly as described in Section 4.2, with an intentionally added constant operation period of 64 days. Due to the limited current density coverage,

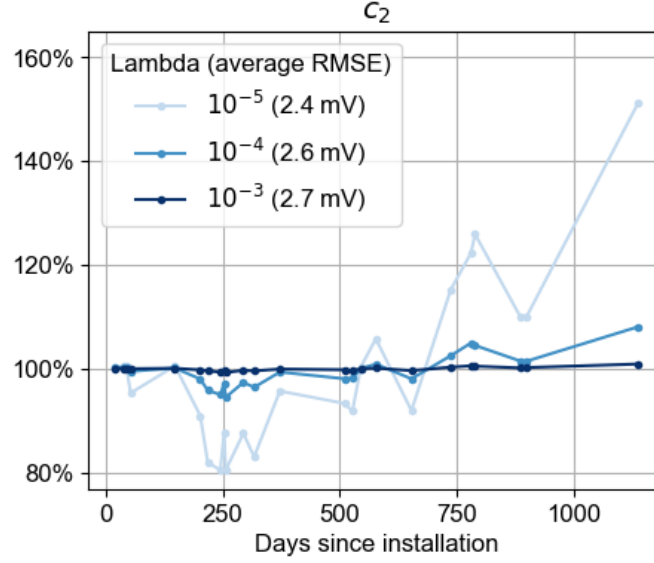


Figure 4.14: The impact of λ on the model coefficient c_2 and fitting error.

the coefficient c_1 , which describes the voltage-current relationship, drifts away from the ground truth and remains close to the last fitted c_1 . Consequently, the voltage cannot be accurately extrapolated to the reference condition, and the bias gets larger as the constant operation continues.

To overcome this limitation, the second proposed method (Section 5) incorporates prior knowledge of the model coefficients into the fitting process, which is derived from the overall degradation trend of similar electrolyzers. The prior knowledge provides additional information that enables the model coefficients to be updated even during extended periods of limited data coverage.

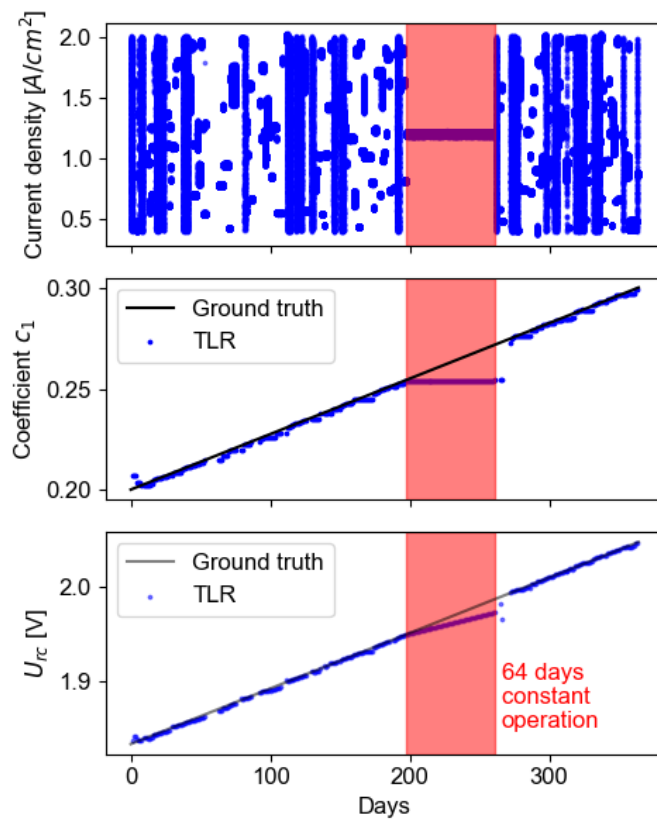


Figure 4.15: Impact of long constant operation period.

5

Method 2: SOH quantification with Bayesian inference

This chapter presents the [SOH](#) quantification method using Bayesian inference. The method is detailed in [Section 5.1](#), then evaluated with both synthetic data ([Section 5.2](#)) and industrial data ([Section 5.3](#)). Limitations are discussed in [Section 5.4](#). This chapter is largely based on the publication: Yan, X., Helmers, L., Zhou, K., Nieße, A.: Fleet-based degradation state quantification for industrial water electrolyzers. *Electrochemical Science Advances* 5(3), 70002 (2025). doi:[10.1002/elsa.70002](#) [[13](#)].

5.1 ALGORITHM

5.1.1 Overview

The underlying principle of this method is the same as the first method: It fits an empirical voltage model for each interval and uses the fitted model to calculate the [SOH](#) indicator U_{rc} ([Fig. 5.1](#)). To facilitate model fitting with limited data coverage, fleet knowledge of how the model parameters develop over time is incorporated with Bayesian inference. The fleet knowledge is derived by aggregating the model parameters from multiple similar electrolyzers, i.e. the electrolyzer fleet.

Next, the empirical voltage model will be detailed in [Section 5.1.2](#) and the model fitting method in [Section 5.1.3](#).

5.1.2 Voltage model

As pointed out in [Section 4.4.1](#), the previous empirical voltage model ([Eq. 4.1](#)) is not able to capture the voltage behavior of degraded electrolyzers. Therefore, the voltage model is updated to incorporate a stronger physical basis and to account for the non-linear voltage-current behavior that occurs in degraded electrolyzers. The updated voltage model follows this structure:

$$\text{Voltage} = \text{Reversible voltage} + \text{Overpotentials (impact of current, temperature, and start)}.$$

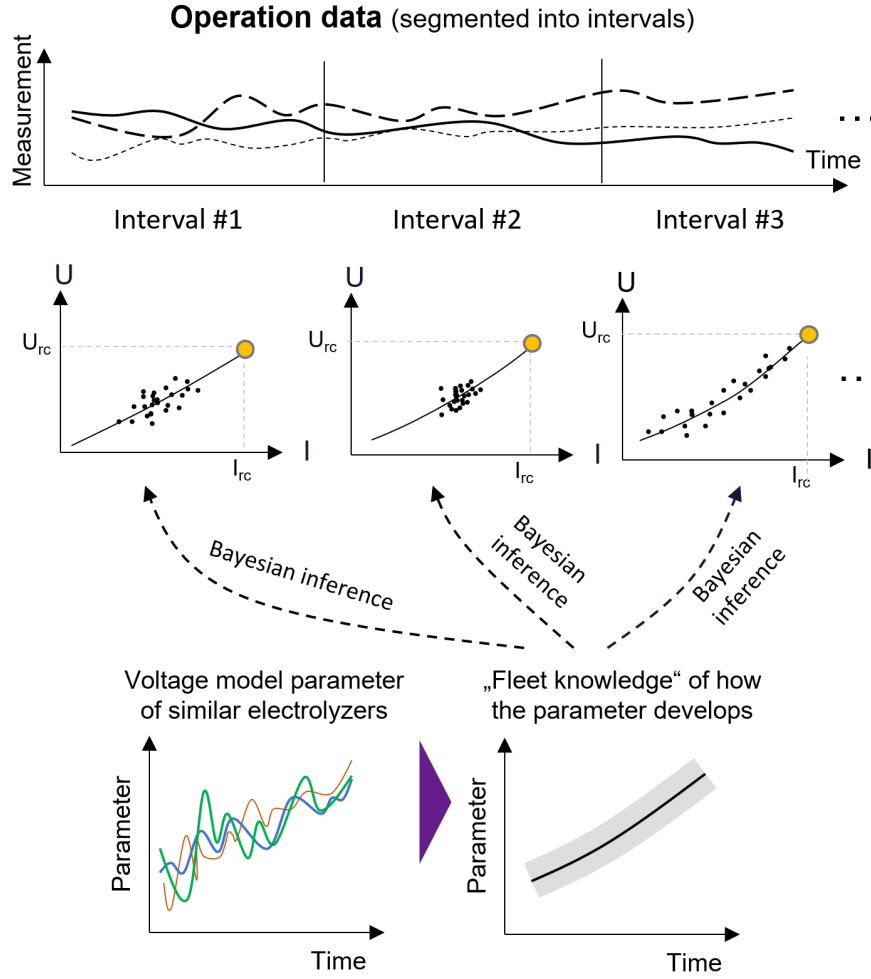


Figure 5.1: Overview of the SOH quantification methods with Bayesian inference

As introduced in Section 2.2.2.1, reversible voltage (U_0) is the minimum voltage required to split 1 mole of water. This physical quantity can be accurately calculated with the Python package CoolProp [115]. But for fast computation, it is approximated with a linear function

$$U_0 = -8.2975 \times 10^{-4}T + 1.24965, \quad (5.1)$$

where T is in $^{\circ}\text{C}$ and P is set at the atmospheric pressure. Fig. 5.2 shows that the approximation aligns well with the CoolProp calculation within the typical operation temperature range of 50 to 60 $^{\circ}\text{C}$. The mean absolute error is 0.002 mV.

The overpotentials are modeled empirically considering the impacts of current, temper-

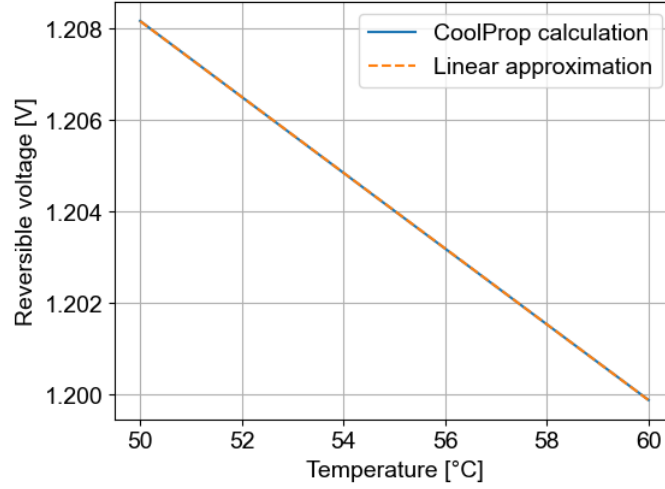


Figure 5.2: Comparison between the reversible voltage calculated with the Python package CoolProp and with the linear approximation (Eq. 5.1).

ature, and starts:

$$U_{\text{overpotentials}} = c_1 i + c_2 iT + c_3 \ln(OH) + c_4 i^2 + c_5 \quad (5.2)$$

where i is the electric current density in A/cm², T is temperature in °C, and OH stands for operating hours since last start. The coefficients c_1 to c_5 are unknown and need to be fitted with operation data. The variables included in the proposed overpotential model are selected for the following reasons.

- i : The investigated electrolyzers are operated mostly under medium current density, where voltage and current show a linear relation (quasi-ohmic behavior).
- iT : This term is to capture the impact of temperature, which is dependent on the current level.
- $\ln(OH)$: This term is to describe the impact of starts (i.e., the electric current supply rises from below the minimum operating current to the normal range). Similar as reported in [57, 59, 83], the investigated electrolyzers typically show a voltage increase after each start, with a pattern resembling a logarithmic curve.
- i^2 : This term is to capture the nonlinear voltage-current relation observed with degraded electrolyzers as described in Section 4.4.1.

Combine Eq. 5.1 and 5.2, the final empirical voltage model is

$$U = (-8.2975 \times 10^{-4} T + 1.24965) + c_1 i + c_2 iT + c_3 \ln(OH) + c_4 i^2 + c_5. \quad (5.3)$$

This model is validated with industrial operation data (described in Section 3.2.1.1) from different operation schemes. Three examples are shown in Fig. 5.3: (1) constant operation with several starts, (2) volatile operation, and (3) staircase operation. These example intervals are 24 hours long. As the data preprocessing step, data beyond the normal operation range (e.g., during standstill) are excluded. The proposed empirical model Eq. 5.3 is trained on 75% of the preprocessed data and RMSE is calculated on the remaining 25% of the data. The results show that the model can accurately describe the voltage with RMSE at a single-digit-mV level, fulfilling the required accuracy defined in Section 3.2.

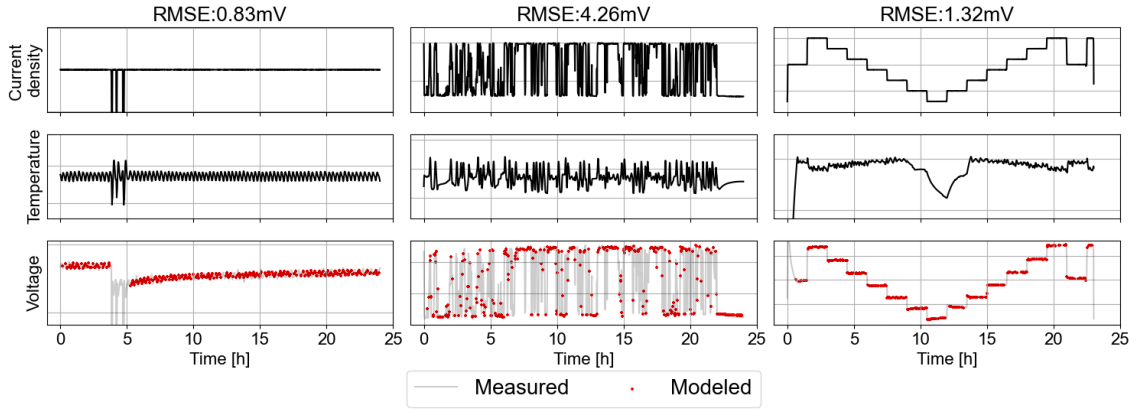


Figure 5.3: Validation of the proposed empirical model Eq. 5.3 with operation data from three different operation schemes. For confidentiality reasons, the y-scales are hidden.

The model Eq. 5.3 can be written in the matrix form as in Section 3.1:

$$Y = X\beta + \epsilon, \quad (5.4)$$

The dependent variable Y is a vector of the overpotentials

$$Y = \begin{bmatrix} U_1 - U_{0,1} \\ U_2 - U_{0,2} \\ \dots \\ U_n - U_{0,n} \end{bmatrix}. \quad (5.5)$$

The independent variable X is a matrix of the impact factors, where the column of ones

corresponds to the constant term c_5

$$X = \begin{bmatrix} i_1 & i_1 T_1 & \ln(OH)_1 & i_1^2 & 1 \\ i_2 & i_2 T_2 & \ln(OH)_2 & i_2^2 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ i_n & i_n T_n & \ln(OH)_n & i_n^2 & 1 \end{bmatrix}. \quad (5.6)$$

The parameter β is a vector of the unknown coefficients

$$\beta = [c_1 \quad c_2 \quad c_3 \quad c_4 \quad c_5]^T, \quad (5.7)$$

and ϵ is a vector of errors

$$\epsilon = [\epsilon_1 \quad \epsilon_2 \quad \dots \quad \epsilon_n]^T. \quad (5.8)$$

The subscript n denotes the number of data points used for model fitting. The parameter β is unknown and needs to be fitted. As the benchmark method, [SLR](#) is used to fit β as described in Section 3.2.3, where no fleet knowledge is incorporated in the model fitting process. Whereas to address model fitting difficulty under limited data coverage (as illustrated in Fig. 3.1), a model fitting method that incorporates fleet knowledge with Bayesian reference is introduced in the next section. To ease the convergence, data is first scaled to a comparable range of 0 to 1 with the MinMaxScaler function from the Python package scikit-learn [116] prior to model fitting.

With the fitted coefficient β , the [SOH](#) indicator U_{rc} under the predefined reference condition X_{rc} can be calculated:

$$U_{rc} = X_{rc} \beta, \quad (5.9)$$

where

$$X_{rc} = [i_{rc} \quad i_{rc} T_{rc} \quad \ln(OH_{rc}) \quad i_{rc}^2 \quad 1]. \quad (5.10)$$

Additionally, the polarization curve can be obtained by calculating U_{rc} for various i_{rc} values while fixing T_{rc} and OH_{rc} .

5.1.3 Model fitting

The proposed empirical model Eq. 5.3 consists of 5 unknown coefficients c_1 to c_5 that need to be fitted with operation data for each interval (Fig. 5.1). This section presents a

model fitting approach that incorporates fleet knowledge through Bayesian inference. The process of extracting and incorporating fleet knowledge is illustrated in Fig. 5.4. It has the following three steps.

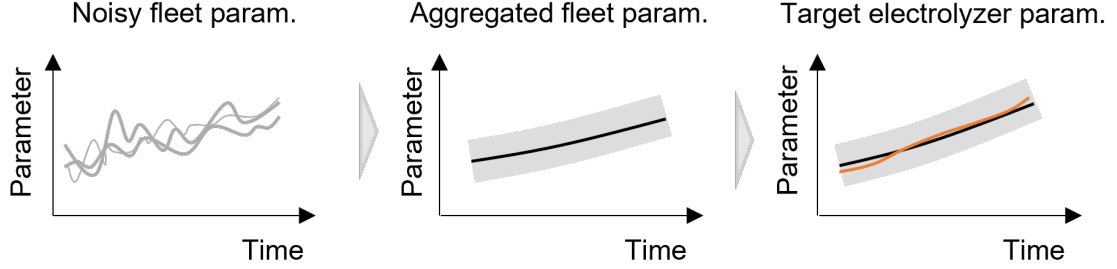


Figure 5.4: Process of constructing and incorporating fleet knowledge in model fitting for the target electrolyzer.

Step 1: Obtain noisy fleet parameters with [SLR](#). In this step, the [SLR](#) method (Section 3.2.3) is applied to get model parameters for the electrolyzer fleet. Due to the model fitting difficulty under limited data coverage (Fig. 3.1 Case B), the obtained parameters are very unstable at this stage.

Step 2: Remove unreliable results and aggregate to fleet knowledge with Gaussian process regression. In this step, unreliable fitting results identified with high condition numbers of the covariance matrix [117, 118] are first removed. Then the remaining fitted parameters are aggregated with Gaussian process regression ([GPR](#)). [GPR](#) is a machine learning technique to obtain a probabilistic model based on observed data [119, 120]. It is used for two reasons: (1) It produces probabilistic prediction, which can be utilized by the Bayesian linear regression in the next step. (2) It is a nonparametric method, which does not rely on any assumed functional form to describe the parameter development trend.

Step 3: Model fitting with Bayesian linear regression [121] using fleet knowledge as the prior probability distribution. The GPR model from Step 2 provides a prior probability distribution on the model parameter distribution for each model fitting interval. It is combined with the measured operation data with Bayesian linear regression to obtain the posterior probability distribution of the model parameter. This is illustrated in Fig. 5.5 using the coefficient c_1 as an example. By applying Bayes' theorem, the posterior model parameter distribution is proportional to the product of the prior distribution and the likelihood function [121, 122]:

$$\underbrace{p(\beta \mid data)}_{\text{posterior}} \propto \underbrace{p(\beta)}_{\text{prior}} \times \underbrace{p(data \mid \beta)}_{\text{likelihood}}. \quad (5.11)$$

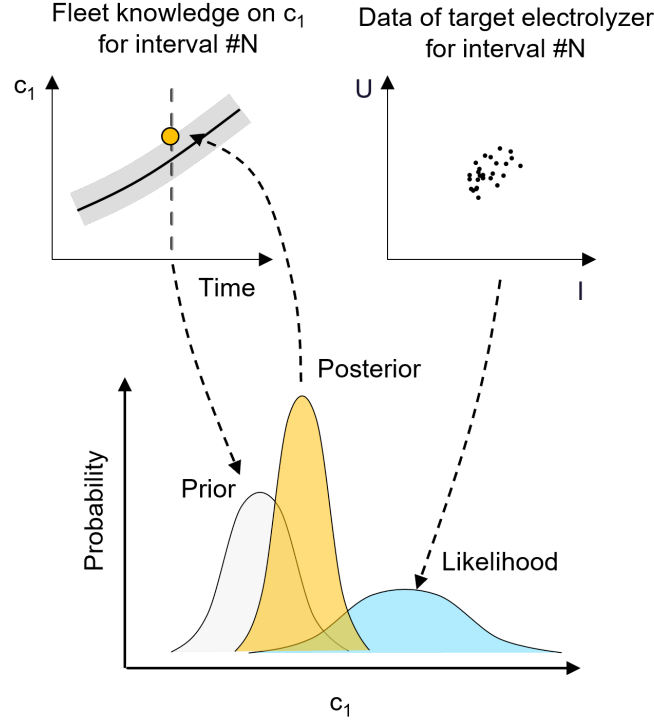


Figure 5.5: Using Bayesian method to obtain the posterior probability distribution of model parameter (c_1 as an example) by combining fleet knowledge and data of the target electrolyzer.

Assume that the prior distribution for the model parameter β , obtained from the GPR process in Step 2, follows a normal distribution with mean m_f and covariance S_f (f denotes that it is from the fleet knowledge):

$$p(\beta) = \mathcal{N}(m_f, S_f). \quad (5.12)$$

Bayesian linear regression gives the posterior distribution of the model parameter β [121]:

$$p(\beta \mid \text{data}) = \mathcal{N}(m, S), \quad (5.13)$$

where

$$S = (S_f^{-1} + \frac{1}{\sigma^2} X^T X)^{-1}, \quad (5.14)$$

$$m = S(S_f^{-1} m_f + \frac{1}{\sigma^2} XY), \quad (5.15)$$

and σ^2 is the noise variance, which represents the inaccuracy of the target linear regression

model. A higher value of σ^2 results in a reduced contribution from the measurement data ($\frac{1}{\sigma^2}X^T X$ and $\frac{1}{\sigma^2}XY$ get smaller), so the posterior is more dominated by the prior. The mean m of the posterior distribution can be used as the point estimator of the parameter β (informally represented as $\beta = m$).

This 3-step model fitting process is demonstrated with synthetic data in Section 5.2.

5.2 METHOD EVALUATION WITH SYNTHETIC DATA

In this section, synthetic data is utilized to validate the proposed method. First, the entire modeling process is demonstrated with a synthetic electrolyzer fleet. Then, three in-depth sensitivity analyses are presented, regarding the impact of data coverage, fleet-target discrepancy, and fleet size.

5.2.1 Evaluation with a synthetic fleet

The synthetic data is generated following the scheme described in Section 3.2.1.2, with the empirical voltage model Eq. 5.3. To assess the fidelity of the synthetic data generation scheme, it is first used to reproduce the industrial data (data of the target electrolyzer used in Section 5.3). Due to the model simplification and unknown model coefficients, reproduction error cannot be avoided. The error distribution is shown in Fig. 5.6. The average error is -1.6 mV, and 87.9% of the data has an error within ± 10 mV - the magnitude of the measurement noise (Section 3.2.1.1). With this, the synthetic data generation scheme is considered satisfactory.

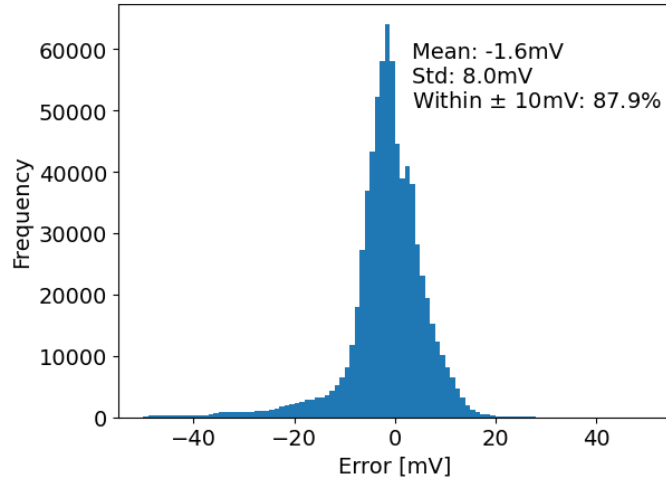


Figure 5.6: Error of using the synthetic data generation scheme to reproduce industrial data.

Then, the following settings are incorporated into the synthetic data generation scheme to produce a synthetic electrolyzer fleet. These settings are inspired by industrial operation data but do not necessarily represent actual values.

- Resolution: 1 min.
- Length: 1 year.
- Full load: 2 A/cm².
- Temperature: Sine wave between 58 and 60°C. Decrease to 20°C during standstill.
- Noise: Current $\pm 1\%$, temperature $\pm 0.1^\circ\text{C}$, voltage $\pm 1\text{ mV}$.
- Operation modes: The entire time frame is randomly segmented into periods with one of the following operation modes. (The percentage values are related to the full load 2 A/cm².)
 - Constant at 100%.
 - Constant at 60%.
 - Constant at 20%.
 - Random fluctuate between 60 and 100%.
 - Random fluctuate between 20 and 100%.
 - Random fluctuate between 20 and 60%.
 - Standstill.
- Number of electrolyzers: 11 (1 target electrolyzer and 10 electrolyzers to form the fleet knowledge). To demonstrate the effectiveness of incorporating fleet knowledge, the target electrolyzer is deliberately designed to have a challenging operation scheme: It is often under constant operation, making the voltage-current relationship difficult to model; it also has few starts, making the start behavior (associated with the coefficient c_3) difficult to model. The rest of the 10 electrolyzers have randomly assigned operation modes, without special treatment.
- Degradation behavior is described with model coefficients $c_1 - c_5$ evolving over time. Their values are set based on the fitted values from the industrial data.
 - c_1 : Linearly increases from 0.3 to 0.36.
 - c_2 : Linearly decreases -0.002 to -0.0024.
 - c_3 : Stays constant at 0.004.
 - c_4 : Exponentially increases from -0.0225 to 0.1.
 - c_5 : Stays constant at 0.35.

The resulting synthetic operation data of the target electrolyzer is shown in Fig. 5.7.

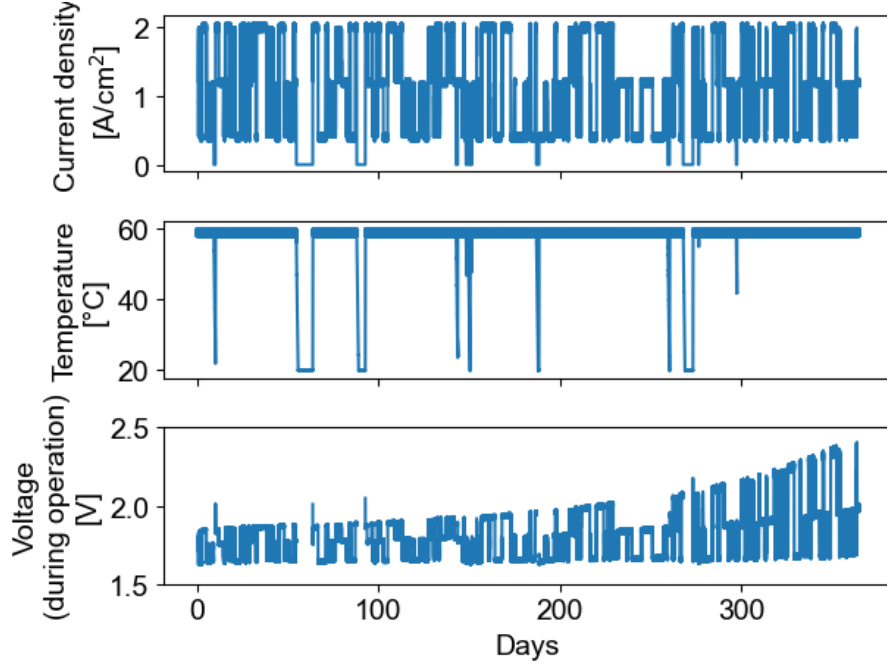


Figure 5.7: Synthetic data of the target electrolyzer.

The proposed method is executed with the following settings.

- Model fitting interval: 1 day.
- Reference condition to calculate U_{rc} : $i_{rc} = 2\text{A/cm}^2$, $T_{rc} = 60^\circ\text{C}$, and $OH_{rc} = 72\text{h}$.
- Condition number threshold to identify unstable fitting results: 10^5 (determined empirically by examining the distribution of the condition numbers from SLR model fitting in Step 1).
- Parameters for GPR [116]: Kernel (a function that computes the similarity between data points) is set as the widely used radial basis function kernel [123]. Alpha (noise of the data) is set as the standard error obtained from SLR in Step 1.
- Parameters for Bayesian linear regression: noise σ is set as the average RMSE from SLR in Step 1.

Fig. 5.8 demonstrates the proposed modeling process described in Section 5.1.3. In Step 1, parameters (taking c_1 as an example) from different electrolyzers obtained with SLR are very unstable. In Step 2, after removing unreliable results based on condition number, the parameters exhibit a clear trend, which is captured with GPR as the fleet knowledge. In Step 3, the fleet knowledge is incorporated into the model fitting process for the target electrolyzer with Bayesian linear regression, resulting in parameters that closely follow the fleet knowledge (the orange line). In contrast, model fitting without fleet knowledge using the benchmark method SLR results in very unstable parameters (the blue line).

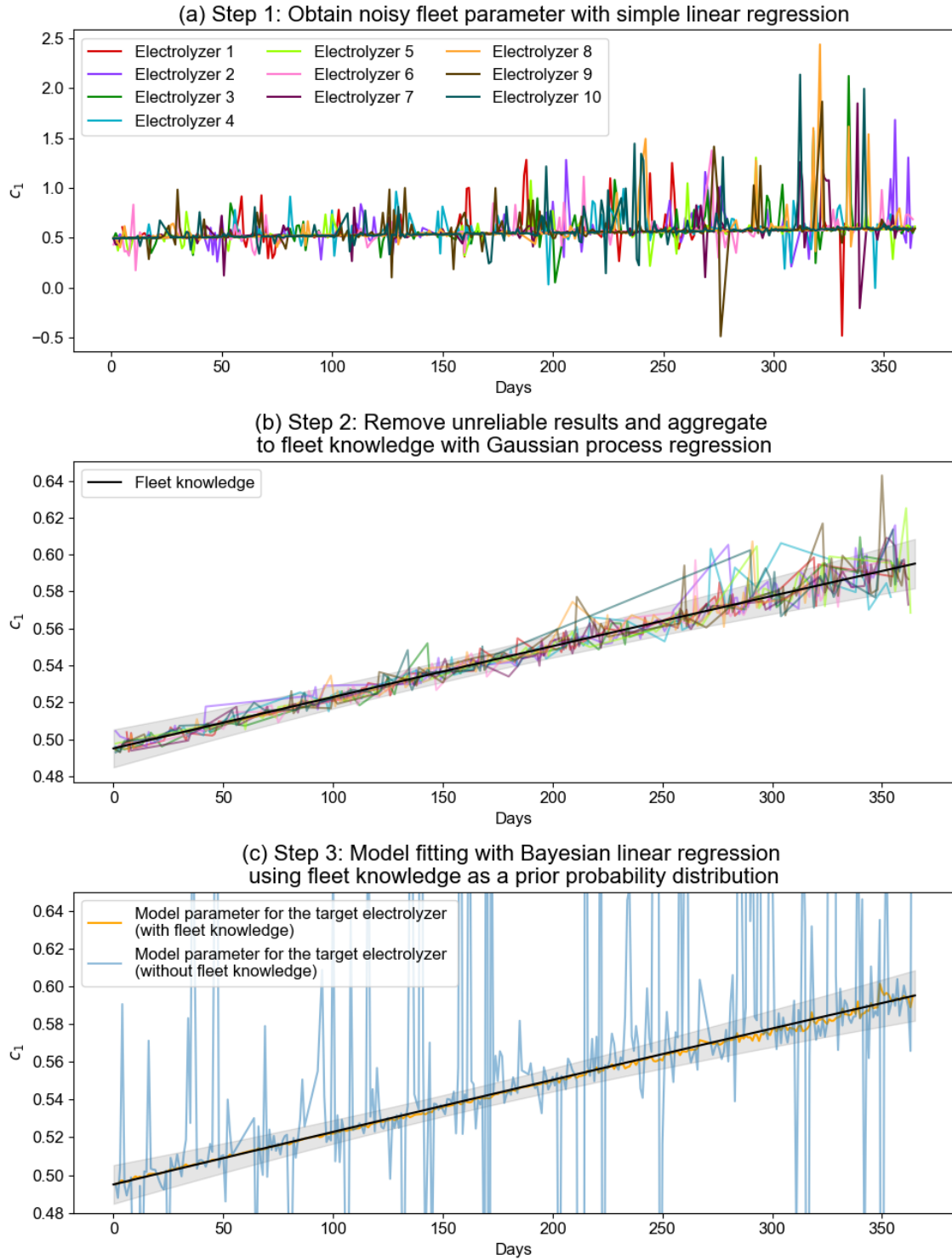


Figure 5.8: Demonstration of the modeling process with synthetic fleet data.

Fig. 5.9 compares the results obtained with the proposed fleet-based method against the benchmark method SLR without fleet knowledge. Fig. 5.9a shows that incorporating fleet knowledge with the proposed method reduces the U_{rc} estimation error from 106.6 mV to 1.7 mV. Regarding the accuracy of polarization curves (generated by applying the fitted voltage model Eq. 5.3 on various current densities between 0.4 A/cm² and 2.0 A/cm², while keeping $T = 60^\circ\text{C}$ and $OH = 72\text{ h}$), Fig. 5.9b shows that incorporating fleet knowledge significantly stabilizes the polarization curves, showing a consistent rise over time, and the RMSE is reduced from 51.9 mV to 0.9 mV. Last, Fig. 5.9c shows that voltage models fitted with the fleet-based method can forecast voltage more accurately than the benchmark method, where RMSE are mostly below 10 mV for all three forecast horizons except a few outliers. These comparisons clearly highlight the superiority of the proposed fleet-based method.

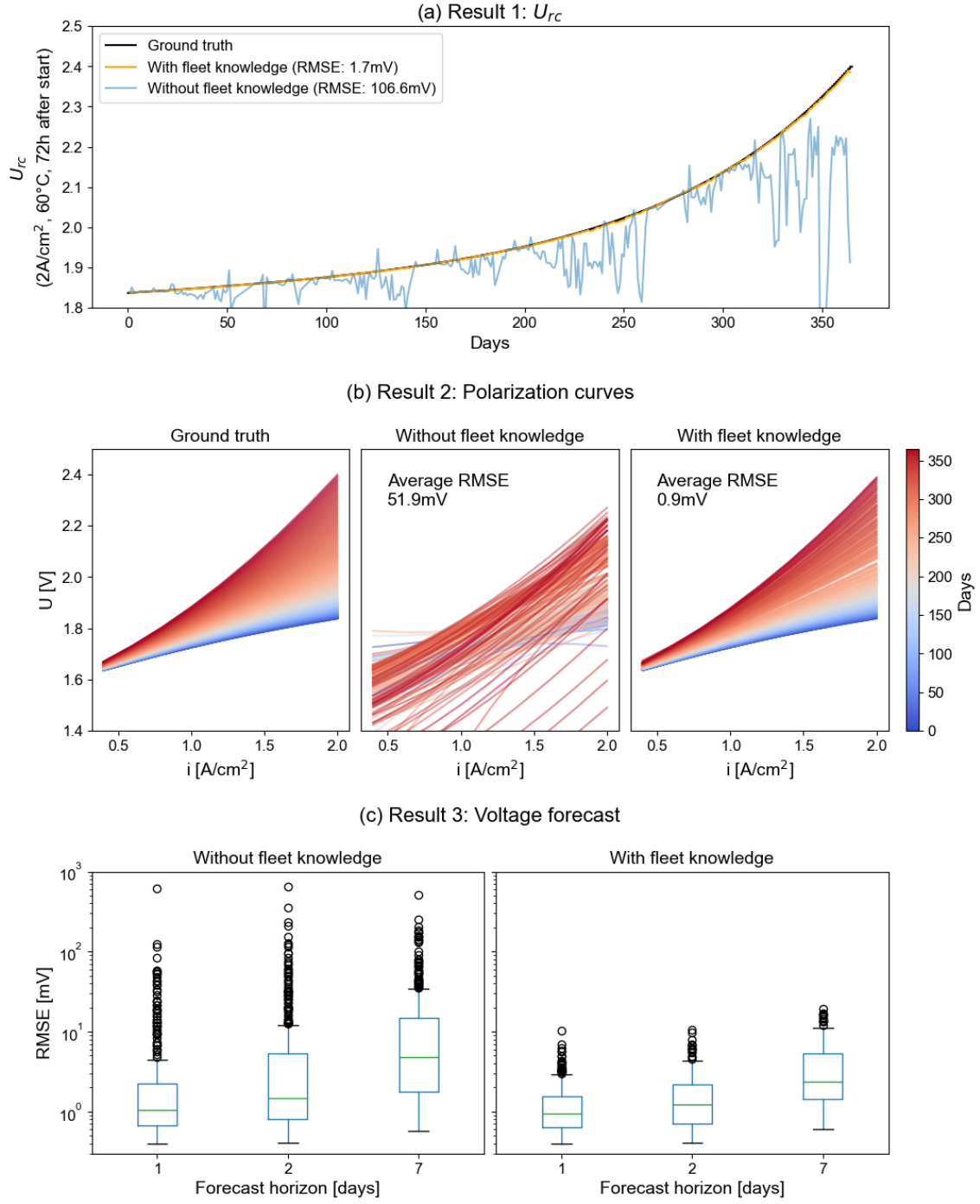


Figure 5.9: Results comparison between the proposed fleet-based method and the benchmark method [SLR](#) without using fleet knowledge.

5.2.2 Sensitivity analysis 1: impact of data coverage

Following the demonstration of the modeling process with a synthetic electrolyzer fleet, synthetic data is now utilized to conduct three in-depth sensitivity analyses. The first sensitivity analysis focuses on the impact of data coverage. Specifically, the operating range (i.e., the coverage of current density) is varied to observe how it affects the accuracy of U_{rc} estimation (Fig. 5.10).

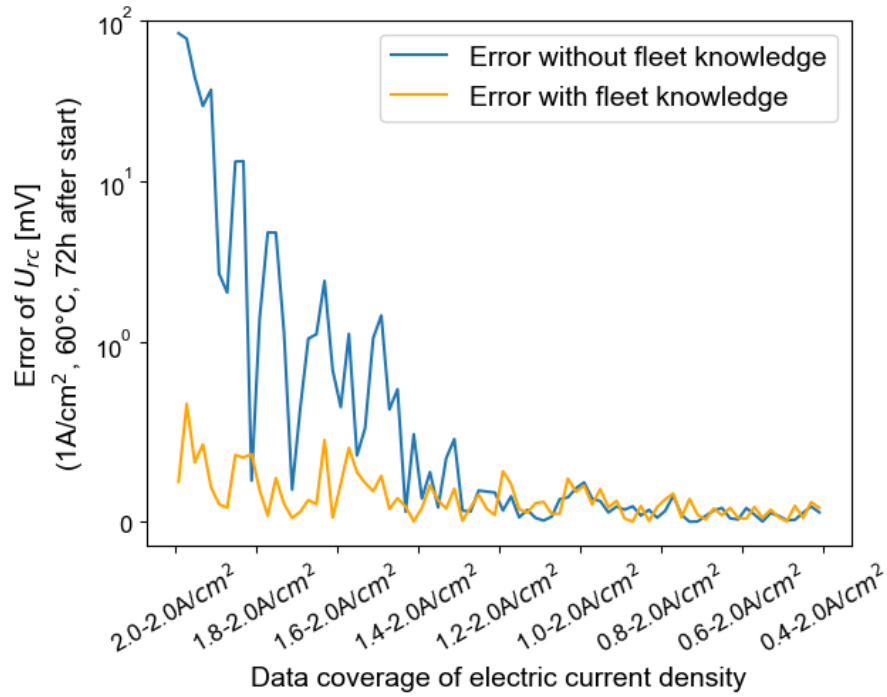


Figure 5.10: Impact of data coverage.

The data is generated using the following setup.

- Resolution: 1 min.
- Length: 1 day.
- Number of electrolyzers: 1.
- Operation range: Current density randomly fluctuates within X to 2 A/cm^2 . The lower bound X decreases from 2 A/cm^2 to 0.4 A/cm^2 (the x-axis in Fig. 5.10), corresponding to an increase in data coverage.
- Temperature: Sine wave between 58 and 60°C .
- Noise: The noise of voltage is randomly generated from a normal distribution with a mean of 0 V and a standard deviation of 0.001 V .
- Coefficients $c_1 - c_5$: Mean of the fleet knowledge derived in Section 5.2 at the first

timestamp.

For each dataset with a different operation range, both model fitting methods with and without fleet knowledge are applied. The proposed fleet-based method is conducted with the following parameters.

- Model fitting interval: 1 day.
- Reference condition to calculate U_{rc} : $i_{rc} = 1\text{A/cm}^2$, $T_{rc} = 60^\circ\text{C}$, and $OH_{rc} = 72\text{h}$.
- Parameters for Bayesian linear regression: The prior probability distribution is the fleet knowledge derived in Section 5.2 at the first timestamp, which is the basis for data generation as mentioned above. The noise parameter σ is set to be 0.001V , matching the standard deviation of the synthetic noise.

The resulting U_{rc} estimation errors are shown in Fig. 5.10. It can be observed that with low data coverage, the benchmark method without using fleet knowledge produces errors of up to 100 mV and the fluctuation shows that the method is unstable, while including fleet knowledge significantly reduces the error to less than 1 mV with high stability. On the other hand, with high data coverage, both methods show comparable accuracy. This underscores the necessity of incorporating fleet knowledge especially when handling limited data coverage.

Note that the issue of data coverage applies to all variables on the right-hand side of the voltage model Eq. 5.3. Here only the coverage of the electric current is demonstrated. But in practice, the term $\ln(OH)$ is also problematic when an electrolyzer keeps operating for a long time. This is because the coverage of $\ln(OH)$ gradually decreases as OH increases. For example, from operating hour 1 to 2 after a start corresponds to $\ln(2) - \ln(1) = 0.69$ coverage in $\ln(OH)$; while from operating hour 101 to 102 corresponds only to $\ln(102) - \ln(101) = 0.01$ coverage in $\ln(OH)$.

5.2.3 Sensitivity analysis 2: impact of fleet-target discrepancy

The second sensitivity analysis focuses on the impact of fleet-target discrepancy on the accuracy of U_{rc} estimation by varying the coefficient c_1 (Fig. 5.11).

The data generation scheme and modeling setup are mostly the same as in Section 5.2.2, with the following differences.

- Operation range: 1.98 to 2 A/cm^2 .
- For data generation, coefficient c_1 is additionally increased by 0 to 25% based on the fleet knowledge, to represent the discrepancy between the fleet and the target (the x-axis in Fig. 5.11).

For each dataset, both model fitting methods with and without fleet knowledge are applied. The resulting errors of U_{rc} estimations are shown in Fig. 5.11. It can be observed

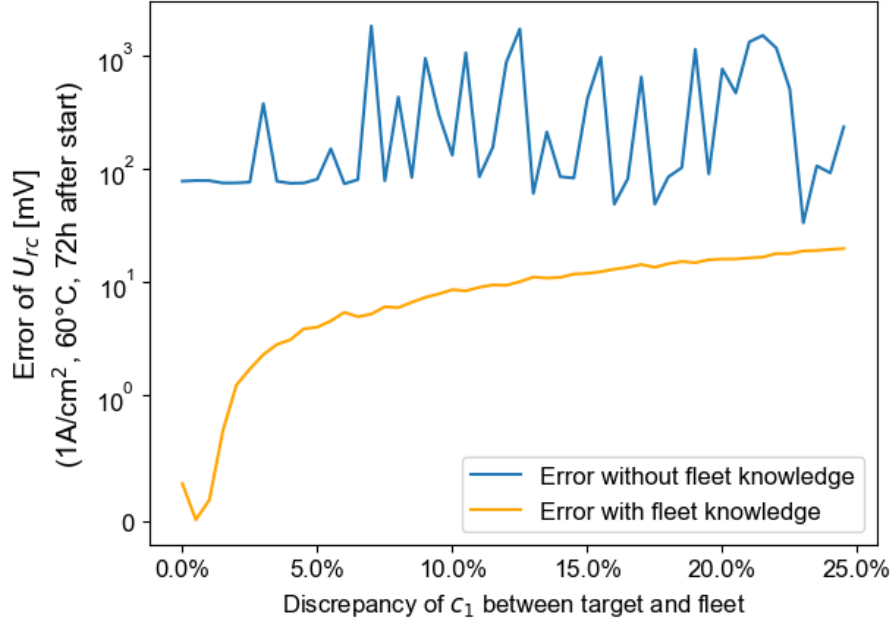


Figure 5.11: Impact of the discrepancy between the fleet and the target electrolyzer.

that model fitting without using fleet knowledge produces errors fluctuating between 10^2 and 10^3 mV, while incorporating fleet knowledge reduces the error by nearly an order of magnitude or even more. Nevertheless, as the discrepancy increases, the error of the fleet-based U_{rc} also increases, indicating the advantage of having an electrolyzer fleet that is similar to the target electrolyzer.

5.2.4 Sensitivity analysis 3: impact of fleet size

The third sensitivity analysis focuses on the impact of fleet size (i.e., the number of electrolyzers used to construct the fleet knowledge) on the quality of the fleet knowledge by varying the fleet size from 1 to 10 electrolyzers (Fig. 5.12a, b). Furthermore, the subsequent impact on the accuracy of U_{rc} estimation is investigated using multiple scenarios with various data coverage (Fig. 5.12c) and fleet-target discrepancy (Fig. 5.12d), with similar data generation schemes as in Section 5.2.2 and Section 5.2.3, respectively. Model fitting with fleet knowledge is applied.

Fig. 5.12a and 5.12b show that increasing the fleet size reduces the error and uncertainty of the fleet knowledge. The benefit of error reduction gets more evident as the data coverage decreases (Fig. 5.12c): If deliberately introducing errors to the fleet knowledge of c_1 from 0.1 to 0, the U_{rc} estimation error decreases significantly with a small data coverage of 1.8–2.0 mA/cm²; in contrast, with a large data coverage of 0.6–2.0 mA/cm², the U_{rc} estimation

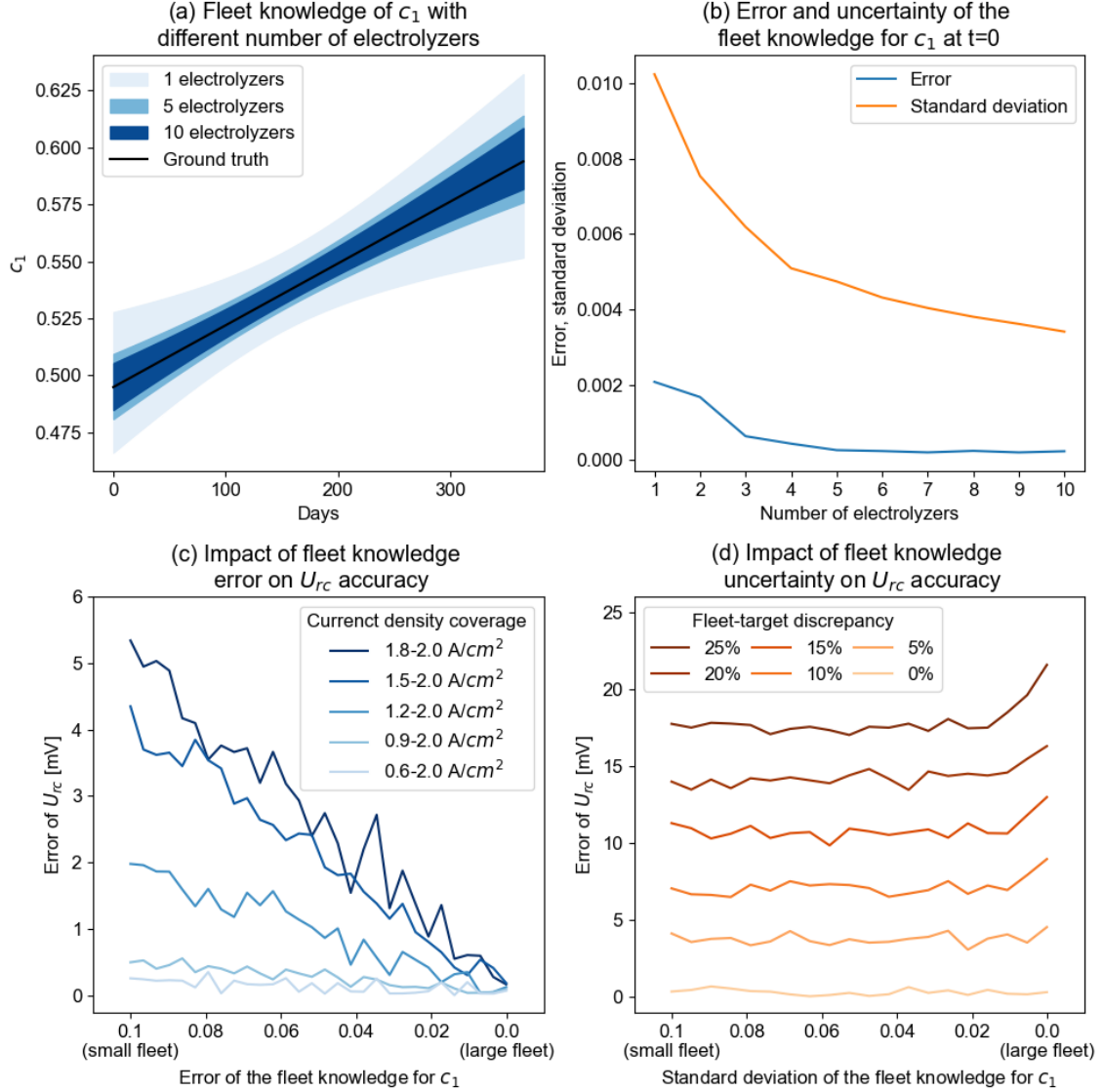


Figure 5.12: Impact of fleet size.

error stays low. This is because, under a small data coverage, the fleet knowledge gains more importance in the model fitting, thus the error of the fleet knowledge has a more evident impact.

The uncertainty reduction effect of an increasing fleet size has a negative consequence under large fleet-target discrepancy (Fig. 5.12d): If altering the standard deviation of the fleet knowledge for c_1 from 0.1 to 0, the U_{rc} estimation error increases in the presence of a fleet-target discrepancy. The reason is, referring to Fig. 5.5, as the uncertainty of fleet knowledge reduces, the resulting posterior estimation tends to converge more closely with

the prior fleet knowledge, thus diverging from the data of the target electrolyzer. However, despite the potential error increase with a very large fleet, it is crucial to acknowledge the substantial benefit of incorporating fleet knowledge (demonstrated in Fig. 5.11): The inclusion of fleet knowledge can reduce the U_{rc} error by almost tenfold or more given fleet-target discrepancies. In comparison, the potential error introduced by using a large fleet remains acceptable.

5.3 METHOD EVALUATION WITH INDUSTRIAL DATA

Next, the proposed method is applied to real-world operation data from an industrial fleet of electrolyzers. The data has been introduced in Section 3.2.1.1. The fleet comprises 12 electrolyzers, one of which is regarded as the target electrolyzer, while data from the remaining 11 electrolyzers are used to form the fleet knowledge. The operation profiles are not standardized but business-driven, including periods with limited data coverage. For confidentiality reasons, all figures in this section do not show the voltage and current values.

The proposed method is executed with the following settings, which are mostly the same as in the evaluation with synthetic data (Section 5.2), except for the reference conditions.

- Model fitting interval: 1 day.
- Reference condition to calculate U_{rc} : Three different reference conditions. One is the nominal operating condition, and two are operation conditions that occur relatively often with abundant data to facilitate model validation. For confidentiality reasons, the exact conditions cannot be provided.
- Condition number threshold to identify unstable fitting results: 10^5 (determined empirically by examining the distribution of the condition numbers from SLR model fitting in Step 1).
- Parameters for GPR [116]: Kernel (a function that computes the similarity between data points) is set as the widely used Radial basis function kernel [123]. Alpha (noise of the data) is set as the standard error obtained from SLR in Step 1.
- Parameters for Bayesian linear regression: noise σ is set as the average RMSE from SLR in Step 1.

Fig. 5.13 presents the calculated U_{rc} for the target electrolyzer under three different reference conditions, alongside the measured raw data filtered at these conditions. It should be noted that the raw data are only sporadically available (especially at the high current level, i.e., reference condition 1) because the operating conditions are arbitrary and rarely match the reference conditions. This is the motivation to derive the degradation indicator U_{rc} for continuous condition monitoring.

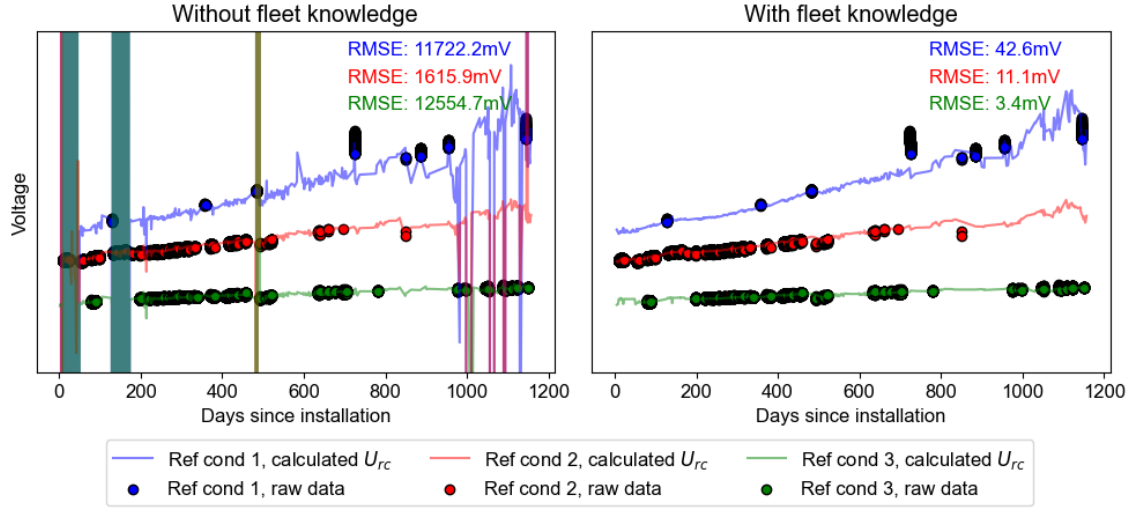


Figure 5.13: U_{rc} calculated with the benchmark method (without fleet knowledge, left) and with the proposed method (with fleet knowledge, right) under three reference conditions. They are compared with the raw measured data filtered at these conditions.

The accuracy of U_{rc} is quantified with the **RMSE** between the calculated U_{rc} and the raw data, as described in Section 3.2.2.1. The comparison between the **RMSE** of the benchmark method **SLR** without using fleet knowledge (Fig. 5.13, left) and the proposed fleet-based method (Fig. 5.13, right) shows that incorporating fleet knowledge significantly improves the U_{rc} estimation accuracy. However, the **RMSE** at a high current level (reference condition 1) is much higher than the other two reference conditions. This is because the investigated electrolyzer rarely operates at this current level. The lack of operation experience makes the model less accurate in describing the voltage behavior in the high current region.

By calculating U_{rc} at multiple current levels, polarization curves can be obtained (Fig. 5.14). The benchmark method without using fleet knowledge produces noisy curves. In contrast, with the support of fleet knowledge, the proposed method yields stable model parameters, leading to consistent polarization curves. This comparison further proves the superiority of the proposed method.

The polarization curves obtained with the fleet-based method are further compared with the raw measurement data, as described in Section 3.2.2.2. Fig. 5.15 shows that five polarization curve tests have been performed with the target electrolyzer, where the input current varies from the minimum operating current to beyond the nominal current. The **RMSE** ranges from 1 mV to 12 mV, which is a significant improvement over the first proposed method (up to 54 mV as shown in Fig. 4.13).

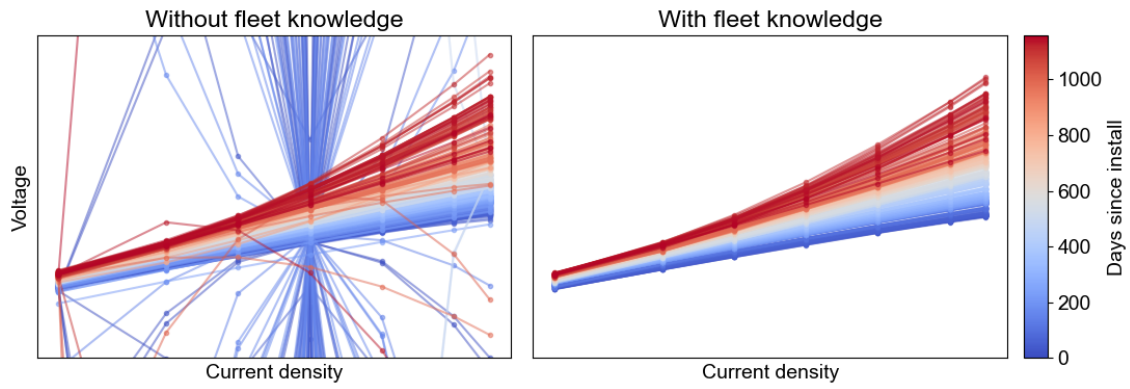


Figure 5.14: Polarization curves obtained with the benchmark method (without fleet knowledge, left) and with the proposed method (with fleet knowledge, right).

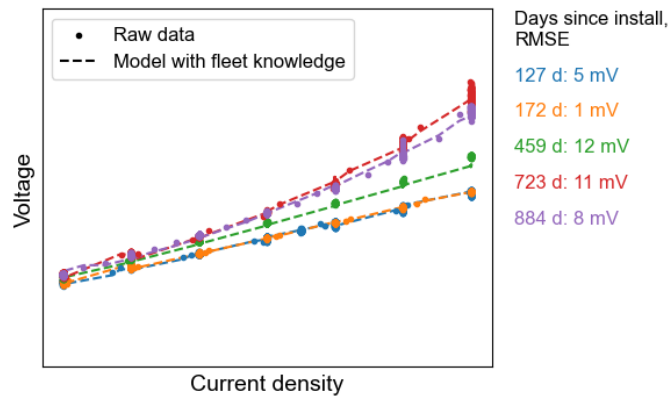


Figure 5.15: Comparison of polarization curves modeled using the fleet-based method with the measured data.

The proposed method is also evaluated regarding the forecast capability of the fitted voltage models and is compared with the benchmark method. **RMSE** is calculated as described in Section 3.2.2.3. For both benchmark (Fig. 5.16, left) and the proposed fleet-based method (Fig. 5.16, right), the forecast error increases as the forecast horizon extends from 1 day to 7 days, due to the changing degradation state over time. The voltage models fitted with the benchmark method produce errors of up to 10^5 mV, whereas with the fleet-based method, the errors consistently remain below 100 mV, with the majority (75th percentile, marked with the upper edge of the box) below 10 mV. The error reduction effect of the fleet-based method is particularly pronounced for the outliers (dots beyond the whiskers), showing the robustness of the algorithm.

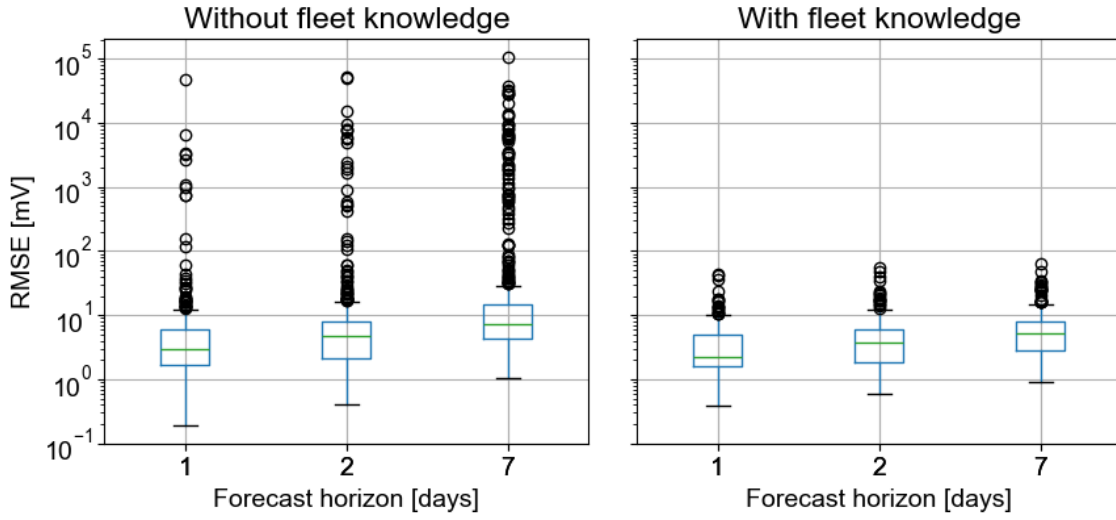


Figure 5.16: Voltage forecast accuracy of the voltage models fitted with the benchmark method (left) and the proposed fleet-based method (right).

5.4 LIMITATIONS

5.4.1 Similarity between the fleet and the target electrolyzer

The sensitivity analysis in Section 5.2.3 shows that larger fleet-target discrepancy causes larger U_{rc} estimation error with the proposed method. To account for this, future research could develop a metric for quantifying the degree of similarity between the fleet and the target electrolyzer. This metric may include the similarity of both operational data [93] and static metadata, such as manufacturing parameters. This similarity information can then be incorporated into the Bayesian process. For example, if the target electrolyzer has a large difference from the fleet, one can assign a higher uncertainty value to the fleet

knowledge (i.e., the prior probability distribution), so the Bayesian process attributes more weight to the data from the target electrolyzer.

5.4.2 Influence of operation modes

The fleet knowledge used in this study is constructed with Gaussian process regression using calendar time as the independent variable. This implies that all electrolyzers in the fleet degrade similarly relative to calendar time. However, degradation is also influenced by operation modes. For example, studies have shown that dynamic operation with frequent cycling [82] and high current density [124] might accelerate degradation. Therefore, future research could consider including operation modes in the fleet knowledge construction process, to broaden the applicability of the derived fleet knowledge.

5.4.3 Nonlinear voltage models

The voltage model used in this method (Eq. 5.3) has a linear structure, which is proven to be effective with Fig. 5.3. However, due to the model's empirical nature, it might be invalid for other electrolyzers. In cases where the voltage model cannot be linearized, the analytical solution for the posterior probability distribution (Eq. 5.14 and 5.15) based on Bayesian linear regression does not apply anymore. Instead, Markov chain Monte Carlo (MCMC) sampling can be deployed to derive a numerical approximation of the posterior probability distribution of the model parameters [125].

6

Discussions on the methods

This section discusses additional aspects of the proposed methods, including the impact of using different temperature sensors (Section 6.1), the impact of different lengths of model fitting intervals (Section 6.2), other impact factors on voltage (Section 6.3), and the transferrability of the method to other health indicators (Section 6.4).

6.1 IMPACT OF USING DIFFERENT TEMPERATURE DATA

The investigated industrial electrolyzers (Section 3.2.1.1) have multiple temperature sensors for each stack: H₂-side inlet, O₂-side inlet, H₂-side outlet, and O₂-side outlet. The average temperature of these 4 sensors is regarded as the average stack temperature and is used in the modeling process.

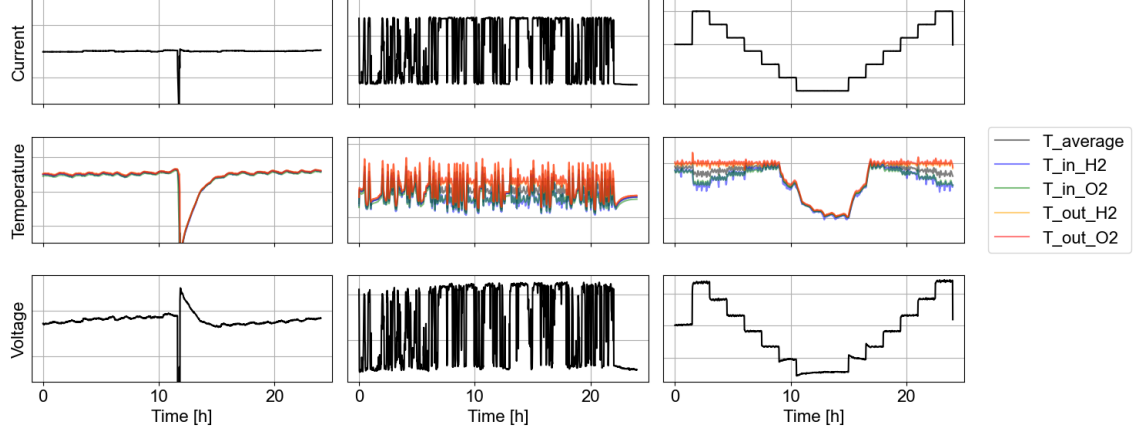
Fig. 6.1a illustrates the temperature data for three exemplary intervals, each featuring a different operation scheme: constant, volatile, and staircase operation. The temperature differences compared with the average temperature (baseline) are quantified in Fig. 6.1b: The inlet temperatures are consistently lower than the average, by up to 0.8 °C, while the outlet temperatures are the opposite.

The different temperature data are used in fitting the proposed voltage model Eq. 5.3 for these three intervals, with the benchmark model fitting method SLR. The resulting fitting errors RMSE are summarized in Fig. 6.1b. Overall, all errors are at a single-digit-mV level, fulfilling the accuracy requirement defined in Section 3.2.1.1. Furthermore, the difference in using different temperature data is marginal, indicating that the model fitting quality is insensitive to the different temperature data.

6.2 IMPACT OF MODEL FITTING INTERVAL LENGTH

The proposed methods are applied to industrial data on a daily basis: A voltage model is fitted for each day, thereby yielding a daily U_{rc} . Setting the model fitting interval involves a trade-off between the amount of data and the model accuracy. Choosing a longer interval provides more data for model fitting, but also encompasses more degradation within the interval, leading to an inaccurate SOH estimation.

- (a) Different temperature measurements for 3 exemplary periods with different operation schemes: constant, volatile, and staircase operation. For confidentiality reasons, the y-scales are hidden.



- (b) Temperature differences and impact on model fitting errors.

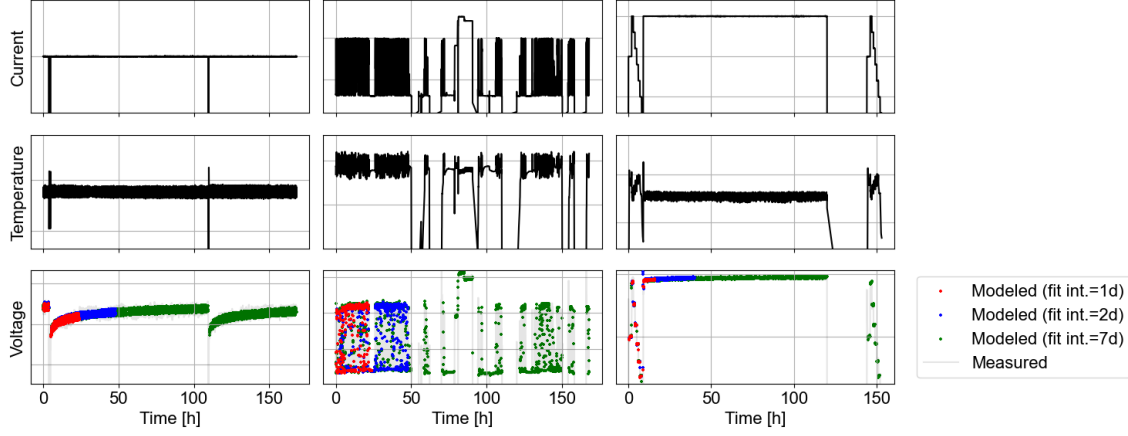
	Example interval 1		Example interval 2		Example interval 3	
Temperature data used in model fitting	Temperature difference	RMSE [mV]	Temperature difference	RMSE [mV]	Temperature difference	RMSE [mV]
Average stack temperature	Baseline	0.26	Baseline	2.62	Baseline	1.09
H ₂ -side inlet temperature	Baseline-0.0°C	0.27	Baseline-0.7°C	1.90	Baseline-0.8°C	1.06
O ₂ -side inlet temperature	Baseline-0.2°C	0.29	Baseline-0.8°C	2.05	Baseline-0.7°C	0.99
H ₂ -side outlet temperature	Baseline+0.1°C	0.24	Baseline+0.7°C	3.05	Baseline+0.7°C	1.25
O ₂ -side outlet temperature	Baseline+0.1°C	0.24	Baseline+0.8°C	3.16	Baseline+0.8°C	1.26

Figure 6.1: Different temperature data and their impact on model fitting error [RMSE](#).

Fig. 6.2a presents three exemplary 7-day intervals, each featuring a different operation scheme: constant, volatile, and staircase & constant operation. For each interval, the proposed voltage model Eq. 5.3 is fitted with 1, 2, or 7 days of data using the benchmark method [SLR](#). The modeled voltage is plotted in the last row of figures, aligning well with the measured voltage. The model fitting errors are summarized in Fig. 6.2b. In all cases, the [RMSE](#) is at a single-digit-mV level, fulfilling the accuracy requirement defined in Section 3.2.1.1. As the model fitting interval increases, two different effects can be observed: For examples 1 and 2, the error increases, likely due to the degradation over a longer interval; while for example 3, the error decreases, which can be attributed to the larger number of data points used in the fitting process.

In practice, the modeling interval can be determined based on factors such as the degradation rate (a shorter interval for electrolyzers with a higher degradation rate), data coverage (a shorter interval if there is abundant data with sufficient coverage), and business routines (e.g., aligning with the frequency of scheduled condition monitoring activities).

(a) 7-day intervals of 3 exemplary periods with different operation schemes: constant, volatile, and staircase & constant operation. For confidentiality reasons, the y-scales are hidden.



(b) Impact of model fitting interval on fitting errors.

	Example interval 1 (Constant operation)	Example interval 2 (Volatile operation)	Example interval 3 (Staircase & constant operation)
Model fitting interval	RMSE [mV]	RMSE [mV]	RMSE [mV]
1 day	0.80	2.62	4.44
2 days	0.93	2.67	3.65
7 days	1.98	3.59	2.96

Figure 6.2: Different model fitting intervals and their impact on model fitting error [RMSE](#).

6.3 OTHER IMPACT FACTORS ON VOLTAGE

The voltage models used in this research (Eq. 4.1 and 5.3) contain three variables that influence the voltage: current density, temperature, and time since the last start. These variables have been demonstrated to effectively capture the voltage behavior of the investigated electrolyzers (Section 3.2.1.1). Yet, for different electrolyzer designs, the impact factors on the voltage behavior may extend beyond these three parameters.

For example, for electrolyzer systems with fluctuating pressures, the impact of pressure should also be considered. [38] and [126] both show that higher pressure leads to higher voltage. To capture this behavior, pressure should be added into the voltage mode, and a reference condition regarding pressure needs to be defined to calculate U_{rc} .

However, it should be noted that including more variables in the model will also increase the difficulty of determining its coefficients. Overfitting and instability might occur. Therefore, careful consideration must be given to the trade-off between model complexity and its predictive accuracy, ensuring that the model remains both robust and practically

applicable.

6.4 TRANSFERABILITY OF THE METHOD TO OTHER HEALTH INDICATORS

The essence of the proposed methods is to correct the impact of operation mode on degradation-related indicators. This research focuses on voltage because it is commonly used, easy to measure, and has high economic relevance in the industry. Nevertheless, the underlying principle of the proposed methods can be applied to other degradation-related indicators that are impacted by operational modes.

For example, hydrogen permeation rate can be another use case. The crossover of hydrogen gas from the cathode to the anode through the membrane not only serves as an indicator of membrane degradation, but also poses a safety concern due to the formation of explosive gas mixtures [27]. Trinke et al. [127] reported that the hydrogen permeation rate is linearly related to current density and increases with elevated temperatures. By employing the proposed methods, the influence of operational modes can be isolated, allowing for the comparison of the permeation rate over time as a SOH indicator.

Similarly, the fluoride release rate, an indicator of the chemical degradation of the membrane, is also impacted by operating conditions. As demonstrated in [87], the fluoride release rate peaks at a moderate current density of approximately 0.4 A/cm^2 , decreases at lower or higher current densities, and increases with rising temperatures. These impacts can also be addressed with the proposed methods.

7

Practical guide for fleet-based algorithm development

Chapter 3 to 6 presented two fleet-based SOH quantification methods for PEM electrolyzers developed in this research. This chapter distills the algorithm development process and provides a framework for industrial practitioners. First, core steps of fleet-base algorithms are summarized (Section 7.1). Then, a step-by-step guide on developing fleet-based algorithms is provided (Section 7.2). This chapter is based on the publication: Yan, X.⁼, Woelke, J.⁼, Bensmann, B., Eckert, C., Hanke-Rauschenbach, R., Nieße, A.: Cross-method overview of fleet-based machine health estimation and prediction: a practical guide for industrial applications. IEEE Access 13, 60131–60147 (2025). doi:[10.1109/access.2025.3556251](https://doi.org/10.1109/access.2025.3556251) (= equal contribution) [14].

7.1 CORE STEPS OF FLEET-BASED ALGORITHMS

Through reviewing various fleet-based algorithms (Section 2.4), three core steps shared across these algorithms are identified: data selection, model development, and model adjustment (Table 7.1).

Data selection refers to selecting the most relevant machine or data segment as the basis for model training to mitigate the potential negative impact of irrelevant data (also referred to as *negative transfer* in the context of transfer learning [128]). It is included in the sufleet

Table 7.1: Core steps of fleet-based algorithms

	Fleet modeling	Subfleet modeling	Similarity-based modeling	Transfer learning		
				Parameter-based	Feature-based	Instance-based/hybrid
Data selection		X	X			X
Model development	X	X	X	X	X	X
Model adjustment				X	X	X

modeling, similarity-based modeling, and instance-based/hybrid transfer learning methods. In subfleet modeling, data selection refers to identifying machines that belong to the same subfleet (i.e., a subgroup of fleet units sharing some similarities). The most straightforward approach is to leverage engineering knowledge. For example in [129], milling machines of different materials (cast iron and steel) are separated and different SOH quantification models are developed for each material type. To facilitate such engineering knowledge-based selection more systematically, machines' technical features and operational context can be structured semantically into an ontology model [102]. When conducting any modeling task, one can search for similar machines based on this ontology model. Besides using engineering knowledge to identify subfleets, another approach is to measure statistical similarities. For example, [105] uses the Euclidean distance of the degradation trajectories to quantify the similarity between the target machines and other machines. If the distance is lower than a predefined threshold, the units are regarded as similar, then their data are pooled together to train a degradation forecast model. Such data selection methods using statistical similarity measurements are widely employed in similarity-based modeling and instance-based/hybrid transfer learning. The most commonly used metric to quantify similarity is Euclidean distance, adopted in [76–81, 129–142]. Moreover, some existing algorithms inherently contain a data selection process, such as TrAdaBoostR2 [143]. It is an instance-based transfer learning algorithm for regression tasks, which dynamically adjusts the weights of the training instances based on the prediction error [144].

Model development refers to establishing a model that maps input data with the desired output. It is a necessary step in all algorithms. The most popular model architecture is a neural network-based model, including long short-term memory (LSTM) network [93, 145–159], convolutional neural network (CNN) [107, 160–167], autoencoders (AE) [168–170], etc. Other models that are not based on neural networks, can also be found in the literature, including regularized linear regression [68, 73], Gaussian process regression [171], Gaussian mixture model [172], random forest [173], kernel ridge regression [174], support vector machine [170], etc. Some studies incorporate physical knowledge (i.e., empirical degradation models) into the data-driven model, such as in [143, 175].

Model adjustment is a transfer learning-specific step, which tailors a source model (mostly neural network-based models) toward the target machine. This can be achieved by fine-tuning the parameters of the source model: either parameters of all the neural network layers [162, 163, 168, 169] or parameters of the last layers [147–152, 161]. Besides, model adjustment can be done by aligning features between the source and target domain, through adversarial training [153, 164, 165, 176], incorporating the feature difference into the loss function [107, 166, 167, 177, 178], or transfer component analysis [170, 174].

7.2 FLEET-BASED ALGORITHM DEVELOPMENT PROCESS

The above-mentioned three core steps are the building blocks of developing fleet-based algorithms for industrial applications. This section suggests a step-by-step guide for industrial practitioners to integrate these steps, ensuring the developed algorithms are well-suited to their specific requirements. The guide (Fig. 7.1) is developed with the following practical considerations.

First, the guide starts with a simple method and gradually adds complexity as needed, because simple methods tend to be more robust and understandable than complicated alternatives [179], which is highly desired by industrial users.

Second, the development of algorithms is considered as an iterative process, not as a one-off decision. Therefore, the guide offers the flexibility for the modeler to evaluate the model, subsequently enrich it with additional components, or even refine their objectives and restart the modeling process.

Third, it is assumed that a predefined requirement on the accuracy of the algorithm exists based on the business need. For example, if a remaining useful life prediction algorithm is developed for maintenance planning, and a typical time buffer for preparing a maintenance activity is half a year (e.g., for gathering the required spare parts or arranging onsite workers), then the required accuracy for the algorithm might be at month-level. In the proposed guide, the question regarding the satisfaction of the model quality refers to the predefined requirement.

The guide consists of questions (blue), branches of possible answers (red), and suggestions (grey), leading practitioners through the hierarchically structured levels that are processed one after the other.

Level 1: Scope of application The guide starts with the main question regarding the use case (scope of application) of the model, because this significantly limits the suitable methods. If a model for the whole fleet is required, one can directly proceed to build a fleet model (branch 1.1). If a model for a subfleet (e.g., for a specific product generation) is required, one can select the data from the desired subfleet and construct a corresponding subfleet model (branch 1.2). In contrast, if a model for one specific machine is of interest and the data of the target machine itself is insufficient to obtain a satisfactory model, the next level can be entered to check the similarity of the target machine compared to the rest of the fleet (branch 1.3).

Level 2: Similarity Depending on the similarity of the target machine to the rest of the fleet, a decision is made as to whether and, if so, which data from the fleet should be used to assist in the modeling of the target machine. If the target machine is similar to the entire fleet, a fleet model can be built (branch 2.1). If it is only similar to a

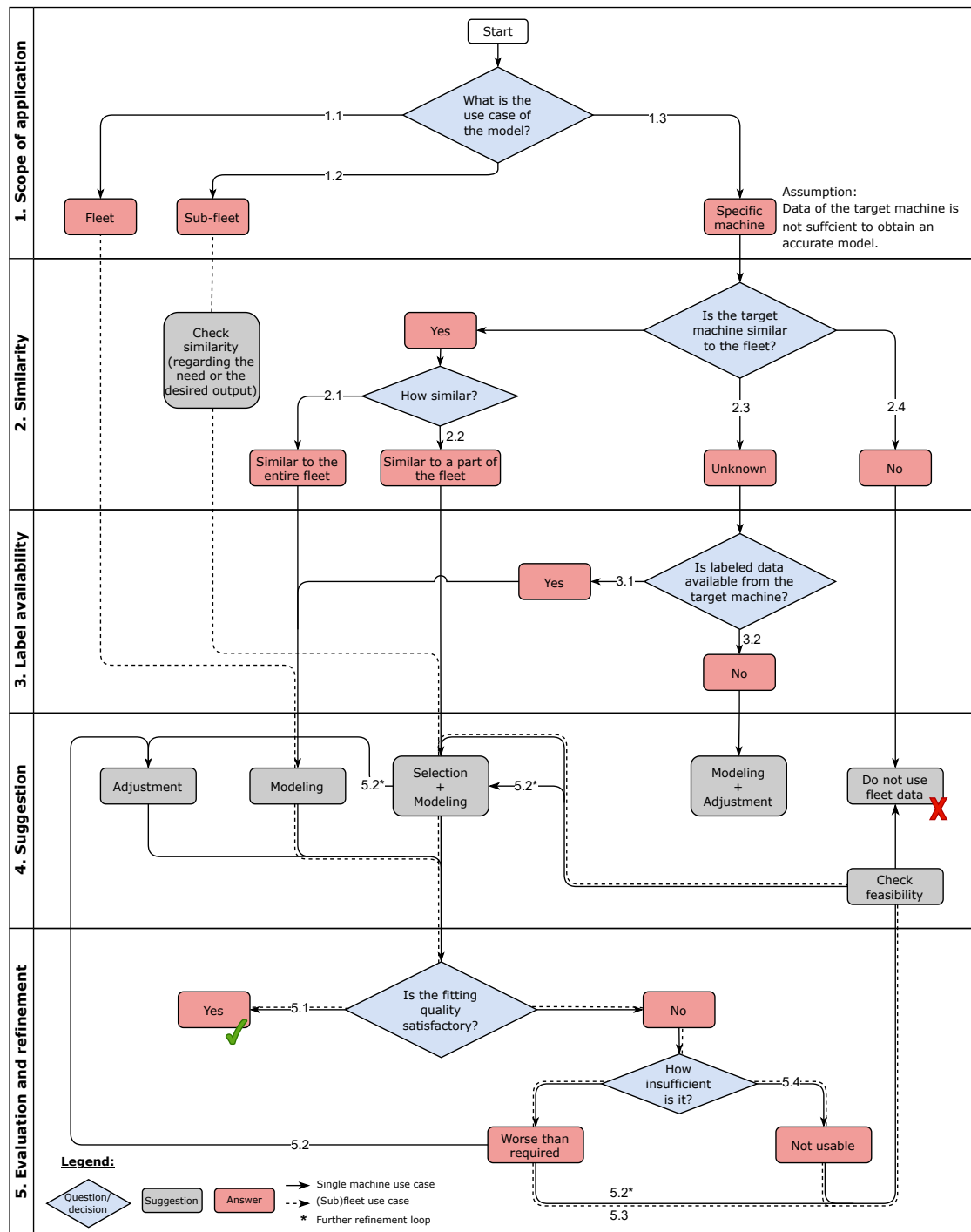


Figure 7.1: Practical guide for incorporating fleet knowledge in industrial applications

part of the fleet, similar machines can be selected in order to use their data to build a corresponding subfleet model (branch 2.2). In cases where the similarity is unknown, the next level can be entered to check the label availability (branch 2.3). However, the use of fleet data is not recommended if it is known that the target machine is not similar to the fleet (e.g., if the machine belongs to a new product generation) to avoid a negative influence from the fleet (branch 2.4).

Level 3: Label availability This level checks whether the target machine has labeled data. This determines whether training a supervised model and model quality quantification are possible. Having at least partly labeled data from the target machine, one can start with building a basic model (branch 3.1) and check the quality of this model using the labeled data in the next level. Otherwise, if the available dataset is completely unlabeled (branch 3.2), only transductive transfer learning is applicable [104]. Among all transductive transfer learning techniques, feature-based transfer learning is the most common [104]. Therefore, incorporating feature-based transfer learning into the modeling process is recommended.

Level 4: Suggestion On this level, first-order suggestions are made based on the previous answers. The suggestions have been justified above. For the modeling part, the full potential of general data-driven techniques such as comprehensive feature engineering and hyperparameter optimization should be fully explored, in order to elaborate the additional benefit of using fleet knowledge.

Level 5: Evaluation and refinement As mentioned at the beginning of this section, the method development is designed as an iterative process. Therefore, the first-order suggestions are followed by an evaluation step, to check the need for further improvement of the method. The resulting fitting quality (e.g., [RMSE](#)) should be compared against the predefined business requirement. If the quality is satisfactory for the required business use cases, no further improvement is needed (branch 5.1). If the quality is unsatisfactory and the modeling scope is the (sub)fleet (branch 5.3), it is perhaps not appropriate to build one overarching model for the (sub)fleet. Therefore, reconsidering the modeling scope is needed, e.g., by downscaling the modeling scope from fleet to subfleet, from subfleet to a specific machine, or by regrouping the subfleets. If the quality is unsatisfactory and the modeling scope is a target machine (branch 5.2), a model adjustment in the form of an additional transfer learning step can be added to address slight fleet-target deviation. The primary recommended method is parameter-based transfer learning as it is the most common inductive transfer learning method [104]. If adding a model adjustment step still provides unsatisfactory modeling quality, one can reconsider which data to use in the modeling process (branch 5.2*), e.g., by enhancing exploratory data analysis to identify machines that

share the most commonalities with the target machine. In case the quality deviates far from the requirements, there is a high risk of negative transfer, i.e., using fleet data might deteriorate the modeling quality. It is necessary to reconsider which data should be used in the modeling process (branch 5.4), just as mentioned above.

All branches leading to the refinement of the selection and modeling steps (5.2*, 5.3, and 5.4) are continuously evaluated for feasibility. This ensures that the refinement process does not get trapped in an infinite loop.

7.3 EXAMPLE OF USING THE GUIDE

Now, the SOH quantification methods developed in this research will be used as examples to navigate through the guide.

Level 1: Scope of application As PEM water electrolyzer is an emerging technology with quick technology iterations, machine-specific models are desired to accommodate the differences among the electrolyzers (Section 2.4). Therefore, branch 1.3 should be followed.

Level 2: Similarity The 12 electrolyzers investigated in this research (Section 3.2.1) have various designs, operating history, and degradation states, resulting in ambiguous similarities. Additionally, to ensure that the developed methods are generally applicable, this research intentionally avoids making assumptions on the similarities. Therefore, branch 2.3 is followed.

Level 3: Label availability Labels are available (e.g., the ground truth of U_{rc} , as shown in Fig. 4.9 and 5.13). Therefore, branch 3.1 is followed.

Level 4: Suggestion Following branch 3.1, it is suggested to begin with constructing a basic model. This corresponds to the benchmark method SLR (Section 3.2.3), which fits the voltage models without any additional treatment.

Level 5: Evaluation and refinement As demonstrated in Sections 4.3 and 5.3, the benchmark method SLR yields inaccurate and unstable results, failing to fulfill the accuracy target of a single-digit-mV error as defined in Section 3.2.2. Despite these shortcomings, there are instances where the SLR results align well with the ground truth (e.g., when the data coverage is sufficient). This suggests that although SLR does not fully meet the requirement, it retains some value. Therefore, branch 5.2 is followed and model adjustment toward the target electrolyzer is conducted. Model adjustment is conducted with either the transfer linear regression algorithm (Chapter 4) or the Bayesian approach (Chapter 5). Both methods yield results fulfilling the accuracy criteria and thus end at branch 5.1.

8

Conclusion

The global movement towards decarbonization is catalyzing an unprecedented surge in the demand for green hydrogen, produced via water electrolysis. PEM water electrolysis is one of the mature technologies. However, its market expansion is hindered by high costs and a relatively short lifetime. Therefore, operators of PEM electrolyzers must closely monitor their SOH to optimize the utilization. Quantifying the SOH of industrial electrolyzers faces several challenges, such as limited testing opportunities and restricted data coverage. The central goal of this research is to provide solutions to quantifying SOH for industrial PEM water electrolyzers.

In this chapter, a brief recap of the research outcome is provided (Section 8.1), followed by a reflection on the developed methods including their limitations (Section 8.2). Finally, Section 8.3 closes this dissertation with an outlook on future research.

8.1 ANSWER TO THE RESEARCH QUESTION

The research question raised in Chapter 1 is:

How to quantify the state-of-health of industrial PEM water electrolyzers?

This question is answered from the following perspectives.

Suitable SOH indicators for industrial PEM water electrolyzers (Section 2.3). Voltage under reference condition U_{rc} can be used as a SOH indicator for industrial PEM water electrolyzers. U_{rc} has the following advantages:

- It is based on voltage, which is an easily measurable quantity in industrial environments and has direct economic relevance.
- It does not require a specific operation mode, but uses statistical methods to correct the voltage measured under arbitrary operating conditions to the reference condition.
- The above-mentioned statistical correction does not rely on any predefined degradation model.

Besides voltage, other **SOH** indicators found in the literature (such as foreign gas concentration and fluoride release rate, Section 2.3.2) also have the potential to be applied to industrial electrolyzers. They provide additional insights into the degradation mechanism. Nevertheless, for these indicators to be viable in industrial environments, the measurement devices must further mature with lower costs, and the impact of operating conditions needs to be corrected, e.g., with the help of the proposed methods in this research.

Suitable fleet-based modeling methods (Section 2.4). The coverage of industrial operation data can be limited, which poses challenges in establishing voltage models. Fleet data can be leveraged to tackle this problem. Six types of fleet-based modeling approaches are identified in the literature: fleet modeling, subfleet modeling, similarity-based modeling, parameter-based transfer learning, feature-based transfer learning, and instance-based/hybrid transfer learning. However, none of these six types fully fulfills the requirements of providing machine-specific models, having an interpretable model structure, and offering the capability for quick updates. Therefore, the following two methods are proposed to address the gap.

A method with transfer linear regression (Section 4). The transfer linear regression algorithm is developed in this research. It builds upon the strengths of parameter-based transfer learning, providing machine-specific models that can be quickly updated through fine-tuning of model parameters. In addition, it features a linear model structure, which is easily interpretable and thus well-suited for industrial applications. This algorithm is used to fit voltage models with operation data in a sequential manner, then the fitted models are used to correct the measured voltage to the reference condition, thereby yielding U_{rc} .

A method with Bayesian inference (Section 5). The second proposed method for **SOH** quantification leverages Bayesian inference and is characterized by the utilization of an aggregated fleet knowledge – the degradation trajectories of multiple electrolyzers are aggregated to a prior probability distribution. This prior probability distribution is then used in Bayesian inference to derive the posterior probability distribution of the parameters in the voltage model. With these inferred parameters, U_{rc} can be calculated.

Customize the method based on the business need (Section 7). The proposed methods are not the only solutions. Depending on the intended application scope (e.g., for the entire electrolyzer fleet or for only one electrolyzer), the fleet characteristic (e.g., how similar are the electrolyzers), the data situation (e.g., whether labels are available), and the accuracy requirement, customizing the method is advisable to achieve a suitable complexity and accuracy in the given context. This research suggests a gen-

eral framework to guide industrial practitioners in developing fleet-based algorithms. Through a cross-method literature review, three core steps can be identified: data selection, model development, and model adjustment. In the proposed practical guide, these steps are integrated in an iterative manner based on the business purpose and accuracy requirement.

8.2 METHODS REFLECTION

This research develops two methods to quantify the [SOH](#) of industrial [PEM](#) water electrolyzers. The highlights of these two methods are: (1) They have simple linear structures that are easy to interpret. (2) The models are updated regularly over time to capture the evolving degradation states of individual electrolyzers. (3) Fleet knowledge is incorporated to assist the model fitting.

The first method is based on the transfer linear regression algorithm, which adjusts a pre-existing fitted voltage model based on new data in a sequential manner. The voltage model is an empirical model developed based on industrial data, which has a linear voltage-current relationship. However, as more industrial data is collected, a new degradation behavior is revealed: The linear voltage-current relationship no longer holds. Another drawback of this method is the need to determine hyperparameters before model fitting. These hyperparameters largely influence the fitted coefficients but their correctness is hard to judge. This poses potential risks for modeling error. In addition, this method relies on the last fitted voltage model. If the new data have a limited coverage, the model tends to stay unchanged. All these limitations motivate a further improvement of the method, which leads to the second proposed method.

The second method uses a revised voltage model, which captures the nonlinear voltage-current relationship. The fleet knowledge is constructed by aggregating the model parameters from multiple electrolyzers. The voltage model is fitted with Bayesian inference and the fleet knowledge is incorporated as the prior probability distribution. In comparison to the first method, this method is based on established algorithms that do not require customized hyperparameters; and it does not rely on the last fitted model, but is supported by the degradation trajectory of other electrolyzers. While the revised voltage model with more unknown coefficients increases the model fitting difficulty, incorporating fleet knowledge with Bayesian inference is proven to effectively capture the ground truth. Sensitivity analyses show that this method is particularly helpful for small data coverage and small fleet-target discrepancy.

The limitations of these two methods have been presented in [Section 4.4](#) and [5.4](#). Besides the method-specific limitations, there is one generic limitation to both methods: Although

they facilitate the monitoring of the temporal evolution of the overall voltage, they do not provide physical insights into the degradation mechanisms. While it may be tempting to attach physical meanings to the model parameters, such as interpreting c_1 in Eq. 4.1 or Eq. 5.3 as the Ohmic resistance, it should be avoided. Due to the empirical nature of these models, they are significantly simplified in comparison to physics-based voltage models (Section 2.2.2.1), and their coefficients do not directly correspond to physical parameters. Consequently, even when the proposed methods are employed to monitor the SOH, opportunities to conduct experimental tests should still be seized, as they are invaluable to obtain a deeper understanding of the degradation processes.

8.3 OUTLOOK

Based on the learnings from this research, an outlook on future research is suggested in this section.

First, closely monitor the degradation behavior and refine the model accordingly. This research (in particular Section 4.4) already demonstrates the necessity of closely tracking the suitability of the voltage model with long-term operation data. As PEM water electrolyzer just recently gained industrial interest, the long-term behavior is not yet well known and is likely design-specific. Therefore, there is no one-size-fits-all solution. Researchers and industrial practitioners must keep learning from the data, to ensure the model's suitability over time.

Second, consider other degradation indicators beyond voltage. As mentioned in Section 2.1, the degradation of PEM water electrolyzers can be observed on a macroscopic level through the increase of voltage, decrease of gas quality, and increase of fluoride release. Considering multiple degradation indicators offers a more comprehensive view of the degradation state. While the proposed methods in this research are applied to quantify the voltage degradation, their fundamental principle – correcting the impact of operating conditions – can be adapted to other degradation indicators (Section 6.4).

Finally, identify degradation contributors with a large amount of industrial data. One research direction is to investigate the temporal evolution of the SOH, to identify operational factors related to degradation. The methods proposed in this research enable tracking the SOH of industrial electrolyzers, which lays a foundation for analyzing long-term degradation patterns and revealing the operational stressors in real-world industrial settings. Another research direction is to explore differences among electrolyzers. Owing to the advancement in data acquisition and management technologies, the industry can now systematically collect lifecycle data of electrolyzers, from procurement to manufacturing and beyond. With this data asset, one can, for example, correlate the degradation behavior

with the manufacturing parameters (such as the pressure for the press belt and the speed for direct coating [180]), which can be a basis for improving the manufacturing process to increase the lifespan of PEM water electrolyzers.

A | Appendix

A.1 DERIVATION OF EQUATION 4.15

The minimization in Eq. 4.14 can be computed by setting the partial derivative of the loss function with respect to γ to zero:

$$\frac{\partial(\|Y - X(\beta + \gamma)\|^2 + \lambda\|X_T\gamma\|^2)}{\partial\gamma} = 0.$$

The loss function can be written as

$$\begin{aligned} & \|Y - X(\beta + \gamma)\|^2 + \lambda\|X_T\gamma\|^2 \\ &= (Y - X(\beta + \gamma))^T(Y - X(\beta + \gamma)) + \lambda(X_T\gamma)^T(X_T\gamma) \\ &= Y^TY - 2(\beta + \gamma)^TX^TY + (\beta + \gamma)^TX^TX(\beta + \gamma) \\ &+ \lambda\gamma^TX_T^TX_T\gamma. \end{aligned}$$

Its derivative with respect to γ is

$$\begin{aligned} & \frac{\partial(\|Y - X(\beta + \gamma)\|^2 + \lambda\|X_T\gamma\|^2)}{\partial\gamma} \\ &= -2X^TY + 2X^TX(\beta + \gamma) + 2\lambda X_T^TX_T\gamma. \end{aligned}$$

Setting it to zero yields

$$\begin{aligned} & -2X^TY + 2X^TX(\beta + \gamma) + 2\lambda X_T^TX_T\gamma = 0 \\ & -X^TY + X^TX\beta + (X^TX + \lambda X_T^TX_T)\gamma = 0 \\ & \gamma = (X^TX + \lambda X_T^TX_T)^{-1}X^T(Y - X\beta) \end{aligned}$$

This is Eq. 4.15 in the main text.

List of Figures

1.1	Global hydrogen demand by sector in the Net Zero Emission scenario [6].	1
1.2	Global electrolyzer capacity by technology based on announced projects [4]. The expected capacity in 2030 includes early-stage projects (e.g., projects with only a stakeholder cooperation agreement).	2
1.3	Comparison between alkaline and PEM electrolysis technologies [9–11].	3
1.4	Dissertation structure.	7
2.1	PEM water electrolyzer cell, stack, and array (source: Siemens Energy).	9
2.2	An exemplary polarization curve with the individual voltage components (adapted from [17]).	16
2.3	Categories of fleet-based algorithms.	24
3.1	Case A: Voltage model fitting and U_{rc} calculation based on arbitrary oper- ation data. Case B: Model fitting ambiguity under limited data coverage. (Note: Voltage is a function of several variables. To simplify the visualization, only the electric current is plotted on the x-axis here.)	30
3.2	The proposed methods involve fitting voltage models to convert measured voltage under arbitrary operating conditions to a reference condition (de- noted by the subscript "rc"). The fitted models can be used for calculating the SOH indicator U_{rc} , polarization curves, and voltage forecast.	32
3.3	Synthetic data generation scheme.	33
4.1	Overview of the SOH quantification method using TLR.	38
4.2	Validation of the empirical model Eq. 4.1 with operation data from three different operation schemes. For confidentiality reasons, the y-scales are hidden.	39
4.3	Illustration for (a) data-enriched linear regression algorithm and (b) transfer linear regression algorithm.	41
4.4	Error of using the synthetic data generation scheme to reproduce industrial data.	46
4.5	Synthetic data.	47

4.6	Different λ s ranging from 10^{-8} to 10^2 are tested for their impact on the coefficient stability (represented with the standard deviations of c_1 to c_4) and on the fitting error.	48
4.7	TLR captures ground truth with significantly smaller errors than the benchmark method SLR.	50
4.8	Different λ s ranging from 10^{-8} to 10^2 are tested for their impact on the coefficient stability (represented with the standard deviations of c_1 to c_4) and on the fitting error.	52
4.9	U_{rc} calculated with SLR (left) and TLR (right) under 3 different reference conditions, compared with measured voltage at roughly these conditions. For confidentiality reasons, the y-scales are hidden.	53
4.10	TLR produces more stable polarization curves than the SLR algorithm. For confidentiality reasons, the x- and y-scales are hidden.	54
4.11	Voltage forecast accuracy of the voltage models fitted with SLR and TLR.	54
4.12	U_{rc} calculated with TLR under 3 different reference conditions, compared with measured voltage at roughly these conditions, over roughly 1200 days.	55
4.13	Fitted voltage model compared with measured polarization curves.	56
4.14	The impact of λ on the model coefficient c_2 and fitting error.	57
4.15	Impact of long constant operation period.	58
5.1	Overview of the SOH quantification methods with Bayesian inference	60
5.2	Comparison between the reversible voltage calculated with the Python package CoolProp and with the linear approximation (Eq. 5.1).	61
5.3	Validation of the proposed empirical model Eq. 5.3 with operation data from three different operation schemes. For confidentiality reasons, the y-scales are hidden.	62
5.4	Process of constructing and incorporating fleet knowledge in model fitting for the target electrolyzer.	64
5.5	Using Bayesian method to obtain the posterior probability distribution of model parameter (c_1 as an example) by combining fleet knowledge and data of the target electrolyzer.	65
5.6	Error of using the synthetic data generation scheme to reproduce industrial data.	66
5.7	Synthetic data of the target electrolyzer.	68
5.8	Demonstration of the modeling process with synthetic fleet data.	69
5.9	Results comparison between the proposed fleet-based method and the benchmark method SLR without using fleet knowledge.	71

5.10	Impact of data coverage.	72
5.11	Impact of the discrepancy between the fleet and the target electrolyzer. . .	74
5.12	Impact of fleet size.	75
5.13	U_{rc} calculated with the benchmark method (without fleet knowledge, left) and with the proposed method (with fleet knowledge, right) under three reference conditions. They are compared with the raw measured data filtered at these conditions.	77
5.14	Polarization curves obtained with the benchmark method (without fleet knowledge, left) and with the proposed method (with fleet knowledge, right). . .	78
5.15	Comparison of polarization curves modeled using the fleet-based method with the measured data.	78
5.16	Voltage forecast accuracy of the voltage models fitted with the benchmark method (left) and the proposed fleet-based method (right).	79
6.1	Different temperature data and their impact on model fitting error RMSE. . .	82
6.2	Different model fitting intervals and their impact on model fitting error RMSE. . .	83
7.1	Practical guide for incorporating fleet knowledge in industrial applications . .	88

List of Tables

2.1 Two types of SOH quantification methods evaluated against requirements:
(a) in-situ measurements for PEM water electrolyzers that are commonly
used in the literature and (b) statistical methods that mitigate the impact
of operation modes. 20

2.2 Fleet-based algorithms evaluated against requirements. Only one exemplary
reference is given for each category. More references can be found in [14]. . 26

4.1 Impacts of setting extreme values for hyperparameters λ and X_T 44

7.1 Core steps of fleet-based algorithms 85

Bibliography

- [1] European Commission: A hydrogen strategy for a climate-neutral Europe (2020). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0301>
- [2] IEA: Global Hydrogen Review 2024, Paris (2024). <https://www.iea.org/reports/global-hydrogen-review-2024>
- [3] International Renewable Energy Agency: Decarbonising hard-to-abate sectors with renewables: Perspectives for the G7, Abu Dhabi (2024). <https://www.irena.org/Publications/2024/Apr/Decarbonising-hard-to-abate-sectors-with-renewables-Perspectives-for-the-G7>
- [4] IEA: Global Hydrogen Review 2023, Paris (2023). <https://www.iea.org/reports/global-hydrogen-review-2023>
- [5] IEA: Net Zero Roadmap: A Global Pathway to Keep the 1.5°C Goal in Reach, Paris (2023). <https://www.iea.org/reports/net-zero-roadmap-a-global-pathway-to-keep-the-15-0c-goal-in-reach>
- [6] IEA: Global hydrogen demand in the Net Zero Scenario, 2022-2050 (2023). <https://www.iea.org/data-and-statistics/charts/global-hydrogen-demand-in-the-net-zero-scenario-2022-2050>
- [7] IEA: Global Hydrogen Review 2021, Paris (2021). <https://www.iea.org/reports/global-hydrogen-review-2021>
- [8] IEA: Energy Statistics Data Browser (2023). <https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser>
- [9] Guo, Y., Li, G., Zhou, J., Liu, Y.: Comparison between hydrogen production by alkaline water electrolysis and hydrogen production by PEM electrolysis. IOP Conference Series: Earth and Environmental Science **371**(4), 042022 (2019). doi:[10.1088/1755-1315/371/4/042022](https://doi.org/10.1088/1755-1315/371/4/042022)

- [10] Smolinka, T., Wiebe, N., Sterchele, P., Palzer, A., Lehner, F., Jansen, M., Kiemel, S., Mieke, R., Wahren, S., Zimmermann, F.: Study IndWEDe–Brief Overview: Industrialisation of water electrolysis in Germany: Opportunities and challenges for sustainable hydrogen for transport, electricity and heat (2018). <https://www.ise.fraunhofer.de/en/publications/studies/indwede.html>
- [11] Buttler, A., Spliethoff, H.: Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review. *Renewable and Sustainable Energy Reviews* **82**, 2440–2454 (2018). doi:10.1016/j.rser.2017.09.003
- [12] Yan, X., Locci, C., Hiss, F., Nieße, A.: State-of-health estimation for industrial H₂ electrolyzers with transfer linear regression. *Energies* **17**(6), 1374 (2024). doi:10.3390/en17061374
- [13] Yan, X., Helmers, L., Zhou, K., Nieße, A.: Fleet-based degradation state quantification for industrial water electrolyzers. *Electrochemical Science Advances* **5**(3), 70002 (2025). doi:10.1002/elsa.70002
- [14] Yan, X., Woelke, J., Bensmann, B., Eckert, C., Hanke-Rauschenbach, R., Nieße, A.: Cross-method overview of fleet-based machine health estimation and prediction: A practical guide for industrial applications. *IEEE Access* **13**, 60131–60147 (2025). doi:10.1109/access.2025.3556251
- [15] Yan, X., Helmers, L., Lettenmeier, P.: Fleet-based performance model and degradation estimation. In: *Book of Abstracts of the 4th International Conference on Electrolysis*, p. 160 (2023). <https://engineering.nwu.ac.za/hysa/book-abstracts>
- [16] Smolinka, T., Ojong, E.T., Garche, J.: Hydrogen production from renewable energies—electrolyzer technologies. In: *Electrochemical Energy Storage for Renewable Sources and Grid Balancing*, pp. 103–128. Elsevier, (2015). doi:10.1016/b978-0-444-62616-5.00008-5
- [17] Smolinka, T., Ojong, E.T., Lickert, T.: Fundamentals of PEM water electrolysis. In: *PEM Electrolysis for Hydrogen Production: Principles and Applications*, pp. 11–33. CRC Press, (2015). doi:10.1201/b19096-6
- [18] Feng, Q., Liu, G., Wei, B., Zhang, Z., Li, H., Wang, H., *et al.*: A review of proton exchange membrane water electrolysis on degradation mechanisms and mitigation strategies. *Journal of Power Sources* **366**, 33–55 (2017). doi:10.1016/j.jpowsour.2017.09.006

- [19] Paciok, P., Schalenbach, M., Carmo, M., Stolten, D.: On the mobility of carbon-supported platinum nanoparticles towards unveiling cathode degradation in water electrolysis. *Journal of power sources* **365**, 53–60 (2017). doi:[10.1016/j.jpowsour.2017.07.033](https://doi.org/10.1016/j.jpowsour.2017.07.033)
- [20] Lettenmeier, P., Wang, R., Abouatallah, R., Helmly, S., Morawietz, T., Hiesgen, R., Kolb, S., Burggraf, F., Kallo, J., Gago, A.S., Friedrich, K.A.: Durable membrane electrode assemblies for proton exchange membrane electrolyzer systems operating at high current densities. *Electrochimica Acta* **210**, 502–511 (2016). doi:[10.1016/j.electacta.2016.04.164](https://doi.org/10.1016/j.electacta.2016.04.164)
- [21] Olivier, P., Bourasseau, C., Bouamama, P.B.: Low-temperature electrolysis system modelling: A review. *Renewable and Sustainable Energy Reviews* **78**, 280–300 (2017). doi:[10.1016/j.rser.2017.03.099](https://doi.org/10.1016/j.rser.2017.03.099)
- [22] Abomazid, A.M., El-Taweel, N.A., Farag, H.E.: Novel analytical approach for parameters identification of PEM electrolyzer. *IEEE Transactions on Industrial Informatics* **18**(9), 5870–5881 (2021). doi:[10.1109/tii.2021.3132941](https://doi.org/10.1109/tii.2021.3132941)
- [23] Nafeh, A.E.-S.A.: Hydrogen production from a PV/PEM electrolyzer system using a neural-network-based MPPT algorithm. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* **24**(3), 282–297 (2011). doi:[10.1002/jnm.778](https://doi.org/10.1002/jnm.778)
- [24] Krenz, T., Weiland, O., Trinke, P., Helmers, L., Eckert, C., Bensmann, B., Hanke-Rauschenbach, R.: Temperature and performance inhomogeneities in PEM electrolysis stacks with industrial scale cells. *Journal of The Electrochemical Society* **170**(4), 044508 (2023). doi:[10.1149/1945-7111/accb68](https://doi.org/10.1149/1945-7111/accb68)
- [25] Lebbal, M., Lecœuche, S.: Identification and monitoring of a PEM electrolyser based on dynamical modelling. *International journal of hydrogen energy* **34**(14), 5992–5999 (2009). doi:[10.1016/j.ijhydene.2009.02.003](https://doi.org/10.1016/j.ijhydene.2009.02.003)
- [26] Marangio, F., Santarelli, M., Cali, M.: Theoretical model and experimental analysis of a high pressure PEM water electrolyser for hydrogen production. *International journal of hydrogen energy* **34**(3), 1143–1158 (2009). doi:[10.1016/j.ijhydene.2008.11.083](https://doi.org/10.1016/j.ijhydene.2008.11.083)
- [27] Trinke, P., Haug, P., Brauns, J., Bensmann, B., Hanke-Rauschenbach, R., Turek, T.: Hydrogen crossover in PEM and alkaline water electrolysis: mechanisms, direct comparison and mitigation strategies. *Journal of The Electrochemical Society* **165**(7), 502 (2018). doi:[10.1149/2.0541807jes](https://doi.org/10.1149/2.0541807jes)

- [28] Nie, J., Chen, Y., Cohen, S., Carter, B.D., Boehm, R.F.: Numerical and experimental study of three-dimensional fluid flow in the bipolar plate of a PEM electrolysis cell. *International journal of thermal sciences* **48**(10), 1914–1922 (2009). doi:[10.1016/j.ijthermalsci.2009.02.017](https://doi.org/10.1016/j.ijthermalsci.2009.02.017)
- [29] Grigoriev, S., Kalinnikov, A., Millet, P., Porembsky, V., Fateev, V.: Mathematical modeling of high-pressure PEM water electrolysis. *Journal of applied electrochemistry* **40**, 921–932 (2010). doi:[10.1007/s10800-009-0031-z](https://doi.org/10.1007/s10800-009-0031-z)
- [30] Bensmann, B., Rex, A., Hanke-Rauschenbach, R.: An engineering perspective on the future role of modelling in proton exchange membrane water electrolysis development. *Current Opinion in Chemical Engineering* **36**, 100829 (2022). doi:[10.1016/j.coche.2022.100829](https://doi.org/10.1016/j.coche.2022.100829)
- [31] Bensmann, B., Hanke-Rauschenbach, R., Müller-Syring, G., Henel, M., Sundmacher, K.: Optimal configuration and pressure levels of electrolyzer plants in context of power-to-gas applications. *Applied energy* **167**, 107–124 (2016). doi:[10.1016/j.apenergy.2016.01.038](https://doi.org/10.1016/j.apenergy.2016.01.038)
- [32] Bensmann, B., Hanke-Rauschenbach, R.: Engineering modeling of PEM water electrolysis: a survey. *ECS Transactions* **75**(14), 1065 (2016). doi:[10.1149/07514.1065ecst](https://doi.org/10.1149/07514.1065ecst)
- [33] Rogler, M., Suermann, M., Wagner, R., Thiele, S., Straub, J.: Advanced method for voltage breakdown analysis of PEM water electrolysis cells with low iridium loadings. *Journal of The Electrochemical Society* **170**(11), 114521 (2023). doi:[10.1149/1945-7111/ad0b74](https://doi.org/10.1149/1945-7111/ad0b74)
- [34] Fritz, D.L., Mergel, J., Stolten, D.: PEM electrolysis simulation and validation. *ECS Transactions* **58**(19), 1 (2014). doi:[10.1149/05819.0001ecst](https://doi.org/10.1149/05819.0001ecst)
- [35] Dörner, S., Kinzl, M., Mitsos, A., Bongartz, D.: Dynamic lumped cell-level model of chemical membrane degradation in PEM electrolysis: Impact of pressure and time dependence (2023). doi:[10.1149/osf.io/ea8qmq](https://doi.org/10.1149/osf.io/ea8qmq)
- [36] Kalinnikov, A.A., Grigoriev, S.A., Bessarabov, D.G.: Numerical analysis of the electrochemical dissolution of iridium catalyst and evaluation of its effect on the performance of polymer electrolyte membrane water electrolyzers. *International Journal of Hydrogen Energy* **48**(59), 22342–22365 (2023). doi:[10.1016/j.ijhydene.2023.03.136](https://doi.org/10.1016/j.ijhydene.2023.03.136)
- [37] Falcão, D., Pinto, A.: A review on PEM electrolyzer modelling: Guidelines for beginners. *Journal of Cleaner Production* **261**, 121184 (2020). doi:[10.1016/j.jclepro.2020.121184](https://doi.org/10.1016/j.jclepro.2020.121184)

- [38] Yigit, T., Selamet, O.F.: Mathematical modeling and dynamic simulink simulation of high-pressure PEM electrolyzer system. *International Journal of Hydrogen Energy* **41**(32), 13901–13914 (2016). doi:[10.1016/j.ijhydene.2016.06.022](https://doi.org/10.1016/j.ijhydene.2016.06.022)
- [39] Krenz, T., Gottschalk, T., Helmers, L., Trinke, P., Bensmann, B., Hanke-Rauschenbach, R.: Current interrupt technique to fully characterize PEMWE cells. *Journal of The Electrochemical Society* **171**(3), 034509 (2024). doi:[10.1149/1945-7111/ad3057](https://doi.org/10.1149/1945-7111/ad3057)
- [40] Atlam, O., Kolhe, M.: Equivalent electrical model for a proton exchange membrane (PEM) electrolyser. *Energy Conversion and management* **52**(8-9), 2952–2957 (2011). doi:[10.1016/j.enconman.2011.04.007](https://doi.org/10.1016/j.enconman.2011.04.007)
- [41] Sossan, F., Bindner, H., Madsen, H., Torregrossa, D., Chamorro, L.R., Paolone, M.: A model predictive control strategy for the space heating of a smart building including cogeneration of a fuel cell-electrolyzer system. *International Journal of Electrical Power & Energy Systems* **62**, 879–889 (2014). doi:[10.1016/j.ijepes.2014.05.040](https://doi.org/10.1016/j.ijepes.2014.05.040)
- [42] Chavez-Ramirez, A.U., Munoz-Guerrero, R., Sanchez-Huerta, V., Ramirez-Arredondo, J.M., Ornelas, R., Arriaga, L.G., Siracusano, S., Brunaccini, G., Napoli, G., Antonucci, V., Arico, A.S.: Dynamic model of a PEM electrolyzer based on artificial neural networks. *Journal of New Materials for Electrochemical Systems* **14**(2), 113–119 (2011). doi:[10.14447/jnmes.v14i2.119](https://doi.org/10.14447/jnmes.v14i2.119)
- [43] Mohamed, A., Ibrahim, H., Kim, K.: Machine learning-based simulation for proton exchange membrane electrolyzer cell. *Energy Reports* **8**, 13425–13437 (2022). doi:[10.1016/j.egyr.2022.09.135](https://doi.org/10.1016/j.egyr.2022.09.135)
- [44] Chen, X., Rex, A., Woelke, J., Eckert, C., Bensmann, B., Hanke-Rauschenbach, R., Geyer, P.: Machine learning in proton exchange membrane water electrolysis—a knowledge-integrated framework. *Applied Energy* **371**, 123550 (2024). doi:[10.1016/j.apenergy.2024.123550](https://doi.org/10.1016/j.apenergy.2024.123550)
- [45] Carmo, M., Fritz, D.L., Mergel, J., Stolten, D.: A comprehensive review on pem water electrolysis. *International journal of hydrogen energy* **38**(12), 4901–4934 (2013). doi:[10.1016/j.ijhydene.2013.01.151](https://doi.org/10.1016/j.ijhydene.2013.01.151)
- [46] Millet, P.: Characterization tools for polymer electrolyte membrane (PEM) water electrolyzers. In: *PEM Electrolysis for Hydrogen Production*, pp. 191–230. CRC Press, (2016). doi:[10.1201/b19096-14](https://doi.org/10.1201/b19096-14)

- [47] García-Valverde, R., Espinosa, N., Urbina, A.: Simple PEM water electrolyser model and experimental validation. *international journal of hydrogen energy* **37**(2), 1927–1938 (2012). doi:[10.1016/j.ijhydene.2011.09.027](https://doi.org/10.1016/j.ijhydene.2011.09.027)
- [48] Görgün, H.: Dynamic modelling of a proton exchange membrane (PEM) electrolyzer. *International journal of hydrogen energy* **31**(1), 29–38 (2006). doi:[10.1016/j.ijhydene.2005.04.001](https://doi.org/10.1016/j.ijhydene.2005.04.001)
- [49] Bühre, L.V., Bullerdiek, S., Trinke, P., Bensmann, B., Deutsch, A.-L.E., Behrens, P., Hanke-Rauschenbach, R.: Application and analysis of a salt bridge reference electrode setup for PEM water electrolysis: towards an extended voltage loss break down. *Journal of The Electrochemical Society* **169**(12), 124513 (2022). doi:[10.1149/1945-7111/ac9ee1](https://doi.org/10.1149/1945-7111/ac9ee1)
- [50] Järvinen, L., Puranen, P., Kosonen, A., Ruuskanen, V., Ahola, J., Kauranen, P., Hehemann, M.: Automized parametrization of PEM and alkaline water electrolyzer polarisation curves. *International Journal of Hydrogen Energy* **47**(75), 31985–32003 (2022). doi:[10.1016/j.ijhydene.2022.07.085](https://doi.org/10.1016/j.ijhydene.2022.07.085)
- [51] Han, B., Steen III, S.M., Mo, J., Zhang, F.-Y.: Electrochemical performance modeling of a proton exchange membrane electrolyzer cell for hydrogen energy. *International Journal of Hydrogen Energy* **40**(22), 7006–7016 (2015). doi:[10.1016/j.ijhydene.2015.03.164](https://doi.org/10.1016/j.ijhydene.2015.03.164)
- [52] Abdin, Z., Webb, C., Gray, E.M.: Modelling and simulation of a proton exchange membrane (PEM) electrolyser cell. *International journal of hydrogen energy* **40**(39), 13243–13257 (2015). doi:[10.1016/j.ijhydene.2015.07.129](https://doi.org/10.1016/j.ijhydene.2015.07.129)
- [53] Kim, H., Park, M., Lee, K.S.: One-dimensional dynamic modeling of a high-pressure water electrolysis system for hydrogen production. *International Journal of Hydrogen Energy* **38**(6), 2596–2609 (2013). doi:[10.1016/j.ijhydene.2012.12.006](https://doi.org/10.1016/j.ijhydene.2012.12.006)
- [54] Maier, M., Smith, K., Dodwell, J., Hinds, G., Shearing, P., Brett, D.: Mass transport in PEM water electrolyzers: A review. *International Journal of Hydrogen Energy* **47**(1), 30–56 (2022). doi:[10.1016/j.ijhydene.2021.10.013](https://doi.org/10.1016/j.ijhydene.2021.10.013)
- [55] Trinke, P., Keeley, G.P., Carmo, M., Bensmann, B., Hanke-Rauschenbach, R.: Elucidating the effect of mass transport resistances on hydrogen crossover and cell performance in pem water electrolyzers by varying the cathode ionomer content. *Journal of The Electrochemical Society* **166**(8), 465–471 (2019). doi:[10.1149/2.0171908jes](https://doi.org/10.1149/2.0171908jes)

- [56] Malkow, T., Pilenga, A., Tsotridis, G., De Marco, G.: EU Harmonised Polarisation Curve Test Method for Low-temperature Water Electrolysis. Publications Office of the European Union, (2018). doi:[10.2760/179509](https://doi.org/10.2760/179509)
- [57] Li, N., Araya, S.S., Kær, S.K.: Long-term contamination effect of iron ions on cell performance degradation of proton exchange membrane water electrolyser. *Journal of Power Sources* **434**, 226755 (2019). doi:[10.1016/j.jpowsour.2019.226755](https://doi.org/10.1016/j.jpowsour.2019.226755)
- [58] Weiß, A., Siebel, A., Bernt, M., Shen, T.-H., Tileli, V., Gasteiger, H.: Impact of intermittent operation on lifetime and performance of a PEM water electrolyzer. *Journal of the electrochemical society* **166**(8), 487 (2019). doi:[10.1149/ma2018-02/46/1606](https://doi.org/10.1149/ma2018-02/46/1606)
- [59] Suermann, M., Bensmann, B., Hanke-Rauschenbach, R.: Degradation of proton exchange membrane (PEM) water electrolysis cells: looking beyond the cell voltage increase. *Journal of The Electrochemical Society* **166**(10), 645 (2019). doi:[10.1149/2.1451910jes](https://doi.org/10.1149/2.1451910jes)
- [60] Bernt, M., Siebel, A., Gasteiger, H.A.: Analysis of voltage losses in PEM water electrolyzers with low platinum group metal loadings. *Journal of The Electrochemical Society* **165**(5), 305–314 (2018). doi:[10.1149/ma2017-01/30/1382](https://doi.org/10.1149/ma2017-01/30/1382)
- [61] Ginsberg, M.J., Venkatraman, M., Esposito, D.V., Fthenakis, V.M.: Minimizing the cost of hydrogen production through dynamic polymer electrolyte membrane electrolyzer operation. *Cell Reports Physical Science* **3**(6) (2022). doi:[10.1016/j.xcrp.2022.100935](https://doi.org/10.1016/j.xcrp.2022.100935)
- [62] Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V.: Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on knowledge and data engineering* **29**(10), 2318–2331 (2017). doi:[10.1109/tkde.2017.2720168](https://doi.org/10.1109/tkde.2017.2720168)
- [63] Zhou, T., Francois, B.: Modeling and control design of hydrogen production process for an active hydrogen/wind hybrid power system. *International journal of hydrogen energy* **34**(1), 21–30 (2009). doi:[10.1016/j.ijhydene.2008.10.030](https://doi.org/10.1016/j.ijhydene.2008.10.030)
- [64] Ulleberg, Ø.: Modeling of advanced alkaline electrolyzers: a system simulation approach. *International journal of hydrogen energy* **28**(1), 21–33 (2003). doi:[10.1016/s0360-3199\(02\)00033-2](https://doi.org/10.1016/s0360-3199(02)00033-2)
- [65] Hayatzadeh, A., Fattahi, M., Rezaveisi, A.: Machine learning algorithms for operating parameters predictions in proton exchange membrane water electrolyzers: An-

- ode side catalyst. *International Journal of Hydrogen Energy* **56**, 302–314 (2024). doi:[10.1016/j.ijhydene.2023.12.149](https://doi.org/10.1016/j.ijhydene.2023.12.149)
- [66] Xue, B., Xu, H., Huang, X., Zhu, K., Xu, Z., Pei, H.: Similarity-based prediction method for machinery remaining useful life: A review. *The International Journal of Advanced Manufacturing Technology* **121**(3-4), 1501–1531 (2022). doi:[10.1007/s00170-022-09280-3](https://doi.org/10.1007/s00170-022-09280-3)
- [67] Zhang, Y., Wik, T., Bergström, J., Pecht, M., Zou, C.: A machine learning-based framework for online prediction of battery ageing trajectory and lifetime using histogram data. *Journal of Power Sources* **526**, 231110 (2022). doi:[10.1016/j.jpowsour.2022.231110](https://doi.org/10.1016/j.jpowsour.2022.231110)
- [68] Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., Fraggedakis, D., *et al.*: Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy* **4**(5), 383–391 (2019). doi:[10.1002/eom2.12213](https://doi.org/10.1002/eom2.12213)
- [69] Shu, X., Shen, J., Li, G., Zhang, Y., Chen, Z., Liu, Y.: A flexible state-of-health prediction scheme for lithium-ion battery packs with long short-term memory network and transfer learning. *IEEE Transactions on Transportation Electrification* **7**(4), 2238–2248 (2021). doi:[10.1109/tte.2021.3074638](https://doi.org/10.1109/tte.2021.3074638)
- [70] Che, Y., Deng, Z., Lin, X., Hu, L., Hu, X.: Predictive battery health management with transfer learning and online model correction. *IEEE Transactions on Vehicular Technology* **70**(2), 1269–1277 (2021). doi:[10.1109/tvt.2021.3055811](https://doi.org/10.1109/tvt.2021.3055811)
- [71] Sønderskov, S.H., Ilsøe, J.W., Rasmussen, D., Blom-Hansen, D., Munk-Nielsen, S.: State of health estimation and prediction of fuel cell stacks in backup power systems (2019). doi:[10.31224/osf.io/6b75s](https://doi.org/10.31224/osf.io/6b75s)
- [72] Zhang, D., Baraldi, P., Cadet, C., Yousfi-Steiner, N., Bérenguer, C., Zio, E.: An ensemble of models for integrating dependent sources of information for the prognosis of the remaining useful life of proton exchange membrane fuel cells. *Mechanical Systems and Signal Processing* **124**, 479–501 (2019). doi:[10.1016/j.ymssp.2019.01.060](https://doi.org/10.1016/j.ymssp.2019.01.060)
- [73] Olsson, T., Ramentol, E., Rahman, M., Oostveen, M., Kyprianidis, K.: A data-driven approach for predicting long-term degradation of a fleet of micro gas turbines. *Energy and AI* **4**, 100064 (2021). doi:[10.1016/j.egyai.2021.100064](https://doi.org/10.1016/j.egyai.2021.100064)

- [74] Mao, W., He, J., Sun, B., Wang, L.: Prediction of bearings remaining useful life across working conditions based on transfer learning and time series clustering. *IEEE Access* **9**, 135285–135303 (2021). doi:[10.1109/access.2021.3117002](https://doi.org/10.1109/access.2021.3117002)
- [75] Zhao, H., Liu, H., Jin, Y., Dang, X., Deng, W.: Feature extraction for data-driven remaining useful life prediction of rolling bearings. *IEEE Transactions on Instrumentation and Measurement* **70**, 1–10 (2021). doi:[10.1109/tim.2021.3059500](https://doi.org/10.1109/tim.2021.3059500)
- [76] Yang, X., Fang, Z., Li, X., Yang, Y., Mba, D.: Similarity-based information fusion grey model for remaining useful life prediction of aircraft engines. *Grey Systems: Theory and Application* **11**(3), 463–483 (2021). doi:[10.1108/gs-05-2020-0066](https://doi.org/10.1108/gs-05-2020-0066)
- [77] Huang, C.-G., Huang, H.-Z., Peng, W., Huang, T.: Improved trajectory similarity-based approach for turbofan engine prognostics. *Journal of Mechanical Science and Technology* **33**, 4877–4890 (2019). doi:[10.1007/s12206-019-0928-3](https://doi.org/10.1007/s12206-019-0928-3)
- [78] Liu, Y., Hu, X., Zhang, W.: Remaining useful life prediction based on health index similarity. *Reliability Engineering & System Safety* **185**, 502–510 (2019). doi:[10.1016/j.ress.2019.02.002](https://doi.org/10.1016/j.ress.2019.02.002)
- [79] Wang, H., Ni, G., Chen, J., Qu, J.: Research on rolling bearing state health monitoring and life prediction based on pca and internet of things with multi-sensor. *Measurement* **157**, 107657 (2020). doi:[10.1016/j.measurement.2020.107657](https://doi.org/10.1016/j.measurement.2020.107657)
- [80] Yu, W., Kim, I.Y., Mechefske, C.: Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme. *Mechanical Systems and Signal Processing* **129**, 764–780 (2019). doi:[10.1016/j.ymssp.2019.05.005](https://doi.org/10.1016/j.ymssp.2019.05.005)
- [81] Yu, W., Kim, I.Y., Mechefske, C.: Analysis of different rnn autoencoder variants for time series classification and machine prognostics. *Mechanical Systems and Signal Processing* **149**, 107322 (2021). doi:[10.1016/j.ymssp.2020.107322](https://doi.org/10.1016/j.ymssp.2020.107322)
- [82] Alia, S.M., Stariha, S., Borup, R.L.: Electrolyzer durability at low catalyst loading and with dynamic operation. *Journal of The Electrochemical Society* **166**(15), 1164–1172 (2019). doi:[10.1149/2.0231915jes](https://doi.org/10.1149/2.0231915jes)
- [83] Siracusano, S., Van Dijk, N., Backhouse, R., Merlo, L., Baglio, V., Aricò, A.: Degradation issues of PEM electrolysis MEAs. *Renewable Energy* **123**, 52–57 (2018). doi:[10.1016/j.renene.2018.02.024](https://doi.org/10.1016/j.renene.2018.02.024)

- [84] Martinson, C.A., Van Schoor, G., Uren, K., Bessarabov, D.: Characterisation of a PEM electrolyser using the current interrupt method. *International journal of hydrogen energy* **39**(36), 20865–20878 (2014). doi:[10.1016/j.ijhydene.2014.09.153](https://doi.org/10.1016/j.ijhydene.2014.09.153)
- [85] Voronova, A., Kim, S., Kim, D., Park, H.-Y., Jang, J.H., Seo, B.: Systematic degradation analysis in renewable energy-powered proton exchange membrane water electrolysis. *Energy & Environmental Science* **16**(11), 5170–5184 (2023). doi:[10.1039/d3ee01959d](https://doi.org/10.1039/d3ee01959d)
- [86] Frensch, S.H., Fouda-Onana, F., Serre, G., Thoby, D., Araya, S.S., Kær, S.K.: Influence of the operation mode on PEM water electrolysis degradation. *International Journal of Hydrogen Energy* **44**(57), 29889–29898 (2019). doi:[10.1016/j.ijhydene.2019.09.169](https://doi.org/10.1016/j.ijhydene.2019.09.169)
- [87] Chandesris, M., Médeau, V., Guillet, N., Chelghoum, S., Thoby, D., Fouda-Onana, F.: Membrane degradation in PEM water electrolyzer: Numerical modeling and experimental evidence of the influence of temperature and current density. *International Journal of Hydrogen Energy* **40**(3), 1353–1366 (2015). doi:[10.1016/j.ijhydene.2014.11.111](https://doi.org/10.1016/j.ijhydene.2014.11.111)
- [88] Kuhnert, E., Heidinger, M., Bernroither, A., Kiziltan, Ö., Berger, E., Hacker, V., Bodner, M.: Fluoride emission rate analysis in proton exchange membrane water electrolyzer cells. *Frontiers in Energy Research* **12**, 1457310 (2024). doi:[10.3389/fenrg.2024.1457310](https://doi.org/10.3389/fenrg.2024.1457310)
- [89] Van der Merwe, J., Uren, K., Van Schoor, G., Bessarabov, D.: Characterisation tools development for PEM electrolyzers. *International Journal of Hydrogen Energy* **39**(26), 14212–14221 (2014). doi:[10.1016/j.ijhydene.2014.02.096](https://doi.org/10.1016/j.ijhydene.2014.02.096)
- [90] Kozlova, M., Pushkareva, I., Butrim, S., Solovyev, M., Simkin, D., Grigoriev, S., Pushkarev, A.: Investigation of the degradation of the membrane electrode assembly for a proton exchange membrane water electrolyzer by accelerated stress tests. *Nanobiotechnology Reports* **18**(Suppl 2), 375–388 (2023). doi:[10.1134/s2635167624600135](https://doi.org/10.1134/s2635167624600135)
- [91] Bensmann, B., Hanke-Rauschenbach, R., Sundmacher, K.: In-situ measurement of hydrogen crossover in polymer electrolyte membrane water electrolysis. *International journal of hydrogen energy* **39**(1), 49–53 (2014). doi:[10.1016/j.ijhydene.2013.10.085](https://doi.org/10.1016/j.ijhydene.2013.10.085)
- [92] Bressel, M., Hilairet, M., Hissel, D., Bouamama, B.O.: Remaining useful life prediction and uncertainty quantification of proton exchange membrane fuel cell under

- variable load. *IEEE Transactions on Industrial Electronics* **63**(4), 2569–2577 (2016). doi:[10.1109/tie.2016.2519328](https://doi.org/10.1109/tie.2016.2519328)
- [93] Ma, R., Xie, R., Xu, L., Huangfu, Y., Li, Y.: A hybrid prognostic method for PEMFC with aging parameter prediction. *IEEE Transactions on Transportation Electrification* **7**(4), 2318–2331 (2021). doi:[10.1109/tte.2021.3075531](https://doi.org/10.1109/tte.2021.3075531)
- [94] Lv, J., Yu, Z., Zhang, H., Sun, G., Muhl, P., Liu, J.: Transformer based long-term prognostics for dynamic operating PEM fuel cells. *IEEE Transactions on Transportation Electrification* **10**(1), 1747–1757 (2023). doi:[10.1109/tte.2023.3266803](https://doi.org/10.1109/tte.2023.3266803)
- [95] Hua, Z., Zheng, Z., Pahon, E., Péra, M.-C., Gao, F.: Multi-timescale lifespan prediction for PEMFC systems under dynamic operating conditions. *IEEE Transactions on Transportation Electrification* **8**(1), 345–355 (2021). doi:[10.1109/tte.2021.3103921](https://doi.org/10.1109/tte.2021.3103921)
- [96] Remmlinger, J., Buchholz, M., Meiler, M., Bernreuter, P., Dietmayer, K.: State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation. *Journal of power sources* **196**(12), 5357–5363 (2011). doi:[10.1016/j.jpowsour.2010.08.035](https://doi.org/10.1016/j.jpowsour.2010.08.035)
- [97] Asudeh, A., Jin, Z., Jagadish, H.: Assessing and remedying coverage for a given dataset. In: 2019 IEEE 35th International Conference on Data Engineering (ICDE), pp. 554–565. IEEE, (2019). doi:[10.1109/icde.2019.00056](https://doi.org/10.1109/icde.2019.00056)
- [98] Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *Journal of big data* **6**(1), 1–48 (2019). doi:[10.1186/s40537-019-0197-0](https://doi.org/10.1186/s40537-019-0197-0)
- [99] Wen, Q., Sun, L., Yang, F., Song, X., Gao, J., Wang, X., Xu, H.: Time series data augmentation for deep learning: A survey, 4653–4660 (2021). doi:[10.24963/ijcai.2021/631](https://doi.org/10.24963/ijcai.2021/631)
- [100] Demir, S., Mincev, K., Kok, K., Paterakis, N.G.: Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting. *Applied Energy* **304**, 117695 (2021). doi:[10.1016/j.apenergy.2021.117695](https://doi.org/10.1016/j.apenergy.2021.117695)
- [101] Wang, Y., Yao, Q., Kwok, J.T., Ni, L.M.: Generalizing from a few examples: A survey on few-shot learning. *ACM computing surveys* **53**(3), 1–34 (2020). doi:[10.1145/3386252](https://doi.org/10.1145/3386252)
- [102] Medina-Oliva, G., Voisin, A., Monnin, M., Leger, J.-B.: Predictive diagnosis based on a fleet-wide ontology approach. *Knowledge-Based Systems* **68**, 40–57 (2014). doi:[10.1016/j.knosys.2013.12.020](https://doi.org/10.1016/j.knosys.2013.12.020)

- [103] Wagner, C., Hellingrath, B.: Fleet knowledge for prognostics and health management—identifying fleet dimensions and characteristics for the categorization of fleets. In: Annual Conference of the PHM Society, vol. 9 (2017). doi:[10.36001/phmconf.2017.v9i1.2395](https://doi.org/10.36001/phmconf.2017.v9i1.2395)
- [104] Braig, M., Zeiler, P.: Using data from similar systems for data-driven condition diagnosis and prognosis of engineering systems: A review and an outline of future research challenges. IEEE Access (2022). doi:[10.1109/access.2022.3233220](https://doi.org/10.1109/access.2022.3233220)
- [105] Hou, M., Pi, D., Li, B.: Similarity-based deep learning approach for remaining useful life prediction. Measurement **159**, 107788 (2020). doi:[10.1016/j.measurement.2020.107788](https://doi.org/10.1016/j.measurement.2020.107788)
- [106] Cai, H., Jia, X., Feng, J., Li, W., Pahren, L., Lee, J.: A similarity based methodology for machine prognostics by using kernel two sample test. ISA transactions **103**, 112–121 (2020). doi:[10.1016/j.isatra.2020.03.007](https://doi.org/10.1016/j.isatra.2020.03.007)
- [107] Lu, J., Xiong, R., Tian, J., Wang, C., Sun, F.: Deep learning to estimate lithium-ion battery state of health without additional degradation experiments. Nature Communications **14**(1), 2760 (2023). doi:[10.1038/s41467-023-38458-w](https://doi.org/10.1038/s41467-023-38458-w)
- [108] Ma, J., Shang, P., Zou, X., Ma, N., Ding, Y., Sun, J., Cheng, Y., Tao, L., Lu, C., Su, Y., *et al.*: A hybrid transfer learning scheme for remaining useful life prediction and cycle life test optimization of different formulation li-ion power batteries. Applied Energy **282**, 116167 (2021). doi:[10.1016/j.apenergy.2020.116167](https://doi.org/10.1016/j.apenergy.2020.116167)
- [109] Molnar, C.: Interpretable Machine Learning, 2nd edn. (2022). <https://christophm.github.io/interpretable-ml-book>
- [110] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A survey of methods for explaining black box models. ACM computing surveys **51**(5), 1–42 (2018). doi:[10.1145/3236009](https://doi.org/10.1145/3236009)
- [111] pandas development team, T.: pandas-dev/pandas: Pandas (2024). doi:[10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134)
- [112] Bühlmann, P., Mächler, M.: Computational Statistics. ETH Zürich, Zürich, Switzerland (2016). <https://stat.ethz.ch/lectures/ss17/comp-stats.php>
- [113] Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., *et al.*: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods **17**, 261–272 (2020). doi:[10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)

- [114] Chen, A., Owen, A.B., Shi, M.: Data enriched linear regression. *Electronic journal of statistics* **9**(1), 1078–1112 (2015). doi:[10.1214/15-ejs1027](https://doi.org/10.1214/15-ejs1027)
- [115] Bell, I.H., Wronski, J., Quoilin, S., Lemort, V.: Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library coolprop. *Industrial & Engineering Chemistry Research* **53**(6), 2498–2508 (2014). doi:[10.1021/ie4033999](https://doi.org/10.1021/ie4033999)
- [116] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Édouard Duchesnay: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**(85), 2825–2830 (2011)
- [117] Montgomery, D.C., Peck, E.A., Vining, G.G.: *Introduction to Linear Regression Analysis*. John Wiley & Sons, (2021)
- [118] Belsley, D.A., Kuh, E., Welsch, R.E.: *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, (2005)
- [119] Rasmussen, C.E., Williams, C.K.: *Gaussian Processes for Machine Learning* vol. 1. MIT Press, (2005). doi:[10.7551/mitpress/3206.001.0001](https://doi.org/10.7551/mitpress/3206.001.0001)
- [120] Williams, C.K.: Prediction with gaussian processes: From linear regression to linear prediction and beyond. In: *Learning and Inference in Graphical Models*, pp. 599–621. Springer, (1998). doi:[10.1007/978-94-011-5014-9_23](https://doi.org/10.1007/978-94-011-5014-9_23)
- [121] Bishop, C.M.: *Pattern Recognition and Machine Learning*. Springer, (2006). doi:[10.1007/978-0-387-45528-0](https://doi.org/10.1007/978-0-387-45528-0)
- [122] Tipping, M.E.: Bayesian inference: An introduction to principles and practice in machine learning. In: *Advanced Lectures on Machine Learning*, pp. 41–62. Springer, (2003). doi:[10.1007/978-3-540-28650-9_3](https://doi.org/10.1007/978-3-540-28650-9_3)
- [123] Wang, J.: An intuitive tutorial to gaussian processes regression. *Computing in Science & Engineering* (2023). doi:[10.1109/mcse.2023.3342149](https://doi.org/10.1109/mcse.2023.3342149)
- [124] Rakousky, C., Reimer, U., Wippermann, K., Kuhri, S., Carmo, M., Lueke, W., Stolten, D.: Polymer electrolyte membrane water electrolysis: Restraining degradation in the presence of fluctuating power. *Journal of Power Sources* **342**, 38–47 (2017). doi:[10.1016/j.jpowsour.2016.11.118](https://doi.org/10.1016/j.jpowsour.2016.11.118)

- [125] Salvatier, J., Wiecki, T.V., Fonnesbeck, C.: Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* **2**, 55 (2016). doi:[10.7717/peerj-cs.55](https://doi.org/10.7717/peerj-cs.55)
- [126] Selamet, Ö.F., Acar, M.C., Mat, M.D., Kaplan, Y.: Effects of operating parameters on the performance of a high-pressure proton exchange membrane electrolyzer. *International Journal of Energy Research* **37**(5), 457–467 (2013). doi:[10.1002/er.2942](https://doi.org/10.1002/er.2942)
- [127] Trinke, P., Bensmann, B., Hanke-Rauschenbach, R.: Current density effect on hydrogen permeation in PEM water electrolyzers. *International Journal of Hydrogen Energy* **42**(21), 14355–14366 (2017). doi:[10.1016/j.ijhydene.2017.03.231](https://doi.org/10.1016/j.ijhydene.2017.03.231)
- [128] Zhang, W., Deng, L., Zhang, L., Wu, D.: A survey on negative transfer. *IEEE/CAA Journal of Automatica Sinica* **10**(2), 305–329 (2022). doi:[10.1109/jas.2022.106004](https://doi.org/10.1109/jas.2022.106004)
- [129] Duan, Y., Li, H., He, M., Zhao, D.: A BiGRU autoencoder remaining useful life prediction scheme with attention mechanism and skip connection. *IEEE Sensors Journal* **21**(9), 10905–10914 (2021). doi:[10.1109/jsen.2021.3060395](https://doi.org/10.1109/jsen.2021.3060395)
- [130] Xue, B., Xu, F., Huang, X., Xu, Z., Zhang, X.: Improved similarity based prognostics method for turbine engine degradation with degradation consistency test. *Applied Intelligence*, 10181–10201 (2022). doi:[10.1007/s10489-021-03034-6](https://doi.org/10.1007/s10489-021-03034-6)
- [131] Cui, L., Wang, X., Wang, H., Jiang, H.: Remaining useful life prediction of rolling element bearings based on simulated performance degradation dictionary. *Mechanism and Machine Theory* **153**, 103967 (2020). doi:[10.1016/j.mechmachtheory.2020.103967](https://doi.org/10.1016/j.mechmachtheory.2020.103967)
- [132] Yang, Y., Guo, Y., Huang, Z., Chen, N., Li, L., Jiang, Y., He, N.: Research on the milling tool wear and life prediction by establishing an integrated predictive model. *Measurement* **145**, 178–189 (2019). doi:[10.1016/j.measurement.2019.05.009](https://doi.org/10.1016/j.measurement.2019.05.009)
- [133] Liang, Z., Gao, J., Jiang, H., Gao, X., Gao, Z., Wang, R.: A degradation degree considered method for remaining useful life prediction based on similarity. *Computing in Science & Engineering* **21**(1), 50–64 (2018). doi:[10.1109/mcse.2018.110145829](https://doi.org/10.1109/mcse.2018.110145829)
- [134] Yu, W., Kim, I.Y., Mechefske, C.: An improved similarity-based prognostic algorithm for rul estimation using an rnn autoencoder scheme. *Reliability Engineering & System Safety* **199**, 106926 (2020). doi:[10.1016/j.ress.2020.106926](https://doi.org/10.1016/j.ress.2020.106926)
- [135] Soons, Y., Dijkman, R., Jilderda, M., Duivesteijn, W.: Predicting remaining useful life with similarity-based priors. In: *Advances in Intelligent Data Analysis XVIII*, pp. 483–495. Springer, (2020). doi:[10.1007/978-3-030-44584-3_38](https://doi.org/10.1007/978-3-030-44584-3_38)

- [136] Bektas, O., Jones, J.A., Sankararaman, S., Roychoudhury, I., Goebel, K.: A neural network filtering approach for similarity-based remaining useful life estimation. *The International Journal of Advanced Manufacturing Technology* **101**, 87–103 (2019). doi:[10.1007/s00170-018-2874-0](https://doi.org/10.1007/s00170-018-2874-0)
- [137] Li, X., Mba, D., Lin, T.: A similarity-based and model-based fusion prognostics framework for remaining useful life prediction. In: 2019 Prognostics and System Health Management Conference (PHM-Qingdao), pp. 1–5. IEEE, (2019). doi:[10.1109/phm-qingdao46334.2019.8943006](https://doi.org/10.1109/phm-qingdao46334.2019.8943006)
- [138] Ding, P., Qian, Q., Wang, H., Yao, J.: A symbolic regression based residual useful life model for slewing bearings. *IEEE Access* **7**, 72076–72089 (2019). doi:[10.1109/access.2019.2919663](https://doi.org/10.1109/access.2019.2919663)
- [139] Cannarile, F., Baraldi, P., Zio, E.: An evidential similarity-based regression method for the prediction of equipment remaining useful life in presence of incomplete degradation trajectories. *Fuzzy Sets and Systems* **367**, 36–50 (2019). doi:[10.1016/j.fss.2018.10.008](https://doi.org/10.1016/j.fss.2018.10.008)
- [140] Wang, H., Chen, J., Qu, J., Ni, G.: A new approach for safety life prediction of industrial rolling bearing based on state recognition and similarity analysis. *Safety science* **122**, 104530 (2020). doi:[10.1016/j.ssci.2019.104530](https://doi.org/10.1016/j.ssci.2019.104530)
- [141] Zeming, L., Jianmin, G., Hongquan, J.: A maintenance support framework based on dynamic reliability and remaining useful life. *Measurement* **147**, 106835 (2019). doi:[10.1016/j.measurement.2019.07.063](https://doi.org/10.1016/j.measurement.2019.07.063)
- [142] Wen, L., Ping, T.J., Wei, S.Y.: An evidential similarity-based regression method for the prediction of equipment remaining useful life under complex conditions. In: 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 362–366. IEEE, (2021). doi:[10.1109/iccece51280.2021.9342372](https://doi.org/10.1109/iccece51280.2021.9342372)
- [143] Qiang, B., Shi, K., Liu, N., Ren, J., Shi, Y.: Integrating physics-informed recurrent gaussian process regression into instance transfer for predicting tool wear in milling process. *Journal of Manufacturing Systems* **68**, 42–55 (2023). doi:[10.1016/j.jmsy.2023.02.019](https://doi.org/10.1016/j.jmsy.2023.02.019)
- [144] Pardoe, D., Stone, P.: Boosting for regression transfer. In: Proceedings of the 27th International Conference on International Conference on Machine Learning. ICML'10, pp. 863–870 (2010). <https://dl.acm.org/doi/10.5555/3104322.3104432>

- [145] Dhada, M., Parlikad, A.K., Palau, A.S.: Federated learning for collaborative prognosis. In: International Conference on Precision, Meso, Micro and Nano Engineering (COPEN 2019) (2019). doi:[10.17863/CAM.50577](https://doi.org/10.17863/CAM.50577)
- [146] Arunan, A., Qin, Y., Li, X., Yuen, C.: A federated learning-based industrial health prognostics for heterogeneous edge devices using matched feature extraction. *IEEE Transactions on Automation Science and Engineering* (2023). doi:[10.1109/tase.2023.3274648](https://doi.org/10.1109/tase.2023.3274648)
- [147] Lu, S., Wang, F., Piao, C., Ma, Y.: Health state prediction of lithium ion battery based on deep learning method. *IOP Conference Series: Materials Science and Engineering* **782**(3), 032083 (2020). doi:[10.1088/1757-899x/782/3/032083](https://doi.org/10.1088/1757-899x/782/3/032083)
- [148] Pan, D., Li, H., Wang, S.: Transfer learning-based hybrid remaining useful life prediction for lithium-ion batteries under different stresses. *IEEE Transactions on Instrumentation and Measurement* **71**, 1–10 (2022). doi:[10.1109/tim.2022.3142757](https://doi.org/10.1109/tim.2022.3142757)
- [149] Shu, X., Shen, J., Li, G., Zhang, Y., Chen, Z., Liu, Y.: A flexible state-of-health prediction scheme for lithium-ion battery packs with long short-term memory network and transfer learning. *IEEE Transactions on Transportation Electrification* **7**(4), 2238–2248 (2021). doi:[10.1109/TTE.2021.3074638](https://doi.org/10.1109/TTE.2021.3074638)
- [150] Kim, S., Choi, Y.Y., Kim, K.J., Choi, J.-I.: Forecasting state-of-health of lithium-ion batteries using variational long short-term memory with transfer learning. *Journal of Energy Storage* **41**, 102893 (2021). doi:[10.1016/j.est.2021.102893](https://doi.org/10.1016/j.est.2021.102893)
- [151] Gupta, A., Sheikh, M., Tripathy, Y., Widanage, W.D.: Transfer learning LSTM model for battery useful capacity fade prediction. In: 2021 24th International Conference on Mechatronics Technology (ICMT), pp. 1–6. IEEE, (2021). doi:[10.1109/icmt53429.2021.9687230](https://doi.org/10.1109/icmt53429.2021.9687230)
- [152] Chen, Z., Chen, L., Shen, W., Xu, K.: Remaining useful life prediction of lithium-ion battery via a sequence decomposition and deep learning integrated approach. *IEEE Transactions on Vehicular Technology* **71**(2), 1466–1479 (2022). doi:[10.1109/TVT.2021.3134312](https://doi.org/10.1109/TVT.2021.3134312)
- [153] Ragab, M., Chen, Z., Wu, M., Kwoh, C.K., Li, X.: Adversarial transfer learning for machine remaining useful life prediction. In: 2020 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 1–7. IEEE, (2020). doi:[10.1109/icphm49022.2020.9187053](https://doi.org/10.1109/icphm49022.2020.9187053)

- [154] Liu, C., Mauricio, A., Qi, J., Peng, D., Gryllias, K.: Domain adaptation digital twin for rolling element bearing prognostics. *Annual Conference of the PHM Society* **12**(1), 10 (2020). doi:[10.36001/phmconf.2020.v12i1.1294](https://doi.org/10.36001/phmconf.2020.v12i1.1294)
- [155] Fu, S., Zhang, Y., Lin, L., Zhao, M., Zhong, S.-s.: Deep residual LSTM with domain-invariance for remaining useful life prediction across domains. *Reliability Engineering & System Safety* **216**, 108012 (2021). doi:[10.1016/j.ress.2021.108012](https://doi.org/10.1016/j.ress.2021.108012)
- [156] Ma, J., Liu, X., Zou, X., Yue, M., Shang, P., Kang, L., Jemei, S., Lu, C., Ding, Y., Zerhouni, N., Cheng, Y.: Degradation prognosis for proton exchange membrane fuel cell based on hybrid transfer learning and intercell differences. *ISA Transactions* **113**, 149–165 (2021). doi:[10.1016/j.isatra.2020.06.005](https://doi.org/10.1016/j.isatra.2020.06.005)
- [157] Amogne, Z.E., Wang, F.-K., Chou, J.-H.: Transfer learning based on transferability measures for state of health prediction of lithium-ion batteries. *Batteries* **9**(5), 280 (2023). doi:[10.3390/batteries9050280](https://doi.org/10.3390/batteries9050280)
- [158] Tan, Y., Zhao, G.: Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries. *IEEE Transactions on Industrial Electronics* **67**(10), 8723–8731 (2020). doi:[10.1109/TIE.2019.2946551](https://doi.org/10.1109/TIE.2019.2946551)
- [159] Deng, Z., Lin, X., Cai, J., Hu, X.: Battery health estimation with degradation pattern recognition and transfer learning. *Journal of Power Sources* **525**, 231027 (2022). doi:[10.1016/j.jpowsour.2022.231027](https://doi.org/10.1016/j.jpowsour.2022.231027)
- [160] Shen, S., Sadoughi, M., Li, M., Wang, Z., Hu, C.: Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. *Applied Energy* **260**, 114296 (2020). doi:[10.1016/j.apenergy.2019.114296](https://doi.org/10.1016/j.apenergy.2019.114296)
- [161] Li, Y., Li, K., Liu, X., Wang, Y., Zhang, L.: Lithium-ion battery capacity estimation — a pruned convolutional neural network approach assisted with transfer learning. *Applied Energy* **285**, 116410 (2021). doi:[10.1016/j.apenergy.2020.116410](https://doi.org/10.1016/j.apenergy.2020.116410)
- [162] Khaled, B., Yue, M., Samir, J., Noureddine, Z.: A knowledge transfer approach for online pemfc degradation prediction with uncertainty quantification. In: *2022 12th International Conference on Power, Energy and Electrical Engineering (CPEEE)*, pp. 65–70. IEEE, (2022). doi:[10.1109/cpeee54404.2022.9738717](https://doi.org/10.1109/cpeee54404.2022.9738717)
- [163] Li, Y., Tao, J.: CNN and transfer learning based online soh estimation for lithium-ion battery. In: *2020 Chinese Control And Decision Conference (CCDC)*, pp. 5489–5494 (2020). doi:[10.1109/CCDC49329.2020.9164208](https://doi.org/10.1109/CCDC49329.2020.9164208)

- [164] Zeng, F., Li, Y., Jiang, Y., Song, G.: An online transfer learning-based remaining useful life prediction method of ball bearings. *Measurement* **176**, 109201 (2021). doi:[10.1016/j.measurement.2021.109201](https://doi.org/10.1016/j.measurement.2021.109201)
- [165] Li, X., Zhang, W., Ma, H., Luo, Z., Li, X.: Data alignments in machinery remaining useful life prediction using deep adversarial neural networks. *Knowledge-Based Systems* **197**, 105843 (2020). doi:[10.1016/j.knosys.2020.105843](https://doi.org/10.1016/j.knosys.2020.105843)
- [166] Zhuang, J., Jia, M., Ding, Y., Ding, P.: Temporal convolution-based transferable cross-domain adaptation approach for remaining useful life estimation under variable failure behaviors. *Reliability Engineering & System Safety* **216**, 107946 (2021). doi:[10.1016/j.ress.2021.107946](https://doi.org/10.1016/j.ress.2021.107946)
- [167] Ding, Y., Ding, P., Zhao, X., Cao, Y., Jia, M.: Transfer learning for remaining useful life prediction across operating conditions based on multisource domain adaptation. *IEEE/ASME Transactions on Mechatronics* **27**(5), 4143–4152 (2022). doi:[10.1109/tmech.2022.3147534](https://doi.org/10.1109/tmech.2022.3147534)
- [168] Meraghni, S., Terrissa, L.S., Yue, M., Ma, J., Jemei, S., Zerhouni, N.: A data-driven digital-twin prognostics method for proton exchange membrane fuel cell remaining useful life prediction. *International Journal of Hydrogen Energy* **46**(2), 2555–2564 (2021). doi:[10.1016/j.ijhydene.2020.10.108](https://doi.org/10.1016/j.ijhydene.2020.10.108)
- [169] He, Z., Shao, H., Ding, Z., Jiang, H., Cheng, J.: Modified deep autoencoder driven by multisource parameters for fault transfer prognosis of aeroengine. *IEEE Transactions on Industrial Electronics* **69**(1), 845–855 (2022). doi:[10.1109/tie.2021.3050382](https://doi.org/10.1109/tie.2021.3050382)
- [170] Mao, W., He, J., Zuo, M.J.: Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning. *IEEE Transactions on Instrumentation and Measurement* **69**(4), 1594–1608 (2019). doi:[10.1109/tim.2019.2917735](https://doi.org/10.1109/tim.2019.2917735)
- [171] Lucu, M., Martinez-Laserna, E., Gandiaga, I., Liu, K., Camblong, H., Widanage, W.D., Marco, J.: Data-driven nonparametric li-ion battery ageing model aiming at learning from real operation data—part a: Storage operation. *Journal of Energy Storage* **30**, 101409 (2020). doi:[10.1016/j.est.2020.101409](https://doi.org/10.1016/j.est.2020.101409)
- [172] Pandhare, V., Jia, X., Lee, J.: Collaborative prognostics for machine fleets using a novel federated baseline learner. In: *Annual Conference of the PHM Society*, vol. 13 (2021). doi:[10.36001/phmconf.2021.v13i1.2989](https://doi.org/10.36001/phmconf.2021.v13i1.2989)

- [173] Fan, Y., Nowaczyk, S., Rögnvaldsson, T.: Transfer learning for remaining useful life prediction based on consensus self-organizing models. *Reliability Engineering & System Safety* **203**, 107098 (2020). doi:[10.1016/j.ress.2020.107098](https://doi.org/10.1016/j.ress.2020.107098)
- [174] Li, Y., Sheng, H., Cheng, Y., Stroe, D.-I., Teodorescu, R.: State-of-health estimation of lithium-ion batteries based on semi-supervised transfer component analysis. *Applied Energy* **277**, 115504 (2020). doi:[10.1016/j.apenergy.2020.115504](https://doi.org/10.1016/j.apenergy.2020.115504)
- [175] Kim, S.W., Oh, K.-Y., Lee, S.: Novel informed deep learning-based prognostics framework for on-board health monitoring of lithium-ion batteries. *Applied Energy* **315**, 119011 (2022). doi:[10.1016/j.apenergy.2022.119011](https://doi.org/10.1016/j.apenergy.2022.119011)
- [176] Liu, W., Xu, Y.: Data-driven online health estimation of li-ion batteries using a novel energy-based health indicator. *IEEE Transactions on Energy Conversion* **35**(3), 1715–1718 (2020). doi:[10.1109/tec.2020.2995112](https://doi.org/10.1109/tec.2020.2995112)
- [177] Cao, Y., Jia, M., Ding, P., Ding, Y.: Transfer learning for remaining useful life prediction of multi-conditions bearings based on bidirectional-GRU network. *Measurement* **178**, 109287 (2021). doi:[10.1016/j.measurement.2021.109287](https://doi.org/10.1016/j.measurement.2021.109287)
- [178] Zhang, W., Li, X., Ma, H., Luo, Z., Li, X.: Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions. *Reliability Engineering & System Safety* **211**, 107556 (2021). doi:[10.1016/j.ress.2021.107556](https://doi.org/10.1016/j.ress.2021.107556)
- [179] Skiena, S.S.: *The Data Science Design Manual*. Springer, (2017). doi:[10.1007/978-3-319-55444-0](https://doi.org/10.1007/978-3-319-55444-0)
- [180] Locci, C., Mertens, M., Höyng, S., Schmid, G., Bagus, T., Lettenmeier, P.: Scaling-up PEM electrolysis production: Challenges and perspectives. *Chemie Ingenieur Technik* **96**(1-2), 22–29 (2024). doi:[10.1002/cite.202300111](https://doi.org/10.1002/cite.202300111)

Statement of Originality

I hereby confirm in lieu of oath, that I have written the accompanying thesis by myself, without contributions from any sources other than those cited in the text and acknowledgments. I certify, that I have followed the general principles of scientific work and publications as written in the guidelines of good research of the Carl von Ossietzky University of Oldenburg. This work has not yet been submitted to any examination office in the same or similar form.



Ottobrunn, June 30, 2025

Xuqian Yan