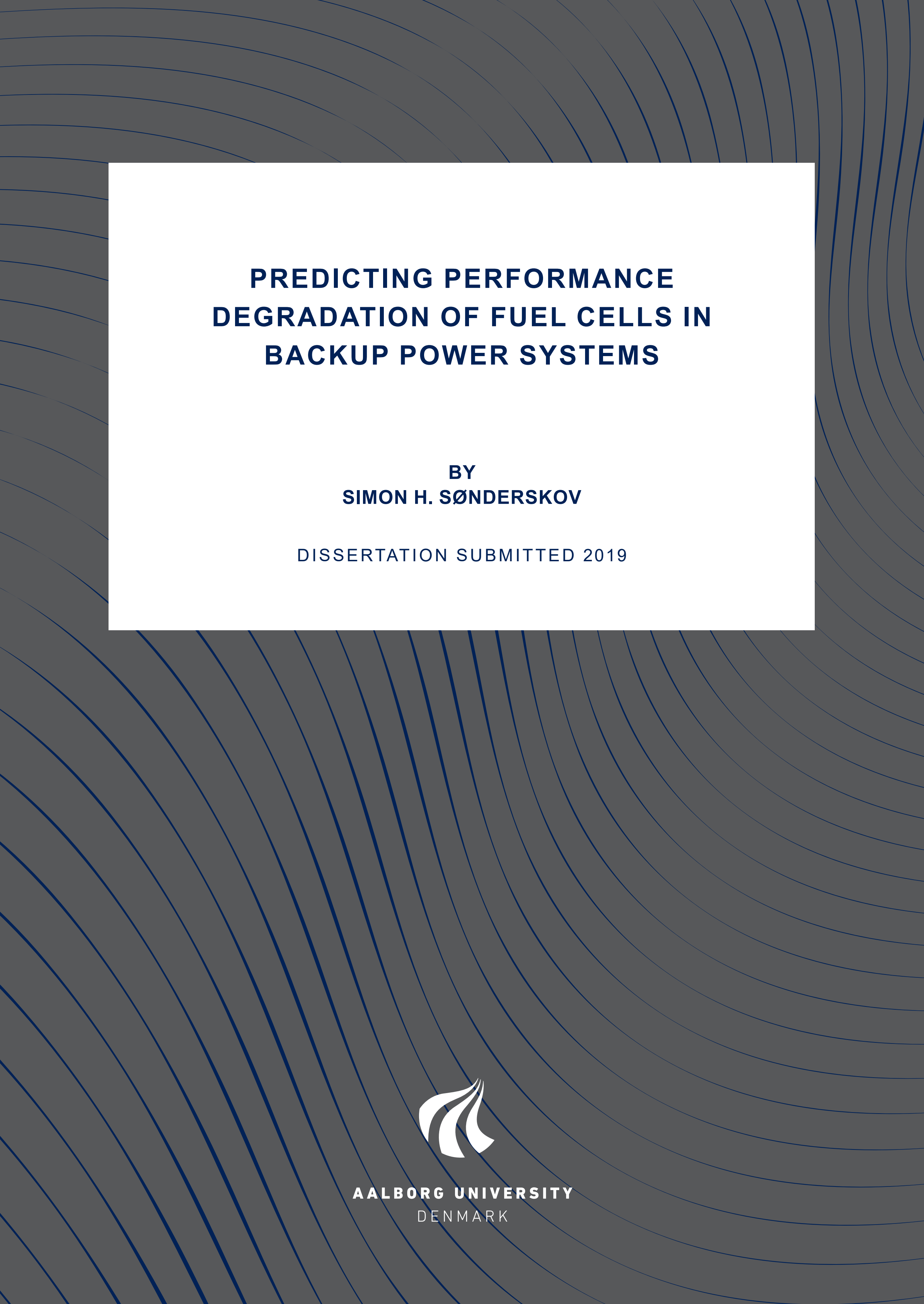
Predicting Performance Degradation of Fuel Cells in Backup Power Systems

Heindorf Sønderskov, Simon

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Heindorf Sønderskov, S. (2019). *Predicting Performance Degradation of Fuel Cells in Backup Power Systems*. Aalborg Universitetsforlag.



PREDICTING PERFORMANCE DEGRADATION OF FUEL CELLS IN BACKUP POWER SYSTEMS

BY
SIMON H. SØNDERSKOV

DISSERTATION SUBMITTED 2019

AALBORG UNIVERSITY
DENMARK

PREDICTING PERFORMANCE DEGRADATION OF FUEL CELLS IN BACKUP POWER SYSTEMS

Ph.D. Thesis

Simon H. Sønderskov

Thesis submitted August, 2019

Dissertation submitted: August, 2019

PhD supervisor: Prof. Stig Munk-Nielsen
Aalborg University

PhD committee: Associate Professor Erik Schaltz (chairman)
Aalborg University

Professor Cesare Pianese
University of Salerno

Senior Lecturer Suresh Perinpanayagam
Cranfield University

PhD Series: Faculty of Engineering and Science, Aalborg University

Department: Department of Energy Technology

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-466-9

Published by:
Aalborg University Press
Langagervej 2
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2019

List of Publications

The following publications were published and/or submitted during the period of the Ph.D. study, and are appended in Part II of this thesis. The papers will be referred to throughout the thesis using the alphabetic reference as listed below, to distinguish the appended paper from the works of other authors. Secondary references will be cited using a numeric reference system.

[A] **S. D. Sønderskov**, M. J. Swierczynski and S. Munk-Nielsen, "Lifetime prognostics of hybrid backup power system: State-of-the-art," *2017 IEEE International Telecommunications Energy Conference (INTELEC),* Broadbeach, QLD, 2017.

[B] **S. D. Sønderskov**, L. Török and S. Munk-Nielsen, "Monitoring Performance Indicators of PEM Fuel Cell Backup Power Systems," *2018 IEEE International Telecommunications Energy Conference (INTELEC),* Turino, Italy, 2018.

[C] **S. H. Sønderskov**, D. Rasmussen, J. Ilsøe, D. Blom-Hansen and S. Munk-Nielsen. "Detecting performance outliers in fuel cell backup power systems". Submitted to *Power Electronics and Applications (EPE'19 ECCE Europe), 2019 21st European Conference on,.* Accepted.

[D] **S. H. Sønderskov**, J. Ilsøe, D. Rasmussen, D. Blom-Hansen and S. Munk-Nielsen. "State of Health Estimation and Prediction of Fuel Cell Stacks in Backup Power Systems". Submitted to *IEEE Transactions on Industrial Electronics.* 2019. In review.

[E] **S. H. Sønderskov**, J. Ilsøe, D. Rasmussen, D. Blom-Hansen and S. Munk-Nielsen. "Predicting Performance Indicators of Fuel Cell Stacks in Backup Power Systems". Submitted to *Elsevier Journal of Power Sources.* 2019. In review.

Furthermore, the following papers have been published and/or submitted during the period of the Ph.D. study, but are not included in the content of this thesis.

- D. N. Dalal, N. Christensen, A. B. Jørgensen, **S. D. Sønderskov**, S. Bęczkowski, C. Uhrenfeldt and S. Munk-Nielsen, “Gate driver with high common mode rejection and self turn-on mitigation for a 10 kV SiC MOSFET enabled MV converter,” *2017 19th European Conference on Power Electronics and Applications (EPE’17 ECCE Europe),* Warsaw, 2017.

- N. Christensen, A. B. Jørgensen, D. Dalal, **S. D. Sønderskov**, S. Bęczkowski, C. Uhrenfeldt and S. Munk-Nielsen, “Common mode current mitigation for medium voltage half bridge SiC modules,” *2017 19th European Conference on Power Electronics and Applications (EPE’17 ECCE Europe),* Warsaw, 2017.

- A. B. Jørgensen, N. Christensen, D. N. Dalal, **S. D. Sønderskov**, S. Bęczkowski, C. Uhrenfeldt and S. Munk-Nielsen, “Reduction of parasitic capacitance in 10 kV SiC MOSFET power modules using 3D FEM,” *2017 19th European Conference on Power Electronics and Applications (EPE’17 ECCE Europe),* Warsaw, 2017.

- L. Török, **S. D. Sønderskov** and S. Munk-Nielsen, “Bidirectional Operation of High Efficiency Isolated DC-DC Converter in Fuel Cell Telecom Back-up Systems,” *2018 IEEE International Telecommunications Energy Conference (INTELEC),* Turino, Italy, 2018.

- A. B. Jørgensen, **S. H. Sønderskov**, S. Beczkowski, B. Bidoggia, S. Munk Nielsen, “Analysis of Cascaded Silicon Carbide MOSFETs Using a Single Gate Driver for Medium Voltage Applications,” Submitted to *IET Power Electronics.* 2019. Accepted.

Abstract

Backup power systems provide critical systems with a continuous supply of electricity, even through disturbances and interruptions in the power grid. One application, that relies on backup power systems is telecommunication substations, which provide the critical infrastructure for cellular as well as internet communication.

Backup power systems rely on energy storage elements to compensate the disturbances in the power grid, to supply the load with a smooth and continuous power profile. Normally, batteries are used as the storage element. However, batteries might be infeasible in applications that require prolonged backup-time capabilities from the backup system. To extend the backup-time, diesel generators have been used in the past. However, diesel generators suffer from a number of downsides, including high pollution levels, noise, and high maintenance requirements. Instead of diesel generators, the telecommunication industry has largely turned to fuel cells as the main source of backup power.

Fuel cells use hydrogen and oxygen to generate electrical power and leaves only water and heat as waste products. However, fuel cell technology is still in the early market phase and relies on some expensive materials in its construction. This makes fuel cells relatively expensive to purchase.

The high initial cost can be compensated by reducing the maintenance costs and extending the lifetime thorough predictive maintenance. The prognostics and health management (PHM) framework describes several steps towards this end. Two key parts of PHM, which are addressed in this work, are i) the assessment of system condition and ii) prediction of future system condition. The latter is often referred to as prognostics, and is a key step to enable predictive maintenance, as maintenance efforts can be planned in advance based on the estimated future system condition.

The studies in this work are based on a unique dataset, which comprises field measurements from numerous fuel cell based backup power systems, which have been operating in telecommunication substations. The main contributions of this work are i) the extraction and analysis of key performance indicators based on field data from a fleet of fuel cell based backup power systems. ii)

Estimation of the degradation level of the fuel cell stacks. iii) A method for detecting outlier fuel cell stacks based on the extracted performance indicators. iv) A method for predicting the future degradation levels of the fuel cell stacks using a recurrent neural network.

Resumé

Nødstrømssystemer forsyner kritiske systemer med kontinuer elektricitet, selv under forstyrrelser og afbrydelser i elnettet. En applikation, som afhænger af nødstrømssystemer, er telekommunikations stationer, som udgør en kritisk infrastruktur for mobil-, så vel som internetkommunikation.

Nødstrømssystemer bruger energilagringselementer til at kompensere de forstyrrelser, der forekommer i elnettet, således, at lasten forsynes med en jævn og kontinuer strøm. Normalt bliver batterier brugt som energilagringselement. Men, batterier kan være upraktiske i applikationer hvor langvarig nødstrømsforsyning er nødvendigt. For at forlænge nødstrømsforsyningstiden, har diesel generatorer tidligere været brugt. Men diesel generatorer har en række ulemper, deriblandt højt forureningsniveau, støj, og høje vedligeholdelseskrav. I stedet for diesel generatorer, anvender telekommunikationsindustrien i høj grad brændselsceller som det primære energilagringselement.

Brændselsceller anvender hydrogen og oxygen til at generere elektricitet og efterlader kun vand og varme som biprodukter. Men, brændselscelle teknologi er stadig i et tidligt-marked stadie og afhænger af dyre materialer i dets konstruktion. Dette gør at brændselsceller er relativt dyre i indkøbspris.

Den høje indkøbspris kan kompenseres ved at reducere vedligeholdelsesomkostninger og ved at forlænge levetiden igennem forudsigende vedligeholdelsesstrategier. Prognostics and Health Management (PHM) modellen beskriver adskillige trin hvormed dette kan opnås. To af hovedelementerne i PHM, hvilke er behandlet i dette arbejde, er i) estimering af systemets tilstand og ii) forudsigelse af systemets fremtidige tilstand. Sidstnævnte er også ofte kaldet Prognostics og er et vigtigt skridt mod en forudsigende vedligeholdelsesstrategi, da vedligeholdelsesindsatser kan planlægges baseret på den estimerede fremtidige tilstand af systemet.

Studierne i dette arbejde er baserede på et unikt datasæt, hvilket indeholder målinger fra talrige brændselscelle baserede nødstrømssystemer, som har været i drift i telekommunikations stationer. Hovedbidragene i dette arbejde er i) ekstraktion og analyse af nøgle ydeevneindikatorer baserede på feltdata fra en flåde af brændselscelle baserede nødstrømssystemer. ii) Estimering af degraderingsniveauet af brændselscellestakkene. iii) En metode til detektering af afvigende brændselscellestakke baseret på de ekstrakterede ydeevneindikatorer. iv) En metode til forudsigelse af fremtidige degraderingsniveau af brændselscellestakkene ved brug af Recurrent Neural Networks.

Preface

This thesis has been submitted to the Faculty of Engineering and Science at Aalborg University in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Energy Technology. The thesis is submitted in the form of a collection papers with an extended summary. The research in this thesis was conducted in the period from January 2017 to August 2019 under the supervision of Professor Stig Munk-Nielsen. The research has been funded partly by the CrossCut project and partly by the Department of Energy Technology. The PhD project has been carried out in collaboration with Ballard Power Systems Europe A/S.

I would like to extend my gratitude to my supervisor Stig Munk-Nielsen for the opportunity of participating in this project, for the support during my studies, and for his “everything is possible” attitude. I am grateful for the collaborating company Ballard Power Systems A/S for lending me the dataset that became the basis for this thesis. Specifically, I extend by thanks to Daniel Blom-Hansen, Jakob Ilsøe, and Dean Rasmussen. My appreciation goes to my colleagues at the Department of Energy Technology for a generally good working environment.

Finally I thank my family for their support and forbearance. Especially, I thank my wife, Henriette and my son, Oskar for providing a loving home and reminding Me what is truly important.

Simon Sønderskov

Simon H. Sønderskov
Aalborg University, August 5, 2019

Contents

Part I

Extended Summary

1

Introduction

This chapter will introduce the background and motivation behind this project as well as an overview of the three topics that constitute the basis of this work, namely backup power systems, fuel cells, and prognostics and health management. Furthermore, the project objectives are presented, and the thesis content is outlined.

1.1 Background and Motivation

With an increased effort to bring down green house gas emissions and air pollution in urban areas, countries around the world strive for an increasingly electrified energy system based on renewable energy technologies. This trend will not diminish in the future, as people and governments become more aware of the irreversible effects of climate change.

A complete reformation of the energy system on a global scale, from steady and reliable fossil fuel based power generation to fluctuating and distributed renewable generation, such as wind power and solar power, is not without challenges. With an increasingly electricity-driven world, critical infrastructures, such as telecommunication substations, will continue to depend on continuous and reliable electricity supply. And without central fossil fuel power plants, electrical grids might be more vulnerable and prone to faults. Therefore, backup power systems are essential to the continuous supply of electrical power to critical infrastructure.

Backup power systems require some form of energy storage, to continuously supply the load with power, especially through prolonged grid outages. Traditionally, a combination of batteries and diesel generators have been used. However, diesel generators are pollution heavy, which makes them unsuitable in a modern renewable energy system. Batteries, on the other hand, are not associated with pollution during operation. But batteries' backup time is limited by the amount of energy they can store electrochemically inside the cell. When the energy capacity is exhausted, the battery can not be recharged until grid power is restored.

Fuel cells use hydrogen and oxygen to produce electricity with water as a bi-product. This allows for the energy capacity to be refueled simply by adding more or refueling the hydrogen supply. The oxygen is normally supplied from the surrounding air. Furthermore, fuel cells cause no pollution in operation and have no moving parts. This makes fuel cell technology a promising candidate for backup power applications, where prolonged backup-times are required.

However, fuel cell technology is still in the developing stage and rely on relative expensive materials. This makes fuel cells expensive in terms of their initial cost. Therefore, reduction of operation and maintenance cost and extension of system lifetime are important factors in bringing down the total cost of ownership of such systems.

1.2 Backup Power Systems

Interruption of certain electrical systems, such as data centers, telecommunication equipment and hospital equipment, can mean loss of data, productivity and, in the worst case, lives. These interruptions might be caused by utility grid outages or even by minor disturbances in the voltage waveform. Therefore, it is important to have systems that can adequately mitigate these grid disturbances and maintain a continuous, high quality power supply to the critical loads. For this purpose uninterruptible power supply (UPS) systems are widely used, as they provide power from a separate source when disturbances occur on the grid.

1.2.1 Power Disturbances

Ideally, the grid voltage is a smooth sinusoidal waveform with constant amplitude and frequency. However, in reality there is a number of natural and man-made phenomena that affect and distort the grid voltage. This section will describe some of these disturbances, which are typically treated by UPS systems. The eight most common grid disturbances are: 1) line failure; 2) voltage sag; 3) voltage surge; 4) under-voltage; 5) over-voltage; 6) voltage spike; 7) frequency variation; 8) EMI. [1]

The disturbances are visualized in Fig. 1.1. Line failure is when the grid power is completely lost for an extended time period, causing an outage. A voltage sag is when the voltage level decreases for a short period of time, after which it resumes its normal level. A voltage surge is, similarly, when the voltage level increases for a short period of time, after which it resumes its normal level. Under-voltage is when the grid voltage is low for an extended period of time. Similarly, over-voltage when the grid voltage is high for an extended period of time. A voltage spike is when a very short pulse occurs on the voltage. Frequency variation is simply when the frequency of the voltage waveform deviates from the intended value. EMI is when superimposed higher

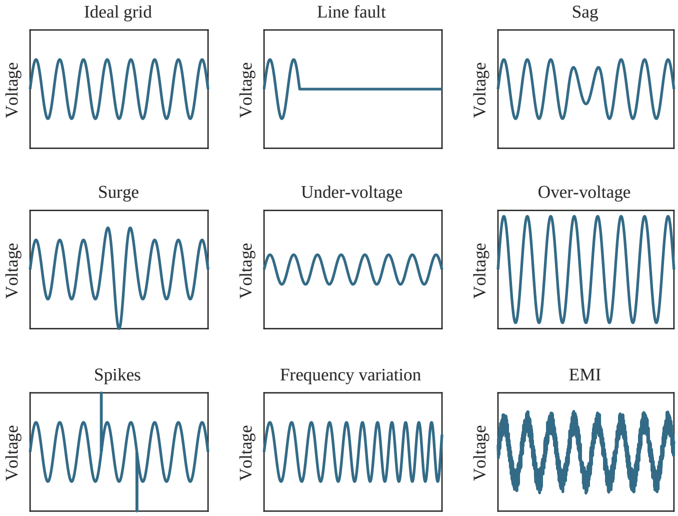


Fig. 1.1: Representation of common grid disturbances

frequency waveforms distort the smooth sinusoidal voltage. Harmonic distortion is a special case of EMI, where the frequency of the distorting component is a multiple of the fundamental frequency.

1.2.2 Uninterruptible Power Supply Architectures

UPS systems come in a wide variety of architectures, but are mainly grouped in three categories depending on the grid disturbances they address: offline, online, and line-interactive [2]. Basic block diagrams of the three architectures are shown in Fig. 1.2.

The offline architecture normally supplies the load directly from the grid through a static bypass switch. In case of grid failure, the switch is turned off and the load is supplied from the energy storage element through a DC/AC converter. In case of battery energy storage elements, the batteries are charged from the grid in periods of good grid condition. The direct supply of energy from the grid to the load gives the advantage of lossless supply in normal operation, but also means that there is no isolation between the grid and the load. Furthermore, the architecture provides no voltage regulation and it requires some time to switch to the backup operation in case of grid fault. The main advantages is the simplicity and low cost of the system. Line-interactive UPS systems feed the grid power directly to the load through a static switch and a filter. If grid power is absent, the storage element provides power to the load through a bidirectional power converter, which is in parallel connection with the load. This way, the storage element can also provide reactive power for power factor correction (PFC). However, it provides no isolation or voltage

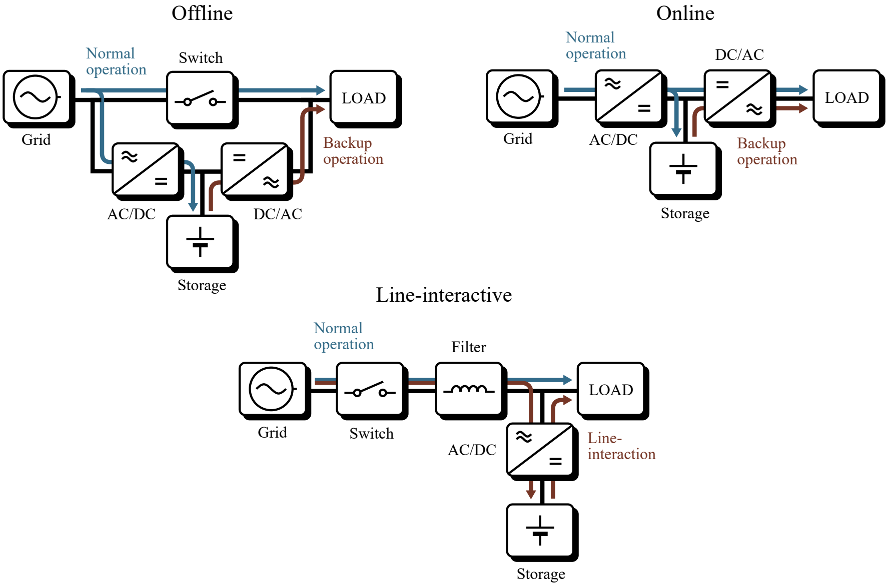


Fig. 1.2: The three main classes of uninterruptible power supply architectures: offline, line-interactive, and online

regulation capabilities. Online UPS architectures has two conversion stages: AC to DC and DC to AC. The storage element is interfaced to the intermediate DC stage often referred to as DC-link or DC-bus. Grid power is always fed through both converters, which gives rise to some losses. However, the load voltage is completely decoupled from the grid voltage, meaning that precise control of the load voltage is possible regardless of the grid voltage instabilities. Furthermore, there is no transition time between normal operation and the supply of backup power. This comes at the cost of reduced efficiency, increased complexity and cost. [1], [3]

A summary of the grid disturbance mitigation abilities of each of the three main UPS architectures is shown in Table 1.1 [1], [2]. Although more expensive, the online class of UPS architectures can handle all of the listed grid disturbances, which makes them suitable in applications, such as telecommunications, where equipment is sensitive to disturbances such as voltage spikes and EMI.

1.2.3 TELECOMMUNICATION BACKUP POWER SYSTEMS

Telecommunication facilities are critical to the infrastructure of modern society, providing the backbone of cellular and internet communication. Therefore, their continuous operation, regardless of utility grid condition, should be ensured through appropriate backup power systems. Telecommunication equipment is often sensitive to even small disturbances in the utility grid, which makes online

Table 1.1: Power disturbances handled by different classes of UPS

Line disturbance	Offline	Line-interactive	Online
Line failure	✓	✓	✓
Voltage sag	✓	✓	✓
Voltage surge	✓	✓	✓
Under-voltage		✓	✓
Over-voltage		✓	✓
Voltage spike			✓
Frequency variation			✓
EMI			✓

UPS architectures the most suitable choice for a telecommunication power supply. Also, the telecommunication usually requires DC power, meaning that the AC to DC conversion is required, even during normal operation. Hence, offline and line-interactive solutions would require an additional AC/DC conversion step, undermining their advantage of simplicity and lack of conversion loss.

Telecommunication sites are situated in a wide variety of locations. From urban areas with reliable utility grid connection to rural areas where the utility grid can be unreliable. Extreme weather conditions and natural catastrophes can further impair grid availability, which is often when communication services are most critical. Therefore backup power systems for telecommunication facilities are often required to ensure extended periods of backup power. [4]–[6]

Extended backup times has traditionally been achieved through the use of diesel generators. However, increased environmental concerns and the promise of reduced maintenance effort has shifted the focus to fuel cell technology [6]–[10]. Fuel cells, unlike diesel generators, produce no pollution when converting their fuel to electricity, do not emit noise, and have no moving parts which translates to better reliability and less maintenance effort [11], [12]. Like diesel generators, fuel cells have some startup time [13], meaning that an additional small storage element is needed, which can provide the load power during fuel cell startup. Often ultracapacitors or a small battery pack is used.

A typical fuel cell UPS system for telecommunication is shown in Fig. 1.3. The load is normally supplied from the grid through the AC/DC converter. When the grid fails, the fuel cell starts up to provide backup power. In the meantime, the ultracapacitor module provides the load power. Both storage elements are interfaced to the DC-bus through a DC/DC converter, which can be shared by the two elements. [5], [14], [15]

The load usually requires DC power, but in some cases the load runs on AC power or a combination of the two. Therefore a shaded DC/AC converter is included in Fig. 1.3. Also, some key elements, like hydrogen supply, ultracapacitor charger, and fuel cell voltage-booster, are not included in the figure for the sake of simplicity.

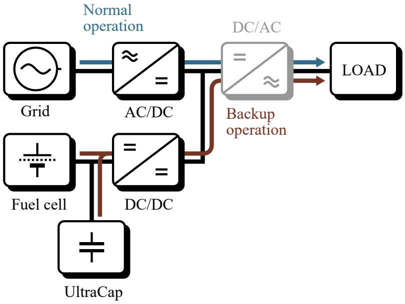


Fig. 1.3: Basic architecture of fuel cell based uninterruptible power supply for telecommunication applications

1.3 FUEL CELLS

Fuel cells are electrochemical devices, that converts a fuel directly to electrical power. Unlike batteries, which are also electrochemical devices that produce electrical power, fuel cells store their fuel externally. This way, the performance of the cell is not limited by the amount of reactants that can be fitted within the cell. As long as fuel is available, a fuel cell is able to produce electrical power indefinitely. Therefore fuel cells are rated in terms of their power level, rather than their energy capacity. [16]

Although the principal of the fuel cell was described already in 1843 [17], the technology is only recently gaining commercial traction as an energy storage device. The interest in fuel cell technology has seen a rise with the increasing focus on energy storage in an increasingly renewable and distributed energy system [18]. The relatively high initial cost of fuel cells remain a challenge to their wide adaptation. However, this cost is projected to decrease as the technology matures further and as the production scale increases [19].

1.3.1 BASIC WORKING PRINCIPAL

Fuel cells come in many different types, mainly distinguished by their operating temperature and the fuel they use. The most commercially successful type is the proton exchange membrane (PEM) fuel cell, which is a low-temperature type operating at 30-100°C. They require a pure hydrogen fuel and use oxygen as the oxidant, which is normally supplied from the surrounding air.

The PEM fuel cell consists of anode and cathode electrodes, and a solid proton conducting electrolyte membrane in-between. Each electrode consists of a catalyst layer and a gas diffusion layer (GDL). The basic structure of a PEM fuel cell is depicted in Fig. 1.4 which also sketches the operating principle of the fuel cell. At the anode side, hydrogen diffuses through the GDL and undergoes oxidization, i.e. the hydrogen atoms lose an electron and effectively becomes hydrogen ions (protons). The released electrons are free to migrate through the

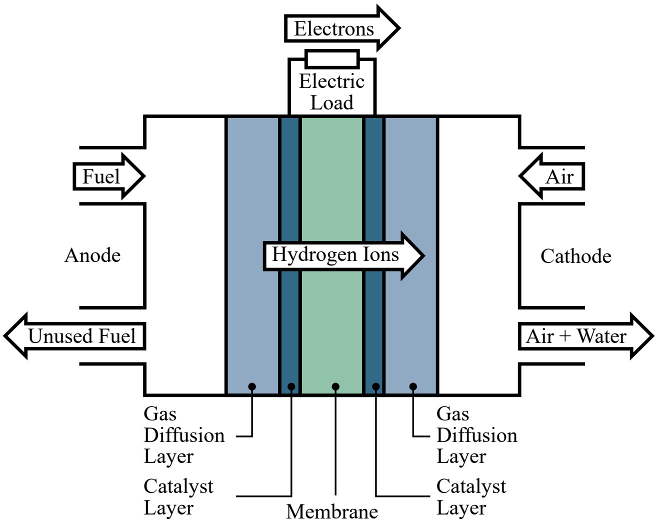


Fig. 1.4: Basic structure and working principal of a proton exchange membrane fuel cell [B]

GDL to the external electrical circuit and the protons can move through the membrane. This reaction is shown in (1.1). [16], [20]

$$2\mathrm{H}_2 \rightarrow 4\mathrm{H}^+ + 4e^- \tag{1.1}$$

At the cathode, oxygen from the air supply meets the electrons from the electrode and the protons, that have migrated from the anode side through the membrane, to form water. This reaction is shown in (1.2).

$$\mathrm{O}_2 + 4\mathrm{H}^+ + 4e^- \rightarrow 2\mathrm{H}_2\mathrm{O} \tag{1.2}$$

The overall reaction of the fuel cell is then described by (1.3). Some heat is also produced in the process.

$$2\mathrm{H}_2 + \mathrm{O}_2 \rightarrow 2\mathrm{H}_2\mathrm{O} + \mathrm{heat} \tag{1.3}$$

The amount of electricity produced in a single PEM fuel cell is usually around a single volt or less. Therefore, fuel cells are combined in series to increase the voltage level. These combinations are referred to as fuel cell stacks. The basic combination of fuel cells into fuel cell stacks is illustrated in Fig. 1.5. The shown stack consists of membrane-electrode assemblies and bipolar plates, which act as anode for one cell and cathode for the neighboring cell. The bipolar plates also have flow channels for the fuel and oxidant.

1.3.2 Electrical Properties

Under lossless conditions, the produced voltage in the fuel cell is referred to as the open-circuit voltage (V_r). However, real fuel cell operation is associated

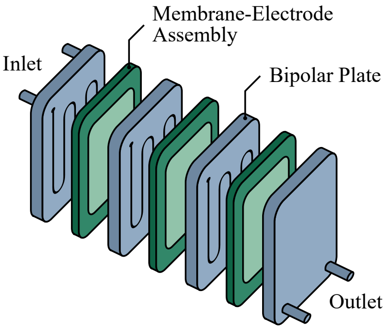


Fig. 1.5: Illustration of fuel cell stack assembly [B]

with some major voltage losses, also known as irreversibilities. The resulting produced voltage is described by the open-circuit voltage minus the voltage losses as shown in (1.4). The losses are: activation loss (ΔV_{act}), ohmic loss (ΔV_{ohmic}), and mass transportation loss (ΔV_{mass}).

$$V_c = V_r - \Delta V_{act} - \Delta V_{ohmic} - \Delta V_{mass} \tag{1.4}$$

The activation loss is a nonlinear effect related to the chemical reactions that transfer electrons from the anode and to the cathode, respectively. The voltage loss caused by this effect is described in (1.5), where $\mathcal{R}$ is the ideal gas constant, $\mathcal{F}$ is the Faraday constant, T is the cell temperature, α is the charge-transfer coefficient, I is the produced current, and I_o is the exchange current, i.e. the continuous backwards and forwards current caused by the equilibrium reactions when no current is drawn from the cell.

$$\Delta V_{act} = \frac{\mathcal{R}T}{2\alpha\mathcal{F}} \ln\left(\frac{I}{I_o}\right) \tag{1.5}$$

The ohmic loss is caused by the electrical resistance in the electrodes and the resistance to ion flow in the membrane. This loss is described in (1.6), where R is the combined electrical resistance of these effects.

$$\Delta V_{ohmic} = RI \tag{1.6}$$

The mass transport loss is a consequence of the falling concentration of reactant in the supply gas, as these reactants are used to produce current. This effect mainly happens at the cathode side as the oxygen is supplied from air, which has a limited concentration of oxygen. The mass transport loss is described by (1.7), where I_L is the limiting current, i.e. the current at which the fuel is consumed at the same rate as it can be supplied.

$$\Delta V_{mass} = -\frac{\mathcal{R}T}{2\mathcal{F}} \ln\left(1 - \frac{I}{I_L}\right) \tag{1.7}$$

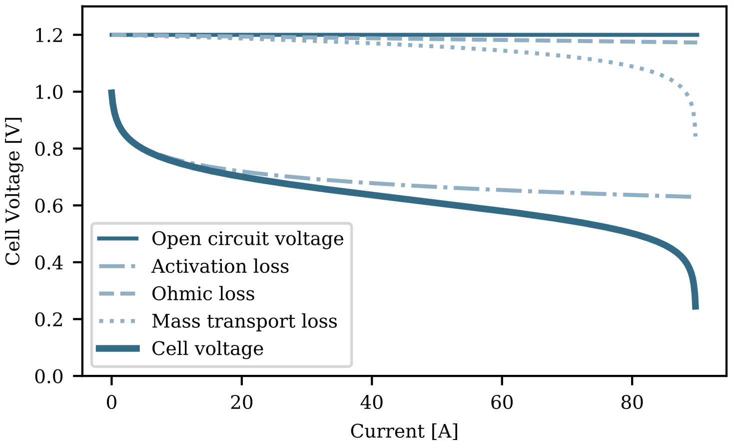


Fig. 1.6: Illustration of fuel cell current-voltage characteristics (polarization curve)

An additional cause of loss in the fuel cell is the fuel crossover. Although the membrane is designed to not let hydrogen atoms through, in practice, some atoms do get through the membrane. Another, smaller loss is caused by the conductivity of the membrane, which leads to a small internal current. These two effects can be accounted for by subtracting a current (I_n) from the current drawn from the cell. Hence, the cell voltage when accounting for activation, ohmic, and mass transport loss, as well as internal currents and fuel crossover is described by (1.8).

$$V_c = V_r - \frac{\mathcal{R}T}{2\alpha\mathcal{F}} \ln\left(\frac{I + I_n}{I_o}\right) - R(I + I_n) + \frac{\mathcal{R}T}{2\mathcal{F}} \ln\left(1 - \frac{I + I_n}{I_L}\right) \tag{1.8}$$

The fuel cell voltage and current characteristics are illustrated in Fig. 1.6, such a curve is often called the polarization curve. The losses are shown as differences from the open circuit voltage. At low currents, the activation loss is dominant, the ohmic loss increases linearly with the current, and the mass transport loss becomes dominant at high currents, where the fuel supply is exhausted.

1.4 Prognostics and Health Management

A suitable maintenance strategy is essential to ensure continuous availability of the backup system [21]. Generally there are two approaches to maintenance: reactive maintenance and proactive maintenance [22], [23]. In broad terms, reactive maintenance waits for something to break and then replaces or repairs the broken component or system. Whereas proactive maintenance aims to

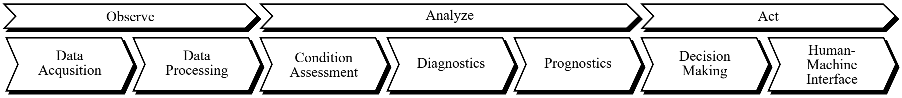


Fig. 1.7: Prognostics and health management framework [A]

schedule maintenance to prevent sudden system failures and to avoid unscheduled downtime. The schedule for proactive maintenance can be predetermined, for example during the design phase of the system, by estimated lifetimes of components or subsystems. This is generally referred to as preventive maintenance. However, this approach does not consider the variability of real field conditions where external factors and abnormalities can affect the condition of the system. Therefore, predictive maintenance strategies, that use the actual stress loading to continuously assess the system condition and forecasting remaining lifetime, is increasingly adopted.

A popular framework for predictive maintenance, which has been applied to fuel cell systems, is prognostics and health management (PHM). PHM consists of three main phases: Observation, Analysis, and Action, as outlined in Fig. 1.7. Observation includes measurement and logging of system parameters (data acquisition) and data processing, which transforms the raw data to more compact and useful information, for example storing data in a database and extracting useful features for later PHM steps. The analysis phase involves assessment of system condition, detection and localization of faults (diagnostics), as well as prediction of future system condition (prognostics). The action phase involves the planning of a suitable maintenance schedule based on the findings in the analysis phase.

1.4.1 Prognostics of Fuel Cell Systems

Within fuel cell system research, PHM has been an active field of research in the recent years [24]–[26]. Especially the topics of diagnostics and prognostics have been addressed. Diagnostics of fuel cell systems seeks to detect and isolate faults in the fuel cell system, which can then be mitigated to prevent further degradation. On the other hand, the aim of prognostics is to predict the future degradation level of the system, which can then be used to plan maintenance activities for example through a remaining useful lifetime (RUL) metric. The focus in this work is on the prognostics aspect. Hence, diagnostics will not be covered further in this thesis.

Prognostics approaches described in literature can be categorized by the type of measurement that is used as the health indicator and by the type of prediction method. For the health indicator, two main approaches are used: i) a continuous measurement, such as stack voltage or power and ii) intermediate test sequences such as polarization curves or electrochemical impedance spectroscopy (EIS). The continuous measurements such as voltage are convenient because they are often measured for control purposes, hence no additional sensors are required.

Polarization curves or EIS characteristics can provide more information on the fuel cell condition, than the simple continuous measurements. EIS is especially popular and have been very successful in diagnostics studies [27]–[29]. However, to obtain these characteristics, it requires dedicated measuring routines, the operation of the fuel cell is interrupted, and additional sensors or measuring circuitry may be required. The main categories of prediction methods are model-based, data-driven and a combination of the two (hybrid or fusion methods) [30]. The pros and cons of model-based versus data-driven methods are outlined from a diagnostics perspective in [31] and [32], respectively.

Model-Based Prognostics

Model-based prognostics methods rely on accurate analytical descriptions of the systems. Mostly empirical models are used rather than theoretical models, since the fuel cell is a multi-physics domain system and the degradation mechanisms are not yet well understood [33].

In [34] and [35] the continuous fuel cell stack voltage under constant load conditions is used in a model-based particle filtering framework to estimate degradation level and degradation rate. The degradation parameters are then used to predict the RUL. In [36], the fuel cell voltage under constant load current is used to estimate a state-of-health (SOH) metric. The SOH is determined in a degradation model by an observer-based extended Kalman filter (EKF) which is also used to for forecasting and consequently RUL estimation. The degradation model is obtained by parameter estimation from experimental polarization and EIS characterizations. Reference [37] uses EIS characterizations to fit an electrical equivalent model, which can be used for prognostics.

These model-based methods are able to accurately predict the degradation in the reported studies described above. The downside to model-based methods is that in-depth knowledge of the fuel cell physics and degradation mechanisms or extensive empirical testing is needed to parameterize the models. Furthermore, a model of a specific system may not apply to other systems experiencing different operating conditions or load profiles [38].

Data-Driven Prognostics

In data-driven prognostics methods, no knowledge of the physical properties of the operation or degradation of the fuel cell is needed. However, example experimental data of the degradation phenomena is required.

In [39], an artificial neural network (ANN) is used to model the polarization characteristics from current density, anode inlet temperature, and mass flow measurements. The method is applied throughout several thousand hours of operation, where it is shown to be capable of capturing the degradation of the fuel cell. In [40], wavelet analysis and several different degradation models are used for RUL prediction from stack voltage data. Both static and dynamic load conditions are investigated. References [41], [42] both use variations of ANNs, namely adaptive neuro fuzzy interface system (ANFIS) and long short-term

Table 1.2: Web of Science literature topic search results (as of July 8 2019)

Set	Search statement	Results
#1	TOPIC: ("backup power" OR "uninter* power" OR "emergency power")	45 377
#2	TOPIC: ("fuel cell*")	112 808
#3	TOPIC: (prognos* OR PHM)	628 594
#4	#1 AND #2	453
#5	#1 AND #3	889
#6	#2 AND #3	123
#7	#1 AND #2 AND #3	2

memory (LSTM) recurrent neural network (RNN). Both use stack voltage as health indicator and study both constant and dynamic load conditions. References [38], [43], [44] all use an echo state network (ESN), which is a variation of an RNN. In [43] the ESN is used to predict the mean cell voltage under constant load current. In [38], [44] the ESN predicts a virtual steady state stack voltage.

1.5 Literature in Numbers

Although prognostics of fuel cell systems is an ongoing topic of research, as outlined in the previous section, very few studies have addressed prognostics of fuel cells in a backup power system, where the operating conditions are very different from those presented in literature. Common for the literature studies is that the fuel cell stacks are continuously operating for extended periods which allows for continuously assessing the degradation level. In the backup power systems, the stacks are normally inactive and only operate sporadically for short periods. This makes the approaches described in literature difficult to apply.

To illustrate the amount of existing literature on the topics of this work, a literature search has been conducted on the research database Web of Science. The searched topics are backup power systems, fuel cells, and prognostics. The topics and their overlap are shown in the Venn diagram in Fig. 1.8. Some additional search terms are included to account for variations in terminology by different researchers. The complete search terms are shown in Table 1.2. The results show that there are roughly 45 000 publications on backup power systems, 113 000 publications on fuel cells, and 629 000 publications on prognostics. However, prognostics is a term which is used in many other fields such as medicine, which skews this number considerably. Within engineering categories there is only around 6 000 publications on prognostics.

Combining the search terms reveals the amount of research on the overlapping topics. For example Web of Science has cataloged 453 publications on the combination of backup power systems and fuel cells. Similarly the search shows 123 publications on prognostics of fuel cells. Combining all three topics, i.e.

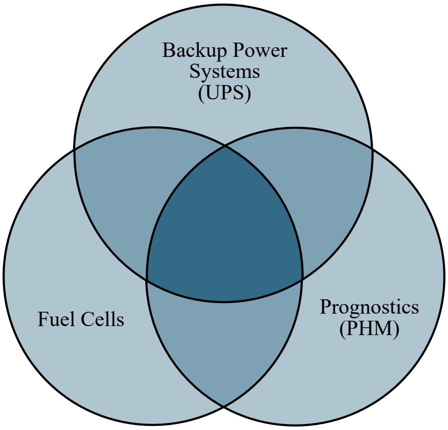


Fig. 1.8: Topics of the thesis

prognostics of fuel cell based backup power systems, only two publications are found, where of one is written by the author of this thesis.

This search is not exhaustive, but represents a large portion of existing literature and therefore illustrates the gap in research on the combined topic of prognostics of fuel cell based backup power systems.

1.6 Project Objectives

A good prognostics and health management (PHM) strategy is a key development towards more reliable, available, and economic fuel cell systems. PHM consists of three main parts: observation, analysis, and action. Whereof the subtasks of condition assessment and prognostics of the analysis part is the main focus in this work.

Fuel cell backup power systems have fundamentally different operating patterns than most other fuel cell-based systems investigated in literature. The backup systems often only operate for short bursts with irregular intervals. This means that the established monitoring and prognostics approaches, that investigate continuous or semi-continuous operating systems, are not directly applicable.

Many monitoring and prognostics approaches consider only constant operating conditions in a laboratory environment or are based on expensive and time-consuming tests of the fuel cells, such as electrochemical impedance spectroscopy (EIS). Although effective, this is usually infeasible in commercial systems.

A unique dataset consisting of measurements and system logs from numerous fuel cell based backup power systems has been made available for this work. The measurements originates from systems installed in the field, experiencing normal backup system operating conditions and therefore provides a unique picture of system performance in real field application. The data is further explained in Chapter 2.

This leads to the main objective of this project, which is:

To investigate how to assess and predict performance degradation of fuel cell based backup power systems based on historical data from normal backup power system operation.

To achieve this objective, the following research questions will be addressed in the thesis:

1. How to establish a set of performance metrics for quick and easy comparison of fuel cell stack performance and usage?
2. How to assess fuel cell stack degradation levels?
3. How to detect fuel cell stacks with abnormal performance levels or operating patterns?
4. How to predict future degradation levels of the fuel cell stacks?

The studies of each of the research questions take offset in the available dataset of normal in-field operation of a fleet of fuel cell based backup power systems.

1.7 Thesis Outline

The thesis contains two parts: i) an extended summary consisting of five chapters summarizing the project, and ii) the appended papers, which have been submitted for publication within the project period and which makeup the foundation of this thesis. The structure of the extended summary is as follows.

In Chapter 1, the background and motivation of the project was presented along with background knowledge on each topic areas that are relevant to this work, i.e. backup power systems, fuel cells, and the prognostics and health management framework. This led to the formulation of the project objectives.

Chapter 2 presents the system architecture and the dataset under investigation and how performance indicators can be extracted to provide a basis of comparison between stacks in the fleet of backup power systems. Both single value metrics and time series are extracted.

Chapter 3 presents methods of detecting abnormally performing stacks based on the single value metrics extracted in Chapter 2, and to detecting groups of similar stacks based on the time series metrics.

In Chapter 4 methods of predicting future stack performance levels are presented. Two approaches are explored: i) predicting stack degradation level from examples of historic stack degradation levels and ii) predicting the underlying stack parameters from examples of several historic stack parameters.

Finally, Chapter 5 concludes the extended summary of the thesis.

2

Data Foundation and Performance Indicators

This chapter starts by introducing the fuel cell based backup power systems that are investigated in this work. Raw data from measurements and system logs are available from a number of backup systems that have been running in the field. The data is structured in a database and key performance indicators are extracted to provide a simplistic overview of the performance and usage of each stack. These metrics allow the comparison between stack performances and provides a holistic view of the performance of the fleet as a whole. Stacks that experience widely different performance or operating conditions compared to other stacks in the fleet can be detected as outliers as described in Chapter 3.

2.1 System Description and Measurements

The system under investigation is a 5 kW backup system based on two 2.8 kW PEM fuel cell stacks. The two fuel cell stacks are each supplied with hydrogen and air to produce electrical energy, which is supplied to the load through DCDC converters. The electrical power produced by each stack is handled by three parallel DCDC converters. Two ultracapacitor modules, one for each fuel cell stack, supplies the load with power during the startup of the fuel cells. An illustration of this system architecture is depicted in Fig. 2.1.

Other than these mentioned components, the system contains several essential components including valves, booster circuits, controller units, test loads, and many more. The configuration of which will not be addressed further.

The described system is primarily used for backup in telecommunication sites, where they provide backup power to mitigate fluctuations and failures in the electrical grid. The partner company has many such systems in operation in

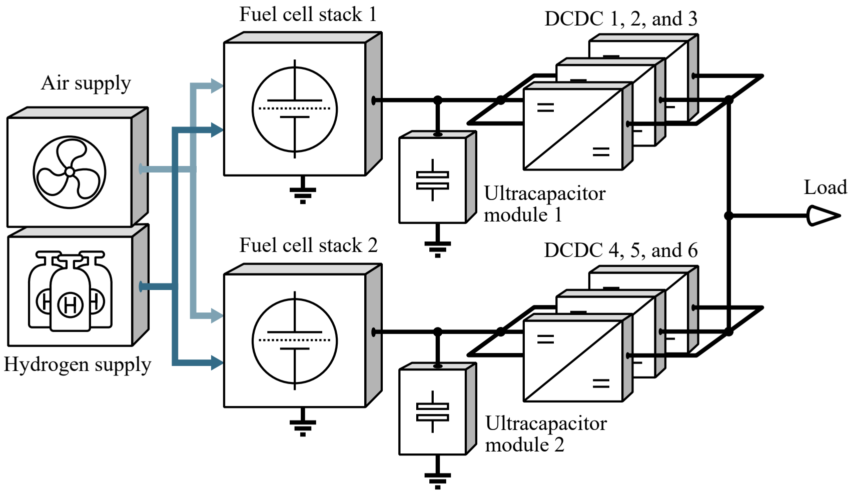


Fig. 2.1: Backup system architecture

Denmark as well as abroad. During the lifetime of each system, measurements of various system parameters has been performed and collected.

2.1.1 MEASUREMENTS

The measurements, that are performed on the systems in operation include electrical and temperature measurements on each fuel cell stack as well as on each of the DCDC converters; set points of the cathode fan speed, stack heater, and the proportional valve; temperatures of the valve block, air inlet, air outlet, controller box, and externally between the two stacks; inlet temperature of the hydrogen; system output voltage. A complete list of the measured parameters is shown in Table 2.1.

Table 2.1: Measured and logged system parameters

No.	Parameter	Description	Unit
1	`Date`	Date	[D-M-Y]
2	`Time`	Time	[H:m:s]
3	`Stack_heat`	Stack heater setpoint	[%]
4	`Fan_speed`	Cathode fan setpoint	[%]
5	`VALVE_PWM`	Proportional valve setpoint	[%]
6-11	`Iin_DCDC{1-6}`	DCDC converter {1-6} input current	[A]
12-17	`Uin_DCDC{1-6}`	DCDC converter {1-6} input voltage	[V]

continued . . .

...continued

No.	Parameter	Description	Unit
18-23	`Iout_DCDC{1-6}`	DCDC converter {1-6} output current	[A]
24-29	`Uout_DCDC{1-6}`	DCDC converter {1-6} output voltage	[V]
30-35	`T_DCDC{1-6}`	DCDC converter {1-6} temperature	[°C]
36	`T_Valves`	Valve block temperature	[°C]
37	`T_Stack_1_1`	Stack 1 temperature (sensor 1)	[°C]
38	`T_Stack_1_2`	Stack 1 temperature (sensor 2)	[°C]
39	`T_Stack_2_1`	Stack 2 temperature (sensor 1)	[°C]
40	`T_Stack_2_2`	Stack 2 temperature (sensor 2)	[°C]
41	`T_Air_Inlet`	Air inlet temperature	[°C]
42	`T_Air_Outlet`	Exhaust air temperature	[°C]
43	`T_Contr_Box`	Controller box temperature	[°C]
44	`V_stack_1`	Stack 1 voltage	[V]
45	`V_stack_2`	Stack 2 voltage	[V]
46	`I_stack_1`	Stack 1 current	[A]
47	`I_stack_2`	Stack 2 current	[A]
48	`H2_P_in_LP`	Hydrogen inlet pressure (low pressure side)	[mBar]
49	`H2_P_in_HP`	Hydrogen inlet pressure (high pressure side)	[mBar]
50	`U_line`	System output voltage	[V]
51	`T_FCC_Room`	Temperature in the area between the fuel cell stacks	[°C]
52	`Cath_fan_tacho`	Cathode fan tacho feedback	[Hz]

2.1.2 OPERATING MODES

The system can operate in a number of different modes numbering 0 to 10. The most relevant modes of operation are listed in Table 2.2. Modes 0, 1, and 7-10 are related to service and safety, and are not used in this work. The modes are used for the overall system and for each of the fuel cell stacks individually. Hence, the system can be in backup mode, i.e. providing power to the load, where one fuel cell stack is supplying all the power, while the other remains in standby mode. That is, the system mode is logged as 4, stack 1 mode as 4, and stack 2 mode as 2 or vice versa. The system and stack modes are logged together with the measured parameters as `SYS_MODE`, `ST1_MODE`, and `ST2_MODE`, respectively.

Table 2.2: System and stack modes

No.	Mode	Description
2	Standby	When the system is inactive but available
3	Powerup	Powering up the fuel cells before taking over the loads
4	Backup	Powering the load from the fuel cells
5	Powerdown	Ramping down power after backup or selftest event
6	Selftest	A power ramp-up used for testing and exercising the fuel cells

Typical Operating Pattern

During normal operation, the system is in a standby mode. That is, the fuel cell stack is turned off and the backup system does not provide power to the load. This mode is used when the grid is supplying power to the load without interruptions. A backup system supplying a normal telecommunication site will spend the majority of its lifetime in standby mode.

When a grid failure occurs, the fuel cells must ramp up its produced current to take over the load power. In the mean time, the ultracapacitors provide the load power. This operating mode is called power-up. When the fuel cells are fully powered up, they take over the entire load power and the system has entered its backup mode. When grid power is restored, the fuel cell current is ramped down in the power-down mode.

The self-test mode, is used to regularly exercise the fuel cell stacks during prolonged standby periods. During self-test, the stack is powered up to a constant power level and kept at this level for a few minutes after which it is powered down again using the power-down mode. The power produced during self-tests is dissipated in a test load. One stack is tested at a time: stack 1 followed by stack 2. When the self-test is being performed on one stack, the other remains in standby mode.

2.1.3 Structuring the Data

All of the logged parameters are stored on local SD cards at each system site and are collected manually on service visits. The data is stored in CSV files, one or more for each day, such that file `01021601.CSV` is the parameters logged on 1st of February 2016, the last ‘01’ denotes that it is the first file from that day. Hence, the data from each site is stored in at least 365 separate CSV files per year. This data structure makes it difficult to analyze the complete dataset of a specific site, not to mention analyzing data across the entire fleet of systems. Therefore, the first step is to store all the data in a database structure, where it can be accessed and segmented more conveniently.

The programing language Python was used to write a script, which automatically goes through a data folder containing a subfolder for each system. These system folders may contain one or more additional layers of subfolders, in which the CSV files with raw data is located. The script uses Python's PyMySQL package to set up a MySQL database table containing all the raw data from all the systems. Additionally to the raw data parameters, a system id column is created in the database.

From the MySQL database, subsets of the raw data can be easily accessed using SQL queries such as shown in Listing 2.1, which would extract the columns `datetime`, `sys_mode`, and `st1_v` (stack 1 voltage) from the `raw_data` table, but only for the system with `system_id` of 1 and for dates later than January 1st 2017. This extraction of data, without an SQL database, would be very tedious and time consuming.

Listing 2.1: SQL query example

```sql
SELECT datetime, sys_mode, st1_v
FROM raw_data
WHERE system_id = 1 AND datetime >= "2017-01-01"
```

2.2 Extracting Performance Indicators

With the amount of raw data, it can be difficult to get an overview of how each system performs in comparison to the fleet of systems. Therefore, it is advantageous to extract some simple metrics, that convey the performance and usage of each system. These metrics are referred to as performance indicators or key performance indicators (KPI). Two types of performance indicators are extracted from the raw data set: counters and totalizers - measures of how much the system has been used and in which modes it has operated; and self-test characteristics - metrics extracted from each self-test, which allows for tracking the dynamics in the performance over the system's lifetime.

2.2.1 Counters and Totalizers

The counters and totalizers include various operating times, production and consumption levels, and number of system startups. The definition of each indicator is described in the following paragraphs.

System Startups

A system startup is the process of going from an inactive stack to an active stack, i.e. providing power. The definition of system startup is when the stack current has increased from 0 A to above 5 A, i.e. 5 A is the threshold for the stack being active. Each system startup is associated with some degradation of the fuel cell. The startups are extracted by counting events where the current is increased from 0 A to above 5 A.

When the stack has been inactive for some time, there is air present on the anode side. The air needs to be replaced with fresh hydrogen during startup. The filling of hydrogen into the anode will create some internal current that can lead to corrosion and thereby irreversible degradation. The corrosion currents are reduced by quickly purging the hydrogen into the anode as well as applying a load during startup to draw down the cell voltage. These events are called air-air startups and are extracted by counting the startups where the average cell voltage is below 0.1 V.

OPERATING TIME

Active operation of the fuel cell is another cause of performance degradation. The active operating time (`Runtime`) of the fuel cell stack can be calculated by looking at when the stack has supplied more than a certain threshold level of current (5 A). The active operating time is derived by finding the instances where the current rises above and falls below the threshold level.

To give an overview of the usage profile of the systems, the time spent in each of the operating modes (Table 2.2) can be investigated. The mode operating times are calculated by finding the instances where a certain mode is entered (changes from a different mode) and exited (changes to a different mode). An example of the mode logs and mode times of a specific stack is shown in Fig. 2.2. The figure shows the system modes versus the date and the total time spent in each of the system modes during the total operating time of an example system. The 'unknown' mode is not an actual operating mode, but simply indicates that no other system mode has been logged during a time interval. This might be caused by a number of things, including transport time to the system site, missing or corrupted data points, or downtime in relation to service of the system. The figure clearly indicates that the normal state of the system is in standby mode and only a small fraction of the time, is the system actually actively operating.

The fuel cells have an optimum operating temperature approximated by a linear relationship with the produced current (i): $T_{\text{opt}} = a \cdot i + b$. The time the fuel cell stack spends at non-optimum temperatures can also have an influence on the performance level. This is split into two indicators: under- and over-temperature:

$$T_{\text{under}} \leq T_{\text{opt}} - \Delta T \tag{2.1}$$

$$T_{\text{over}} \geq T_{\text{opt}} + \Delta T \tag{2.2}$$

where ΔT defines the tolerable deviation from the optimum operating temperature. The under/over-temperature operating times are extracted for each of the systems modes (`Runtime_undertemp_{mode}` and `Runtime_overtemp_{mode}`) as well as for the active operating time (`Runtime_overtemp` and `Runtime_undertemp`).

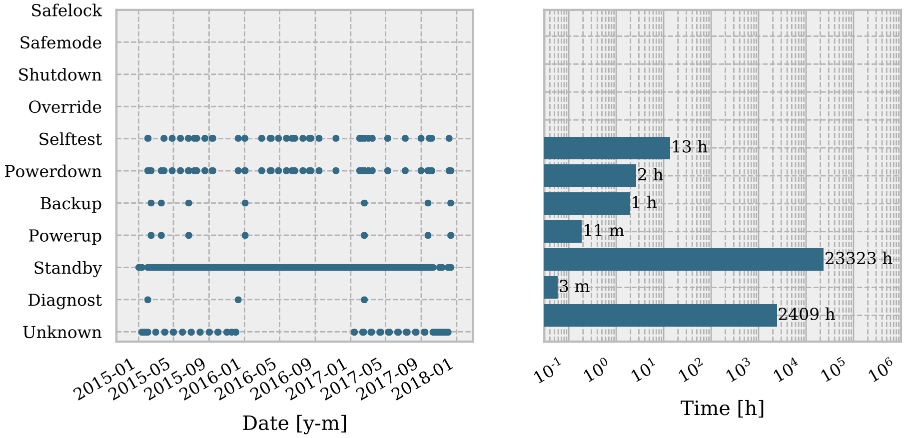


Fig. 2.2: Example system operating modes. Left: dots indicate instances of the different system modes versus date. Right: Total time spent in each mode.

Production and Consumption

Another way of quantifying the usage of the stacks is to calculate the amount of energy and charge it has produced over its lifetime. The energy produced by the fuel cell stack is the integral of the fuel cell power as such

$$E = \int_0^{t_{end}} V(t) \cdot I(t) \mathrm{d}t \tag{2.3}$$

for the sampled measurements, trapezoidal integration is used to approximate the energy:

$$E = \sum_{i=1}^{L} \frac{V_{i-1} \cdot I_{i-1} + V_i \cdot I_i}{2} (t_i - t_{i-1}) \tag{2.4}$$

where L is the number of samples in the dataset. Similarly, the produced charge, which is the integral of the produced current, is calculated by

$$Q = \sum_{i=1}^{L} \frac{I_{i-1} + I_i}{2} (t_i - t_{i-1}) \tag{2.5}$$

The charge can then be used to approximate the consumed hydrogen and oxygen, respectively. The reaction levels are calculated from (2.6) and (2.7), where K_{H2} and K_{O2} are empirical constants.

$$\dot{m}_{H2} = K_{H2} \cdot Q \tag{2.6}$$

$$\dot{m}_{O2} = K_{O2} \cdot Q \tag{2.7}$$

Table 2.3: Counters and totalizers

KPI	Description
`Startups`	Number of startups of each stack
`Airair_startups`	Number of startups of each stack when there is no hydrogen at the anode
`Runtime`	Operating time above current threshold
`Runtime_{mode}`	Operating time in each mode
`Runtime_overtemp`	Operating time at above optimum temperatures above current threshold
`Runtime_overtemp_{mode}`	Operating time at above optimum temperatures for each mode
`Runtime_undertemp`	Operating time at under optimum temperatures above current threshold
`Runtime_undertemp_{mode}`	Operating time at under optimum temperatures for each mode
`Charge_produced`	The amount of electric charge produced
`Energy_produced`	The amount of electric energy produced
`Oxygen_reacted`	The amount of reacted oxygen
`Hydrogen_reacted`	The amount of reacted hydrogen

EXTRACTED VALUES

The list of KPIs are summarized in Table 2.3 and the extracted KPIs for each stack in the available systems' data is shown in the boxplot of Fig. 2.3. This provides an overview of each stack in the fleet of backup power systems and the distribution of each KPI for the fleet. This is the basis for detecting abnormally performing stacks, which is presented in Chapter 3.

2.2.2 SELF-TEST CHARACTERISTICS

As the systems spend most of the time in standby mode, where it is difficult to extract useful performance indications, the active periods can be used for getting a picture of the change in performance over the stack lifetimes. However, the backup events are not predictable and vary in their load profile. The self-tests, on the other hand, are more consistent and typically more numerous. Hence, these will be used as the basis for extracting performance indicators.

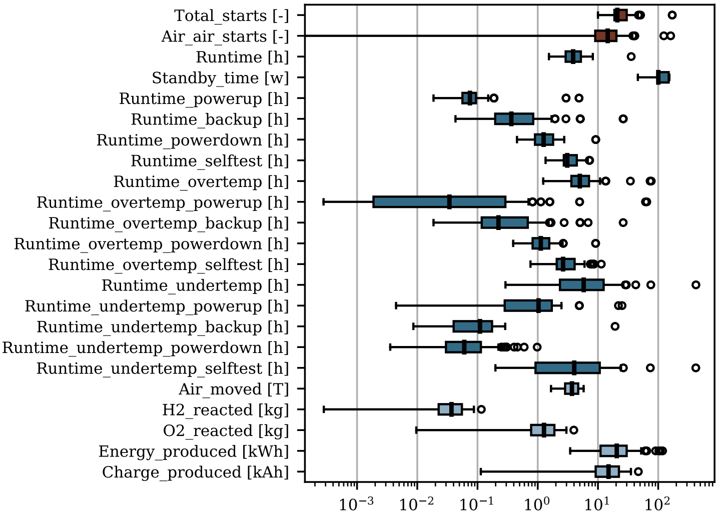


Fig. 2.3: Counter and totalizer performance indicators boxplot

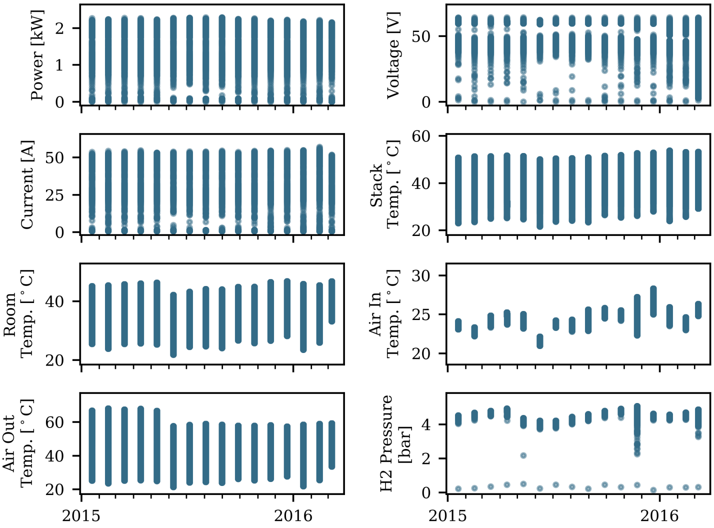


Fig. 2.4: Self-test data of a specific stack

The investigated systems parameters are the stack power, stack voltage, current, stack temperature, room temperature (between the two stacks), air inlet temperature, air outlet temperature, and the hydrogen inlet pressure. These parameters during stack self-tests are shown for a single stack in Fig. 2.4.

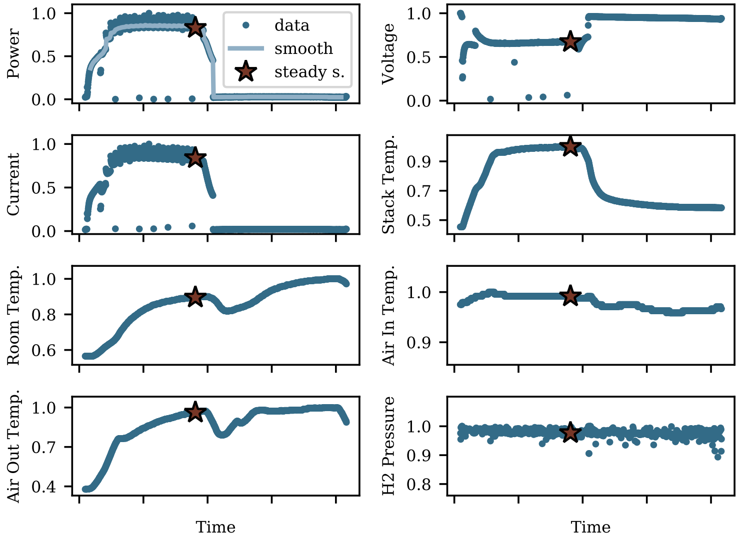


Fig. 2.5: Example of extracting steady state values of a self-test of a specific stack

A zoomed view of a single self-test is shown in Fig. 2.5 where the data has been normalized. This shows the self-test procedure of ramping up the current until a certain power level, which is then kept constant for some time before ramping down again. During the constant power segment, the voltage slightly increases and the current decreases accordingly. This is due to a recovery effect, where some of the reversible degradation from the standby period is recovered.

The fuel cell temperature increases to a steady value along with the power increase and falls when the test is done. The room temperature increases throughout the test as the fuel cell produces heat. The continuous increase after the end of the test is caused by the subsequent self-test in the other stack in the system. The air inlet temperature is approximately constant, whereas the outlet temperature increases as the produced heat is removed through the outlet airstream. The hydrogen pressure is approximately constant.

The extracted self-test characteristics of each stack are shown in Fig. 2.6 where three stacks have been highlighted in different colors to show the progression of values for individual stacks.

STATE OF HEALTH

State of health (SOH) is a single parameter which indicates the degradation level of a system. The SOH is usually given as a percentage between 0% and 100%, where 100% SOH indicates a completely healthy system and 0% SOH indicates a fully degraded system. SOH is commonly used in battery systems [45], [46].

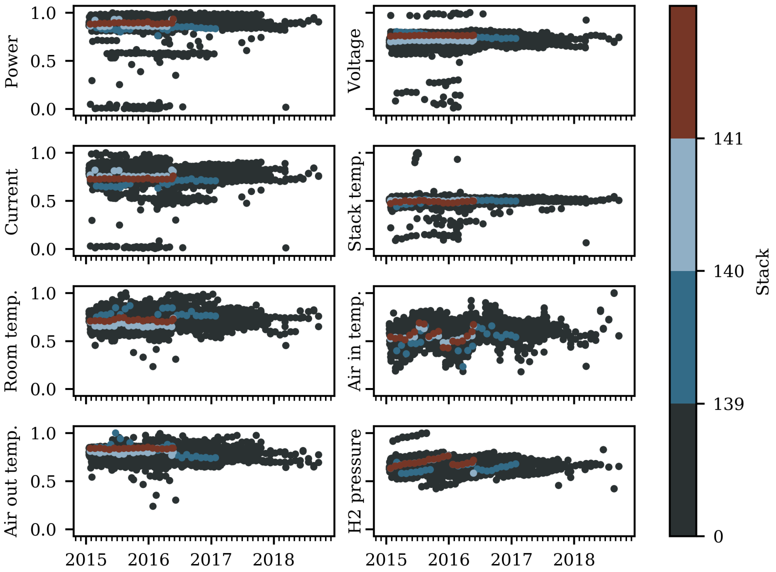


Fig. 2.6: Extracted self-test characteristics of all fuel cell stacks. Three stacks have been highlighted in blue, light-blue and red, respectively.

There is no standard in calculating the state of health of fuel cell stacks. The most common metric is to look at decline in stack voltage at a given load current [34], [35], [43]. However, the self-tests of this study are conducted in a way that the power is ramped to a certain value. Hence, as the fuel cells degrade and the voltage drops, the needed current to maintain the power level must increase. Therefore, the drop in voltage can not be directly used as SOH indicators, as they are taken at different load levels. In [47], Dolenc et al. uses voltage and temperature measurements to estimate the area specific resistance (ASR) of a solid oxide fuel cell. A threshold for the ASR is used as an end of life indicator.

Another way of determining the SOH, is using characterizations of stacks at their beginning of life (BOL) and at their end of life (EOL). This method was presented in [48], where the so called representative loss, in the form of resistance, is compared to the loss of a new cell (at BOL) and the loss at which the cell can no longer provide the power demanded by the application (at EOL). The formula for this comparison takes the general form:

$$SOH = \frac{HI - HI_{EOL}}{HI_{BOL} - HI_{EOL}} \tag{2.8}$$

where HI is the health indicator - in the case of [48], the representative loss. For the fuel cell under investigation in this work, BOL and EOL criteria are given by the manufacturer in the form of polarization curves as shown in

Fig. 2.7. The points of the BOL and EOL criteria are fitted to the polarization equation of (1.8) to obtain the polarization curve functions, $V_{BOL}(I)$ and $V_{EOL}(I)$, respectively.

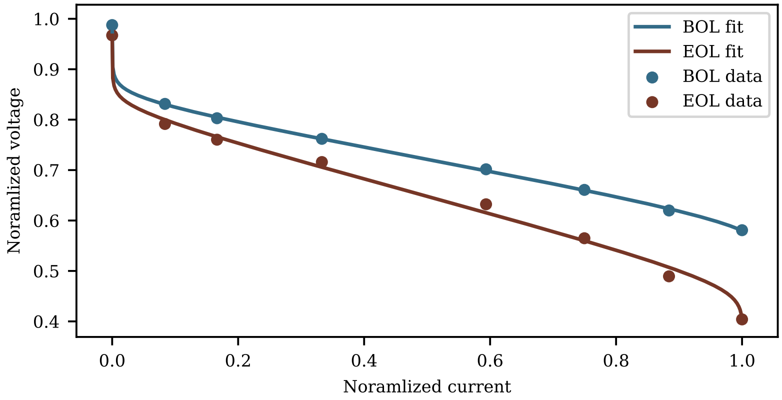


Fig. 2.7: Beginning of life and end of life polarization curves

These BOL and EOL functions are then used to estimate the SOH from the steady state voltage and current measurements of each self-test using the relation from (2.8):

$$SOH = \frac{V - V_{\mathrm{EOL}}(I)}{V_{\mathrm{BOL}}(I) - V_{\mathrm{EOL}}(I)} \tag{2.9}$$

where V is the extracted steady state self-test voltage and I is the extracted steady state self-test current.

Calculating the SOH for each stack, the curves of Fig. 2.8 are obtained. The plot shows the wide range of where the stacks are in their life cycle. Some are near to the end of life criteria ($SOH \approx 0$) while some still perform better than expected for a newly installed stack ($SOH > 1$).

It should be noted that the BOL and EOL criteria are idealized thresholds obtained in a laboratory setting and do not consider the specific application. The EOL limit has been determined as the point when the peak power point is reached at the maximum current of the stack. A more suitable EOL criteria could be set for this specific application, if the SOH values should be used directly for determining the sufficiency of the fuel cell performance. This SOH does provide a relative measure of the degradation level of the fuel cell and will be used as such.

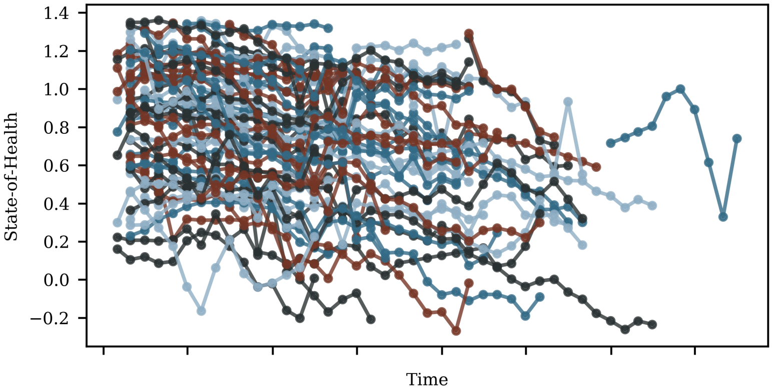


Fig. 2.8: State of health of each stack in the dataset

3

Outlier Detection and Time Series Clustering

Having extracted key performance indicators, describing the usage and performance levels of each stack in the available dataset, a method of comparing the fuel cell stacks is needed. Section 3.1.1 presents a method of using the counters and totalizer KPIs, presented in Section 2.2.1, to detect stacks that perform significantly different compared to the fleet of systems as a whole. Also, the self-test characteristics, presented in Section 2.2.2, are analyzed. A method of assessing similarities in the time series and grouping the stack based on these similarities is presented in Section 3.2.

3.1 Performance Outliers

The fuel cell stacks in the dataset undergo widely varying operating conditions and load profiles. The extracted counter and totalizer KPIs provide simple metrics to evaluate the usage and performance of the stacks. However, the varying age, operating time, etc. of the fuel cell stacks makes it difficult to directly compare the KPI values between the stacks. Therefore, a method of detecting outlier stacks, considering all the KPI metrics simultaneously, is needed.

Since there is no prior examples of which stacks can be considered normal or outliers, an unsupervised method is used. Furthermore, the distribution of the KPIs is not guaranteed to being Gaussian and several distinct groups may be present within the dataset. Hence, a density based method, namely local outlier factor (LOF) is chosen for this task.

3.1.1 LOCAL OUTLIER FACTOR

The local outlier factor algorithm is an unsupervised machine learning method of detecting outliers in a dataset. LOF compares the local density of a given point and compares it to that of the neighboring points to attribute an outlier factor to the point. The point can then be classified as an inlier or an outlier based on this outlier factor. [49], [50]

The LOF algorithm uses the distances between points in the dataset to establish a set of k nearest neighbors for each point. For a given point A, the set is denoted as $N_k(A)$, which contains all points that are within the k-distance (d_k), i.e. the distance between A and its k^{th} nearest point. The local density of a point is estimated by the local reachability distance (lrd), which is defined as the ratio between the number of points in the set of k nearest neighbors and the sum of reachability distances (rd) between point A and each point in the set of k nearest neighbors:

$$lrd_k(A, B) = \frac{|N_k(A)|}{\sum\limits_{B \in N_k(A)} rd(A, B)} \tag{3.1}$$

The reachability distance is the maximum of the k-distance of point B and the true distance between points A and B (d):

$$rd(A, B) = \max\{d(A, B), d_k(B)\} \tag{3.2}$$

Finally, the LOF is calculated as in (3.3).

$$LOF_k(A) = \frac{\sum\limits_{B \in N_k(A)} \frac{lrd_k(B)}{lrd_k(A)}}{|N_k(A)|} \tag{3.3}$$

The LOF value gives an indication of how similar a given point is to other points in its neighborhood. An LOF score below or close to 1 indicates an inlier point, i.e. a point with high degree of similarity to its neighbors, and an LOF score much greater than 1 indicates an outlier. Where exactly the threshold between inliers and outliers should be located depends on the dataset. However, in the case of the performance levels of fuel cell stacks, it makes sense to look at the LOF values directly and address the stack with the highest LOF value first.

Calculating the LOF values of the KPIs of each stack, as presented in [C], the distribution of the LOF values are shown in Fig. 3.1. The majority of the stacks score an LOF value at around 1. A few stacks score LOF values above 4, which are definitely abnormal.

Looking at one of the abnormally performing stacks: stack 2 of system 82, which has an outlier factor of 6.6, it has abnormally high values in the number of startups and runtime in both powerup and powerdown modes. If the runtime in backup mode would have been equally elevated, this might have been a sign that the system was simply providing backup power more frequently than other

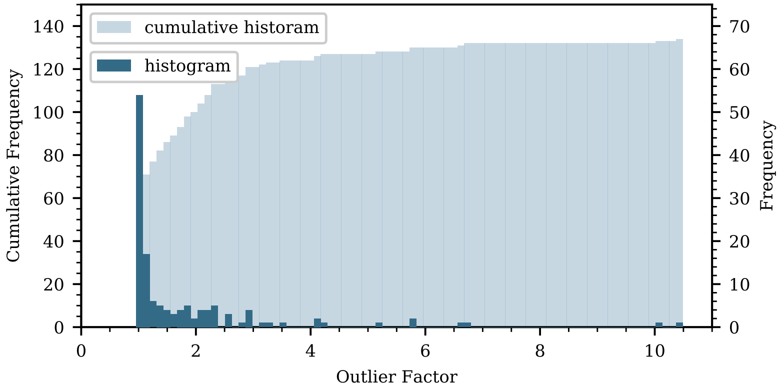


Fig. 3.1: Local outlier factor histogram [C]

systems. However, the backup runtime is low, indicating that some error might have occurred at the site.

This is confirmed when looking at the operating pattern of the stack (Fig. 3.2). For a few days, the stack continuously powers up to enter backup mode for only a few seconds after which it powers down again. This leads to the low backup time and high powerup/powerdown time and number of startups, which is detected as an outlier.

3.2 CLUSTERING SELF-TEST TIME SERIES

Similarly to comparing stacks based on their single-value KPIs, as described in the previous section, it would be advantageous to have a way of comparing the stacks based on the self-test time series described in Section 2.2.2. However, a direct comparison is not straight forward. Especially, the variation in sequence length makes the time series of different stacks difficult to compare.

3.2.1 DYNAMIC TIME WARPING

Usual similarity measures, such as Euclidian distance, is not well suited for the case of asynchronous time series of variable length. Hence, an alternative method is needed. A popular method of shape similarity quantification is the dynamic time warping (DTW) algorithm [51], which has been successfully applied in areas such as speech recognition [52], data mining [53], and prognostics of battery systems [54]. The DTW algorithm warps the time series in the time dimension in order to achieve an optimal alignment between them and thereafter calculating the distance.

For two time series of different length: $Q = (q_1, q_2, ..., q_n)$ and $C = (c_1, c_2, ..., c_m)$, an $n \times m$ matrix (D) is constructed. The matrix is filled so

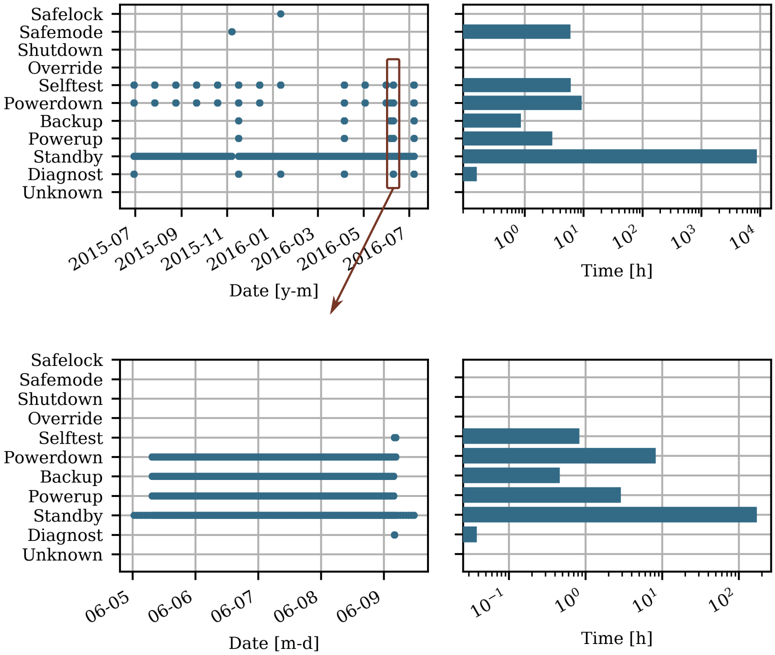


Fig. 3.2: Operating pattern of a detected abnormal stack. Top: full view, bottom: zoomed view of problem period.

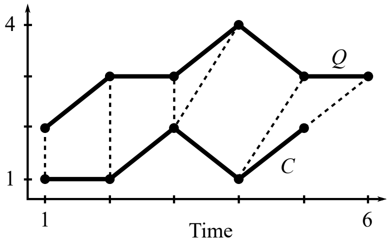
that the entry $d_{(i,j)}$ is the absolute difference between q_i and c_j plus the minimum value of the previous adjacent matrix entries, i.e.

$$d_{(i,j)} = |q_i - c_i| + \min\{d_{(i-1,j)}, d_{(i-1,j-1)}, d_{(i,j-1)}\} \tag{3.4}$$

This procedure is illustrated in the toy example of Fig. 3.3. The two series, Q and C are of length 6 and 5, respectively and are depicted on the left half of the figure. The right side of the figure shows the resulting 6×5 matrix.

The DTW algorithm tries to find the optimum path going from the starting point of each series (bottom left corner of the matrix) to the end of both series (top right corner of the matrix). The optimum path is the path giving the lowest score when summing the numbers along the path. In the shown example, the optimal path is shown in bold numbers and equates to a similarity score for the two series of 31.

In effect, the DTW algorithm allows the comparison of two arbitrary time series by a single similarity metric. In paper [E], the DTW algorithm is used to estimate the similarities between the fuel cell stacks in the dataset based on a number of features. The features used for estimating similarity are the time series of voltage, current, stack temperature, room temperature, air-in



		Q-series →					
		2	**3**	**3**	**4**	**3**	**3**
C-series →	**2**	3	4	5	7	8	**9**
	1	3	4	5	7	**8**	9
	2	2	3	**4**	**6**	7	8
	1	2	**3**	5	8	10	12
	1	**1**	3	5	8	10	12

Fig. 3.3: Dynamic time warping example of two time series of unequal length

temperature, air-out temperature, and hydrogen inlet pressure, extracted from the fuel cell self-tests as presented in Section 2.2.2. The results of the DTW algorithm on these time series is depicted in the matrix plot in Fig. 3.4. The color of each pixel in the matrix represents the similarity between the stacks indicated on each of the axes. A darker pixel indicates larger DTW distance and thereby lower similarity. It is clear that some stack are more alike than others. In the following section, it is described how these DTW distances can be used to detecting groups of similarly behaving stacks.

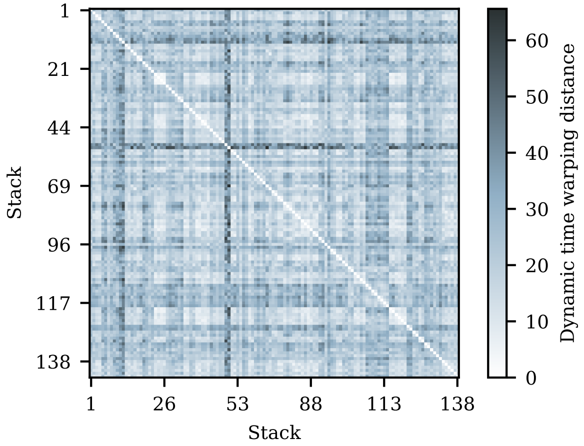


Fig. 3.4: Dynamic time warping distances between stack feature time series [E]

3.2.2 DBSCAN

Density-based spatial clustering algorithm for applications with noise, more commonly known as DBSCAN, is a popular machine learning algorithm for detecting clusters within datasets [55]. DBSCAN builds on some of the same concepts as the previously described LOF algorithm for detecting outliers.

The algorithm takes two parameters: the maximum radius of the neighborhood (ϵ), and the minimum number of points that constitute a cluster ($MinPts$). The number of points within the ϵ-neighborhood for each point is a measure of that point's density. $MinPts$ can be thought of as a threshold to the number of points in the neighborhood that defines the allocation of points to clusters. For a given point A, point B is in A's neighborhood if the distance between A and B is less than or equal to ϵ. The neighborhood of A, i.e. the set of points within ϵ distance of A, is denoted as $N_\epsilon(A)$.

If there are $MinPts$ or more points in $N_\epsilon(A)$, A is considered a core point. If $N_\epsilon(A)$ contains fewer than $MinPts$ points, but A is reachable (within ϵ distance) of another core point, A's is categorized as a border point. If neither of the previous statements are true, A is a noise point.

Clusters are detected by looking at each detected core point. Any chain of points that are sequentially reachable from the given core point (A), is considered as being within the same cluster as A. Hence, clusters can contain both core and border points, but never noise points as these can only be reachable from other noise points.

Fig. 3.5 shows the workings of the DBSCAN algorithm on a toy dataset. On the left hand side, the data points are shown along with their ϵ-distance radius. In the example, the $MinPts$ parameter is set to 4. The detected point categories are shown in the different colors - blue for core points, gray for border points, and red for noise points. On the right hand side of the figure, the resulting clusters are shown in blue and light-blue, respectively.

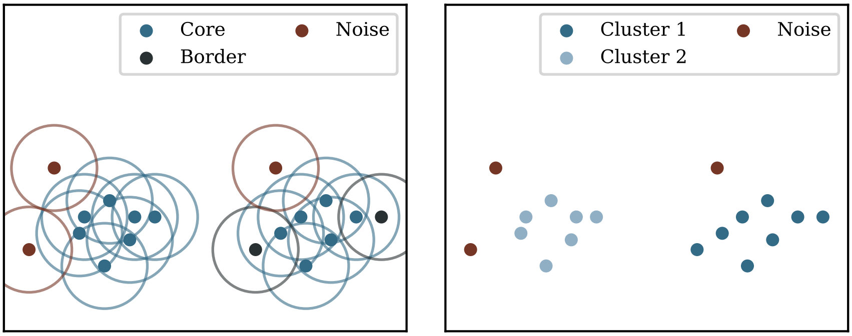


Fig. 3.5: Example of DBSCAN algorithm with $MinPts = 4$. Left: ϵ-distance and determination of core, border, and noise points. Right: resulting clusters.

The DBSCAN algorithm is advantageous in that it can detect clusters of any shape and it does not need a prior specification of the expected number of clusters, which makes it well suited for applications with an unknown number of clusters. Furthermore, the algorithm is robust to noise, as these points are detected and ignored in the cluster formation.

In paper [E], the DBSCAN algorithm is used to detect clusters in the similarity measures of the stack time series extracted by the DTW algorithm. To reduce the complexity and improve the visualization of the DTW data, a dimensionality reduction method is first applied to reduce the number of features from more than 140 to 4. The used algorithm is the widely used principal component analysis (PCA) [56]–[58], which is presented in more detail in [C], [E].

The result of applying DBSCAN to the PCA transformed DTW data is shown in Fig, 3.6. Each subplot shows one transformed-space feature (principal component) plotted against another feature. Instead of plotting each feature against itself in the diagonal subplots, these show the data distributions in a histogram. Each point in the scatter plots represent one stack. The parameters of the algorithm was set to $MinPts = 25$ and $\epsilon = 0.85$. Two clusters of 67 and 33 stacks, respectively were detected in the data as well as 26 noise points. The clusters are mainly separated in one of the four dimensions.

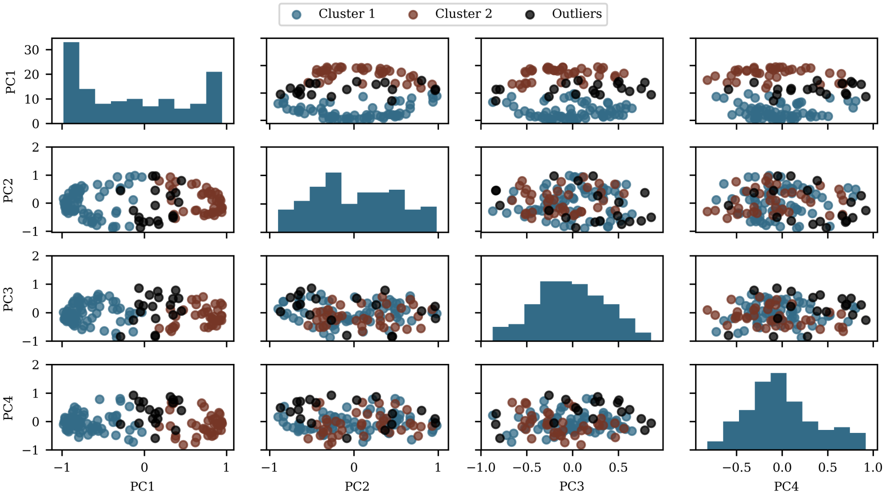


Fig. 3.6: Scatter matrix plot of each principal component of the transformed dynamic time warping distance calculations on the stack time series [E]

These detected clusters of stacks are used in paper [E] to test the accuracy of prediction models trained on groups of data versus the whole dataset. This aspect of the predictions will not be addressed further in this thesis. For more details on this, please refer to paper [E].

4

Predicting Performance Degradation

This chapter presents the chosen approach for predicting the future performance and degradation level of the fuel cell stacks in the backup power systems. Because the data of individual stacks is relatively sparse, a single machine learning model is trained on multiple stacks. This gives more training data for the model and results in a more general model that can predict future values of any stack in a similar system.

Two approaches to predicting the performance degradation of the fuel cell stacks has been investigated: i) forecasting SOH values directly as presented in [D] and ii) forecasting the underlying stack voltage, current, and temperature from varies stack measurements as presented in [E]. Common for both approaches is that they use variants of artificial neural networks (ANN) trained on historical data from all stacks in the dataset.

The following section will present, the basics of artificial neural networks and its adaptation to temporal applications, the used architectures of the networks and their implementation. Finally, the results of the predictions are presented.

4.1 ARTIFICIAL NEURAL NETWORKS

The basis for most ANNs is the neuron model, which takes an input vector ($\boldsymbol{x}$), multiplies it by a weight matrix ($\boldsymbol{W}$), adds biases ($\boldsymbol{b}$), and applies an activation function (σ):

$$\boldsymbol{y} = \sigma(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b}) \tag{4.1}$$

Interconnecting multiple artificial neurons in networks such as shown in Fig. 4.1, gives the ability to model complex nonlinear phenomena. In the shown ANN, the white nodes represent individual neurons whose output is the nonlinear function (4.1) of the collection of its inputs. In the input layer, there is one neuron for each of the input features ($\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$ in Fig. 4.1) and the output layer has the number of neurons corresponding to the wanted output features ($\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ in Fig. 4.1). The layers in between are the hidden layers and can have any number of neurons depending on the complexity of the phenomena which should be modeled. In the hidden layers, the input of each neuron is the collection of all neuron outputs from the preceding layer.

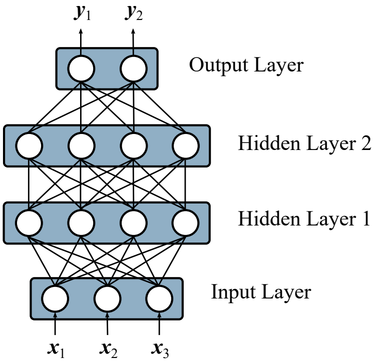


Fig. 4.1: Artificial neural network with four layers consisting of three, four, four, and two neurons for the input layer, first hidden layer, second hidden layer, and output layer, respectively

ANNs are trained on historical examples of input and output features. The weights and biases are updated iteratively for each example of training data using a gradient descent method known as backpropagation. The backpropagation calculates the gradient of the chosen loss function with respect to the weights and updates the weights based on the calculated gradient.

4.1.1 RECURRENT NEURAL NETWORKS

A variation of ANN is the recurrent neural network (RNN) which is adapted for temporal data modeling. This makes RNNs well suited for predicting time series, such as the fuel cell stack state of health progression over time.

The basic recurrent neuron is depicted in Fig. 4.2. The recurrent neuron takes the input ($\boldsymbol{x}$) as well as the previous output of the neuron, as a “memory” input, and performs the weight-bias-activation operation described previously.

$$\boldsymbol{y}_t = \sigma(\boldsymbol{W}[\boldsymbol{x}_t, \boldsymbol{y}_{t-1}] + \boldsymbol{b}) \tag{4.2}$$

When the neuron processes the second temporal input ($\boldsymbol{x}_2$ of the unfolded recurrent neuron; Fig. 4.2), the neuron also receives an input with information from the previously processed input ($\boldsymbol{x}_1$). Note that the unfolded recurrent neurons all share the same weights and biases, i.e. they are essentially the same neuron, but receive different inputs at different times.

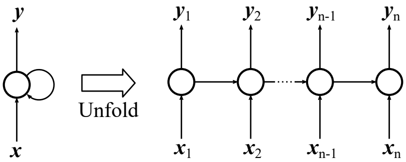


Fig. 4.2: Recurrent neural network cell and its unfolding

When feeding the neuron outputs forward in the temporal dimension, weights and biases are used to scale and offset the input similarly to the operation in the non-temporal dimension. One problem with this structure is that the memory of the first input is multiplied with the weight in each temporal step. For small weights, this can lead to the vanishing gradient problem, where the memory from the early inputs is washed out over time. Or, conversely, an exploding gradient problem for large weights. There are several adaptations of the RNN that solve this issue, the most used is the long short-term memory (LSTM) cell.

LONG SHORT-TERM MEMORY

The LSTM cell provides an internal cell-state, which is passed along the temporal dimension together with the output [59], [60]. The LSTM uses gates that are regulated by activation vectors of the form shown in (4.1) and (4.2). The gates are used to regulate which information should be stored in the cell-state from the previous cell-state as well as from the input and previous output. The structure of the LSTM cell is depicted in Fig. 4.3.

Equations (4.3)-(4.8) describe the LSTM cell. In words, the input ($\boldsymbol{x}_t$) and previous output ($\boldsymbol{h}_{t-1}$) are used to generate the forget-gate activation vector ($\boldsymbol{f}_t$), input-gate activation vector ($\boldsymbol{i}_t$), output-gate activation vector ($\boldsymbol{o}_t$), and the candidate cell-state ($\boldsymbol{g}_t$). The previous cell-state ($\boldsymbol{c}_{t-1}$) is gated by the forget gate activation vector, the candidate cell-state is gated by the input gate activation vector, and the sum of these makeup the new cell-state ($\boldsymbol{c}_t$). Finally, the LSTM output ($\boldsymbol{y}_t$) is calculated by gating the hyperbolic tangent of the call-state. Hence, the output depends on the input, previous output, and previous memory (cell-state).

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{\mathrm{f}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{f}}\right) \tag{4.3}$$

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{\mathrm{i}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{i}}\right) \tag{4.4}$$

$$\boldsymbol{g}_t = \tanh\left(\boldsymbol{W}_{\mathrm{g}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{g}}\right) \tag{4.5}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \otimes c_{t-1} + \boldsymbol{i}_t \otimes \boldsymbol{g}_t \tag{4.6}$$

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_{\mathrm{o}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{o}}\right) \tag{4.7}$$

$$\boldsymbol{y}_t = \boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh\left(\boldsymbol{c}_t\right) \tag{4.8}$$

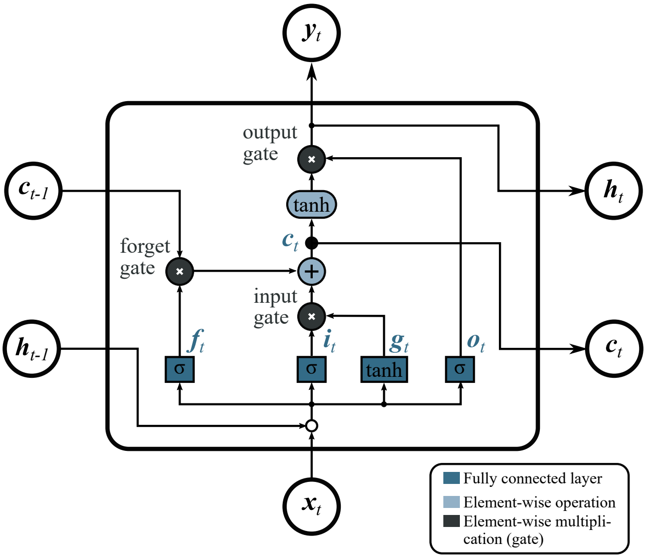


Fig. 4.3: Long short-term memory cell [D]

4.1.2 LSTM RNN Architectures

Combining multiple LSTM cells in layers and stacking these layers, a deep LSTM RNN can be constructed as shown unfolded in Fig. 4.4. This architecture allows to predict sequences of equal length to the input sequence. Alternatively, this architecture can be combined with a standard ANN layer on the output as is done in [D]. This allows for predicting sequences of different length from the input.

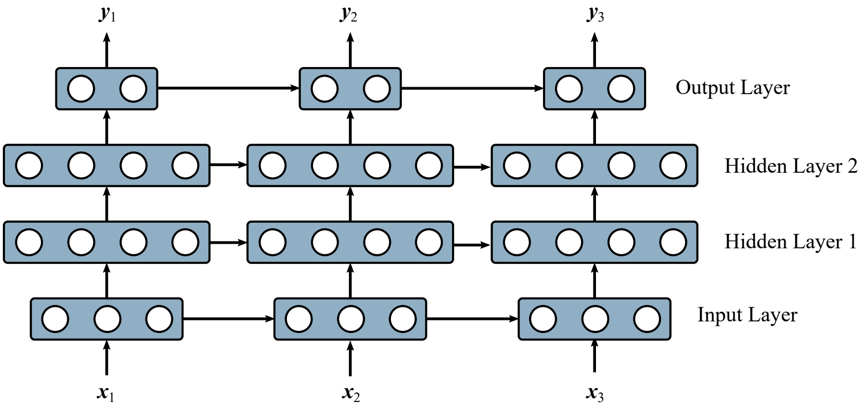


Fig. 4.4: Recurrent neural network with four layers and three temporal steps

Another architecture, that is better adapted for predicting multiple features where the input and output sequence lengths are not the same, is the encoder-decoder architecture [61]. An encoder translates the input sequence to a single vector, that contains the essence of the input sequence that allows the decoder to predict the output sequence. This architecture is depicted in Fig. 4.5 and is the approach used in [E] where multiple features are predicted.

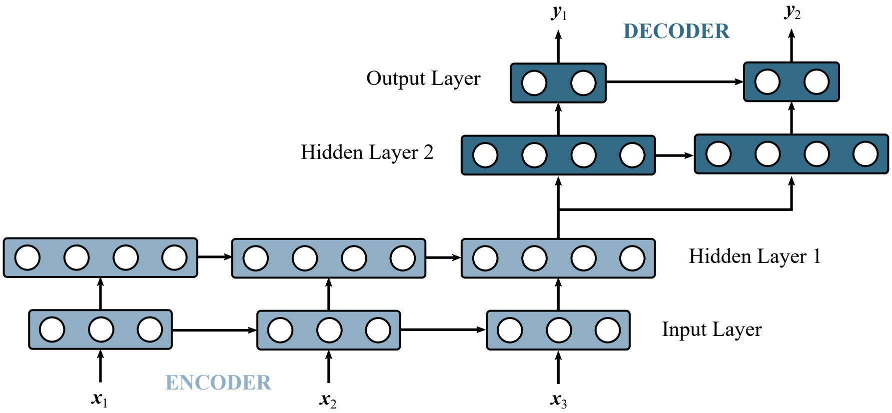


Fig. 4.5: Encoder-decoder recurrent neural network consisting of two encoder layers and two decoder layers with input sequence length of three and output sequence length of two

4.1.3 IMPLEMENTATIONS

The implementation of the recurrent neural networks are done in Python using the Keras [62] module for high level deep learning implementation. Keras provides an intuitive interface to the neural networks, while allowing for user configurations of the networks and the training procedure.

The implementation of the deep RNN is shown in Listing 4.1. The model is initialized using the `Sequential`, onto which layers can be added. The model consists of three bidirectional LSTM RNN layers of 64, 128, and 64 neurons, respectively. The `return_sequences` argument specifies wether each temporal output should be passed to the following layer or just the last output of the sequence. The `input_shape` specifies the dimensions of the input sequence(s) as (`bacth_size, number_of_features`). The `None` entry implies a variable `batch_size`.

Listing 4.1: deep RNN model implementation using Keras [D]

```
from keras.models import Sequential
from keras.layers import Dense, LSTM, Dropout, Bidirectional

# Initiate a sequential model
model = Sequential()
```

```python
# Add first LSTM layer and dropout
model.add(Bidirectional(LSTM(64, return_sequences=True),
    input_shape=(None, 1)))
model.add(Dropout(0.2))
# Add second LSTM layer and dropout
model.add(Bidirectional(LSTM(128, activation='relu',
    return_sequences=True)))
model.add(Dropout(0.2))
# Add third LSTM layer
model.add(Bidirectional(LSTM(64, activation='relu')))
# Add output layer
model.add(Dense(6))
```

After the two first LSTM layers, a dropout layer is added to the model. The dropout layer randomly drops (removes) inputs to the following layer in the model during the training phase. This helps making the network more robust and reduces the risk of overfitting. The final dense layer is a fully connected ANN layer, i.e. non-recurrent, which uses the last instance in the output sequence of the previous LSTM layer to predict six values. These six values are each considered a time step. This way, the network is configured to predict six time steps into the future.

The encoder-decoder LSTM RNN network is implemented as shown in Listing 4.2. The encoder consists of two LSTM layers of 64 and 128 neurons, respectively. The second LSTM layer does not return the sequence, but only the last instance of the sequence. This output vector is repeated to the decoder by the `RepeatVector`. The decoder, similarly, has two LSTM layers of 128 and 64 neurons, respectively. The final layer in the decoder is the output layer, which is a time-distributed dense layer.

Listing 4.2: Encoder-decoder RNN model implementation [E]

```python
from keras.models import Sequential
from keras.layers import Dense, LSTM, RepeatVector,
    TimeDistributed

# Initiate sequential model
model = Sequential()
# Encoder
model.add(LSTM(64, activation='relu', return_sequences=True,
    input_shape=(None, n_features_in)))
model.add(LSTM(128, activation='relu'))
# Decoder
model.add(RepeatVector(n_pred))
model.add(LSTM(128, activation='relu', return_sequences=True)
    )
model.add(LSTM(64, activation='relu', return_sequences=True))
# Output
model.add(TimeDistributed(Dense(n_features_out)))
```

4.2 SINGLE-FEATURE PREDICTION OF STATE OF HEALTH

The bidirectional LSTM RNN, as implemented in Listing 4.1, is trained in paper [D] to predict future SOH values of the stacks. The SOH data, extracted in Chapter 2, is split into training, validation, and testing data (Fig. 4.6). The training data is used for training the RNN model in an iterative process using backpropagation. The validation is used throughout the training phase to evaluate the fitness of the model. The testing data is kept completely separate from the training process. Nine stacks are reserved for validation and three for testing. These twelve systems are not part of the training data.

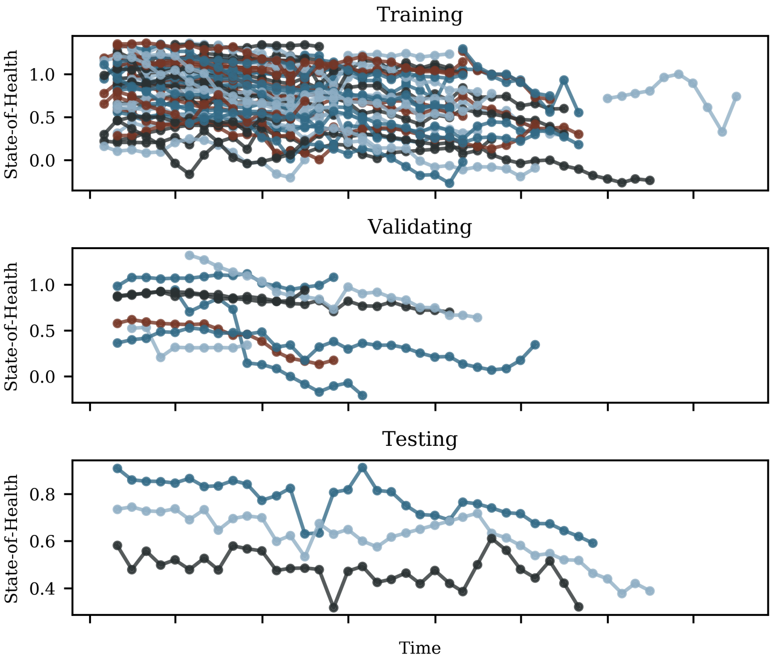


Fig. 4.6: Splitting of training, validation and testing state of health data

The aim of the model is to predict the SOH value for a given stack for each month, six months into the future. Therefore, the training data and validation data of each stack is split into two segments. The last six samples of each stack SOH is reserved as labels, i.e. examples of correct outputs of the RNN model. The remaining samples are used as inputs to the model. In the training process, the weights and biases of the RNN network are iteratively adjusted to minimize the prediction error of the label samples from the input samples.

The model is trained for 200 epochs with a mean absolute error loss function using the Adam [63] optimizer. The training loss after 200 epochs reduced to 0.008 with slightly higher validation loss of 0.018. After training the model, its prediction performance is tested on the validation data and compared to a naive forecasting approach. The results are shown in Fig. 4.7. The model shows good prediction performance for most of the validation stacks, but struggles with some stacks. Stack 84 and 96 show some deviations in the predicted SOH and the actual values. On average, the RNN prediction beats the naive forecasting accuracy considerably with a mean absolute error of 0.025 versus 0.074 of the naive approach.

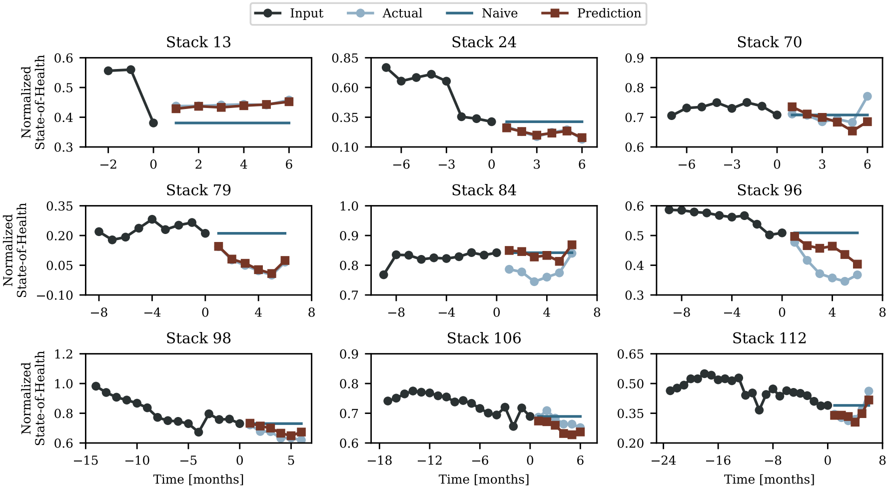


Fig. 4.7: Validation systems state of health prediction results [D]

4.3 Multi-Feature Prediction of Stack Variables

In paper [E], the encoder-decoder LSTM RNN of Listing 4.2, is used to predict voltage, current, stack temperature values of the fuel cell stacks. The input features for predicting these values, are historic examples of voltage, current, and stack temperature, room temperature, air inlet temperature, air outlet temperature, hydrogen pressure, and calendar month at the measurement. The available data is shown in Fig. 4.8.

Similar, to the previously described SOH prediction, the data is split in training and validation stacks. In this case, no test stacks were reserved. Instead, the validation stacks were used for the testing, as these had no influence on the training phase.

The model is trained to predict monthly values, six months into the future. The network is trained for 100 epochs with a mean absolute error loss function. Fifteen separate model were trained, to provide a confidence interval for the predictions. The prediction results on the validation data after training is shown in Fig. 4.9.

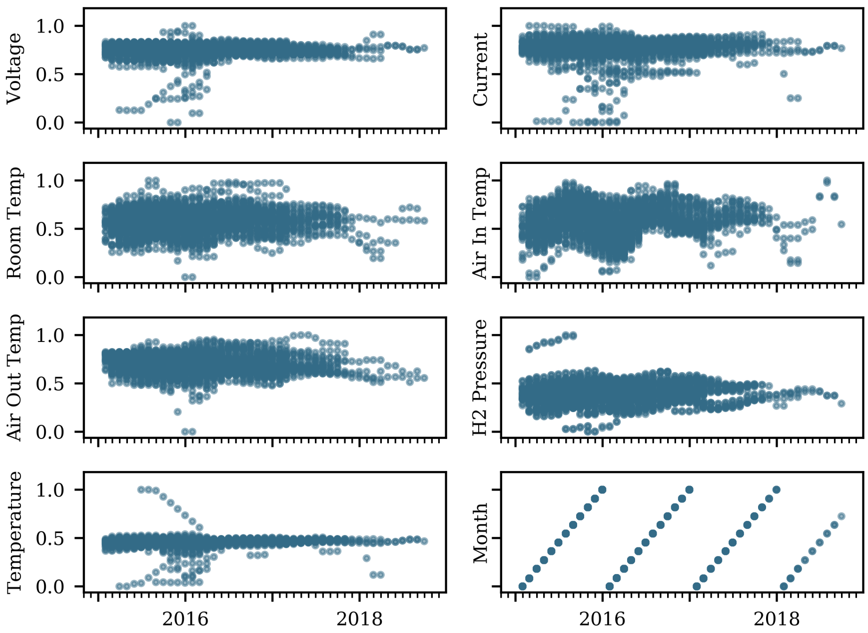


Fig. 4.8: Available data for multi-feature prediction

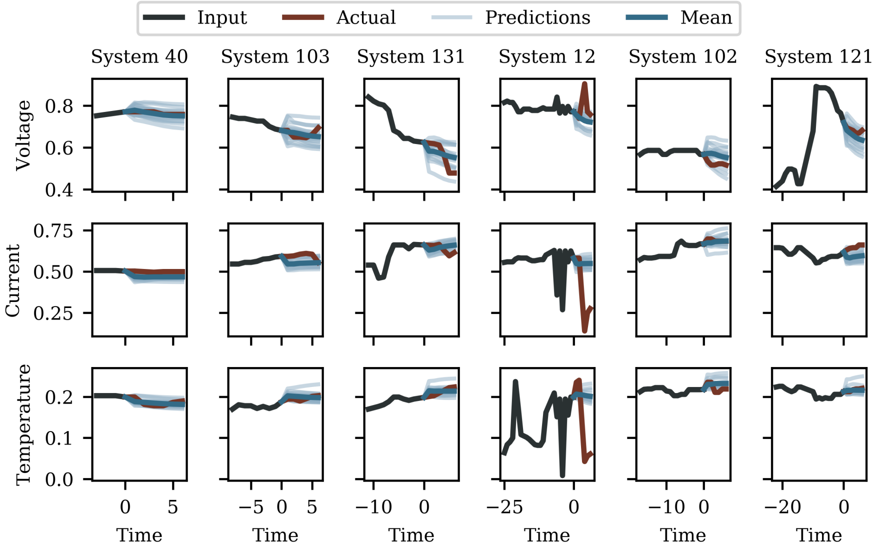


Fig. 4.9: Validation systems voltage, current and temperature prediction results [E]

The accuracy of the predictions are acceptable, on average. However, the confidence interval is a bit wide. For stack 12, the prediction is unable to cope with the fluctuating values, which results in a large prediction error. The mean-absolute-error of mean of the predictions is 0.017 for the voltage, 0.009 for the current, and 0.008 for the temperature. It is expected that the error as well as the confidence interval improves as more data becomes available.

5

Concluding Remarks

This is the first study to apply the Prognostics and Health Management framework to fuel cell stacks in backup power systems based on actual in-field data. Specifically, the assessment of stack degradation level and prediction of future degradation levels where addressed. The use of already-collected in-field data, makes this approach readily applicable to systems already installed in the field without the need for additional sensors or measuring routines.

The main contributions of this work are

- Reducing a dataset of raw measurements from the field operation of fuel cell based backup power systems to a set of performance metrics, i.e. key performance indicators (KPIs). This led to a spectrum of the performance and usage levels of all fuel cell stacks in the fleet of backup power systems and provided the foundation for the following outlier detection and degradation prediction.

- A method for estimating the degradation level as a state of health (SOH) metric was proposed. The method uses the backup system self-tests to extract steady state voltage and current values of the fuel cell stacks. These values are then compared to the manufacturer-defined beginning of life (BOL) and end of life (EOL) criteria in the form of polarization curves. This method provides a load invariant method for estimating the SOH. Hence the SOH can be compared between different systems experiencing different load conditions.

- A method for detecting fuel cell stacks that perform or are used abnormally. It is difficult to make a prior definition of how the normal fuel cell stack should perform under varying conditions. Therefore, the spectrum of KPIs for all stacks in the dataset are used to identify stacks that perform significantly different from the majority of the stacks.

- A method of predicting future SOH values based on historic data. The sparsity of SOH data for each stack, meant that predictions based on individual stacks was infeasible. Therefore, a recurrent neural network was trained on the SOH data of all the stacks in the dataset. In machine learning terminology, this is a single feature sequence to single feature sequence prediction.
- A method for predicting stack voltage, current, and temperature values from a number of parameters. A multi feature sequence to multi feature sequence prediction where the predicted features are a subset of the input features, was constructed in an encoder-decoder recurrent neural network architecture.

5.1 Future Work

If given more time, the next task would be to combine the two recurrent neural network approaches. That is, to construct an RNN, which would take various system measurements as inputs to predict the single SOH metric. The hypothesis is that this approach would be better than either of the tested approaches, as it would have all the information from the various input variables, while predicting the more steady SOH variable instead of the, sometimes, more irregular voltage, current, and temperature variables.

To make for a better suited SOH metric for this specific application, a more accurate definition of stack end of life is needed. This end of life criteria could be based on the minimum requirements of power for the telecommunications load.

As more data is collected on systems in the field, the models can be retrained. It is expected that the prediction accuracy will improve as more data is used in the training phase. This will need to be tested as more data becomes available.

As more stacks reach their end of life time, it would be interesting to test the SOH predictions accuracy in estimating the remaining useful lifetime. This could be done through iteratively predicting SOH values until the EOL criteria is met. The remaining useful life is then the time at the EOL minus the time at prediction. To verify these estimates, data from more stacks that have reached their EOL is needed.

Another interesting research topic would be the health management part of the PHM framework. That is, how to use the obtained information on stack condition and future condition level to optimize the lifetime of the stack. One way might be to load one stack more than the other in a two-stack system, where one stack is less degraded than the other.

References

[1] J. Gurrero, L. D. Vicuna, and J. Uceda, “Uninterruptible power supply systems provide protection”, *IEEE Industrial Electronics Magazine*, vol. 1, pp. 28–38, 2007.

[2] W. Solter, “A new international UPS classification by IEC 62040-3”, *24th Annual International Telecommunications Energy Conference*, pp. 541–545, 2002.

[3] S. Bekiarov and A. Emadi, “Uninterruptible power supplies: classification, operation,dynamics, and control”, *APEC. Seventeenth Annual IEEE Applied Power Electronics Conference and Exposition (Cat. No.02CH37335)*, vol. 1, no. c, pp. 597–604, 2002.

[4] National Renewable Energy Laboratory, “Fuel Cells for Backup Power in Telecommunications Facilities”, *Doe/Go-102009-2709*, 2009.

[5] R. Kaur, V. Krishnasamy, K. Muthusamy, and P. Chinnamuthan, “A novel proton exchange membrane fuel cell based power conversion system for telecom supply with genetic algorithm assisted intelligent interfacing converter”, *Energy Conversion and Management*, vol. 136, pp. 173–183, Mar. 2017.

[6] K. Fosberg, “Fuel cell systems provide backup power in telecom applications”, *Fuel Cells Bulletin*, vol. 2010, no. 12, pp. 12–14, Dec. 2010.

[7] M. J. Vasallo, J. M. Andújar, C. García, and J. J. Brey, “A Methodology for Sizing Backup Fuel-Cell / Battery Hybrid Power Systems”, *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1964–1975, 2010.

[8] Haimin Tao, J. Duarte, and M. Hendrix, “Line-Interactive UPS Using a Fuel Cell as the Primary Source”, *IEEE Transactions on Industrial Electronics*, vol. 55, no. 8, pp. 3012–3021, Aug. 2008.

[9] G. Bruni, S. Cordiner, V. Mulone, A. Giordani, M. Savino, G. Tomarchio, T. Malkow, G. Tsotridis, S. Bodker, J. Jensen, R. Bianchi, and G. Picciotti, "Fuel cell based power systems to supply power to Telecom Stations", *International Journal of Hydrogen Energy*, vol. 39, no. 36, pp. 21 767–21 777, Dec. 2014.

[10] Y. Zhan, Y. Guo, J. Zhu, and H. Wang, "Intelligent uninterruptible power supply system with back-up fuel cell/battery hybrid power source", *Journal of Power Sources*, vol. 179, no. 2, pp. 745–753, May 2008.

[11] H. Paul, "Uninterruptible Power Supply (UPS)", in *Hydrogen and Fuel - Technologies and Market Perspectives*, J. Töpler and J. Lehmann, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, ch. 7.

[12] M. F. Serincan, "Validation of hybridization methodologies of fuel cell backup power systems in real-world telecom applications", *International Journal of Hydrogen Energy*, vol. 41, no. 42, pp. 19 129–19 140, Nov. 2016.

[13] L. Torok, S. L. Sahlin, S. K. Kor, and B. Bidoggia, "Estimation of membrane hydration status for active proton exchange membrane fuel cell systems by impedance measurement: Start-up time measurements", in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, IEEE, Jun. 2016, pp. 1–5.

[14] Z. Zhang, Z. Ouyang, O. C. Thomsen, and M. A. E. Andersen, "Analysis and Design of a Bidirectional Isolated DC–DC Converter for Fuel Cells and Supercapacitors Hybrid System", *IEEE Transactions on Power Electronics*, vol. 27, no. 2, pp. 848–859, Feb. 2012.

[15] L. Torok, S. D. Sonderskov, and S. Munk-Nielsen, "Bidirectional Operation of High Efficiency Isolated DC-DC Converter in Fuel Cell Telecom Back-up Systems", in *2018 IEEE International Telecommunications Energy Conference (INTELEC)*, IEEE, Oct. 2018, pp. 1–5.

[16] A. L. Dicks and D. A. J. Rand, *Fuel Cell Systems Explained.* Chichester, UK: John Wiley & Sons, Ltd, Apr. 2018.

[17] W. R. Grove, "VIII. On the gas voltaic battery. —Experiments made with a view of ascertaining the rationale of its action and its application to eudiometry", *Philosophical Transactions of the Royal Society of London*, vol. 133, pp. 91–112, Jan. 1843.

[18] S. Ould Amrouche, D. Rekioua, T. Rekioua, and S. Bacha, "Overview of energy storage in renewable energy systems", *International Journal of Hydrogen Energy*, vol. 41, no. 45, pp. 20 914–20 927, Dec. 2016.

[19] H. Tsuchiya and O. Kobayashi, "Mass production cost of PEM fuel cell by learning curve", *International Journal of Hydrogen Energy*, vol. 29, no. 10, pp. 985–990, 2004.

[20] M. Knowles, D. Baglee, A. Morris, and Q. Ren, "The state of the art in fuel cell condition monitoring and maintenance", *World Electric Vehicle Journal*, vol. 4, no. 1, pp. 487–494, 2011.

[21] M. F. Serincan, “Reliability considerations of a fuel cell backup power system for telecom applications”, *Journal of Power Sources*, vol. 309, pp. 66–75, 2016.

[22] G. Niu, B.-S. Yang, and M. Pecht, “Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance”, *Reliability Engineering & System Safety*, vol. 95, no. 7, pp. 786–796, Jul. 2010.

[23] M. Ben-Daya, S. O. Duffuaa, A. Raouf, J. Knezevic, and D. Ait-Kadi, Eds., *Handbook of Maintenance Management and Engineering*. London: Springer London, 2009.

[24] M. Jouin, R. Gouriveau, D. Hissel, M. C. Péra, and N. Zerhouni, “Prognostics and Health Management of PEMFC - State of the art and remaining challenges”, *International Journal of Hydrogen Energy*, vol. 38, no. 35, pp. 15 307–15 317, 2013.

[25] T. Sutharssan, D. Montalvao, Y. K. Chen, W.-C. Wang, C. Pisac, and H. Elemara, “A review on prognostics and health monitoring of proton exchange membrane fuel cell”, *Renewable and Sustainable Energy Reviews*, vol. 75, no. November 2015, pp. 440–450, Aug. 2017.

[26] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, “Remaining Useful Life Prediction and Uncertainty Quantification of Proton Exchange Membrane Fuel Cell Under Variable Load”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2569–2577, Apr. 2016.

[27] T. Kim, H. Oh, H. Kim, and B. D. Youn, “An Online-Applicable Model for Predicting Health Degradation of PEM Fuel Cells with Root Cause Analysis”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7094–7103, 2016.

[28] C. Jeppesen, S. S. Araya, S. L. Sahlin, S. Thomas, S. J. Andreasen, and S. K. Kær, “Fault detection and isolation of high temperature proton exchange membrane fuel cell stack under the influence of degradation”, *Journal of Power Sources*, vol. 359, pp. 37–47, 2017.

[29] N. Fouquet, C. Doulet, C. Nouillant, G. Dauphin-Tanguy, and B. Ould-Bouamama, “Model based PEM fuel cell state-of-health monitoring via ac impedance measurements”, *Journal of Power Sources*, vol. 159, no. 2, pp. 905–913, 2006.

[30] T. Sutharssan, D. Montalvao, Y. K. Chen, W.-C. C. Wang, C. Pisac, and H. Elemara, “A review on prognostics and health monitoring of proton exchange membrane fuel cell”, *Renewable and Sustainable Energy Reviews*, vol. 75, no. August 2016, pp. 440–450, Aug. 2017.

[31] R. Petrone, Z. Zheng, D. Hissel, M. C. Péra, C. Pianese, M. Sorrentino, M. Becherif, and N. Yousfi-Steiner, “A review on model-based diagnosis methodologies for PEMFCs”, *International Journal of Hydrogen Energy*, vol. 38, no. 17, pp. 7077–7091, 2013.

[32] Z. Zheng, R. Petrone, M. Péra, D. Hissel, M. Becherif, C. Pianese, N. Yousfi Steiner, and M. Sorrentino, “A review on non-model based diagnosis methodologies for PEM fuel cell stacks and systems”, *International Journal of Hydrogen Energy*, vol. 38, no. 21, pp. 8914–8926, Jul. 2013.

[33] W. Schmittinger and A. Vahidi, “A review of the main parameters influencing long-term performance and durability of PEM fuel cells”, *Journal of Power Sources*, vol. 180, no. 1, pp. 1–14, May 2008.

[34] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, “Prognostics of PEM fuel cell in a particle filtering framework”, *International Journal of Hydrogen Energy*, vol. 39, no. 1, pp. 481–494, Jan. 2014.

[35] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Pera, and N. Zerhouni, “Joint Particle Filters Prognostics for Proton Exchange Membrane Fuel Cell Power Prediction at Constant Current Solicitation”, *IEEE Transactions on Reliability*, vol. 65, no. 1, pp. 336–349, Mar. 2016.

[36] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, “Extended Kalman Filter for prognostic of Proton Exchange Membrane Fuel Cell”, *Applied Energy*, vol. 164, pp. 220–227, Feb. 2016.

[37] E. Lechartier, E. Laffly, M. C. Péra, R. Gouriveau, D. Hissel, and N. Zerhouni, “Proton exchange membrane fuel cell behavioral model suitable for prognostics”, *International Journal of Hydrogen Energy*, vol. 40, no. 26, pp. 8384–8397, 2015.

[38] Z. Li, M. Ieee, Z. Zheng, M. Ieee, and R. Outbib, “Adaptive prognostic of fuel cells by implementing ensemble Echo State Networks in time varying model space”, *IEEE Transactions on Industrial Electronics*, vol. PP, no. c, p. 1, 2019.

[39] D. Marra, M. Sorrentino, C. Pianese, and B. Iwanschitz, “A neural network estimator of Solid Oxide Fuel Cell performance for on-field diagnostics and prognostics applications”, *Journal of Power Sources*, vol. 241, pp. 320–329, Nov. 2013.

[40] M. Ibrahim, N. Steiner, S. Jemei, and D. Hissel, “Wavelets-based approach for online Fuel Cells Remaining Useful lifetime Prediction”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 1–1, Aug. 2016.

[41] R. Silva, R. Gouriveau, S. Jemeï, D. Hissel, L. Boulon, K. Agbossou, and N. Yousfi Steiner, “Proton exchange membrane fuel cell degradation prediction based on Adaptive Neuro-Fuzzy Inference Systems”, *International Journal of Hydrogen Energy*, vol. 39, no. 21, pp. 11 128–11 144, Jul. 2014.

[42] R. Ma, Z. Li, E. Breaz, and C. Liu, “Data - fusion Prognostics of Proton Exchange Membrane Fuel Cell Degradation”, *IEEE Transactions on Industry Applications*, vol. PP, no. c, p. 1, 2019.

[43] S. Morando, S. Jemei, R. Gouriveau, N. Zerhouni, and D. Hissel, "Fuel Cells prognostics using Echo State Network", in *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, ser. IEEE Industrial Electronics Society, IEEE, 2013, pp. 1632–1637.

[44] Z. Li and S. Jeme, "Remaining useful life estimation for PEMFC in dynamic operating conditions", *2016 IEEE Vehicle Power and Propulsion Conference (VPPC)*, no. Cnrs 3539, pp. 1–6, 2016.

[45] S. M. Rezvanizaniani, Z. Liu, Y. Chen, and J. Lee, "Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility", *Journal of Power Sources*, vol. 256, pp. 110–124, 2014.

[46] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van Den Bossche, "Critical review of state of health estimation methods of Li-ion batteries for real applications", *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 572–587, 2016.

[47] B. Dolenc, P. Boškoski, M. Stepančič, A. Pohjoranta, and Đ. Juričić, "State of health estimation and remaining useful life prediction of solid oxide fuel cell stack", *Energy Conversion and Management*, vol. 148, pp. 993–1002, Sep. 2017.

[48] J. Kim, I. Lee, Y. Tak, and B. H. Cho, "State-of-health diagnosis based on hamming neural network using output voltage pattern recognition for a PEM fuel cell", *International Journal of Hydrogen Energy*, vol. 37, no. 5, pp. 4280–4289, 2012.

[49] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying Density-Based Local Outliers", *Proceedings of the 2000 Acm Sigmod International Conference on Management of Data*, pp. 1–12, 2000.

[50] A. Zimek and E. Schubert, "Outlier Detection", in *Encyclopedia of Database Systems*, New York, NY: Springer New York, 2017, pp. 1–5.

[51] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series", *KDD workshop.*, vol. 10, no. 16, pp. 359–370, 1994.

[52] S. Axelrod and B. Maison, "Combination of hidden Markov models with dynamic time warping for speech recognition", in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, May 2004, pp. I–173.

[53] E. J. Keogh and M. J. Pazzani, "Scaling up Dynamic Time Warping for Datamining Applications", pp. 285–289, 2000.

[54] A. Barr, G. Mathias, and D. Riu, "A Real-time Data-driven Method for Battery Health Prognostics in Electric Vehicle Use", pp. 1–8, 2014.

[55] M. Ester, H.-p. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, 1996.

[56] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis”, *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37–52, Aug. 1987.

[57] L. Li, “Dimension Reduction for High-Dimensional Data”, in *Methods in molecular biology (Clifton, N.J.)* Vol. 620, 2010, pp. 417–434.

[58] I. K. Fodor, “A Survey of Dimension Reduction Techniques”, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, Tech. Rep. 1, May 2002, pp. 1–18.

[59] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory”, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[60] F. Gers, “Learning to forget: continual prediction with LSTM”, in *9th International Conference on Artificial Neural Networks: ICANN ’99*, vol. 1999, IEE, 1999, pp. 850–855.

[61] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to Sequence Learning with Neural Networks”, pp. 1–9, 2014.

[62] F. Chollet *et al.*, *Keras*, \url{https://keras.io}, 2015.

[63] D. P. Kingma and J. L. Ba, “ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION”, in *ICLR*, 2015, pp. 1–15.

Part II

Appended Papers

Paper A

Lifetime Prognostics of Hybrid Backup Power System: State-of-the-Art

Simon Dyhr Sønderskov, Maciej Jozef Swierczynski, and Stig Munk-Nielsen

The paper has been published in the
2017 IEEE International Telecommunications Energy Conference (INTELEC),
pp. 574–581, 2017. DOI: 10.1109/INTLEC.2017.8214199

Abstract

Modern telecommunication power supplies are based on renewable solutions, e.g. fuel cell/battery hybrid sys-tems, for immediate and prolonged load support during grid faults. The high demand for power continuity increases the em-phasis on power supply reliability and availability which raises the need for monitoring the system condition for timely mainte-nance and prevention of downtime. Although present on compo-nent level, no current literature addresses the condition monitor-ing from the perspective of a fuel cell/battery hybrid system such as the telecommunication power supply. This paper is a first step towards a condition monitoring approach for such systems. First-ly, the application is defined, thereafter the benefits of predictive maintenance strategies and the prognostics and health manage-ment framework are described. A literature review of condition monitoring of the major system components: fuel cell, battery, and converters, is given. Finally, the paper presents a discussion on the available monitoring techniques from a commercial hybrid system point view.

A.1 Introduction

The telecommunication network is a key part of modern society as it provides the foundation for cellular and internet communication. With many remote sites and in some cases unstable electrical grids, telecommunication sites require backup power to ensure highly reliable and available communication networks. Hence, the power supply of the telecommunication site should protect the sensitive equipment from grid faults ranging from short voltage distortions to complete and sudden absence of grid voltage (from here on referred to as 'grid outage'). This is obtained by using an uninterruptible power supply (UPS), where energy storage devices can absorb the grid fluctuations and provide continuous power during grid outages [1], [2]. UPS systems generally use electrochemical batteries as the storage element. However, regulations require telecommunication sites to be operational for extended time periods without a present grid [3], where batteries are not an economically feasible solution.

The need for long backup times as well as an increased emphasis on environmentalfriendly energy sources has fostered the use of fuel cell (FC) based backup systems [4]. Fuel cells are well suited for backup systems, since they have large and scalable capacity, they are highly reliable and require little maintenance [5]. However, they suffer from high capital cost and slow dynamic performance. The latter is the main reason why FC-based backup systems are hybrid systems consisting of the FC along with some electrical energy storage elements, such as batteries or capacitors [6]. The storage elements provide the telecom load with power during FC power-up and also help to supply extra power during load dynamics.

In some cases, it is feasible to include additional renewable energy sources, such as wind turbines and photovoltaics, in the backup system to further reduce the dependency on the grid [7]. This might be relevant when the grid is especially weak or even to reduce the power drawn from the grid to reduce the cost of electricity.

The availability of the telecommunication service is inherently dependent on the availability of the power supply and its backup functionality in the case of a grid fault. Therefore, the condition of the power supply is of high concern. The system is maintained through appropriate maintenance strategies. However, this often consists of pre-scheduled maintenance which is suboptimal. Also, the condition of the power supply can only be assessed during manual inspection and is thus unknown during normal operation. Manual inspection of the system condition might require taking the UPS system offline.

Condition monitoring (CM) might change this, as it allows for continuous monitoring of the system condition, which can give the operators an assurance of backup functionality. Furthermore, CM paves the way for more optimal maintenance strategies and is a first step towards a cyber-physical system as envisioned in the Industry 4.0 framework [8].

Despite extensive research in various engineering fields, condition monitoring of hybrid power supplies is still an undescribed subject. This paper aims to illuminate the topic of condition monitoring techniques in hybrid power supplies. This is done in the following structure: Firstly, the system under investigation is described and maintenance strategies are introduced. Then condition monitoring and prognostics techniques in fuel cell, battery, and converter subsystems are presented. Finally, a discussion addresses the validity of these methods from a hybrid power supply perspective.

A.2 Telecommunication Power Supply

Telecommunication power supplies rely on energy storage devices to mitigate the consequences of grid faults, such as voltage fluctuations and outages, from affecting the load. Traditionally, two backup energy sources have been dominant: lead-acid batteries and diesel engine-generator sets (gensets) [5]. The lead-acid batteries have been popular due to their low cost and ease of implementation and the gensets due to their extended backup time capabilities. However, by modern standards these technologies are suboptimal. The gensets are polluting, noisy and require relatively high amounts of maintenance. The lead-acid batteries are non-durable and sensitive to variations in temperature and power.

Telecommunication power supplies often supply power in the range of 1 to 10 kilowatts [9]. The most relevant energy storage devices for this power range are batteries (mainly lead-acid and li-ion), supercapacitors, and fuel cells as in the Ragone plot of Fig. A.1. Each of these electrochemical storage elements has its own area of expertise: supercapacitors have high power density but low energy density, meaning that thery are well suited for delivering large, but short bursts of power. Whereas fuel cells are better suited for the long haul, due to their low power density and high energy density. This is highlighted by the constant time lines of Fig. A.1: super capacitors are best suited for delivering its stored power in the range of seconds and fuel cells, in the range of hours

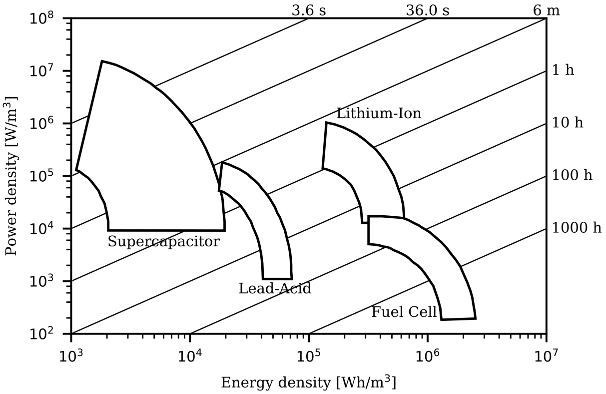


Fig. A.1: Ragone plot showing power density versus energy density of different storage technologies. Adapted from [10].

and days. Batteries bridge the gap between these two technologies, li-ion being slightly higher in both power density and energy density than the lead-acid batteries.

The required backup times of telecommunication power supplies range from a few hours to several days [3]. At these requirements, batteries would be too large and too expensive. The fuel cell, on the other hand, is well suiteed for long backup times, as the energy storage in the form of hydrogen (or hydrogen rich gas) is decoupled from the power stage. The time in which the fuel cell can deliver backup power only depends on the number of hydrogen tanks that are available.

Several types of fuel cells exist, but the most used for the telecommunications backup is the proton exchange membrane (PEM) fuel cell. PEM is preferred due to its relatively high level of maturity and its low-temperature operation which results in relatively short start-up time and fast dynamic response [5]. The response time is in the order of one minute in the case of a cold start-up. This response time is due to the fuel cells inability to deliver its rated power until the temperature in the stack reaches the steady state operating condition [11]. This startup time is the reason why fuel cells for backup applications are often integrated with an additional energy storage, such as batteries, supercapacitors, or a combination of both. These provide power during FC start-up and during fast load transients. Hence, the size of these additional elements is limited to a backup time of a few minutes. This integration of two or more energy storages is often referred to as hybridization.

The backup power supply considered in this work consists of a PEM fuel cell and a li-ion battery as the storage elements. Additionally, a rectifier and converters that interface the grid and the storage elements to a common DC-bus from which the load is supplied. The load is either DC, AC, or a combination of

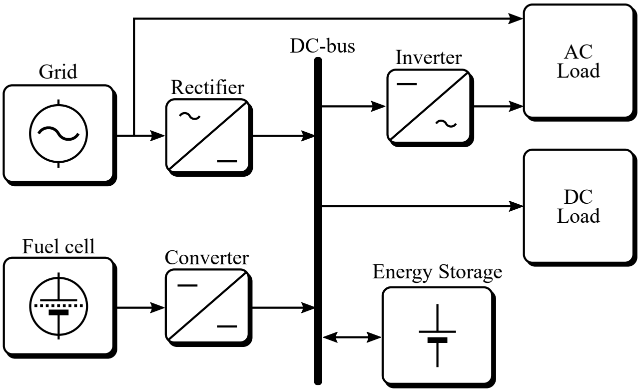


Fig. A.2: Architecture of fuel cell/battery hybrid backup power supply.

the two, depending on the operator, and might include the telecommunication equipment as well as auxiliary equipment such as air conditioning. The DC load may be supplied directly from the DC-bus and the AC load is supplied through an inverter. The architecture of the system is illustrated in Fig. A.2. This architecture is a generic telecommunication power supply similar to those presented in [5], [6], [11], [12].

During normal operation, the grid voltage is supplied directly to the AC load and to the DC load through the rectifier. If the battery is not charged to the desired level, it is charged from the DC-bus. In the case of grid fault, the battery will act as a buffer and supply the required power to the load. If the grid fault exhausts the energy storage of the battery, the fuel cell will have started up and will supply the power to the load and possibly recharge the battery. Hence, the load never ‘sees’ the power disturbance from the grid.

A.3 Maintenance Strategies

Maintenance of the power supply is required to ensure continuous operation of the telecommunication equipment. Parts of the power supply, e.g. the storage elements, spend most of the time in storage mode. However, during grid faults, it is essential that these elements function on demand. This is obtained through appropriate maintenance strategies.

Generally, maintenance strategies are classified in two groups: reactive and proactive methods, each having several sub groups (Fig. A.3). The reactive maintenance strategies, in general, operate from a run-to-failure approach, where components are repaired or replaced when they fail. This approach often leads to unscheduled downtime, which is highly unfortunate in backup applications. Furthermore, this approach may lead to high maintenance costs and safety issues.

Proactive maintenance aims to prevent the sudden failures of components,

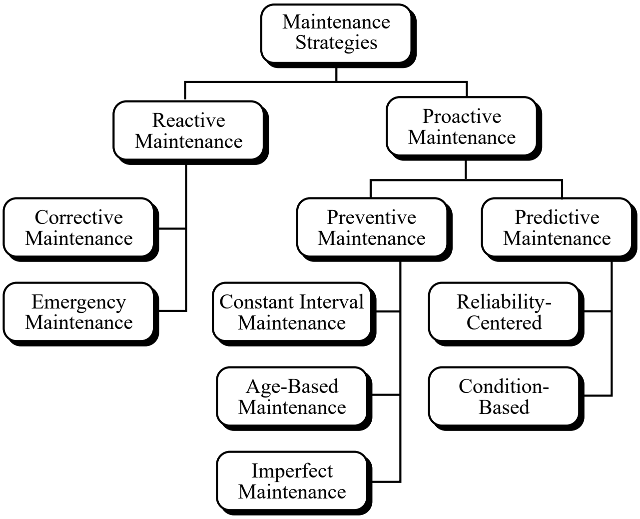


Fig. A.3: Maintenance approaches [13].

either through preventive maintenance or predictive maintenance. In preventive maintenance, routine maintenance is performed in appropriate intervals. These intervals are determined from useful lifetime forecasts during the design phase and expert knowledge [14].

Predictive maintenance does not use fixed schedule, but an adaptively determined schedule. In condition-based maintenance (CBM), a condition monitoring system is used to monitor selected parameters that are indicative of system faults or degradation. Hence, the maintenance schedule is determined from the current condition of the system. A step further in CBM is the utilization of prognostics, i.e. prediction of the future condition of the system. The different maintenance strategies are summarized in Table A.1.

Prognostics and health management (PHM) is a framework, which describes an approach to predictive maintenance. In recent years PHM has become a popular topic of research in fuel cell [15]–[17] and battery applications [18], [19]. PHM describes seven steps which can be used to extend the lifetime of a system. The framework consists of three phases as depicted in Fig. A.4. The observation includes data acquisition, i.e. measurement of relevant parameters such as voltage, current, pressure, temperature, etc.; and data processing, which transforms the signals (e.g. to the frequency domain) and extracts features that may be used for the state of health (SOH) estimation of the system in the condition assessment of the analysis phase. Diagnostics is the act of detecting and localizing a failure in the system. The final step in the analysis phase is prognostics, which aims to estimate the future condition of the system. PHM can aid in two important aspects of system reliability: early warning of developing failures and estimating remaining useful lifetime (RUL), which is a valuable

Table A.1: Principles of maintenance approaches [14].

	Corrective	**Preventive**		
	Run-to-fail (Fix when it brakes)	***Predetermined (Scheduled maintenance)***	***Predictive (Condition-based maintenance)***	
			Diagnostics	*Prognostics*
When scheduled	No scheduled maintenance	Maintenance based on a fixed time schedule for inspect, repair, and overhaul	Maintenance based on current condition	Maintenance based on forecasting of remaining equipment life
Why scheduled	N/A	Intolerable failure effect and possibility of preventing the failure effect through scheduled overhaul or replacement	Maintenance scheduled based on evidence of needs	Maintenance need is projected as probable within mission time
How scheduled	N/A	Based on the useful life of the component forecasted during design and updated through experience	Continuous collection of condition monitoring data	Forecasting of remaining equipment life based on actual stress loading
Type of prediction	None	None	On and off-line, near real-time trend analysis	On and off-line, real-time trend analysis

parameter in the predictive maintenance strategy.

A.4 Condition Monitoring

Condition monitoring (CM) is the process of measuring system parameters and using them to estimate the condition of the system. CM allows for planning optimal maintenance strategies (thorough predictive maintenance or PHM) as well as providing valuable knowledge of the system in a real working environment.

Condition monitoring from the perspective of telecommunication power supply or a fuel cell/battery hybrid system in general is, to the authors' knowledge,

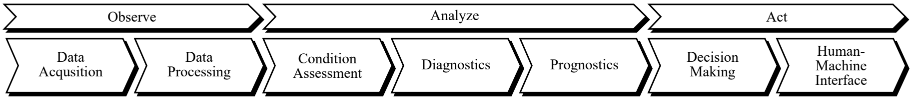


Fig. A.4: Prognostics and health management (PHM) framework.

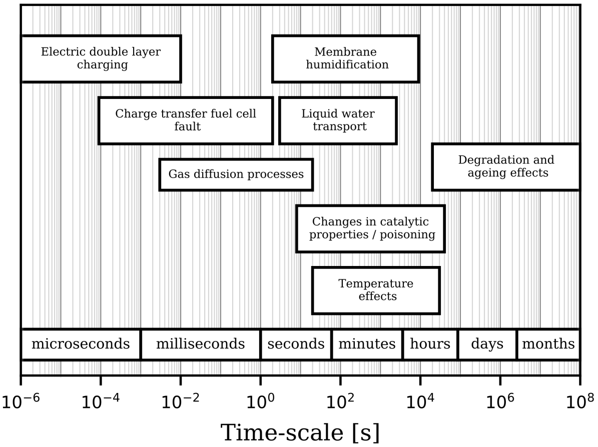


Fig. A.5: Illustration of phenomena and time-scales in fuel cells [20].

not described in literature. However, on subsystem level, fuel cells, batteries, and converters have received much research attention in the fields of condition monitoring, fault detection and prognostics as detailed in the following sections.

A.4.1 Fuel cells

Fuel cells are electrochemical devices, that produce electrical energy from hydrogen and oxygen (or gasses that are rich in these two elements) and leave only water as a waste product. The most adapted fuel cell technology in backup applications and the technology of interest in this paper is the PEM fuel cell. Fuel cells are complex systems, which are affected on multiple timescales by electrical, chemical, and thermal phenomena as illustrated by [20] (Fig. A.5).

PEM fuel cells typically consist of two electrodes separated by a membrane. The electrodes (anode and cathode) are both separated from the membrane by a catalyst. The membrane allows protons to pass, while blocking the gas. On the anode side, hydrogen enters and the catalyst causes the hydrogen to split into free electrons and positively charged hydrogen ions, which are just protons. The electrons travel through the electrical circuit to the cathode side and the protons pass through the membrane to the cathode side. The electrons and protons react with the air or oxygen, which is supplied to the cathode side, to form water. Several of these electrode-membrane cells are combined in fuel cell stacks to increase the voltage and power rating. [21]

Reference [21] summarizes the overall PEM fuel cell fault modes as: membrane dehydration [22], [23], fuel/gas starvation [22], physical defects of membrane [22], [24], and catalyst poisoning [25]. A more exhaustive degradation and ageing analysis is presented in [26]. Each of these faults leads to performance

Table A.2: Faults and their revealing signals in proton exchange membrane fuel cells [27].

Fault on PEMFC	Signal	Reference
Cathode flooding	Cathode pressure drop	[28]
Water present at anode side	Anode pressure drop	
Cathode flooding	Cell voltage variance	[29]
	Mean cell voltages	
Air starvation	Voltage oscillations	[30]
Flooding	Cathode pressure drop	[31]
Drying	Cell resistance	
Anode/cathode crossovers	Cell open circuit voltages	[32]
Flooding	Stack voltage (high frequency)	[33]
All	Electrochemical impedance spectroscopy	[34]

degradation and potentially a failure of the fuel cell. Diagnostics is the process of identifying which failure mode has occurred in the fuel cell. This is typically done by observing system parameters, which are indicative of the failure modes. For this purpose [27] identifies measurable signals that relate to specific failure modes as listed in Table A.2.

The signals of cathode and anode pressure drops, are used for detection of cathode flooding and water at the anode side [28] as well as flooding of the fuel cell [31], require measurement of the pressure in the air and hydrogen supply, respectively. This can be achieved with pressure sensors, but might not be easily implemented in a field application. The cathode flooding detection by cell voltage variance and mean cell voltage [29], requires voltage measurements of each cell in the fuel cell stack. Hence, the terminals of each cell must be accessible, which is not always feasible in the field. The drying degradation [31] and the crossover detection [32] also require measurements of individual cell voltages. In contrast, [30], [33] use voltage measurements on the whole stack to detect air starvation and flooding, respectively. Reference [30] analyzes the voltage fluctuations caused by the lag between load dynamics and air response, to detect temporary air starvation. The approach of [33] uses wavelet transformation to extract information of extensive water accumulation, i.e. flooding, in the electrode from stack voltage measurements. Finally, the electrochemical impedance (EIS) measurements can reveal all of the above faults [34].

The diagnostic approaches presented in literature are often tested on laboratory setups, where there is high access to system variables, which might not be true in a real application. Hence, [27] categorizes fuel cell diagnostics measurements, based on the technical and economic feasibility of implementation (Table A.3). The most easily measured parameters are the stack voltage and current. Also, the cooling water temperature and air compressor speed are easily accessible. More difficult, but still possible, measurements are individual cell voltages, inlet and outlet gas pressures, the rate of gas flow and

Table A.3: Economic and technical evaluation of fuel cell measurements [27].

Possible	Unlikely	Impossible
Stack current	Single cell voltage	Flows in the channels of the bipolar plates
Stack voltage	Gas pressure (inlet/outlet)	Local current density
Stack impedance*	Gas flow	Gas hygrometry
Cooling water temperature	Stack internal temperature	Stack impedance*
Air compressor speed		Gas composition (inlet and outlet)

*The stack impedance measurement is possible if it is done through the output DC/DC converter with a modified control, too cumbersome and expensive through a lab spectrometer.

the internal temperature. In the impossible measurements category, internal or complicatedly measured parameters are found, such as the local current density, gas hygrometry, and gas composition. The powerful measurement of EIS is located both in the possible and impossible categories. Traditionally EIS measurements are carried out by an external spectrum analyzer, which injects sinusoidal voltage perturbations of varying frequency to the fuel cell and measures the corresponding current at each frequency. Thereby, a spectrum of the fuel cell impedance is obtained. These spectrum analyzers are often expensive and require the FC to be disconnected from the load and thereby obstructing the application. This makes the EIS almost impossible to obtain in a real field application. However, recent research is investigating the implementation of the EIS measurement in the DC/DC converter of the fuel cell [35], [36]. This would allow for online EIS measurements, as the small perturbations could be superimposed on the converter control effort. This would, however, require special design of the DC/DC converter and its control, which makes it impractical for implementation in already operating systems.

A.4.2 Batteries

Batteries are electrochemical devices, like fuel cells. However, unlike fuel cells, the energy storage of the battery is internal, i.e. it relies on chemical reactions inside the battery to produce electricity and on electricity to reverse these reactions and thereby store energy. Several battery technologies exist; however, few are suitable and technically mature enough to be considered for backup systems. The focus of this paper will be on lithium-ion batteries as they are gaining traction in a wide variety of applications and are already adapted in backup power supplies [5], [37].

The internal processes of li-ion batteries are complicated to understand and almost impossible to observe [38], [39]. The ageing processes are numerous and

Table A.4: Ageing mechanisms, their effects and accelerators in lithium-ion batteries [40].

Mechanism	**Effect**	**Accelerator**
Electrolyte decomposition → SEI formation	Capacity fade Power fade	High temperature High SOC
Solvent co-intercalation and gas evolution	Capacity fade	Overcharge
Decrease of active surface area	Power fade	High temperature High SOC
Electrode porosity change	Power fade	High cycling rate High SOC
Volume changes during cycling → contact loss of active material particles	Capacity fade	High cycling rate High DOD
Decomposition of binder	Capacity fade	High SOC High temperature
Current collector corrosion	Power fade Enhances other ageing mechanisms	Over-discharge Low SOC
Metallic lithium plating	Capacity fade Power fade	Low temperature High cycling rate Poor cell balancing Geometric misfits

interact with each other in a non-linear way and depend on the load profile of the battery as well as its architecture and the control strategy [39]. However, most of the ageing mechanisms lead to either capacity fade, internal resistance increase, or both. These are both important performance parameters of any battery and their dependability of the ageing processes has made them popular as state-of-health metrics of batteries. Ageing processes are the irreversible changes that occur in the battery during its lifetime. The most common ageing mechanisms identified in [38] are solid-electrolyte interphase (SEI) formation, loss of active material, and lithium plating on the anode side along with chemical structural changes, decomposition, and loss of contact between inactive components on the cathode side. A more exhaustive analysis of ageing mechanisms, their accelerators and effects in li-ion batteries is given in Table A.4 [39].

The state of health of a battery is generally defined as the performance and health condition of the battery compared to a new battery, i.e. how much it has changed since its beginning of life [39]. There is no agreement between researchers or companies on how to define this metric. However, the capacity and internal impedance are popular indicators. A good SOH estimation should consider both, as they both affect the battery performance.

The methods for estimating the battery SOH is split into two main groups [38]: experimental and adaptive methods. The experimental methods combine

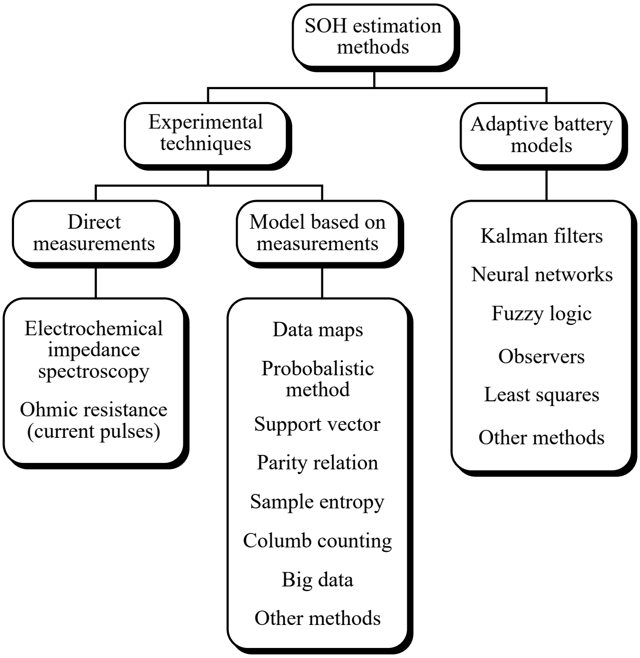


Fig. A.6: State of health estimation methods of lithium-ion batteries [38].

stored data of battery operation with knowledge of how the observed parameters affect the battery lifetime. These two parts are used to estimate the SOH of the battery. The experimental methods can be done in two ways: firstly, the direct measurements, which rely on special measuring techniques to extract the battery SOH. The techniques highlighted in Fig. A.6 are EIS [41] and Ohmic resistance measurements during current pulse injection [42].

Secondly, the models based on measurements. This category includes methods such as Coulomb counting [43], data fitting [44], sample entropy [45], and support vector machine [46]. Opposed to the direct measurements, these methods may not require special battery operation and is thus more applicable for online implementation, assuming the required measurements can be obtained during normal battery operation. Common for all experimental methods, is that they all require prior knowledge of the relationship between SOH and the observed parameters, which is often obtained through extensive laboratory experiments.

The other category of battery SOH estimation techniques is the adaptive battery models. The adaptive models use degradation sensitive parameters of the battery to calculate the SOH. These parameters should be measurable during battery operation and are monitored throughout the battery lifetime. The adaptive models include various Kalman filters [47]–[49], observers [50], fuzzy logic [42], and artificial neural networks [51]. Although highly accurate,

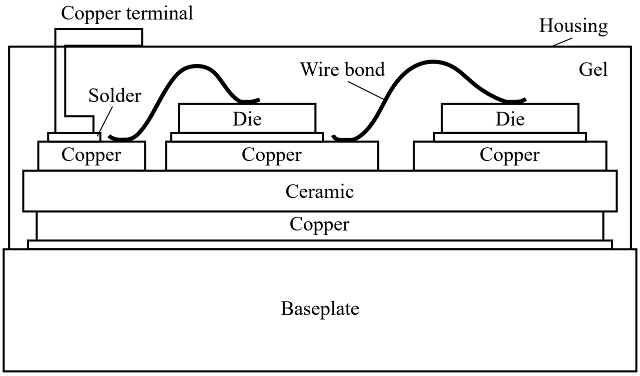


Fig. A.7: Illustration of semiconductor power module.

these methods often require large computation power, which may complicate the online implementation in a real application. [38]

A.4.3 Converters

Converter is a general term for the interfacing electronics systems, which are used to convert the produced voltage of the fuel cell to the DC-bus, to rectify the AC grid voltage and to invert the DC-bus voltage to AC for the AC-load. Additional or different converters may be included in a telecommunication power supply depending on the power sources/storages and the system architecture.

The building blocks of any modern power converter are semiconductor devices such as MOSFETs or IGBTs and passive components such as capacitors and inductors. When several semiconductor devices are integrated in a single package for increased power rating, it is referred to as a power module. These power modules have been subject to extensive reliability studies in recent years. The power modules house the semiconductors, which are typically mounted on direct bonded copper (DBC), connected by aluminum bond wires, and encapsulated in an insulating gel and plastic housing (see Fig. A.7). It is the interconnection of the semiconductors which experience the most failures in such a module and have therefore been studied thoroughly. It has been found that the different coefficient of thermal expansion (CTE) of the many materials in the power module is the main cause of failure modes such as bond wire lift-off and solder delamination. The difference in CTE between two joint materials will cause them to expand and contract differently during the thermal cycling, caused by the switching events of the devices. This will produce shear stress in the interconnections, which will develop into faults over time.

In condition monitoring of power converters the main focus of the literature is on the semiconductor power module [52]–[54], but also the capacitor components are described [55]. In power modules, almost all ageing mechanisms have thermal effects [53], which can be observed in the semiconductor junction temperature, i.e. the temperature internally in the semiconductor. Knowledge of the junction temperature can give valuable information on the power module state-of-health, not only the semiconductor but also the packaging as this also has an effect on the junction temperature. However, the junction temperature is not directly measurable in a real application, which has caused researchers to investigate the estimation of the junction temperature through measurable temperature sensitive electrical parameters (TSEPs).

Common TSEPs described in literature are semiconductor on-state voltage [56], [57], internal gate-resistance [58], and switching parameters [59], [60]. Several other TSEPs exist, as summarized by [61], and might be interesting for condition monitoring of power converters. The monitored TSEP(s) are converted to SOH through a damage models. The SOH parameters can then be extrapolated to the predefined end of life (EOL) criteria, for an RUL estimate.

Alternative to the TSEP monitoring for temperature estimation which then needs to be converted to a health metric, is the monitoring of damage sensitive electrical parameters (DSEPs), which are electrical measurements that do not relate to the temperature, but to the degradation processes directly. Hence, one conversion step and thereby some uncertainty can be omitted. The DSEPs summarized in [61] are the thermal resistance methods [62], [63], and on-state voltage methods [56], [64]. There are two approaches to condition monitoring using DSEPs: past-history approaches and model-based approaches. The past-history approaches use only DSEP measurements on the device, which is under investigation. The model-based approaches rely on DSEP data from other devices to generate a model that can map the measured DSEP to SOH and/or RUL. These models can be based on physical knowledge or on empirical data. The model learning data often stems from accelerated testing of devices, because a complete test in real field environment would be too lengthy.

A.5 Discussion

The topic of condition monitoring and prognostics of fuel cells, batteries, and converters is subject to a lot of recent research. However, the research tends to focus on internal issues in the components, e.g. detection of bond wire lift-off in converters. Hence, monitoring the health condition and failure mechanisms of a complex system, such as a hybrid telecommunication power supply, might require improbable amounts of sensors and algorithms that would drive the cost and effort of implementation to a point where it is not worth the cost [65]. Furthermore, monitoring only subcomponents in a complex system will not reveal failures and ageing of the many auxiliary components and the interconnections between components etc. In a commercial scenario, the system might consist of subsystems from subcontractors, which may not allow the access of internal measurements.

References

[1] M. Aamir, K. Ahmed Kalwar, and S. Mekhilef, “Review: Uninterruptible Power Supply (UPS) system”, *Renewable and Sustainable Energy Reviews*, vol. 58, no. January, pp. 1395–1410, 2016.

[2] J. Gurrero, L. D. Vicuna, and J. Uceda, “Uninterruptible power supply systems provide protection”, *IEEE Industrial Electronics Magazine*, vol. 1, pp. 28–38, 2007.

[3] National Renewable Energy Laboratory, “Fuel Cells for Backup Power in Telecommunications Facilities”, *Doe/Go-102009-2709*, 2009.

[4] K. Fosberg, “Fuel cell systems provide backup power in telecom applications”, *Fuel Cells Bulletin*, vol. 2010, no. 12, pp. 12–14, Dec. 2010.

[5] M. J. Vasallo, J. M. Andújar, C. García, and J. J. Brey, “A Methodology for Sizing Backup Fuel-Cell / Battery Hybrid Power Systems”, *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1964–1975, 2010.

[6] Y. Zhan, Y. Guo, J. Zhu, and H. Wang, “Intelligent uninterruptible power supply system with back-up fuel cell/battery hybrid power source”, *Journal of Power Sources*, vol. 179, no. 2, pp. 745–753, May 2008.

[7] R. Hosseinalizadeh, H. Shakouri G, M. S. Amalnick, and P. Taghipour, “Economic sizing of a hybrid (PV-WT-FC) renewable energy system (HRES) for stand-alone usages by an optimization-simulation model: Case study of Iran”, *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 139–150, 2016.

[8] J. Lee, B. Bagheri, and H.-A. Kao, “A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems”, *Manufacturing Letters*, vol. 3, pp. 18–23, Jan. 2015.

[9] E. Varkaraki, N. Lymberopoulos, and A. Zachariou, “Hydrogen based emergency back-up system for telecommunication applications”, *Journal of Power Sources*, vol. 118, no. 1-2, pp. 14–22, May 2003.

[10] J. H. Pikul, H. Gang Zhang, J. Cho, P. V. Braun, and W. P. King, “High-power lithium ion microbatteries from interdigitated three-dimensional bicontinuous nanoporous electrodes”, *Nature Communications*, vol. 4, p. 1732, 2013.

[11] M. F. Serincan, “Reliability considerations of a fuel cell backup power system for telecom applications”, *Journal of Power Sources*, vol. 309, pp. 66–75, 2016.

[12] B. Bidoggia, G. Spiazzi, S. Buso, T. Lequeu, and L. Ventura, “Design of a fuel cell based backup system for telecoms applications”, in *INTELEC 2008 - 2008 IEEE 30th International Telecommunications Energy Conference*, IEEE, Sep. 2008, pp. 1–8.

[13] M. Ben-Daya, S. O. Duffuaa, A. Raouf, J. Knezevic, and D. Ait-Kadi, Eds., *Handbook of Maintenance Management and Engineering*. London: Springer London, 2009.

[14] G. Niu, B.-S. Yang, and M. Pecht, “Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance”, *Reliability Engineering & System Safety*, vol. 95, no. 7, pp. 786–796, Jul. 2010.

[15] T. Sutharssan, D. Montalvao, Y. K. Chen, W.-C. Wang, C. Pisac, and H. Elemara, “A review on prognostics and health monitoring of proton exchange membrane fuel cell”, *Renewable and Sustainable Energy Reviews*, vol. 75, no. November 2015, pp. 440–450, Aug. 2017.

[16] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, “Remaining Useful Life Prediction and Uncertainty Quantification of Proton Exchange Membrane Fuel Cell Under Variable Load”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2569–2577, Apr. 2016.

[17] M. Jouin, R. Gouriveau, D. Hissel, M. C. Péra, and N. Zerhouni, “Prognostics and Health Management of PEMFC - State of the art and remaining challenges”, *International Journal of Hydrogen Energy*, vol. 38, no. 35, pp. 15 307–15 317, 2013.

[18] J. Zhang and J. Lee, “A review on prognostics and health monitoring of Li-ion battery”, *Journal of Power Sources*, vol. 196, no. 15, pp. 6007–6014, Aug. 2011.

[19] L. Liao and F. Köttig, “Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction”, *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 191–207, 2014.

[20] M. Jouin, M. Bressel, S. Morando, R. Gouriveau, D. Hissel, M.-C. Péra, N. Zerhouni, S. Jemei, M. Hilairet, and B. Ould Bouamama, “Estimating the end-of-life of PEM fuel cells: Guidelines and metrics”, *Applied Energy*, vol. 177, pp. 87–97, Sep. 2016.

[21] M. Knowles, D. Baglee, A. Morris, and Q. Ren, “The state of the art in fuel cell condition monitoring and maintenance”, *World Electric Vehicle Journal*, vol. 4, no. 1, pp. 487–494, 2011.

[22] X. Xue, J. Tang, N. Sammes, and Y. Ding, “Model-based condition monitoring of PEM fuel cell using Hotelling T2 control limit”, *Journal of Power Sources*, vol. 162, no. 1, pp. 388–399, Nov. 2006.

[23] N. Fouquet, C. Doulet, C. Nouillant, G. Dauphin-Tanguy, and B. Ould-Bouamama, “Model based PEM fuel cell state-of-health monitoring via ac impedance measurements”, *Journal of Power Sources*, vol. 159, no. 2, pp. 905–913, 2006.

[24] P. Rodatz, F. Büchi, C. Onder, and L. Guzzella, "Operational aspects of a large PEFC stack under practical conditions", *Journal of Power Sources*, vol. 128, no. 2, pp. 208–217, 2004.

[25] S. Lee, S. Mukerjee, E. Ticianelli, and J. McBreen, "Electrocatalysis of CO tolerance in hydrogen oxidation reaction in PEM fuel cells", *Electrochimica Acta*, vol. 44, no. 19, pp. 3283–3293, 1999.

[26] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Degradations analysis and aging modeling for health assessment and prognostics of PEMFC", *Reliability Engineering & System Safety*, vol. 148, pp. 78–95, Apr. 2016.

[27] D. Hissel and M. Pera, "Diagnostic & health management of fuel cell systems: Issues and solutions", *Annual Reviews in Control*, vol. 42, pp. 201–211, 2016.

[28] J. Chen and B. Zhou, "Diagnosis of PEM fuel cell stack dynamic behaviors", *Journal of Power Sources*, vol. 177, no. 1, pp. 83–95, Feb. 2008.

[29] A. Hernandez, D. Hissel, and R. Outbib, "Fuel cell fault diagnosis: A stochastic approach", in *2006 IEEE International Symposium on Industrial Electronics*, 2006, pp. 1984–1989.

[30] Q. Shen, M. Hou, X. Yan, D. Liang, Z. Zang, L. Hao, Z. Shao, Z. Hou, P. Ming, and B. Yi, "The voltage characteristics of proton exchange membrane fuel cell (PEMFC) under steady and transient states", *Journal of Power Sources*, vol. 179, no. 1, pp. 292–296, Apr. 2008.

[31] F. Barbir, H. Gorgun, and X. Wang, "Relationship between pressure drop and cell resistance as a diagnostic tool for PEM fuel cells", *Journal of Power Sources*, vol. 141, no. 1, pp. 96–101, Feb. 2005.

[32] G. Tian, S. Wasterlain, I. Endichi, D. Candusso, F. Harel, X. François, M.-C. Péra, D. Hissel, and J.-M. Kauffmann, "Diagnosis methods dedicated to the localisation of failed cells within PEMFC stacks", *Journal of Power Sources*, vol. 182, no. 2, pp. 449–461, Aug. 2008.

[33] N. Y. Steiner, D. Hissel, P. Moçotéguy, and D. Candusso, "Non intrusive diagnosis of polymer electrolyte fuel cells by wavelet packet transform", *International Journal of Hydrogen Energy*, vol. 36, no. 1, pp. 740–746, Jan. 2011.

[34] J. Wu, X. Z. Yuan, H. Wang, M. Blanco, J. J. Martin, and J. Zhang, "Diagnostic tools in PEM fuel cell research: Part I Electrochemical techniques", *International Journal of Hydrogen Energy*, vol. 33, no. 6, pp. 1735–1746, Mar. 2008.

[35] M. Ordonez, M. O. Sonnaillon, J. E. Quaicoe, and M. T. Iqbal, "An Embedded Frequency Response Analyzer for Fuel Cell Monitoring and Characterization", *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1925–1934, Jun. 2010.

[36] A. Debenjak, J. Petrovcic, P. Boskoski, B. Musizza, and D. Juricic, “Fuel Cell Condition Monitoring System Based on Interconnected DC-DC Converter and Voltage Monitor”, *IEEE Transactions on Industrial Electronics*, vol. 62, no. 8, pp. 5293–5305, Aug. 2015.

[37] A.-I. Stan, M. Swierczynski, D.-I. Stroe, R. Teodorescu, S. J. Andreasen, and K. Moth, “A comparative study of lithium ion to lead acid batteries for use in UPS applications”, *Telecommunications Energy Conference (INTELEC), 2014 IEEE 36th International*, pp. 1–8, 2014.

[38] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van Den Bossche, “Critical review of state of health estimation methods of Li-ion batteries for real applications”, *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 572–587, 2016.

[39] S. M. Rezvanizaniani, Z. Liu, Y. Chen, and J. Lee, “Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility”, *Journal of Power Sources*, vol. 256, pp. 110–124, 2014.

[40] J. Vetter, P. Novák, M. Wagner, C. Veit, K.-C. Möller, J. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, and A. Hammouche, “Ageing mechanisms in lithium-ion batteries”, *Journal of Power Sources*, vol. 147, no. 1-2, pp. 269–281, Sep. 2005.

[41] D.-i. Stroe, “Lifetime Models for Lithium Ion Batteries used in Virtual Power Plants”, no. September, 2014.

[42] H.-G. Schweiger, O. Obeidi, O. Komesker, A. Raschke, M. Schiemann, C. Zehner, M. Gehnen, M. Keller, and P. Birke, “Comparison of Several Methods for Determining the Internal Resistance of Lithium Ion Cells”, *Sensors*, vol. 10, no. 6, pp. 5604–5625, Jun. 2010.

[43] K. S. Ng, C.-S. Moo, Y.-P. Chen, and Y.-C. Hsieh, “Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries”, *Applied Energy*, vol. 86, no. 9, pp. 1506–1511, Sep. 2009.

[44] T. Sarikurt, M. Ceylan, and A. Balikci, “An analytical battery state of health estimation method”, in *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, IEEE, Jun. 2014, pp. 1605–1609.

[45] X. Hu, S. E. Li, Z. Jia, and B. Egardt, “Enhanced sample entropy-based health management of Li-ion battery for electrified vehicles”, *Energy*, vol. 64, pp. 953–960, Jan. 2014.

[46] A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, “Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods”, *Journal of Power Sources*, vol. 239, pp. 680–688, Oct. 2013.

[47] J. Remmlinger, M. Buchholz, T. Soczka-Guth, and K. Dietmayer, “Onboard state-of-health monitoring of lithium-ion batteries using linear parameter-varying models”, *Journal of Power Sources*, vol. 239, pp. 689–695, Oct. 2013.

[48] G. L. Plett, “Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs”, *Journal of Power Sources*, vol. 134, no. 2, pp. 252–261, Aug. 2004.

[49] Z. Fei, L. Guangjun, and F. Lijin, “Battery state estimation using Unscented Kalman Filter”, in *Robotics and Automation, 2009. ICRA ’09. IEEE International Conference on*, 2009, pp. 1863–1868.

[50] IL-Song Kim, “A Technique for Estimating the State of Health of Lithium Batteries Through a Dual-Sliding-Mode Observer”, *IEEE Transactions on Power Electronics*, vol. 25, no. 4, pp. 1013–1022, Apr. 2010.

[51] A. Eddahech, O. Briat, N. Bertrand, J.-Y. Delétage, and J.-M. Vinassa, “Behavior and state-of-health monitoring of Li-ion batteries using impedance spectroscopy and recurrent neural networks”, *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 487–494, Nov. 2012.

[52] P. Ghimire, A. R. De Vega, S. Beczkowski, B. Rannestad, S. Munk-Nielsen, and P. Thogersen, “Improving power converter reliability: Online monitoring of high-power IGBT modules”, *IEEE Industrial Electronics Magazine*, vol. 8, no. 3, pp. 40–50, 2014.

[53] S. Yang, D. Xiang, A. Bryant, P. Mawby, L. Ran, and P. Tavner, “Condition Monitoring for Device Reliability in Power Electronic Converters: A Review”, *IEEE Transactions on Power Electronics*, vol. 25, no. 11, pp. 2734–2752, Nov. 2010.

[54] B. Rannestad, A. E. Maarbjerg, K. Frederiksen, S. Munk-Nielsen, and K. Gadgaard, “Converter Monitoring Unit for Retrofit of Wind Power Converters”, *IEEE Transactions on Power Electronics*, vol. 8993, no. c, pp. 1–1, 2017.

[55] A. Wechsler, B. C. Mecrow, D. J. Atkinson, J. W. Bennett, and M. Benarous, “Condition Monitoring of DC-Link Capacitors in Aerospace Drives”, *IEEE Transactions on Industry Applications*, vol. 48, no. 6, pp. 1866–1874, Nov. 2012.

[56] S. Beczkowski, P. Ghimre, A. R. de Vega, S. Munk-Nielsen, B. Rannestad, and P. Thogersen, “Online Vce measurement method for wear-out monitoring of high power IGBT modules”, in *2013 15th European Conference on Power Electronics and Applications (EPE)*, IEEE, Sep. 2013, pp. 1–7.

[57] A. B. Jorgensen, S. D. Sonderskov, N. Christensen, K. L. Frederiksen, E. Iciragiye, A. E. Maarbjerg, S. Munk-Nielsen, and B. Rannestad, “Novel screening techniques for wind turbine power converters”, in *2016 18th European Conference on Power Electronics and Applications (EPE’16 ECCE Europe)*, IEEE, Sep. 2016, pp. 1–8.

[58] N. Baker, M. Liserre, L. Dupont, and Y. Avenas, “Improved Reliability of Power Modules: A Review of Online Junction Temperature Measurement Methods”, *IEEE Industrial Electronics Magazine*, vol. 8, no. September, pp. 17–27, 2014.

[59] D. W. Brown, M. Abbas, A. Ginart, I. N. Ali, P. W. Kalgren, and G. J. Vachtsevanos, “Turn-Off Time as an Early Indicator of Insulated Gate Bipolar Transistor Latch-up”, *IEEE Transactions on Power Electronics*, vol. 27, no. 2, pp. 479–489, Feb. 2012.

[60] A. Ginart, D. Brown, P. Kalgren, and M. Roemer, “Online Ringing Characterization as a Diagnostic Technique for IGBTs in Power Drives”, *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 7, pp. 2290–2299, Jul. 2009.

[61] N. Degrenne, J. Ewanchuk, E. David, R. Boldyrjew, S. Mollov, M. E. R, and D. C. Europe, “A Review of Prognostics and Health Management for Power Semiconductor Modules”, in *Annual Conference of the Prognostics and Health Management Society 2015*, M. J. Daigle and A. Bregon, Eds., San Diego, CA, USA: phmsociety, 2015, pp. 1–11.

[62] Bing Ji, Xueguan Song, Wenping Cao, V. Pickert, Yihua Hu, J. W. Mackersie, and G. Pierce, “<italic>In Situ</italic> Diagnostics and Prognostics of Solder Fatigue in IGBT Modules for Electric Vehicle Drives”, *IEEE Transactions on Power Electronics*, vol. 30, no. 3, pp. 1535–1543, Mar. 2015.

[63] B. Saha, J. R. Celaya, P. F. Wysocki, and K. F. Goebel, “Towards prognostics for electronics components”, in *2009 IEEE Aerospace conference*, IEEE, Mar. 2009, pp. 1–7.

[64] J. R. Celaya, A. Saxena, P. Wysocki, S. Saha, and K. Goebel, “Towards Prognostics of Power MOSFETs: Accelerated Aging and Precursors of Failure”, *Phmc 2010*, pp. 0–10, 2010.

[65] A. Sai Sarathi Vasan, C. Chen, and M. Pecht, “A circuit-centric approach to electronic system-level diagnostics and prognostics”, in *2013 IEEE Conference on Prognostics and Health Management (PHM)*, IEEE, Jun. 2013, pp. 1–8.

Paper B

Monitoring Performance Degradation of Proton Exchange Membrane Fuel Cells in Backup Power Systems

Simon Dyhr Sønderskov, Lajos Török, and Stig Munk-Nielsen

The paper has been published in the
2018 IEEE International Telecommunications Energy Conference (INTELEC),
pp. 1–7, 2018. DOI: 10.1109/INTLEC.2018.8612294

ABSTRACT

Proton exchange membrane (PEM) fuel cells are a maturing technology and has in recent years entered the market of backup power systems for e.g. telecommunication applications. Here they provide emission-free and long load support during grid failures and outages. Monitoring performance indicators of PEM fuel cell systems in the field can contribute to improved lifetime, maintenance scheduling and thereby economic competitiveness.

This paper establishes some performance indicators and extracts these from a number of sites, that are all installed and operating in the field. The extracted performance indicators include stack operating time, operating time at non-optimum temperature, consumed reactants, voltage decay, and state of health estimation.

B.1 INTRODUCTION

The power grid is inherently vulnerable to natural or manmade events, that can lead to areas of the network being cut off from the grid, leaving the loads unsupplied. Therefore, critical loads such as hospitals, datacenters, communication infrastructure, etc. rely on backup power systems for continued operation during grid faults. Especially the telecommunication and datacenter equipment is vulnerable to grid faults [1]. Not only will grid failure cause interruptions and loss of data, but also minor grid faults, such as voltage fluctuations, may damage the sensitive equipment. Furthermore, these applications cannot afford any downtime between the grid fault occurring and the backup power is connected. This need is accommodated with backup power systems, which decouple any fluctuations on the power grid from the load and allows for immediate transfer of the load from the failed grid to the backup source [2].

The source of backup power can traditionally be a large variety of technologies with the most dominating being lead-acid batteries and diesel generators [3]. However, in recent years, new technologies have entered the market and especially one has emerged as a viable candidate: proton exchange membrane (PEM) fuel cells. The fuel cell solution benefits from extended backup times compared to the batteries and zero emision, less noise, and less maintenance compared to the diesel generators, which are increasingly prohibited in urban areas.

Yet fuel cells are still a young technology on the market, and still has some hurdles to overcome. Some of the drawbacks of the fuel cell technology are: high cost, low power density, long start-up time, and slow dynamic response. In the case of backup applications, power density is typically less important than energy density. Fuel cells can have very high energy density, as the fuel is stored in separate tanks which can be easily replaced or replenished. The startup time of current PEM fuel cells is in the order of a few minutes [4]. The startup time and the slow dynamic response are in practice mitigated through the inclusion of a small battery or supercapacitor module in the system. Hence, only the cost issue remains as a main challenge for PEM fuel cells in backup power applications.

It is predicted, that the cost of fuel cells will fall as the technology matures and the production quantities increase [5]. Other than this natural cost reduction, there are two main ways of reducing the cost of fuel cells: 1) reduce the cost of materials and manufacturing and 2) reduce the operation and maintenance (O&M) costs.

One way to reduce the O&M cost is through condition monitoring, which is the act of detecting changes in system parameters which indicate developing faults. Condition monitoring can be utilized for optimizing the maintenance schedule (e.g. predictive maintenance), and for mitigating inappropriate conditions that could otherwise lead to degradation in the system. Both seek to extend the lifetime of the system and consequently reduce the total cost of ownership (TCO).

However, the current state of the fuel cell technology means that references for fuel cell performance is largely based on laboratory experiments, and often on single cell or short-stack setups in very controlled environments. This paper seeks to take a first step towards establishing performance references in actual fuel cell backup systems in actual field operation. This is made possible, by the commercial success of the fuel cell technology in backup power systems, which has allowed for collecting data on systems in the field.

In Section B.2 the main performance degradation mechanisms of PEM fuel cells, as reported in literature, is outlined. The system under investigation is presented in Section B.3 and the raw data, collected on the systems is presented in Section E.2. Section B.5 presents how some important system metrics are extracted from the raw dataset. Finally, a conclusion and possible future work is presented in Section B.6.

B.2 Performance Degradation Mechanisms

This section presents a brief review of the performance degradation mechanisms of PEM fuel cells. The structure and operating principle of a PEM fuel cell, as well as the stack assembly, is illustrated in Fig. B.1 and Fig. B.2, respectively.

Degradation of the fuel cell performance can have multiple causes. Some of the major contributors to fuel cell performance degradation are corrosion, contamination, starvation, degradation of the membrane, poor water management, and poor thermal management, which are all briefly explained in the following sections. [6]

B.2.1 Corrosion

Corrosion is a main cause of degradation in PEM fuel cells and can occur in both the electro-catalyst layers, the gas diffusion layers (GDL), and the bipolar plates [7]. The amount of degradation caused by corrosion is linked to the time where cyclic operation of cell voltage occurs. Corrosion of the

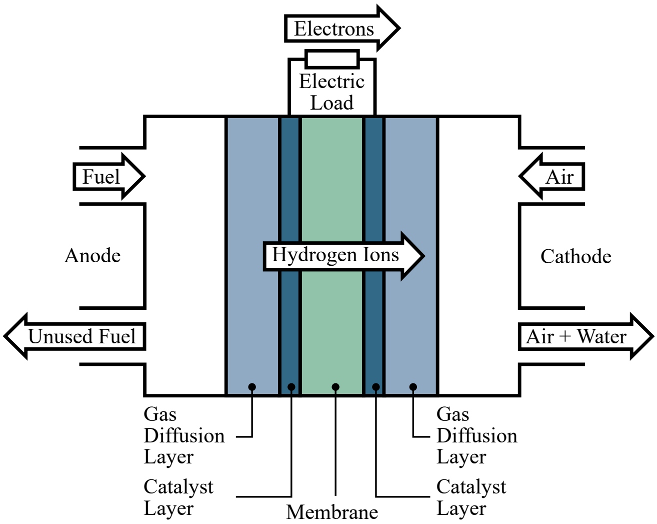


Fig. B.1: Sketch of fuel cell components and operating principle.

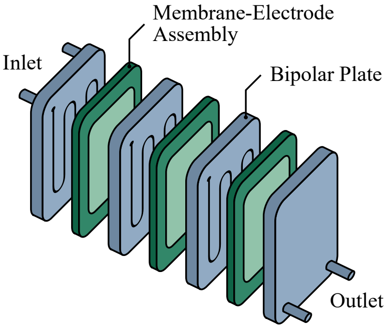


Fig. B.2: Sketch of fuel cell stack assembly.

platinum in the electrodes occurs mainly at the cathode side and causes the loss of electrochemical active surface area (ECSA), increased activation loss and therefore lower cell voltage [7], [8]. This corrosion is accelerated by high relative humidity, cycling of the fuel cell voltage, and high operating temperature [7]. In the GDL, corrosion affects the carbon that supports the catalyst. Corrosion of the GDL will cause lower cell voltage and reduced performance [6] and is accelerated by voltage cycling and low relative humidity of the reactant gases [7]. Corrosion in the bipolar plates can cause the formation of a resistive surface layer, which will increase the ohmic resistance and thereby loss.

B.2.2 CONTAMINATION

Contamination is when impurities pollute or cause chemical attacks and hinder the intended reactions in the cell. Even small amounts of impurities can cause contamination that can seriously degrade the performance and lifetime of the fuel cell [8], [9]. Contamination can occur in both the anode electrode and the membrane. In the anode, carbon monoxide (CO) impurities in the hydrogen supply can lead to contamination (CO-poisoning), which leads to lower cell voltage and reduces energy conversion efficiency [6]. Membrane contamination can be caused by impurities in the hydrogen or air supply, or by impurities resulting from corrosion of fuel cell components and can cause major reduction in performance including conductivity and mass transfer [9].

B.2.3 STARVATION

Starvation of the fuel cell is the lack of sufficient reactant gas and can have multiple causes including: blocked pores in the GDL; poor gas feeding management; imperfect stack and cell design; poor stack assembly; and quick load transients [8], [10]. The starvation of hydrogen will cause a high anode potential, since the current cannot be maintained. This can cause the water at the anode to split. Hence, oxygen will be present at the anode. Similarly, oxygen starvation on the cathode side will cause a reaction where hydrogen is produced. Combined, the oxygen on the anode side and the hydrogen on the cathode side causes voltage reversal of the cell, i.e. a negative voltage between anode and cathode [11]. This voltage reversal causes acceleration of carbon corrosion and eventually damaged components [10], [12].

B.2.4 MEMBRANE DEGRADATION

Membrane degradation is one of the main lifetime reducing factors of PEM fuel cells and is a complicated mechanism, consisting of both mechanical and chemical degradation. The chemical degradation is caused by the formation of radicals – highly chemical reactive molecules, which cause chemical attacks in the membrane [13]. The mechanical degradation is caused by transient operating conditions such as voltage, temperature, and humidification cycling, along with the aforementioned chemical attacks [13]. These degradations can lead to severe membrane degradation such as pinhole formation.

B.2.5 WATER MANAGEMENT

Improper water management can lead to two inappropriate scenarios: fuel cell flooding, and membrane dehydration. Flooding can occur on both the anode and cathode side of the membrane, but occurs in particular on the cathode side [8]. The excess water significantly reduces the transport rate of the reactants causing increasing mass transport loss [8]. The water blocks the reactants from passing through the GDL and thereby causes gas starvation, which leads to an immediate drop in cell voltage. Membrane dehydration is most likely to occur on the anode side as opposed to the cathode side. A dehydrated cell experiences immediate as well as long-term degradation. The decreased water content causes a decrease in the proton conductivity, which leads to higher ionic resistance and therefore ohmic loss [8], [14]. The effect is a voltage drop and temporary power loss [14], [15]. The voltage is usually recoverable through humidification, but prolonged dry cell operation lead to irreversible damage to the membrane, such as development of crazes and cracks [16]. This can start a destructive cycle of mechanical degradation, where the cracks cause gas crossover leading to hotspots, which in turn causes pin holes, that facilitate further gas crossover [17].

B.2.6 Thermal Management

In general, the fuel cell performance is decreased at both low and high temperatures [8], [18]. However, sub-zero temperatures in particular can cause issues in the fuel cell. Mainly due to the freezing of the water content in the cell, which can cause mechanical damage to the components, delamination, and startup issues [6]. High temperature operation has some benefits including higher tolerance to contaminants and enhanced water management and cooling. However, the degradation rate is accelerated and the long-time performance is decreased. [19]

B.2.7 Summary

All of the presented degradation mechanisms, can affect the system performance and ultimately the produced voltage output. Although some of the mentioned mechanisms can be avoided by proper control of the fuel cell system, but many are unavoidable in long term operation in a varying environment.

B.3 Description of the System

The system under investigation is a commercial backup system for telecommunication applications. A sketch of the system architecture is depicted in Fig. B.3. The system consists of two PEM fuel cell stacks, that are supplied with hydrogen and air. The fuel cell stack is interfaced to a DC-bus through parallel DC/DC converters. The DC-bus collects power from the grid through a rectifier during normal operation. Power is delivered to the load directly or through a converter depending on the load type, which can be either DC, AC, or a combination of the two. During grid failure, the fuel cells take over the load. However, the fuel cells have some startup time, in which the bridging components, i.e. batteries or supercapacitors, deliver power to the load. In combination with the bridging components, the fuel cells provide immediate backup, ensuring no interruption of the load.

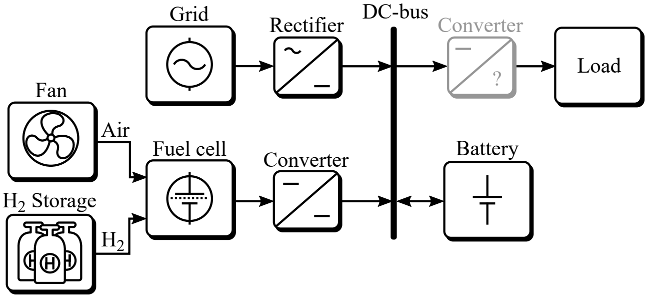


Fig. B.3: System architecture

B.4 Raw Data

The backup system normally resides in standby mode, as the grid is normally supplying the load. Therefore regular self-tests are performed on the system. The purpose of the self-tests is to both ensure that the system is functioning and to exercise the fuel cell, keeping the performance at a high level.

An example of a system use profile is depicted in Fig. B.4. The data originates from a complete fuel cell backup power system, installed in the field, supplying a telecom load.

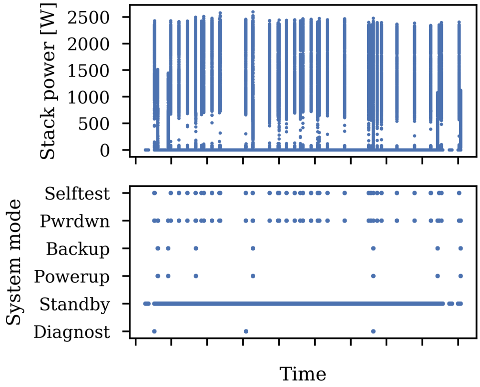


Fig. B.4: System use profile example showing the produced power of a fuel cell stack.

The system has several modes, the main ones being Standby, Backup, and Selftest. The self-tests are performed by a sequence of system excitations, including ramping up the power to a specific level and keeping the power constant. The power profile during a self-test is shown in Fig. B.5. The steep increases in power, that occur periodically throughout the self-test, are caused by the chosen operation strategy.

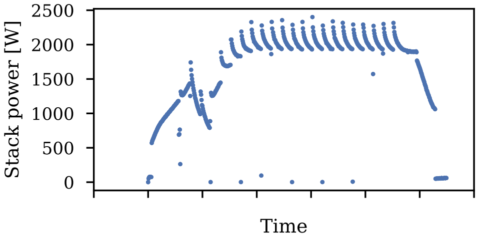


Fig. B.5: Stack power during a self-test.

The collected measurements include various electrical parameters on the stacks and converters, temperatures on stacks and peripheral components, reactant gas pressure, etc. In addition, the system modes and other states are logged. In total 125 parameters are logged with a sampling frequency of 1/h in standby mode and 1/s in all active modes. This data is available for a period of several years for several dozen systems installed in the field.

The raw data is contained in multiple CSV files. One CSV file represents one days worth of data from one system. Hence each system easily comprises 1000 CSV files of raw data. To greatly increase the ease of data access and analysis, the raw CSV files are added to an SQL database.

Specific measurements are used in the following section to extract key performance indicators (KPI) of the systems.

B.5 PERFORMANCE INDICATORS

The amount of collected data and the variability in operating conditions can make it difficult to compare the performance of different systems. Therefore, it is useful to extract some performance indicators from the systems. In section B.5.1, the presented performance indicators are limited to indicators that describe how much the system has been used throughout its lifetime, i.e. counters and totalizers. In Section B.5.2, the voltage decay throughout the system lifetime is extracted of a single example system, and a state of health metric is estimated in section B.5.3.

B.5.1 COUNTERS AND TOTALIZERS

How much the system has been used, is highly relevant for the level of degradation, that can be expected in a given system. Some of the important use-metrics that are presented in this section are: for how long has the system operated above a certain current level; how much time has the system spent in each of the system modes; for how long has the system operated at non-optimum temperatures; how much energy and charge has the system produced; and how many startups has the system made.

Operating Times

The operating time, as defined by the fuel cell manufacturer, is the total time in which the fuel cell stack has delivered above a certain level of current, in this case 5 A. The procedure for extracting the operating time is described in Algorithm 1, where L is the length (number of samples) of the dataset, t is the timestamp, B denotes the beginning of the period in which the current is above 5 A, E denotes the end of the period in which the current is above 5 A, and *operation_time* is the total operating time above 5 A.

Algorithm 1 Calculating stack operating time

```
initialize $j = 1$ and $k = 1$
for $i = 2$ to $L - 1$ do
    if $(stack\_current_i > 5)$ then
        if $(stack\_current_{i-1} < 5)$ then
            $B_j = t_i$
            $j = j + 1$
        end if
        if $(stack\_current_{i+1} < 5)$ then
            $E_k = t_{i+1}$
            $k = k + 1$
        end if
    end if
end for
$operation\_time = \mathrm{sum}(E - B)$
```

Similarly, the time the system spends in each mode (*mode_time*; e.g. *backup_time*) is calculated as shown in Algorithm 2.

Algorithm 2 Calculating time in specific mode

```
initialize $j = 1$ and $k = 1$
for $i = 2$ to $L - 1$ do
    if $(stack\_mode_i = mode)$ then
        if $(stack\_mode_{i-1} \neq mode)$ then
            $B_j = t_i$
            $j = j + 1$
        end if
        if $(stack\_mode_{i+1} \neq mode)$ then
            $E_k = t_{i+1}$
            $k = k + 1$
        end if
    end if
end for
$mode\_time = \mathrm{sum}(E - B)$
```

The optimum temperature of the stack is linearly dependent on the produced current. The algorithm of calculating the time in too high temperature operation is shown in Algorithm 3. parameters a, and b define the relationship of the current and optimum temperature, given in the fuel cell datasheet. ΔT is the allowable deviation from the optimum temperature. The time the system operates below the optimum temperature is obtained similarly.

Algorithm 3 Calculating time above optimum temperature

```
initialize $j = 1$ and $k = 1$
for $i = 2$ to $L - 1$ do
    if ($stack\_mode_i = mode$) then
        if ($stack\_mode_{i-1} \neq mode$) then
            $T_{opt} = a \cdot stack\_current_i + b$
            if ($stack\_temp_i > T_{opt} + \Delta T$) then
                $B_j = t_i$
                $j = j + 1$
            end if
        end if
        if ($stack\_mode_{i+1} \neq mode$) then
            $T_{opt} = a \cdot stack\_current_i + b$
            if ($stack\_temp_i > T_{opt} + \Delta T$) then
                $E_k = t_{i+1}$
                $k = k + 1$
            end if
        end if
    end if
end for
$mode\_time\_overtemp = \mathrm{sum}(E - B)$
```

Production and Consumption

The produced energy of the stack is a useful indication of how much the stack is aged. The produced energy is simply calculated from the trapezoidal approximation of the stack voltage and stack current integral as so

$$E = \sum_{i=1}^{L} \frac{V_{i-1} \cdot I_{i-1} + V_i \cdot I_i}{2} (t_i - t_{i-1}) \tag{B.1}$$

where L is the length of the dataset, v is the stack voltage, and i is the stack current. Similarly the produced charge is:

$$Q = \sum_{i=1}^{L} \frac{I_{i-1} + I_i}{2} (t_i - t_{i-1}) \tag{B.2}$$

For the system to produce power, it consumes reactants, namely hydrogen and oxygen from the air. The amount of hydrogen, consumed is calculated by

$$\dot{m}_{H_2} = K_{H_2} \cdot Q \tag{B.3}$$

where K_{H_2} is a scalar obtained from the fuel cell datasheet. Similarly for the oxygen consumption:

$$\dot{m}_{O_2} = K_{O_2} \cdot Q \tag{B.4}$$

SYSTEM STARTUPS

The number of startups is expected to have an effect on system performance. This number is calculated by counting the number of transitions from a non-active mode (e.g. standby) to an active mode (e.g. backup). Furthermore, air-air starts (when air is initially present at both anode and cathode), is especially important as it can cause corrosion that leads to irreversible damage. The air-air startups as well as total number of startup are counted in Algorithm 4.

Algorithm 4 Counting fuel cell startups

```
$start\_counter = 0$
$air\_to\_air\_start\_counter = 0$
for $i = 1$ to $L$ do
    if ($stack\_mode_i$ changes from non-active to active) then
        $start\_counter += 1$
        if ($stack\_voltage_i < 6.6$) then
            $air\_to\_air\_start\_counter += 1$
        end if
    end if
end for
```

The above performance indicators are extracted on each stack in the available dataset. The resulting indicators are visualized in the boxplots of Fig. B.6, where the data has been normalized to individual parameter mean values.

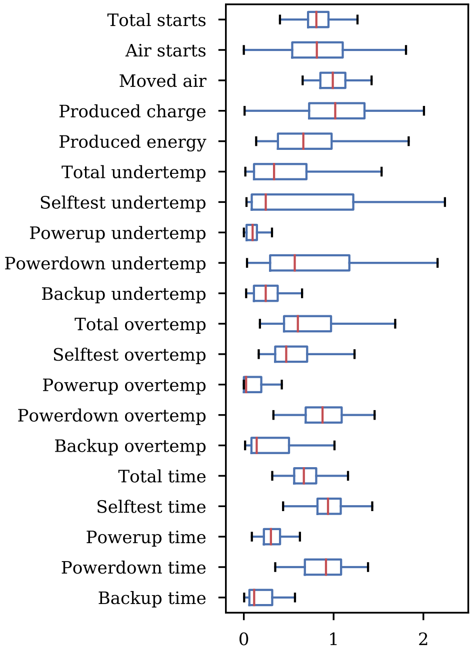


Fig. B.6: Boxplot of performance indicators showing their variation over a fleet of fuel cell stacks. The data of each boxplot has been normalized to its own mean value. In the boxplots, the red line indicates the median value, the blue box indicates the first and third quartile, and the whiskers indicate the minimum and maximum values excluding outliers.

B.5.2 Voltage Decay

The self-tests provide useful information on the voltage decay of the stacks. The useful part of the self-test voltage is when the stack power has settled at a steady value. One example of this voltage profile is shown in Fig. B.7 as the average cell voltage throughout the stack.

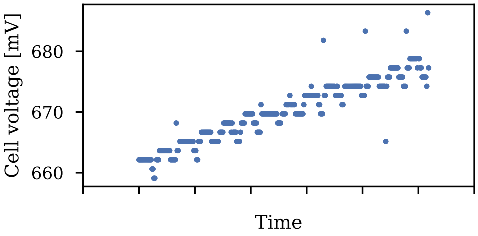


Fig. B.7: Average cell voltage profile during self-test when the stack power has settled.

Two successive self-test voltages are sketched in Fig. B.8 for the purpose of illustrating the calculations in the following subsections.

Fig. B.8: Illustration of two successive self-test voltage profiles.

Standby Decay

In general the fuel cell stacks experience some loss of voltage (voltage decay) during the standby periods between the self-tests, backup, or other active periods. This decay is defined as the decrease in voltage from the end of the voltage profile, as shown in Fig. B.7, to the beginning of the following voltage profile. From the illustration of Fig. B.8, this decay would be calculated as

$$V_{standbydecay,i} = V_{i,2} - V_{i+1,1} \tag{B.5}$$

where i indicates the self-test number (1, 2, ...).

Calculating this standby voltage decay for the two stacks in a single system gives the points presented in Fig. B.9. Due to possible variation and noise in the measured voltage, the voltages used for calculation are the median of the first ten points and last ten points in the self-test voltage, respectively.

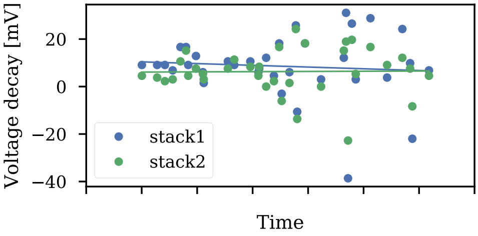


Fig. B.9: Voltage decay during standby periods

Recovered Voltage

Some of the voltage decay from the standby periods is recovered during the self-test, as exemplified by the rise in voltage in Fig. B.7. This recovered voltage is calculated for each self-test (i) by

$$V_{recovered,i} = V_{i,2} - V_{i,1} \tag{B.6}$$

The result from a single system is shown in Fig. B.10. In the shown case anywhere from 1 to 11 mV is recovered per cell each self-test.

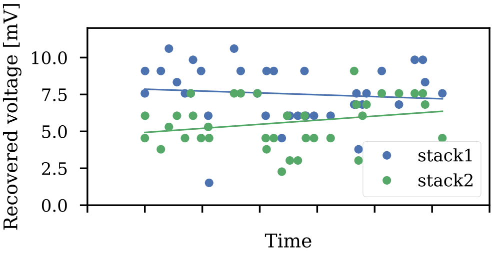


Fig. B.10: Voltage recovery during self-tests

Unrecovered Voltage Decay

However, not all the standby decay is recovered during the self-tests, leading to voltage decay. The unrecovered voltage decay is calculated by

$$V_{decay,i} = V_{i+1,2} - V_{i,2} \tag{B.7}$$

The result is shown in Fig. B.11.

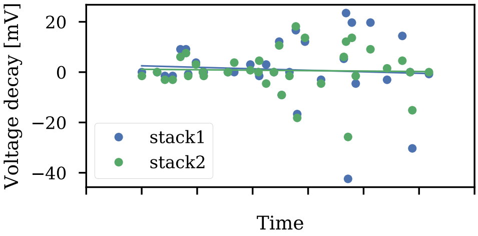


Fig. B.11: Voltage decay including recovered voltage from self-tests

B.5.3 State of Health

At the end of the self-test procedures, the voltage can be considered approximately at steady state value. Plotting these voltages against the measured current at the same instance, indicates how the fuel cell stack is degrading from the beginning of life (BOL) and end of life (EOL) polarization curves (Fig. B.12). The polarization curves are provided in the fuel cell manual and are attained from controlled laboratory experiments, whereas the voltage-current points are taken in a dynamic environment. Therefore, comparing the two should be done with caution, and might not provide accurate information of the actual performance level of the stack, but is nonetheless an indication that might be used to compare fuel cell systems with each other.

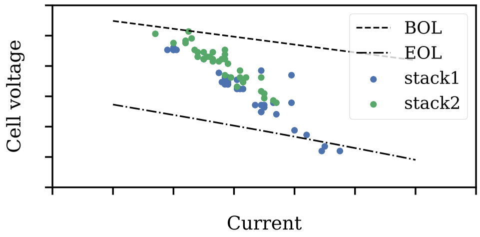


Fig. B.12: Beginning of life (BOL) and end of life (EOL) polarization curves and self-test steady-state voltage-current points

With this in mind, a state of health (SOH) metric of the system is calculated from the BOL and EOL references for each of the voltage current points using the formula

$$SOH_i = \frac{V_i - EOL(I_i)}{BOL(I_i) - EOL(I_i)} \tag{B.8}$$

where V and I are the average cell voltage and current, respectively, and $BOL()$ and $EOL()$ are the beginning of life and end of life functions, respectively.

The resulting SOH evolution of the two stacks in the system is depicted in Fig. B.13. It is clear from the SOH metric, that the system under investigation experiences some degradation in the form of change in voltage-current relationship at a fixed power level throughout its lifetime.

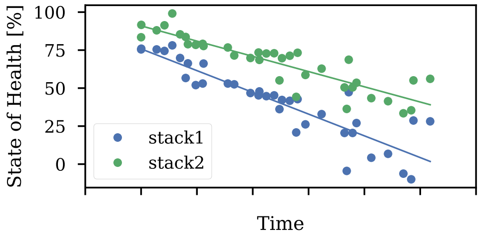


Fig. B.13: State of Health (SOH) throughout the system operation time

B.6 Conclusion

This paper has shown how performance indicators of PEM fuel cell backup systems can be extracted form raw measurements. The extracted performance indicators from fuel cell stacks in the field are presented as well as voltage decay and state of health estimation throughout the lifetime of a single system.

B.6.1 Future Work

This paper presents a first look at field data of commercial fuel cell backup power systems with the purpose of assessing the system performance and its degradation. The following work will go deeper into analyzing the data to establish useful performance assessment criteria and testing influences of specific environmental conditions on system performance and degradation.

References

[1] National Renewable Energy Laboratory, “Fuel Cells for Backup Power in Telecommunications Facilities”, *Doe/Go-102009-2709*, 2009.

[2] J. Gurrero, L. D. Vicuna, and J. Uceda, “Uninterruptible power supply systems provide protection”, *IEEE Industrial Electronics Magazine*, vol. 1, pp. 28–38, 2007.

[3] M. J. Vasallo, J. M. Andújar, C. García, and J. J. Brey, “A Methodology for Sizing Backup Fuel-Cell / Battery Hybrid Power Systems”, *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1964–1975, 2010.

[4] L. Torok, S. L. Sahlin, S. K. Kor, and B. Bidoggia, “Estimation of membrane hydration status for active proton exchange membrane fuel cell systems by impedance measurement: Start-up time measurements”, in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, IEEE, Jun. 2016, pp. 1–5.

[5] H. Tsuchiya and O. Kobayashi, “Mass production cost of PEM fuel cell by learning curve”, *International Journal of Hydrogen Energy*, vol. 29, no. 10, pp. 985–990, 2004.

[6] W. Schmittinger and A. Vahidi, “A review of the main parameters influencing long-term performance and durability of PEM fuel cells”, *Journal of Power Sources*, vol. 180, no. 1, pp. 1–14, May 2008.

[7] R. L. Borup, J. R. Davey, F. H. Garzon, D. L. Wood, and M. A. Inbody, “PEM fuel cell electrocatalyst durability measurements”, *Journal of Power Sources*, vol. 163, no. 1, pp. 76–81, Dec. 2006.

[8] G. Hinds, “Performance and Durability of PEM Fuel Cells: A Review”, *Nat. Phys. Lab,*, no. September, 2004.

[9] X. Cheng, Z. Shi, N. Glass, L. Zhang, J. Zhang, D. Song, Z.-S. Liu, H. Wang, and J. Shen, “A review of PEM hydrogen fuel cell contamination: Impacts, mechanisms, and mitigation”, *Journal of Power Sources*, vol. 165, no. 2, pp. 739–756, Mar. 2007.

[10] Z. Liu, L. Yang, Z. Mao, W. Zhuge, Y. Zhang, and L. Wang, “Behavior of PEMFC in starvation”, *Journal of Power Sources*, vol. 157, no. 1, pp. 166–176, 2006.

[11] T. Van Nguyen and M. W. Knobbe, “A liquid water management strategy for PEM fuel cell stacks”, *Journal of Power Sources*, vol. 114, no. 1, pp. 70–79, Feb. 2003.

[12] T. W. Patterson and R. M. Darling, “Damage to the Cathode Catalyst of a PEM Fuel Cell Caused by Localized Fuel Starvation”, *Electrochemical and Solid-State Letters*, vol. 9, no. 4, A183, 2006.

[13] D. E. Curtin, R. D. Lousenberg, T. J. Henry, P. C. Tangeman, and M. E. Tisack, “Advanced materials for improved PEMFC performance and life”, *Journal of Power Sources*, vol. 131, no. 1-2, pp. 41–48, 2004.

[14] J.-M. Le Canut, R. M. Abouatallah, and D. A. Harrington, “Detection of Membrane Drying, Fuel Cell Flooding, and Anode Catalyst Poisoning on PEMFC Stacks by Electrochemical Impedance Spectroscopy”, *Journal of The Electrochemical Society*, vol. 153, no. 5, A857, 2006.

[15] T. V. Nguyen, “A Water and Heat Management Model for Proton-Exchange-Membrane Fuel Cells”, *Journal of The Electrochemical Society*, vol. 140, no. 8, p. 2178, 1993.

[16] X. Huang, R. Solasi, Y. Zou, M. Feshler, K. Reifsnider, D. Condit, S. Burlatsky, and T. Madden, “Mechanical endurance of polymer electrolyte membrane and PEM fuel cell durability”, *Journal of Polymer Science Part B: Polymer Physics*, vol. 44, no. 16, pp. 2346–2357, Aug. 2006.

[17] J. Yu, T. Matsuura, Y. Yoshikawa, M. N. Islam, and M. Hori, “In Situ Analysis of Performance Degradation of a PEMFC under Nonsaturated Humidification”, *Electrochemical and Solid-State Letters*, vol. 8, no. 3, A156, 2005.

[18] A. Faghri and Z. Guo, “Challenges and opportunities of thermal management issues related to fuel cell technology and modeling”, *International Journal of Heat and Mass Transfer*, vol. 48, no. 19-20, pp. 3891–3920, 2005.

[19] A. S. Arico, A. Stassi, E. Modica, R. Ornelas, I. Gatto, E. Passalacqua, and V. Antonucci, “Evaluation of High Temperature Degradation of Pt/C Catalysts in PEM Fuel Cells”, in *ECS Transactions*, vol. 3, ECS, 2006, pp. 765–774.

Paper C

Detecting Performance Outliers in Fuel Cell Backup Power Systems

Simon Dyhr Sønderskov, Dean Rasmussen, Jakob Ilsøe, Daniel Blom-Hansen, and Stig Munk-Nielsen

The paper has been approved for publication in the
2019 European Conference on Power Electronics and Applications (EPE).

Abstract

Fuel cell based backup power systems are finding application in telecommunication applications. Due to the criticality of the application, as well as the effort of keeping operating costs low, appropriate maintenance strategies are of high importance. This paper investigates key performance indicators, derived from numerous fuel cell based backup systems, installed in the field. The methods of principal component analysis and local outlier factor are applied to the KPIs, in order to identify systems that are performing differently from the majority of the systems. These underperforming systems can then be examined closer to identify potential problems.

C.1 Introduction

Fuel cell technology is a promising candidate in an increasingly renewable and distributed generation based energy system with a growing demand for energy storage [1]. One application where fuel cells are gaining market traction is in backup power systems [2], [3] where they fill a gap where batteries energy storage does not achieve high enough capacity and where diesel generators are too polluting or too noisy [4]. Some of the big advantages of fuel cell systems are their lack of moving parts, ease of fuel replenishment, and pollution free operation. Some setbacks of the technology are the relatively high CAPEX cost and slow dynamic response. The latter can be easily compensated with smaller, faster responding energy storage units such as batteries or ultracapacitors.

The typical fuel cell backup systems, as used widely in the telecommunications, consists of proton exchange membrane (PEM) fuel cell stacks and their peripheral components, a battery and/or ultracapacitor unit, and one or more power electronic converters [3], [5], [6]. An example of a fuel cell based backup power system architecture for telecommunications applications is depicted in Fig. C.1. During normal operation, the telecom load is supplied with power from the grid through appropriate conversion stages. When a fault occurs in the electrical grid, the fuel cell will start up to supply the load power. During the fuel cell start-up, the battery or ultracapacitor unit will supply the load.

The ability of the network provider to provide a reliable connection to its costumers relies on the availability of electrical power. Therefore, the reliability of the backup system is critical, especially in less reliable electrical grids [7]. For this reason, proper maintenance of the backup systems is required, which can often be an expensive task in remote locations. One method for better planning maintenance activities is to adapt a condition monitoring approach, where system critical parameters are remotely monitored and assessed in order to identify potential and emerging faults [8].

This paper will examine historical operating data from multiple fuel cell based backup systems in real field operation in order to identify underperforming systems, for further inspection and possible maintenance actions. To the authors' knowledge, no scientific publications before this, has examined in-field operating data on fuel cells for the purpose of assessing performance levels within the fleet of systems.

C.2 Monitored Parameters

Backup systems based on fuel cells for telecommunications, installed in the field, have collected data on their operating conditions through several years. This collected data is the basis of this study. The data contains logs of the various system states, such as standby, backup, and self-test. As well as performance variables such as produced voltages and currents, and operating conditions such as temperatures of various system components.

In this work, key performance indicators (KPIs) regarding the fuel cell stack's operation is considered. Specifically, the number of stack startups at different conditions and stack runtime in different system modes and system conditions, as listed in Table C.1. For details on how the KPIs are extracted from the raw data, please refer to [9].

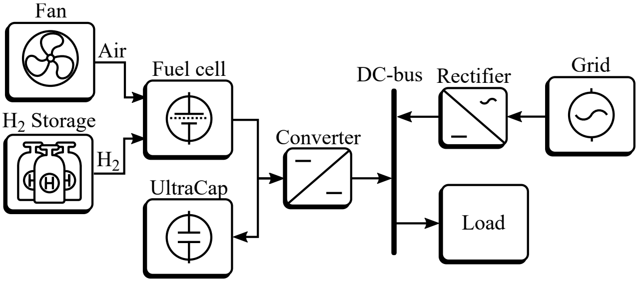


Fig. C.1: Fuel cell backup system architecture. The arrows indicate the direction of energy flow between the components.

Table C.1: Key Performance Indicators

Name	Unit	Description
`Airair_startups`	-	Number of system startups where the stack voltage is below a certain threshold
`Number_Backups`	-	Number of backup events
`Number_Selftests`	-	Number of self-test events
`Startups`	-	Total number of system startups
`Runtime_Backup`	h	Runtime in backup mode
`Runtime_Bad`	h	Runtime in bad stack condition
`Runtime_Diagnost`	h	Runtime in diagnostics mode
`Runtime_Downgraded`	h	Runtime in downgraded stack condition
`Runtime_Power_down`	h	Runtime in stack power-down mode
`Runtime_Powerup`	h	Runtime in stack power-up mode
`Runtime_Safelock`	h	Runtime in safe-lock mode
`Runtime_Safemode`	h	Runtime in safe mode
`Runtime_Self_test`	h	Runtime in self-test mode
`Runtime_Shutdown`	h	Runtime in shutdown mode
`Runtime_Standby`	h	Time in standby mode
`Total_runtime`	h	Total runtime of the stack

Because the systems under investigation have been running in the field for varying amounts of time, the KPIs are all normalized to total runtime of the specific system, effectively giving the number of start-ups per hour of system runtime, etc. The KPIs are further standardized to each have zero mean and unity variance. The resulting KPI distributions are visualized in Fig. C.2.

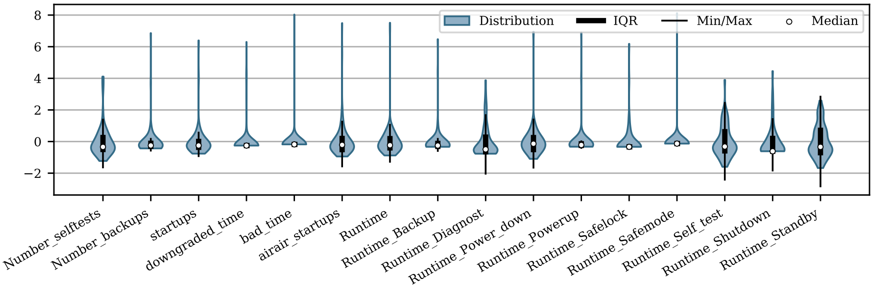


Fig. C.2: Violin plot illustrating the distribution of the standardized performance indicators. The white dot indicates the median value, the black box indicates the interquartile range (from the 25th to the 75th percentile), and the black whiskers indicate the minimum and maximum values, excluding outliers.

C.3 Dimensionality Reduction

Dimensionality reduction is a means of feature extraction or feature selection in machine learning [10]. The purpose is to reduce the number of features (independent variables) while retaining as much information from the original feature space as possible. A widely used method for dimensionality reduction is principal component analysis (PCA) [11]. While the number of features in this case is fairly low, the dimensionality reduction eliminates the covariance between the original features. Also, a reduced number of features simplifies the visualization of the data, and it makes this work more generally applicable to larger feature spaces, should more KPIs be considered in future studies.

C.3.1 Principal Component Analysis

PCA is a linear method of dimensionality reduction, which maps the original data to a lower dimensional space in such a way, that the variance in the new space is maximized. The method is linear and relies on the computation of eigenvectors and eigenvalues. The process of the method as applied on the presented data is outlined in this section. [10]–[12]

Let X be the $n \times p$ matrix of p KPIs of n samples each. The first step is to ensure that X has zero sample (column-wise) mean. This is simply achieved by calculating the mean of each column (u_j) in X and subtracting it from the original matrix:

$$u_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij} \quad \text{for} \quad j = 1, \dots, p \tag{C.1}$$

$$B = X - hu^{\mathrm{T}} \tag{C.2}$$

where X_{ij} is the i^{th} element of the j^{th} column in X, h is a unit column vector of size n and u is the column vector collection of means u_j for all j. B then represents the zero sample mean data.

The covariance matrix (Σ) of B is calculated by

$$\Sigma = \frac{1}{n-1} B^{\mathrm{T}} B \tag{C.3}$$

which gives a $p \times p$ matrix describing the covariance between the KPIs. This is visualized in the matrix plot of the correlation between the KPIs in Fig. C.3 (the correlation matrix can be understood as a normalized covariance matrix).

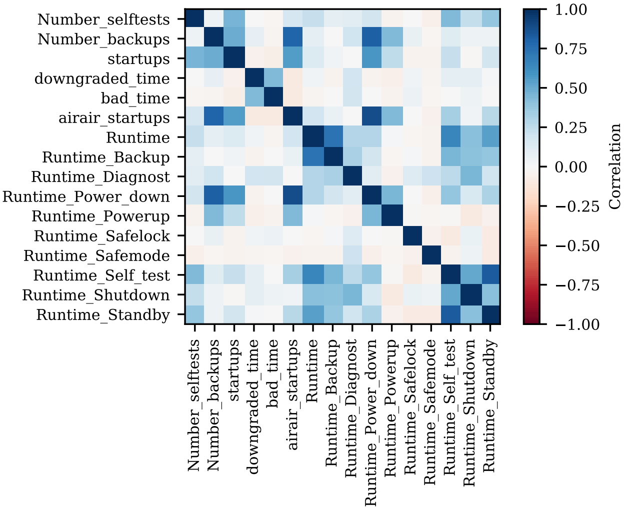


Fig. C.3: Correlation matrix of the examined key performance indicators

Compute the matrix Q of eigenvectors from the following equation

$$Q^{-1}\Sigma Q = \Lambda \tag{C.4}$$

where Λ is the diagonal matrix of eigenvalues of the covariance matrix. Both Q and Λ will have dimensions $p \times p$ with the eigenvalue $\lambda_j = \Lambda_{jk}$, with $k = j$ corresponding to the j^{th} eigenvector of Q for $j = 1, \ldots, p$.

Now a transformation matrix (W) can be obtained. The transformation matrix consists of L eigenvectors, where L is the number of dimensions intended for the transformed data. When the eigenvectors in Q are ordered in descending order of the corresponding eigenvalues, W is obtained by

$$W_{jl} = Q_{jl} \quad \text{for} \quad j = 1, \ldots, p \quad l = 1, \ldots, L \tag{C.5}$$

The absolute value of the eigenvalues is a measure of how much variance is explained by the corresponding eigenvector. Hence the first principal component is the principal components, which explains the most variance, corresponding to the highest absolute eigenvalue. The second principal component, which explains the second most variance, is determined from the second largest eigenvalue and so forth.

How many principal components should be used to represent the original data is a tradeoff between the amount of compression and the loss of variance from the original data. A quantitative way of determining the number of principal components (L), is to look at the proportion of explained variance (PoV) [11], which is the relationship between the explained variance (also called cumulative energy) of the L largest eigenvalues and the sum of all the eigenvalues:

$$PoV = \frac{\sum_{j=1}^{L} \lambda_j}{\sum_{j=1}^{p} \lambda_j} \tag{C.6}$$

A PoV of 1 means that all the variance of the original data will be contained in the transformed data. Hence determining L is done by defining a lower limit of the explained variance, e.g.

$$PoV \geq 0.9 \tag{C.7}$$

Calculating the eigenvectors, eigenvalues and the proportion of explained variance for each, the explained variance ratio (PoV) contribution of each principal component is plotted in Fig. C.4.

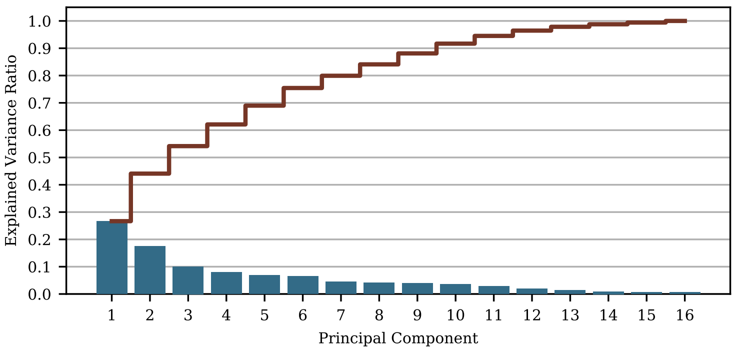


Fig. C.4: Explained variance versus number of included principal components

In this case, to have a PoV above 0.9, nine principal components are needed, giving a PoV of 0.92. Meaning that the nine eigenvectors corresponding to the nine largest eigenvalues are included in the transformation matrix W.

Finally, the original dataset X is transformed to an L dimensional space using the transformation matrix W:

$$T = XW \tag{C.8}$$

The transformed dataset is visualized in Fig. C.5 where each principal component is plotted against each other principal component. The plots in the diagonal show the distribution of the given principal component in a histogram.

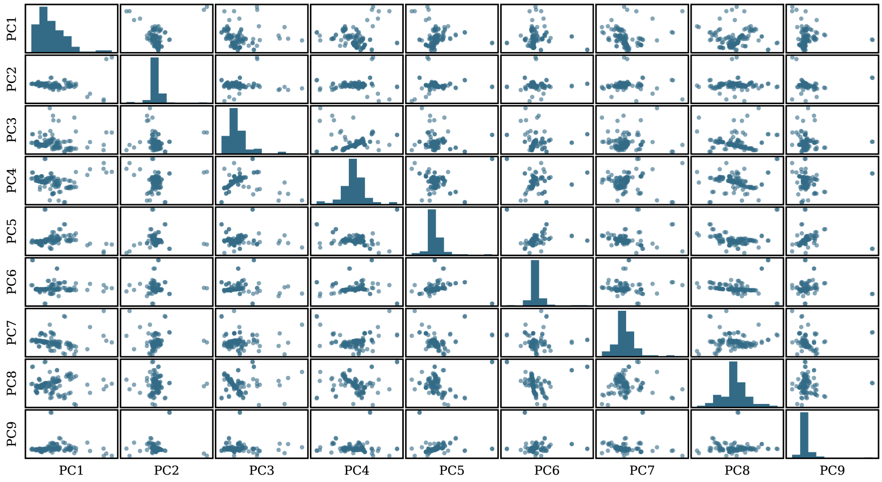


Fig. C.5: Scatter matrix of the transformed dataset.

C.4 ANOMALY DETECTION

There are several methods of anomaly detection in machine learning [13], grouped mainly in supervised and unsupervised methods as well as some hybrid methods. While the supervised methods require examples of known normal and anomaly data, unsupervised methods assume that most of the data is normal and looks for instances that fits the least to the remaining part of the dataset.

Since the dataset considered in this work is unlabeled, an unsupervised anomaly detection method is used. One such method is the local outlier factor (LOF) [14]. LOF works by measuring the relative deviation of a given point in relation to its neighbors. It works by estimating the local density of a point and comparing this density to the density of neighboring points. The locality is determined by the K Nearest Neighbors (KNN) method and the density is described by the distances to these K nearest neighbors. In the end, points that have significantly lower local density than their neighboring points can be considered outliers.

C.4.1 THE LOF ALGORITHM

Let A be a point in the dataset, of which the local outlier factor is to be calculated. The set of K nearest neighbors of point A ($N_K(A)$) is determined by calculating the distances from A to the other points in the dataset and extracting a subset of K points with the lowest distance from A.

The LOF of point A is calculated by comparing its local reachability density

(LRD) to the average LRD of the points in $N_K(A)$:

$$LOF(A) = \frac{\sum\limits_{B \in N_K(A)} LRD(B)}{|N_K(A)| \cdot LRD(A)} \tag{C.9}$$

The LRD of A is calculated by

$$LRD(A) = \frac{|N_K(A)|}{\sum\limits_{B \in N_K(A)} RD_K(A,B)} \tag{C.10}$$

LRD can be considered the inverse of the average reachability distance of the points in the set of K nearest neighbors. Where the reachability distance of point A from point B ($RD_K(A,B)$) is the true distance between the points ($d(A,B)$), but at least the K-distance of point B ($Kd(B)$):

$$RD_K(A,B) = \max\{Kd(B), d(A,B)\} \tag{C.11}$$

Choosing K

Choosing the number of neighboring points (K) can largely affect the calculated LOF vales. Therefore, some rules are applied in choosing this value. In practice, The LOF values are calculated for a range of K values, and the maximum LOF of each point, for any K in the defined rage, is considered the true LOF value of the given point. The range of K values ($[K_{min} : K_{max}]$) is determined from a certain set of guidelines as proposed in the original paper [14]:

- K_{min} is the minimum size of a cluster to which points can be considered outliers
- K_{min} should be no less than 10
- K_{max} is the maximum number of points allowed in a cluster for the points to still be considered outliers

Interpreting the LOF Values

The LOF method is not a binary classification of points being inliers and outliers. Rather, it assigns each point with a degree of being an outlier. On one hand, this gives more information, in the sense, that outliers can be quantified in terms of how much of an outlier it is. On the other hand, it leaves some uncertainty in how to define the threshold of what should be considered inliers and what should be considered outliers. This threshold might be difficult to define and can vary from dataset to dataset.

In general, an LOF value of 1 or less is a good indication of a point being an inlier and points with significantly higher LOF values indicate outliers. Where the threshold for the outliers is located depends on the uniformity, sparseness, and fluctuations of the data.

C.5 RESULTS

Calculating the LOF values on the transformed dataset of Fig. C.5 using the described method gives the distribution of values shown in Fig. C.6. The majority of points are clear inliers with LOF values around 1. There are also some clear outliers with LOF values above 4. From the distribution of the LOF values, a good definition of the outlier threshold could be between 2 and 4 depending on the percentage of points that should be considered outliers. For a threshold of 2, 35 of the stacks would be identified as outliers (26%), whereas a threshold of 3 or 4 would give 13 (10%) and 10 (7%) outlier stacks, respectively. However, the LOF values might be more useful than a binary indication of outliers. With the LOF values, the highest values and therefore the most abnormal systems can be addressed first. The 14 stacks with the highest LOF scores are listed in Table C.2.

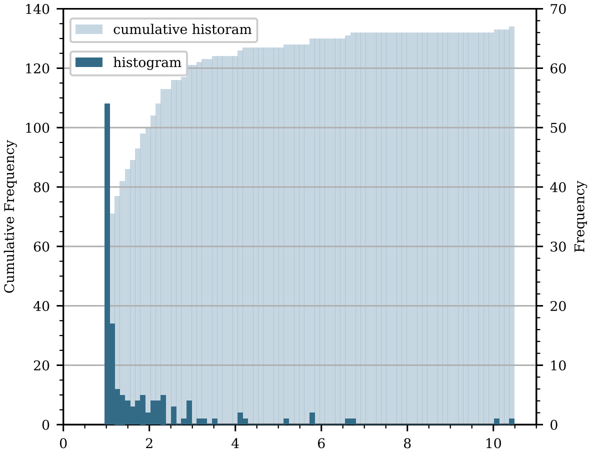


Fig. C.6: Histogram of the local outlier factors (LOF)

Table C.2: LOF scores of detected performance outliers

ID	Stack	LOF
82	2	10.5
82	1	10.1
8	1	6.7
8	2	6.6
45	2	5.7
45	1	5.7
27	1	5.1
28	2	4.2
11	2	4.1
27	2	4.1
25	1	3.5
48	2	3.2
48	1	3.2
16	2	2.9
9	2	2.9

Fig. C.7 shows a visualization of the calculated local outlier factors. The color of the data points shifts from blue for low outlier factors (inliers) to red for high outlier factors (outliers).

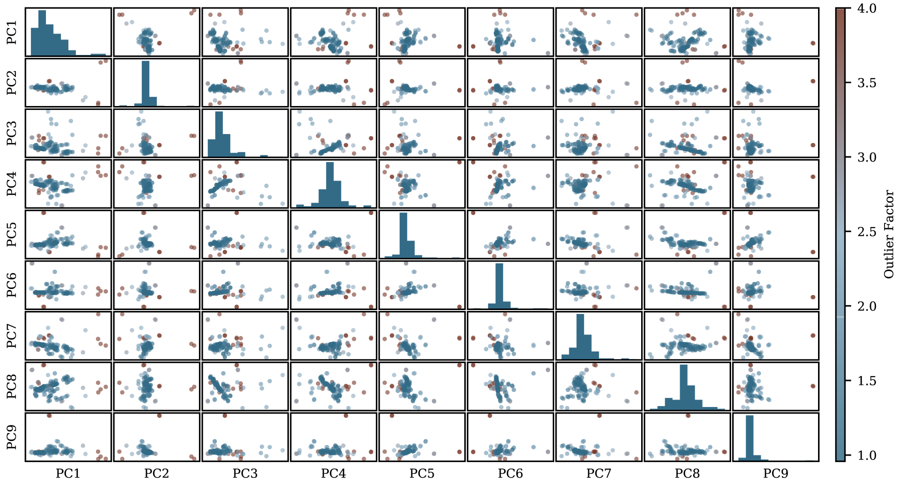


Fig. C.7: Scatter matrix of the nine principal components with indicated local outlier factor by the color scale. The color scale is limited to an outlier factor of 4, whereas the range of outlier factors extends to above 10. In Effect, all points with outlier factors above 4 are colored in the same red color.

From a maintenance perspective, the local outlier factors could indicate where to focus the maintenance effort. That is, stack 1 and 2 of system 82 have the least similarity of their performance indicators to the fleet of systems as a whole, i.e. they have the highest LOF values. Therefore, further investigation of system 82 should have high priority. After system 82 has been addressed, the maintenance effort can be shifted to systems with lower LOF, system 8, then system 64 and so forth.

Fig. C.8 depicts the outlier factors of the stacks in the original (normalized) KPI data space as well as boxplots showing the distribution of each of the KPIs. Stack 2 of system 82, which is the stack with the highest outlier factor (10.98), has been highlighted with a star marker in the figure. The figure shows that, this system is within the normal range for all the KPIs, except for the `Runtime_Powerup`, indicating that the system spends too much time in this mode compared to the time spend in other modes of operation. The figure also indicates that it is possible to have apparent outliers in one KPI without scoring a high LOF, meaning that the stack might have a higher than normal value in one KPI, but still be located within the main nine-dimensional cluster of all the principal components. On the other hand, it is also possible for a stack to be within the normal ranges for all the KPIs, but due to local fluctuations in density, to score a high LOF value.

System 82, stack 1 and 2, both only have abnormal values in the `Runtime_Powerup` KPI, whereas system 8 have abnormal values in several of the KPIs. System 48, stack 1 and 2, lie within the normal range for all the KPIs, yet scores a relatively high LOF of 3.2.

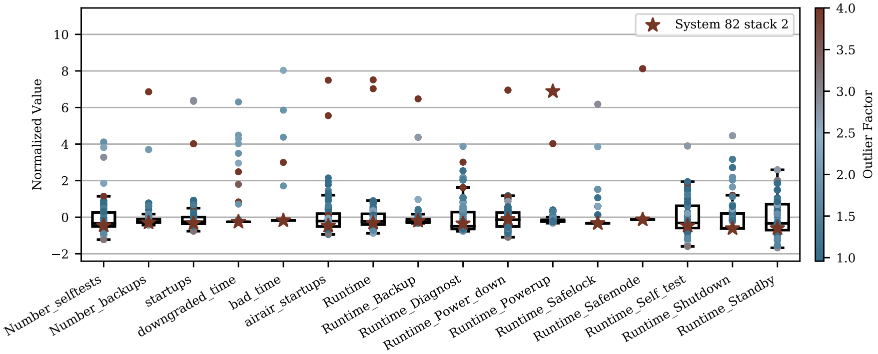


Fig. C.8: Distribution of key performance indicators (KPIs) with indicated outlier factors. Each tick on the x-axis indicate a KPI and the y-axis is the normalized value of each KPI. The box plots show the overall distributions of each KPI and the overlaying points each represent a specific stack. The stack data points are colored by the calculated outlier factor but limited to a score of 4 so that scores above 4 all have the same color.

C.6 CONCLUSION

This paper describes a method for detecting abnormalities in the performance of a fleet of fuel cell based backup systems operating in the field. The abnormality detection is based on dimensionality reduction using principal component analysis (PCA) and the local density-based machine learning method, local outlier factor (LOF).

The results show that performance anomalies can be detected across multiple performance indicators. The method uses information from all the variables in the dataset simultaneously to give a quantitative measure for the degree of a stack being an outlier. In this way, the most abnormally performing stacks can be investigated with highest priority.

If there are many stacks in the fleet, that perform unintentionally, but in the same way, the described method might not be able to detect such misbehavior. The stacks would have high similarity to each other and therefore high local density and would therefore score low LOF values. As such, they would not be detected as outliers, and other methods would be required to detect such unintentional behavior. However, this case is not present in the described dataset.

REFERENCES

[1] S. Ould Amrouche, D. Rekioua, T. Rekioua, and S. Bacha, “Overview of energy storage in renewable energy systems”, *International Journal of Hydrogen Energy*, vol. 41, no. 45, pp. 20 914–20 927, Dec. 2016.

[2] K. Fosberg, “Fuel cell systems provide backup power in telecom applications”, *Fuel Cells Bulletin*, vol. 2010, no. 12, pp. 12–14, Dec. 2010.

[3] Y. Zhan, Y. Guo, J. Zhu, and H. Wang, “Intelligent uninterruptible power supply system with back-up fuel cell/battery hybrid power source”, *Journal of Power Sources*, vol. 179, no. 2, pp. 745–753, May 2008.

[4] M. J. Vasallo, J. M. Andújar, C. García, and J. J. Brey, “A Methodology for Sizing Backup Fuel-Cell / Battery Hybrid Power Systems”, *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1964–1975, 2010.

[5] R. Kaur, V. Krishnasamy, K. Muthusamy, and P. Chinnamuthan, “A novel proton exchange membrane fuel cell based power conversion system for telecom supply with genetic algorithm assisted intelligent interfacing converter”, *Energy Conversion and Management*, vol. 136, pp. 173–183, Mar. 2017.

[6] M. F. Serincan, “Validation of hybridization methodologies of fuel cell backup power systems in real-world telecom applications”, *International Journal of Hydrogen Energy*, vol. 41, no. 42, pp. 19 129–19 140, Nov. 2016.

[7] ——, “Reliability considerations of a fuel cell backup power system for telecom applications”, *Journal of Power Sources*, vol. 309, pp. 66–75, 2016.

[8] G. Niu, B.-S. Yang, and M. Pecht, “Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance”, *Reliability Engineering & System Safety*, vol. 95, no. 7, pp. 786–796, Jul. 2010.

[9] S. D. Sønderskov, “Monitoring Performance Indicators of PEM Fuel Cell Backup Power Systems”, in *2017 IEEE International Telecommunications Energy Conference (INTELEC)*, Torino, Italy: IEEE, 2018.

[10] L. Li, “Dimension Reduction for High-Dimensional Data”, in *Methods in molecular biology (Clifton, N.J.)* Vol. 620, 2010, pp. 417–434.

[11] I. K. Fodor, “A Survey of Dimension Reduction Techniques”, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, Tech. Rep. 1, May 2002, pp. 1–18.

[12] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis”, *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37–52, Aug. 1987.

[13] A. Zimek and E. Schubert, “Outlier Detection”, in *Encyclopedia of Database Systems*, New York, NY: Springer New York, 2017, pp. 1–5.

[14] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, “LOF: Identifying Density-Based Local Outliers”, *Proceedings of the 2000 Acm Sigmod International Conference on Management of Data*, pp. 1–12, 2000.

Paper D

State of Health Estimation and Prediction of Fuel Cell Stacks in Backup Power Systems

Simon Dyhr Sønderskov, Jakob Ilsøe, Dean Rasmussen, Daniel Blom-Hansen, and Stig Munk-Nielsen

The paper has been submitted for publication in the
IEEE Transactions on Industrial Electronics.

Abstract

Fuel cell based backup systems are used as telecommunication power supplies, where availability of power is crucial for a reliable service. State of health (SOH) is a useful metric in aiding predictive maintenance actions, that aim to reduce the system down time as well as the operating cost. In this paper, voltage, current, and temperature data from numerous stacks installed in the field to estimate an SOH metric. A long short term memory (LSTM) recurrent neural network (RNN) is trained to predict SOH values six months into the future. Finally, the RNN performance is evaluated on prediction horizons of six months, as well as longer horizons of twelve, eighteen, and twenty-four months.

D.1 Introduction

Fuel cell based backup systems has become a popular solution in powering telecommunication sites. The fuel cell solution provides a less polluting and less noisy alternative to conventional diesel generators. They are more sustainable for longer runtimes and have easier extendable capacity compared to pure battery solutions. [1]

In backup systems, the availability of power is of paramount importance. Therefore, the system performance should be ensured through appropriate maintenance strategies, monitoring of system parameters, and other measures. Assessing the performance level of a complex system is not trivial and many approaches are described in literature. A popular metric for assessing stack performance is the stack voltage [2], [3]. However, the voltage does not only change with the degradation of the stack, but also with varying operating conditions, such as load and temperature. Therefore, this approach is often used in applications with constant load. Another popular method, especially used for diagnostics and characterization, is based on electrochemical impedance spectroscopy (EIS), which has proven very effective [4]–[6]. However, it requires perturbation of the stack current at a specific operating point and additional hardware for measuring the response.

An alternative method is to estimate a state of health (SOH) metric, i.e. a single number that indicates the condition of the stack. There are no established standards in the way that SoH is derived for a fuel cell. But a number of methods is presented in literature, including voltage based (as mentioned previously), and internal resistance. [7], [8]

In this paper, voltage, current, and temperature measurements from numerous fuel cell stacks installed in the field, are used to estimate the state of health of the stacks. A recurrent neural network (RNN) is trained on the data to predict the state of heath metric six months into the future. The main contributions of this work is a method for estimating SOH of fuel cell stacks in backup systems, where operation is sporadic. As well as a method for predicting future SOH values of individual stacks, based on known SOH metrics from a

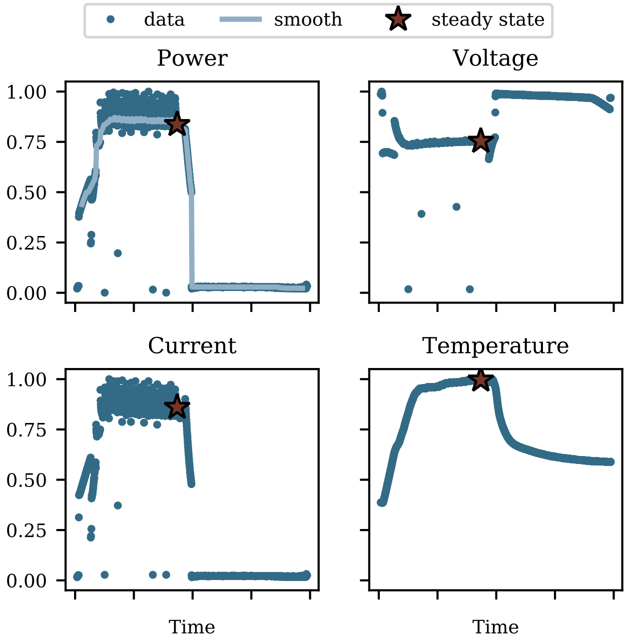


Fig. D.1: Self-test example showing the steady state extraction principle.

fleet of stacks installed in the field, which to the authors' knowledge, has not been seen in the literature before.

D.2 DATA FOUNDATION

Data from 74 stacks is used for the studies in this paper. The stacks are used in telecommunication uninterruptible power supplies, as described in [9]. These systems experience some unique operating conditions, involving long periods of inactive/standby modes and few, sporadic backup scenarios.

To ensure system functionality when backup is needed, the fuel cell stacks are regularly conditioned in so called self-test, where the stacks are powered up to deliver energy to a dummy load, thereby exercising the fuel cells. These self-tests are more regular and consistent, than the backup power scenarios, and are therefore more suited to study the fuel cell degradation over time.

The data, which is used in this work is voltage, current, and temperature of the fuel cell during the before mentioned self-tests. An example of such data from a single self-test of a single stack is shown in Fig. D.1. To simplify the dataset and to allow for easier analysis of performance levels, each self-test is reduced to a single data-point - the steady state value - for each of the three variables.

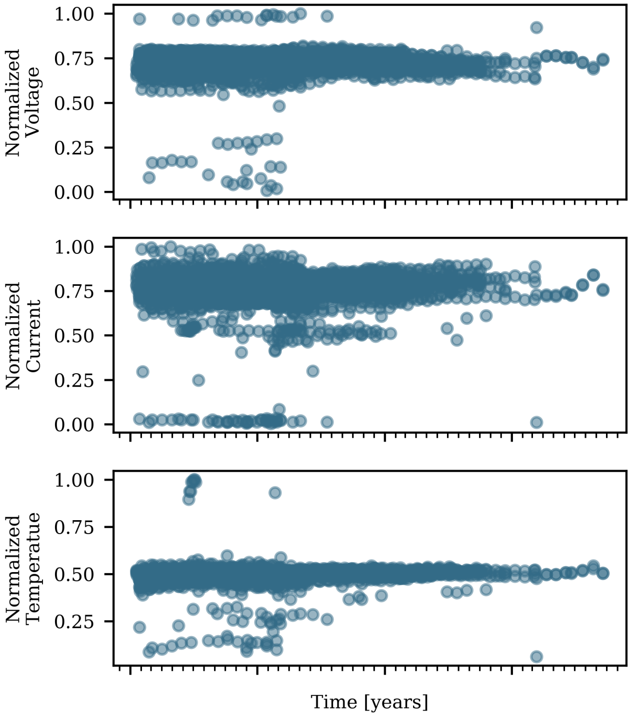


Fig. D.2: Steady state voltage, current, and temperature data for all the stacks in the dataset.

The steady state values are extracted by finding the end of the constant power segment of the data and then taking the variable values at this time-point. The steady state extraction process is described in detail in [10].

Finding these steady state values for each self-test of each stack in the dataset gives the cloud of data, shown in Fig. D.2. The figure does not depict to which stack each data point belongs, but merely illustrates the dataspace. An example of how the voltage, current and temperature steady state values evolve in time for a single stack is shown Fig. D.3.

D.3 State of Health Estimation

To estimate the state of health of any given stack, two reference points are needed; a beginning of life point (BOL; when the stacks is performing as a stack is expected to perform at the beginning of its lifetime), and an end of life point (EOL; when the stack performance has fallen to an unacceptable level).

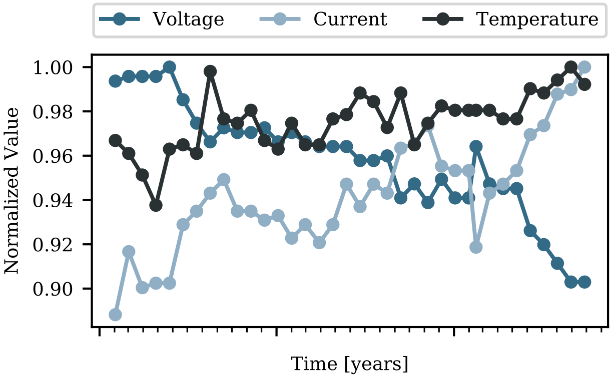


Fig. D.3: Steady state voltage, current, temperature evolution of an example stack.

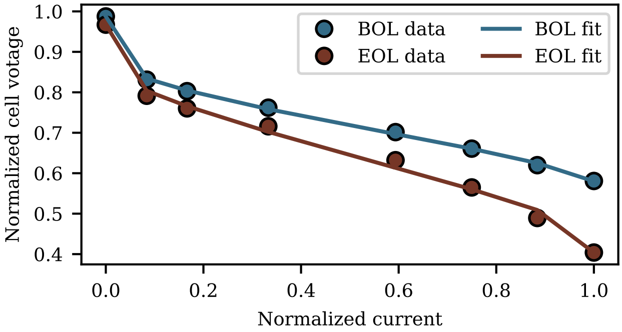


Fig. D.4: Beginning of life (BOL) and end of life (EOL) points as defined in the fuel cell manual and fitted curves.

D.3.1 BEGINNING OF LIFE AND END OF LIFE DEFINITIONS

The fuel cell manufacturer has provided their definitions of the beginning of life and end of life criteria in the datasheet of the fuel cell. These boundaries are given as polarization curves, i.e. voltage-current curves of the fuel cell, as shown in Fig. D.4. The polarization curves are obtained through a specific sequence of current steps in a lab environment with controlled fuel cell temperatures. Throughout the test, the temperature of the fuel cell is kept close to the optimum temperature for that specific current level, i.e. the temperature is lower in the low current steps than at the high current steps.

To allow for comparing values to the polarization curves at any current level, the formula in (D.1) [11] is fitted to the test data. The formula represents the fuel cell voltage as the reversible open circuit voltage minus the different types of voltage loss in the cell.

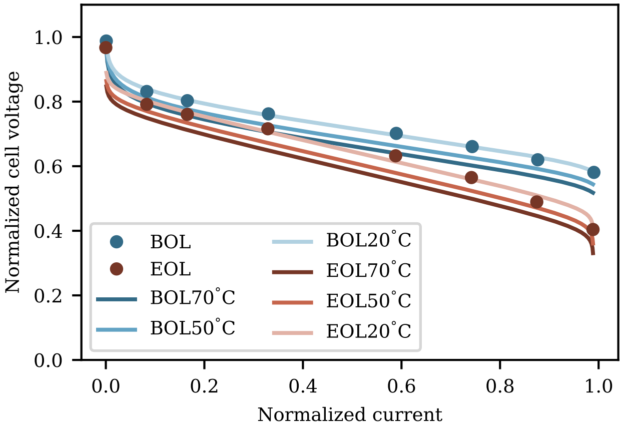


Fig. D.5: Fitted polarization curves for beginning of life and end of life criteria at different temperatures.

$$V_{\mathrm{c}} = V_{\mathrm{r}} - (I + I_{\mathrm{n}})R - \frac{\mathcal{R}T}{2\alpha\mathcal{F}} \ln\left(\frac{I + I_{\mathrm{n}}}{I_{\mathrm{o}}}\right) + \frac{\mathcal{R}T}{2\mathcal{F}} \ln\left(1 - \frac{I + I_{\mathrm{n}}}{I_{\mathrm{L}}}\right) \tag{D.1}$$

where I is the cell current, T is the temperature, $\mathcal{R}$ is the universal gas constant ($\approx 8.314\,\mathrm{J\,mol^{-1}\,K^{-1}}$), and $\mathcal{F}$ is the Faraday constant ($\approx 96\,485.332\,\mathrm{C\,mol^{-1}}$). The remaining parameters are the fitting parameters and are described in Table D.1. The obtained fit is shown for three temperature levels along the datasheet BOL and EOL definitions in Fig. D.5. The graphs show that the cell voltage falls slightly for an increased temperature. This is not usually the case for fuel cells and is mainly due to the fact that the parameters of (D.1), such as i_o and r, have temperature dependencies, which are not modeled here. However, the variation in voltage caused by temperature is relatively small, and the variation in temperature is limited (between 40°C and 60°C for the vast majority of the data points). Therefore, the presented fit is considered applicable for this specific case.

D.3.2 STATE OF HEALTH ESTIMATION

After obtaining an expression for the cell voltage as a function of current and temperature at the beginning of life ($V_{\mathrm{c,bol}}(I, T)$), and an expression at the end of life ($V_{\mathrm{c,eol}}(I, T)$), the state of health of the fuel cell stack is estimated using the formula:

$$SOH(V_{\mathrm{c}}, I, T) = \frac{V_{\mathrm{c}} - V_{\mathrm{c,eol}}(I, T)}{V_{\mathrm{c,bol}}(I, T) - V_{\mathrm{c,eol}}(I, T)} \tag{D.2}$$

Table D.1: Fuel cell polarization curve fitting parameters.

Symbol	Quantity	Unit
V_r	Reversible open circuit voltage	V
I_n	Sum current equivalent of fuel crossover and the internal current	A
I_o	Exchange current	A
I_L	Limiting current	A
R	Resistance	Ω
α	Charge transfer coefficient	-

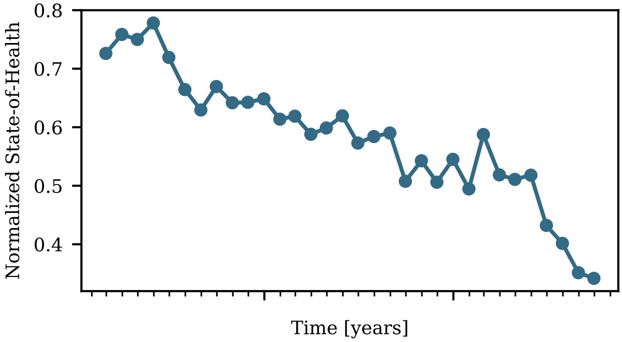


Fig. D.6: State-of-Health estimates of an example stack.

where V_c, I, and T are the self-test steady state data points of voltage, current, and temperature, respectively. The resulting state of health for an example stack is shown in Fig. D.6.

D.4 RECURRENT NEURAL NETWORKS

Recurrent neural networks (RNN) are a variation of artificial neural networks (ANN). The underlying equation in ANN is shown in (D.3), where y represents the output of a neuron and x is the input, W is a weight, and b is a bias. σ is an activation function, which adds nonlinearity to the equation. Popular activation functions include the hyperbolic tangent, Sigmoid, and rectified linear unit (ReLU).

$$y = \sigma\left(Wx + b\right) \tag{D.3}$$

Recurrent neural networks can be viewed as a repetition of multiple ANNs, where the outputs of the ANN layers are used as additional inputs to the

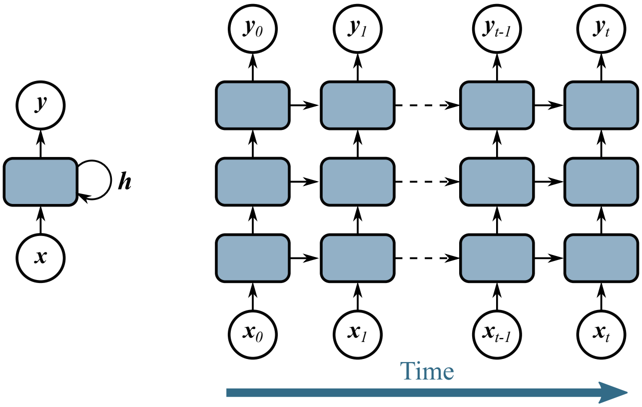


Fig. D.7: Left: single-layer recurrent neural network. Right: three-layer recurrent network unrolled through time. The output and possibly some hidden states are used as additional inputs in the next step of the sequential network.

following ANN layers. This is represented in Fig. D.7 and gives the model a temporal dimension, which makes this type of neural networks ideal for modeling sequential data, such as time-series. In the right hand side of Fig. D.7, a representation of an RNN with three layers is depicted as it is unrolled through time. So that, $\boldsymbol{x}_0$ is the input vector at $t = 0$ and $\boldsymbol{y}_0$ is the resulting output when the input is fed through the three layers. For the following time step, not only is $\boldsymbol{x}_1$ input to the first layer, but also the output of the first layer in the previous time step. This use of previous outputs and hidden states can be considered the models memory.

One shortcoming of the generic RNN as the one discussed above, is that in each time step the memory term is multiplied by the same weight, leading to the vanishing gradient problem, where the memory becomes less significant over time, resulting in improper training of the network. However, this problem is solved in variations of the RNN, such as long short term memory (LSTM) RNN networks [12].

D.4.1 Long Short Term Memory Cell

The long short term memory (LSTM) cell is a variation of the recurrent neural network. The LSTM cell has internal layers, referred to as gates, which allow the network to choose what should be remembered and what should be forgotten. A graphical representation of an LSTM cell is depicted in Fig. D.8. Additional to using the output of the previous layer in the input of the subsequent layer as in the basic RNN, an LSTM network also uses an internal cell state in the input of the following layer. The process and the mathematics of the LSTM cell is described in the following.

The LSTM network has three internal gates: the input gate, the forget gate,

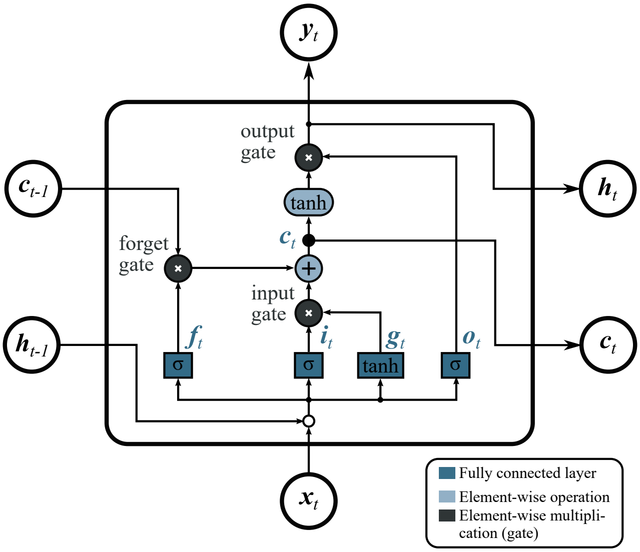


Fig. D.8: Long short term memory (LSTM) cell.

and the output gate. The function of the gates, is to select which parts of the input should be passed through to the next layer. The forget gate selects which part of the cell state from the previous step ($\boldsymbol{c}_{t-1}$) should be passed to the current cell state, and hence which part should be forgotten. The activation vector of the forget gate ($\boldsymbol{f}_t$) is calculated using artificial neuron function on the input ($\boldsymbol{x}_t$) and the output of the previous step ($\boldsymbol{h}_{t-1}$) as shown in (D.4).

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{\mathrm{f}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{f}}\right) \tag{D.4}$$

where $\boldsymbol{W}_f$ is the weight matrix, $\boldsymbol{b}_f$ is the bias, and σ is the activation function, in this case the sigmoid function.

The input gate has two inputs: the new candidate cell state ($\boldsymbol{g}_t$), and the input activation vector ($\boldsymbol{i}_t$). The input gate activation function is calculated similarly to the forget gate activation function:

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{\mathrm{i}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{i}}\right) \tag{D.5}$$

The candidate cell state is also calculated using the artificial neuron function, only the sigmoid activation function is replaced with an hyperbolic tangent (tanh) activation function:

$$\boldsymbol{g}_t = \tanh\left(\boldsymbol{W}_{\mathrm{g}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{g}}\right) \tag{D.6}$$

The actual cell state ($\boldsymbol{c}_t$) is obtained by adding the outputs of the input gate and the forget gate, i.e. the cell state is a combination of a selection of the previous cell state data and a selection of data from the input and foregoing output:

$$\boldsymbol{c}_t = \boldsymbol{f}_t \otimes c_{t-1} + \boldsymbol{i}_t \otimes \boldsymbol{g}_t \tag{D.7}$$

The final gate is the output gate, which determines which parts of the cell state should be output and fed to the next layer and/or next time step. The activation vector of the output gate ($\boldsymbol{o}_t$) is calculated by (D.8).

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_{\mathrm{o}}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_{\mathrm{o}}\right) \tag{D.8}$$

Finally, the output ($\boldsymbol{y}_t$ or $\boldsymbol{h}_t$) is calculated in the output gate:

$$\boldsymbol{y}_t = \boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh\left(\boldsymbol{c}_t\right) \tag{D.9}$$

The output is then used for two purposes: 1) as a final output of the RNN or as an input to a following layer in the model (called $\boldsymbol{y}_t$) and 2) as a hidden memory state to the following time step (called $\boldsymbol{h}_t$).

D.5 CONSTRUCTING AND TRAINING THE MODEL

The LSTM RNN model is constructed using the Python library Keras [13], which is a high-level API for neural network implementation. The network architecture is chosen to consist of three main LSTM layers with 64, 128, and 64 neurons, respectively. Each LSTM layer is put in a bidirectional wrapper, which effectively makes two LSTM layers: One is trained in normal time order, i.e. from past to future; the other is trained in the reverse time order, i.e. from future to past. This allows the model to use both past and future information during the training. This also doubles the size of the layer, so that a bidirectional LSTM layer, with each 64 neurons will amount to a total of 128 neurons. After the two first LSTM layers, a dropout layer is added. The dropout layer randomly zeros a fraction of the input units, thereby ignoring these inputs in the training process. This helps prevent overfitting the model. Finally a standard fully-connected (dense) neural network layer with six neurons is used for the output layer. The output of neuron 1 in the output layer represents the future value one time step forward, neuron 2 is the value two time steps in the future and so forth.

The implementation of these layers using Keras in Python is shown in Listing E.1. And the constructed model is summarized in Table D.2, showing that more than 460,000 parameters (weights and biases) needs to be estimated during the training process. The output shape column shows the shape of the output of the specific layer in the format (*batch_size*, *samples*, *units*). The *batch_size* refers to the number of stacks that is passed through the network at a time, *samples* refers to the number of samples per stack, and *units* is the number of units/neurons in the layer. The First two dimensions are listed as 'None', because these dimensions are of variable size, i.e. different stacks have different number of samples and the number of stacks per batch varies, as explained in Section D.5.1. The output shape of the output layer (Dense_1), only has two

Listing D.1: RNN model implementation

```python
from keras.models import Sequential
from keras.layers import Dense, LSTM, Dropout,
Bidirectional

# Initiate a sequential model
model = Sequential()
# Add first LSTM layer and dropout
model.add(Bidirectional(
    LSTM(64, return_sequences=True),
    input_shape=(None, 1)))
model.add(Dropout(0.2))
# Add second LSTM layer and dropout
model.add(Bidirectional(
    LSTM(128, activation='relu',
    return_sequences=True)))
model.add(Dropout(0.2))
# Add third LSTM layer
model.add(Bidirectional(
    LSTM(64, activation='relu')))
# Add output layer
model.add(Dense(6))

```

dimensions, namely (*batch_size*, *n_pred*), i.e. *n_pred* predictions are made for *batch_size* stacks. *n_pred* is determined by the number of units in the output layer.

D.5.1 PREPROCESSING

Before training the model, the dataset is split in three groups: train, validate, and test. The training data is used for training the model, i.e. tuning the weights and biases of the network so that outputs of the network match the expected based on the training input. The validation data is used to valuate the training process and is not used to update the network parameters. The testing data is kept completely separate from the training process. The training data sample contains 56 stacks (84%), the validation sample contains 9 stacks (12%) and the testing sample is 3 stacks (4%).

Secondly, the training data should be grouped in batches, depending on the number of samples per stack. The number of samples for each stack varies between 5 and 39, with most stacks (10) having 15 samples. Consequently, 24 batches are created, containing the data of between 1 and 10 stacks, each.

The final step before actually training the model is to normalize the data. For this purpose, the MinMaxScaler of the Scikit Learn package [14] in Python

Table D.2: Recurrent neural network model summary.

Layer (type)	**Output Shape**	**Parameters**
LSTM_1 (Bidirectional LSTM)	(None, None, 128)	33 792
Dropout_1 (Dropout)	(None, None, 128)	0
LSTM_2 (Bidirectional LSTM)	(None, None, 256)	263 168
Dropout_2 (Dropout)	(None, None, 256)	0
LSTM_3 (Bidirectional LSTM)	(None, 128)	164 352
Dense_1 (Dense)	(None, 6)	774
Total parameters:		462 086

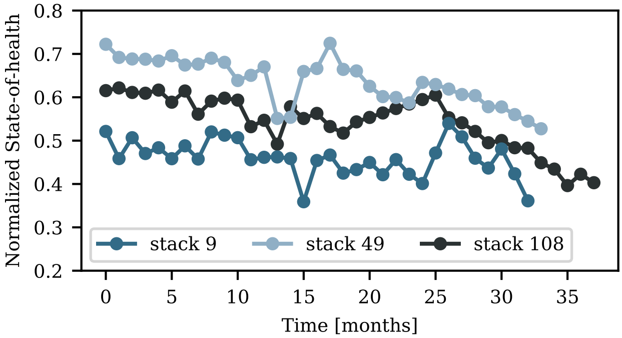


Fig. D.9: Testing data; three randomly selected stacks out of the total set of 74 stacks.

is used, which scales the data to fit between two defined values, in this case, 0 and 1. The normalized data of the three testing stacks is depicted in Fig. D.9.

D.5.2 TRAINING

The last six samples of each stack is used as labels ($\boldsymbol{y}$), i.e. the reference for the prediction, and the remaining samples are used as inputs ($\boldsymbol{X}$):

```
X = batch[:, :-6, :]
y = batch[:, -6:, :]
```

where `batch` is a batch of stacks with the dimensions ($batch\_size$, $samples$, $features$).

The network is trained on the training data for 200 epochs, using the Adam optimizer [15] and mean absolute error (MAE) as the loss function. After each training epoch, the model is used to make predictions on the validations data. The result of the loss function (MAE) for both training and validation data after each epoch is depicted in Fig. D.10 along with the mean-squared-errors

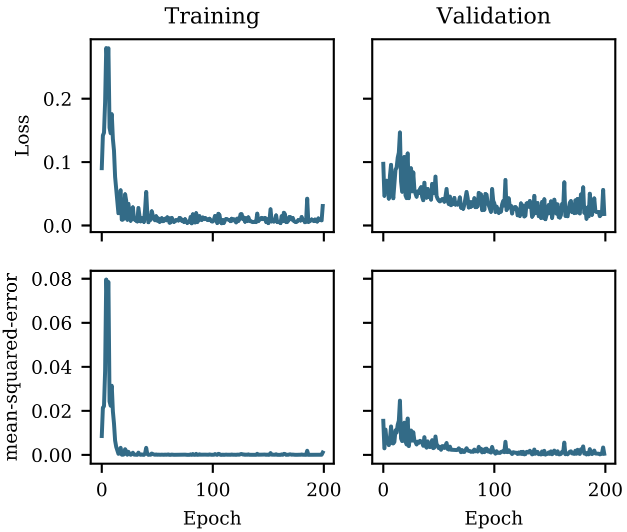


Fig. D.10: Training (left) and validation (right) accuracy (top) and loss (bottom) evolution during the training epochs.

(MSE). After 200 epochs, the training loss has dropped to approximately 0.008 and the validation loss to approximately 0.018. The MSE is used as a secondary metric for the accuracy of the training, which is independent from the MAE used for optimization. After 200 epochs the training MSE is 0.0008 and the validation MSE is 0.0005.

D.6 Model Evaluation

The trained RNN is used to predict the last six samples of each of the validation stacks and the result is shown in Fig. D.11. Points at and below zero months (on the x-axis) represent the data used for input to the model, i.e. considered past data. Points above zero months are considered future values, either true values or predictions.

As a baseline for measuring the performance of the RNN predictions, a naive prediction is made. The naive prediction is that the final data point in the input data is repeated for the future values. Hence, the naive estimate assumes that the SOH at any point in the future ($t > 0$) is the same value as the SOH at time zero ($t = 0$).

Looking at the predictions in Fig. D.11, it seems that the RNN is capable of capturing the overall trends as well as the magnitude of the series. For stacks 13, 24, 70, 79, 98, 106, and 112 the predictions fit very well. Whereas for stacks 84 and 96, the RNN struggles a little to make an accurate prediction.

To compare the performance of the RNN to the naive predictions, the

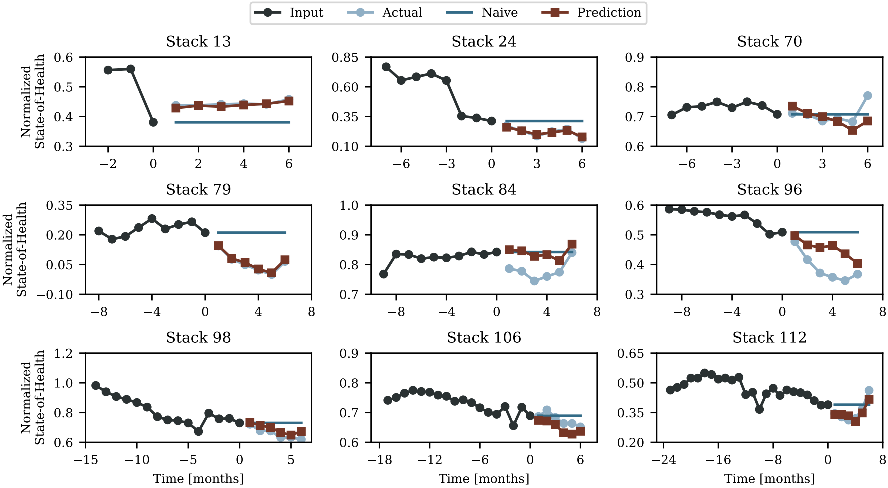


Fig. D.11: Prediction results on the validation data. The x-axis zero-point denotes the time of the prediction, such that any negative x-values are past values and are used as input for the recurrent neural network, and positive x-values are future values, which should be predicted by the outputs of the network. The graph shows the input data, the actual future values, the naive predictions, and the RNN predictions.

absolute errors of both to the true values are computed. The means of the stack prediction errors of the two methods are depicted in Fig. D.12 with the 95% confidence interval indicated by the error bars. For the first prediction, the mean error for the naive and RNN predictions are similar at 0.035 and 0.017, respectively. For the following predictions, however, the RNN performs significantly better at around 30%-40% the mean error of the naive predictions. At the six month prediction, the RNN prediction error is 0.032 - 64% lower than the 0.088 error of the naive prediction. It should be noted, that there is some overlap of the confidence intervals.

D.7 Long-Term State of Health Prediction

In this section, the trained RNN model will be used to make long-term predictions on the testing data. The RNN model is constructed to make monthly predictions six months into the future. To achieve longer prediction horizons, the six months predictions are appended to the original input and used as a new input to the RNN which again predicts six months into the future, making a total of twelve predictions from the starting point. This process can be continued indefinitely. Predictions are made using this approach for each of the test stacks at different prediction horizons.

The number of samples of the test stacks are 33, 34, and 38, respectively. Therefore, the maximum prediction horizon, which can be verified with the

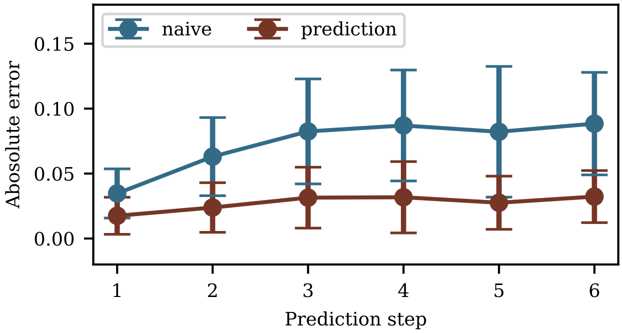


Fig. D.12: Prediction error on the validation data of the naive baseline prediction and the recurrent neural network prediction. The error bars show the 95% confidence interval.

sample data, is 30, i.e. five predictions of six months each.

Firstly, a prediction horizon of six months is calculated using the first 27, 28, and 32 samples for the input, leaving six samples for verification for each of the stacks. The results are shown in the left column of Fig. D.13.

For prediction the SOH twelve months into the future, the first 21, 22, and 26 samples, respectively are used for the initial input to the RNN, leaving twelve samples for verification. The RNN predicts the SOH six months into the future. These six predictions are then appended to the initial input data, so that the new input to the RNN is the initial input samples as well as the first six predictions. The RNN then computes six new predictions, so that twelve predictions are made in total. The result is shown in the center column of Fig. D.13.

The same approach is used for prediction horizons of 18, 24, and 30 months, i.e. three, four, and five iterations of RNN predictions, respectively. The result of the 30 month prediction horizon is shown in the right column of Fig. D.13.

The absolute error at the end of each prediction horizon for each stack is calculated. For stack 9, the six month and twelve month prediction errors are very low, at around 0.069 and 0.068, respectively. The eighteen, twenty-four, and thirty months error all lie steadily between 0.225 and 0.384. The prediction error in stack 49 starts at 0.113 for the six month horizon. For the longer horizons, the error lies between 0.307 and 0.393. The prediction error of stack 108 has a bit more variation than the other two stacks. At six months the error is 0.166 and at twelve months it rises to 0.258 before falling again at eighteen months to 0.113. It increases again for the twenty-four and thirty month horizons.

The prediction errors of each stack for each prediction horizon are summarized in Table D.3 along with the mean prediction error for each of the prediction horizons.

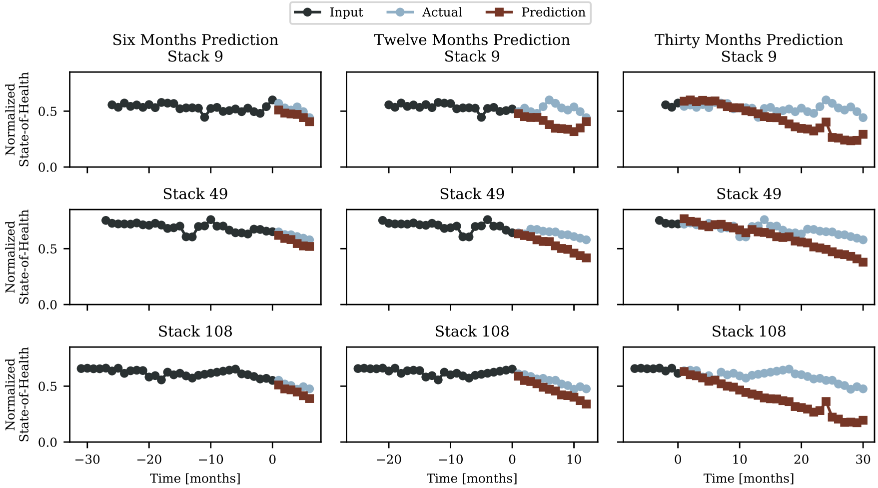


Fig. D.13: Prediction results for the test stacks with a six, twelve, and thirty month prediction horizon, from left to right, respectively.

Table D.3: Prediction errors for test stacks for prediction horizons from six to thirty months.

Prediction horizon	Stack 9	Stack 49	Stack 108	Mean
6	0.069	0.113	0.166	0.116
12	0.068	0.307	0.258	0.211
18	0.225	0.393	0.113	0.243
24	0.384	0.382	0.420	0.395
30	0.284	0.381	0.535	0.400
Mean	0.206	0.315	0.298	0.273

D.8 CONCLUSION

Periodic self-test data has been used to estimate state of health (SOH) values for fuel cell stacks. The data contained voltage, current, and temperature measurements from fuel cell stacks operating in field conditions.

A long short term memory (LSTM) recurrent neural network (RNN) was trained on the SOH data to predict values six months into the future. The RNN was trained for 200 epochs where the mean absolute error loss function was reduced to approximately 0.02 on validation data, with a mean squared error of approximately 0.0005.

The prediction accuracy on the validation data was compared to a naive prediction, where the last SOH value was simply forecast in a repetitive manner. The RNN method showed significant improvement compared to the naive method, with some uncertainty.

Test data was used to test the RNN predictions at longer predictions horizons, by iteratively using RNN predictions in the input of the same RNN. The predictions were made from six to thirty months, in intervals of six months. The error remained relatively low for the longer prediction horizons, although the results vary from stack to stack.

REFERENCES

[1] M. J. Vasallo, J. M. Andújar, C. García, and J. J. Brey, “A Methodology for Sizing Backup Fuel-Cell / Battery Hybrid Power Systems”, *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1964–1975, 2010.

[2] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, “Remaining Useful Life Prediction and Uncertainty Quantification of Proton Exchange Membrane Fuel Cell Under Variable Load”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2569–2577, Apr. 2016.

[3] R. Ma, E. Breaz, C. Liu, H. Bai, P. Briois, and F. Gao, “Data-driven Prognostics for PEM Fuel Cell Degradation by Long Short-term Memory Network”, in *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, IEEE, Jun. 2018, pp. 102–107.

[4] T. Kim, H. Oh, H. Kim, and B. D. Youn, “An Online-Applicable Model for Predicting Health Degradation of PEM Fuel Cells with Root Cause Analysis”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7094–7103, 2016.

[5] C. Jeppesen, S. S. Araya, S. L. Sahlin, S. Thomas, S. J. Andreasen, and S. K. Kær, “Fault detection and isolation of high temperature proton exchange membrane fuel cell stack under the influence of degradation”, *Journal of Power Sources*, vol. 359, pp. 37–47, 2017.

[6] N. Fouquet, C. Doulet, C. Nouillant, G. Dauphin-Tanguy, and B. Ould-Bouamama, “Model based PEM fuel cell state-of-health monitoring via ac impedance measurements”, *Journal of Power Sources*, vol. 159, no. 2, pp. 905–913, 2006.

[7] T. Sutharssan, D. Montalvao, Y. K. Chen, W.-C. Wang, C. Pisac, and H. Elemara, “A review on prognostics and health monitoring of proton exchange membrane fuel cell”, *Renewable and Sustainable Energy Reviews*, vol. 75, no. November 2015, pp. 440–450, Aug. 2017.

[8] B. Dolenc, P. Boškoski, M. Stepančič, A. Pohjoranta, and Đ. Juričić, “State of health estimation and remaining useful life prediction of solid oxide fuel cell stack”, *Energy Conversion and Management*, vol. 148, pp. 993–1002, Sep. 2017.

[9] S. D. Sonderskov, M. J. Swierczynski, and S. Munk-Nielsen, “Lifetime prognostics of hybrid backup power system: State-of-the-art”, in *2017 IEEE International Telecommunications Energy Conference (INTELEC)*, vol. 2017-Octob, IEEE, Oct. 2017, pp. 574–581.

[10] S. D. Sønderskov, “Monitoring Performance Indicators of PEM Fuel Cell Backup Power Systems”, in *2017 IEEE International Telecommunications Energy Conference (INTELEC)*, Torino, Italy: IEEE, 2018.

[11] A. L. Dicks and D. A. J. Rand, *Fuel Cell Systems Explained.* Chichester, UK: John Wiley & Sons, Ltd, Apr. 2018.

[12] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory”, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[13] F. Chollet *et al.*, *Keras*, \url{https://keras.io}, 2015.

[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine Learning in {P}ython”, *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[15] D. Kingma and J. Ba, “Adam: a method for stochastic optimization (2014)”, *arXiv preprint arXiv:1412.6980*, vol. 15, 2015.

Paper E

Predicting Performance Indicators of Fuel Cell Stacks in Backup Power Systems

Simon Dyhr Sønderskov, Jakob Ilsøe, Dean Rasmussen, Daniel Blom-Hansen, and Stig Munk-Nielsen

The paper has been submitted for publication in the
Elsevier Journal of Power Sources.

The layout has been revised.

ABSTRACT

Fuel cell technology is a green alternative to diesel generators in backup power applications that require prolonged backup times. However, fuel cells have high initial cost and therefore rely on reduced operation and maintenance costs to compete commercially. Prognostics is a key enabler of reducing maintenance cost and improving life expectancy. The approach in this work takes offset in a dataset consisting of measured parameters of numerous backup power systems operating in field conditions. Recurrent neural network (RNN)models are trained to predict future evolution of health indicating fuel cell parameters from the historic data. Two approaches are compared. The first uses all available training data to train the RNN. The second separates the dataset into two clusters based on similarities between the stack time series and train separate RNNs for each cluster. It is found that the model that utilizes all the available training data, generally experiences superior performance.

E.1 INTRODUCTION

With many critical infrastructures relying on continuous electricity supply, there is an increased demand for backup power systems and uninterruptible power supplies (UPS). These critical infrastructures may include anything from hospitals and data centers to industrial process where the cost of lost production is high. In this work, the focus is on the telecommunications sector, where critical communication infrastructure need backup power systems in the range of 1-10 kW.

Traditionally, two main technologies have dominated the backup power systems: lead-acid batteries and diesel generators [1]. Although lead-acid batteries are inexpensive and simple to operate, they suffer from short lifetimes and need regular replacement. Also the weight (and size) of batteries is proportional to the energy capacity and therefore available backup time, which makes then unsuitable for applications that require long backup times. Diesel generators, on the other hand, are well suited for prolonged backup times, but require high maintenance due to the many moving parts. Perhaps the biggest drawback of diesel generators, in a modern world with increased focus on emissions and renewable power sources, is the pollution and emissions they produce. Also, diesel generators are noisy, which makes them inappropriate in urban areas. [1]–[3]

Fuel cells, like diesel generators, have decoupled power production and energy storage, meaning that extended backup power periods are possible [1], [3]. Furthermore they are quiet and do not pollute, at least on-site. All this have made fuel cells a promising new technology in the backup power system/UPS field [4], [5], as one of the first fields where fuel cells are commercially viable [3]. However, the initial cost of the fuel cell systems remain a hurdle to wide spread adaptation. Hence, in order for fuel cell systems to be commercially viable, they should compensate for the high initial cost, for example by reduced operations and maintenance cost.

Prognostics and health management (PHM) is a popular framework for extending system lifetime and is a key enabler of predictive maintenance [6]. PHM has received increased attention in the fuel cell research community in recent years [7]–[9]. PHM has three main stages: observation, analysis, and action. Observation involves the acquisition and processing of data from the system. The analysis stages includes condition assessment of the system (some times referred to by the state-of-health term; SOH), diagnostics, and prognostics. Diagnostics is the process of detecting and locating specific faults in the system and prognostics focuses on predicting the future health of the system. Prognostics is therefore a crucial part of predictive maintenance, where actions are planned according to the projected future condition of the system to minimize cost and/or down-time [10]. The maintenance planning and execution, as well as possible changes in the operating strategy of the system is done in the action phase of PHM. In this work, the focus is on the prognostics part of the analysis step of the PHM framework.

Prognostics of fuel cell systems is a difficult task, as the degradation mechanisms involve both chemical, thermal, and mechanical processes [11]. Two main approaches are used in literature: model-based and data-driven. In the model-based approaches, known physical or empirical relationships between measurable system parameters and the degradation process are used to estimate the level of degradation. In [12], the current density and stack voltage measurements are used to estimate the SOH and dynamic degradation with an extended Kalman filter (EKF). An inverse first-order reliability model (IFORM) is then used to forecast the SOH to an end-of-life (EOL) threshold, which is then used to give an estimate for the remaining useful life (RUL). In [13], the active area resistance (ASR) is estimated using an unscented Kalman filter (UKF) and used as a health indicator. A linear Kalman filter and Monte Carlo simulation is used to predict future ASR values, which are compared to a threshold to get an RUL estimate. Although these studies show great results, their model-based approach, limits their use to the specific operating conditions and degradation mechanisms included in the models [14].

Data-driven approaches are mostly limited by the available training data. As such, they can predict any degradation mechanism in any operating condition, provided that there are sufficient historical examples of these events available. Li et al. have demonstrated the use of linear parameter varying (LPV) models to estimate a virtual steady state voltage, which is used as a health indicating feature. Predictions of future feature values are obtained through the use of eco state networks (ESN) [14], [15]. In [8], an adaptive neuro-fuzzy interface system (ANFIS) is used to predict the output voltage under constant operation. In [16], stack current, power, and ageing time are used to predict degradation under dynamic operation using a summation-wavelet extreme learning machine (SW-ELM). Finally, Ma et al. [17], [18] use recurrent neural network (RNN) variants to predict output voltage under constant and dynamic operation.

Common for the prognostics studies of fuel cells, presented in literature, is that they deal with continuous operating data, although some start-stop events may occur throughout the test-time. This is fundamentally different to how the backup power systems under investigation in this study, normally operate, which will be described in Section E.2. Furthermore, the studies in literature often conduct experiments in a controlled lab environment, which is great for testing methods and understanding the processes. However, the conditions and mechanisms experienced in a real-field scenario might not be represented.

In this paper, data collected on a number of fuel cell-based backup power systems installed in real field environments will be presented. The data will be the basis of training a recurrent neural network to predict stack voltage, current, and temperature six months ahead in time. A comparison is made on training the RNN on all the training data and training two separate RNNs: one on each of the clusters detected in the dataset. The specific methods include long short-term memory (LSTM) RNN in an encoder-decoder architecture, dynamic time warping (DTW) for describing similarities between the time series, principal component analysis (PCA) for dimensionality reduction, and density-based spatial clustering of applications with noise (DBSCAN).

E.2 Data

The presented data was collected on backup power systems operating in field conditions, i.e. different and varying load demands, usage profiles, and ambient conditions. The systems normally reside in a standby mode, where the fuel cells produce no power. Occasional grid fault events cause the systems to activate to provide backup power to the load. However, these events are rare and inconsistent, which makes them unsuitable as basis for prognostics. However, the systems are routinely exercised in so-called self-tests, to ensure availability. These self-tests are consistent and more frequent than the backup events. Therefore, these self-tests are the basis for the prognostics approach in this work.

Data from one self-test of one fuel cell stack is presented in Fig. E.1. The presented data contains the stack power, stack voltage, current, room/ambient temperature of the fuel cell, temperature of the air supply at the inlet and outlet, pressure of the hydrogen supply, and the stack temperature. To enable comparison of self-test data across stacks, each self-test event is reduced to a single data point per variable. This is done by finding the end of the steady state power and extracting the variable values at that time instance. The median of ten data points before the detected time is used to mitigate noise in the signals.

The resulting data points for all self-tests of all the fuel cell stacks are shown in Fig. E.2 along with histograms of the data distributions. The month of the data points are added as an additional variable. This will provide the RNN with some seasonal information, which might be associated with the environment of the systems. The data has been further normalized to vary between 0 and 1 and has been resampled to monthly values.

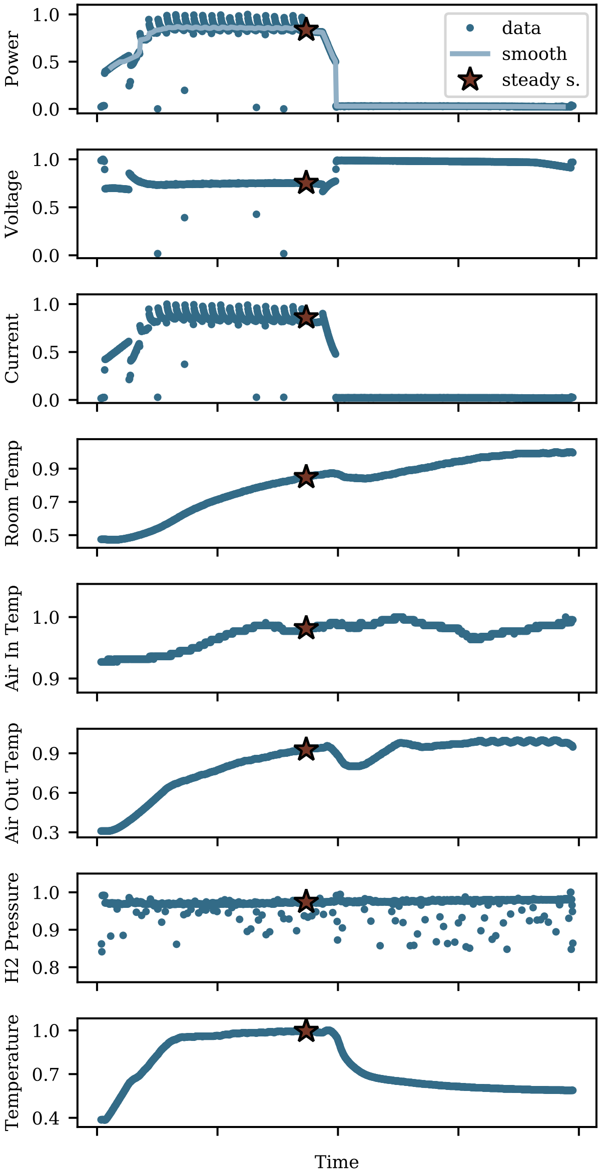


Fig. E.1: Extraction of steady state self-test values from raw data of a single stack. The values are extracted at the point where the stack power (top graph) is at the end of its steady state period. To account for noise in the data, the median of ten points before the steady state point is used.

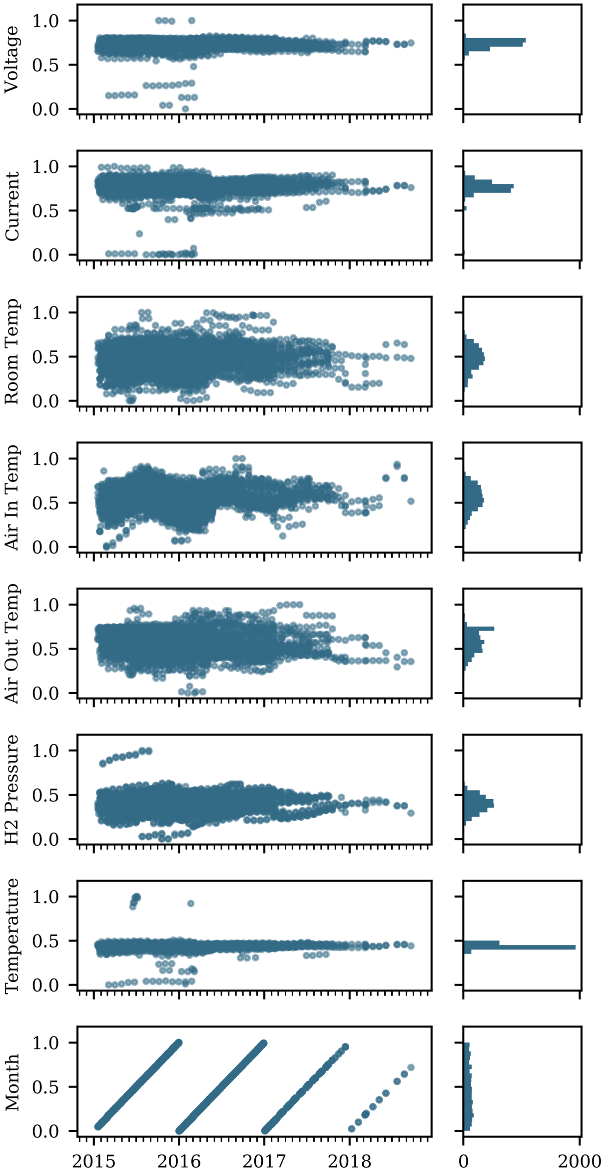


Fig. E.2: Steady state self-test data for all stacks in the dataset. The right hand side column of graphs show histograms of the data distributions.

E.3 Clustering

The stacks in the dataset experience different operating conditions and have different evolutions of their system parameters. The evolution might be similar for some subset of the stacks, while different for other stacks. These difference might separate the stacks in distinct groups of similar behavior, which can be detected through clustering. Clustering is a machine learning methodology of dividing a dataset into groups of data points with high degree of similarity. In this case grouping the stacks based on the similarities in the time series described in the previous section.

However, it is difficult to apply clustering methods on the time series directly. Therefore, a similarity measure between the various time series should be applied to obtain a representation of the stacks that are more easily compared. A simple way of measuring similarity of two time series is to calculate the Euclidian distance between each data point in the series. But this approach is only well suited to series of same length, that experience the same trends at the same time instances, which is far from the case of the dataset in question.

E.3.1 Dynamic Time Warping

Dynamic time warping (DTW) is a distance measure, capable of extracting similarities between asynchronous time series of varying length [19]. It seeks to optimize the alignment of the two series by nonlinear warping of the time dimension before calculating the distances between the points in the series. The DTW algorithm is commonly used in fields such as speech recognition [20] and data mining [21], but has also found applications in health prognostics of engineering systems, such as batteries [22].

Suppose two time series $\boldsymbol{Q}$ and $\boldsymbol{C}$ of length n and m, respectively. To find the DTW alignment, we first construct an $n \times m$ matrix where the entry (i, j) is the distance between the i^{th} and j^{th} element of $\boldsymbol{Q}$ and $\boldsymbol{C}$, i.e. q_i and c_j. The algorithm then seeks to find the path through the matrix, which minimizes the cost of the path, i.e. the sum of distances tracked by the path. Some constraints are defined when finding this path: the path starts at the beginning of both the series, and ends at the ends of both series - the warped series are aligned both at the beginning and the end; every point in each series should be matched to at least one point in the other series; the path must be monotonically increasing. [19]

Calculating the DTW distances between the times series of each of the stacks in the dataset, gives a measure of similarity between each stack and every other stack in the dataset. The result is visualized in Fig. E.3. A low DTW distance represents high similarity - see the diagonal in the matrix plot, that shows that each stack has zero DTW distance to it self. On the other hand high DTW distance implies low similarity.

These similarities will be the basis of the clustering analysis in section E.3.3. But firstly, the high dimensionality of this dataset (one dimension per stack) is

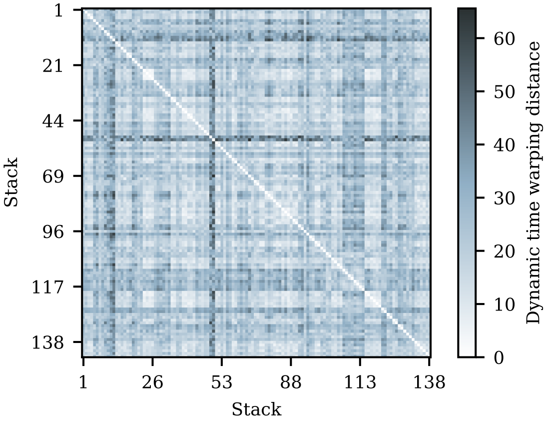


Fig. E.3: Matrix plot of the dynamic time warping distances.

reduced in the following section. The reason for dimensionality reduction is to enable easier inspection and visualization of the dataset.

E.3.2 Principal Component Analysis

Principal component analysis (PCA) is a dimensionality reduction method, which transforms a dataset into a space in which the variables are linearly uncorrelated. The transformed variables are referred to as principal components, which are arranged so that the first principal component has the largest possible variance, the second component has the second largest variance and so forth. That is, the first principal components contain the most information from the original dataset. Hence, by selecting a subset of the first principal components, the dimensionality of the dataset can be reduced while maintaining as much variance, and thereby information, as possible.

The PCA algorithm involves calculating the covariance matrix ($\boldsymbol{\Sigma}$) of the normalized dataset ($\boldsymbol{X}$):

$$\boldsymbol{\Sigma} = \frac{1}{n-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} \tag{E.1}$$

where n is the number of samples per feature.

The eigenvalues (λ) and eigenvectors ($\boldsymbol{\nu}$) of the covariance matrix are obtained by solving the equation

$$\boldsymbol{N}^{-1}\boldsymbol{\Sigma}\boldsymbol{N} = \boldsymbol{\Lambda} \tag{E.2}$$

where $\boldsymbol{N}$ is a matrix comprised of the eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues.

The transformation matrix ($\boldsymbol{M}$) is then comprised of the first L eigenvectors of $\boldsymbol{N}$. L can be determined by choosing an acceptable level of explained variance.

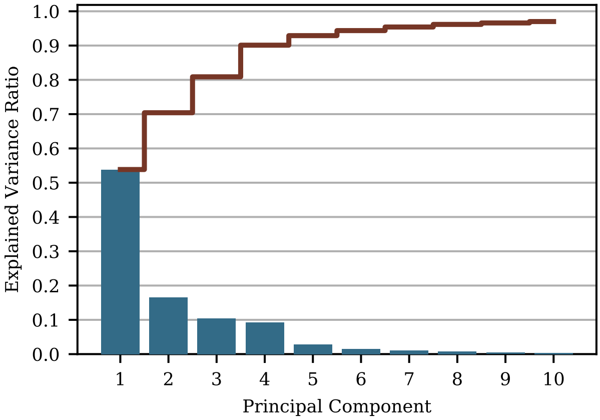


Fig. E.4: Explained variance (information) of the first ten principal components of the dynamic time warping distances.

This can be done through the proportion of variance (PoV), which is calculated from the eigenvalues as:

$$PoV = \frac{\sum_{j=1}^{L} \lambda_j}{\sum_{j=1}^{p} \lambda_j} \tag{E.3}$$

where p is the total number of principal components. The transformed dataset is finally calculated by

$$\boldsymbol{T} = \boldsymbol{M}\boldsymbol{X} \tag{E.4}$$

Applying PCA on the DTW matrix, as is visualized in Fig. E.3, the proportion of explained variance of the first ten principal components are shown in Fig. E.4. The figure shows that the first principal component explains more than half of the variance of the original DTW matrix. The second component explains roughly 15%. Four principal components are included in the transformation matrix, which will represent just over 90% of the variance of the original data, while reducing the dimensionality considerably.

E.3.3 DBSCAN

To detect groups within the transformed dataset, obtained through principal component analysis on the dynamic time warping distance matrix, the DBSCAN (density-based spatial clustering of applications with noise) algorithm is used [23]. DBSCAN looks at the number of neighbors to a given point within a certain distance (ϵ) to determine if the point is part of a cluster or a noise point. The second parameter of the model, $MinPts$ is the minimum number of

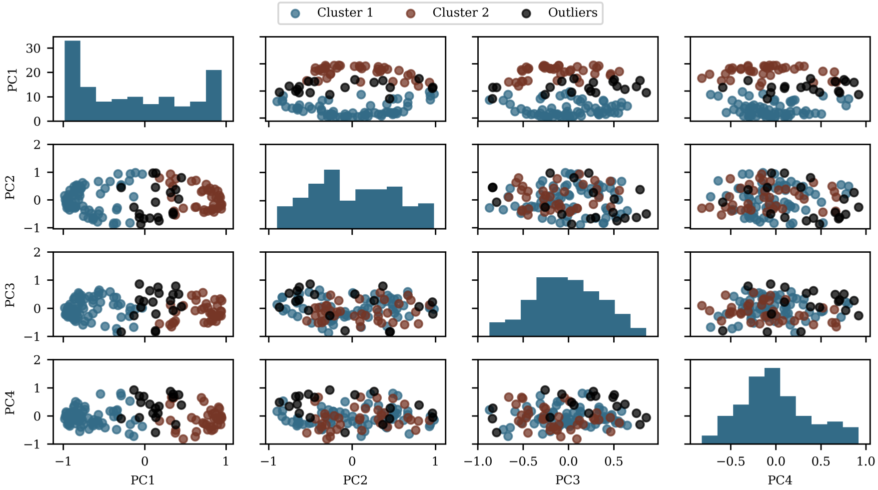


Fig. E.5: Scatter matrix of the transformed and normalized dynamic time warping distances. The clusters, as found by the K-Means method, are indicated with the different colors. The diagonal plots show the histogram of the given principal component. Each tick on the histogram y-axes represent ten points.

points that should be reachable from the point in question, for this point to be considered as a member of a cluster.

The DBSCAN algorithm classifies points in three categories: core points, border points, and noise points. A given point, p, is considered a core point if at least $MinPts$ points are within ϵ distance of p. If p is not a core point, but there is another core point within ϵ distance of p, it is labeled as a border point. Noise points are neither core nor border points and represent outliers in the dataset. Each core and border point are assigned to a cluster by exhaustively searching each unlabeled point. Any point within ϵ distance of a core point is appended to the same cluster.

The advantages of DBSCAN are that the algorithm can detect any number of clusters of any shape, as well as detect and ignore outliers. The drawbacks of the algorithm is mainly its sensitivity to the selection of the parameters ϵ and $MinPts$.

Applying the DBSCAN algorithm to the four-dimensional PCA transformed dataset of DTW distances between fuel cell stacks, gives the clusters as shown in Fig. E.5. The parameters of the algorithm were set to $MinPts = 25$ and $\epsilon = 0.85$. Two clusters were found: cluster 1 consisting of 67 sacks; and cluster 2 of 33 stacks. Also 26 noise points were detected.

The two clusters are mainly separated in one of the four dimensions, i.e. the first principal component, and are more intertwined in the other three dimensions.

E.4 RECURRENT NEURAL NETWORK

The recurrent neural network (RNN) is a popular variation of the artificial neural network (ANN), which is well suited for modeling temporal data, such as time series. This temporal property comes from using a sequence of neural networks where the internal state (memory) is fed to the subsequent neural network. Long short term memory (LSTM) cells are a variation of classic RNN models, which overcomes some of the RNN drawbacks such as the vanishing gradient problem [24]. LSTM RNN is a very popular framework within deep learning for modeling sequential data.

The basic equation of a neural network is

$$\boldsymbol{y} = \sigma\left(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b}\right) \tag{E.5}$$

where the output ($\boldsymbol{y}$) is calculated by scaling the input ($\boldsymbol{x}$) with the weights matrix ($\boldsymbol{W}$) and offsetting it by the bias vector ($\boldsymbol{b}$) and finally applying the activation function (σ). It is the activation function, which gives the neuron its nonlinear properties. The activation function can be any nonlinear function, but popular functions include the sigmoid, rectified linear unit (ReLU), and the hyperbolic tangent (tanh).

The LSTM cell has internal gates, which can regulate what information should be kept in memory and what should be forgotten. Both the output and the memory of the LSTM cell is fed as inputs to the following LSTM cell in the sequence. The mathematic representation of the LSTM cell is shown in (E.6)-(E.11).

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_\mathrm{f}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_\mathrm{f}\right) \tag{E.6}$$

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_\mathrm{i}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_\mathrm{i}\right) \tag{E.7}$$

$$\boldsymbol{g}_t = \tanh\left(\boldsymbol{W}_\mathrm{g}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_\mathrm{g}\right) \tag{E.8}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \otimes c_{t-1} + \boldsymbol{i}_t \otimes \boldsymbol{g}_t \tag{E.9}$$

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_\mathrm{o}\left[\boldsymbol{x}_t, \boldsymbol{h}_{t-1}\right] + \boldsymbol{b}_\mathrm{o}\right) \tag{E.10}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh\left(\boldsymbol{c}_t\right) \tag{E.11}$$

where $\boldsymbol{f}_t$ is the activation function of the forget-gate, $\boldsymbol{i}_t$ is the activation function of the input-gate, $\boldsymbol{g}_t$ is the candidate cell-state, $\boldsymbol{c}_t$ is the actual cell-state, $\boldsymbol{o}_t$ is the activation function of the output-gate, and $\boldsymbol{h}_t$ is the output also known as the hidden state.

Since the problem is to predict multiple time steps of multiple features from multiple input features of different dimensionality (number of features as well as number of time steps), an encoder-decoder model architecture is chosen. The encoder-decoder architecture is developed specifically to enable sequence-to-sequence predictions [25]. The architecture is comprised of two models: the encoder for representing the variable-length input sequence as a fixed-length vector; and the decoder that predicts the output sequence from the fixed-length vector as given by the encoder.

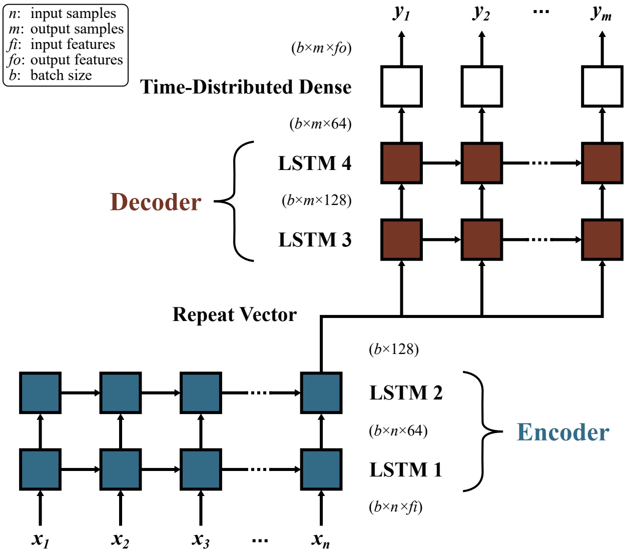


Fig. E.6: Encoder-decoder long short term memory recurrent neural network architecture. The output dimensions of each of the layers as well as the input dimensions are denoted in the parentheses.

The model structure chosen in this paper is an encoder layer consisting of two successive LSTM layers, the first having 64 neurons and the second with 128 neurons. The decoder, similarly, consists of two LSTM layers, the first has 128 neurons and the second has 64 neurons. A time-distributed fully connected (dense) output layer is used to generate the output sequence. The interface between the encoder and the decoder segments is achieved by repeating the output vector of the second LSTM layer of the encoder to match the size of the first LSTM layer in the decoder. The architecture is also represented in Fig. E.6.

The described encoder-decoder LSTM RNN is implemented using the Python module, Keras [26]. Keras is a high-level API for implementing deep learning models and uses Google's TensorFlow framework for the low-level tensor operations. A code segment of the implementation is shown in Listing E.1. ReLU activation functions are used for the LSTM layers. The `return_sequences` argument of the LSTM layers determines whether the LSTM gives an output for each of the temporal steps, or only for the last step. For LSTM layers 1, 3, and 4 (Fig. E.6) the layers should return the sequences, as these are used in the following layers. But for LSTM layer 2, only the output of the last temporal step is used as the encoded feature, hence this layer should not return the sequences.

A summary of the implemented model is shown in Table E.1. LSTM_1 and LSTM_2 comprise the encoder and LSTM_3 and LSTM_4 comprise the

Listing E.1: RNN model implementation

```python
from keras.models import Sequential
from keras.layers import Dense, LSTM, RepeatVector,
    TimeDistributed

# Initiate sequential model
model = Sequential()
# Encoder
model.add(LSTM(64, activation='relu',
    input_shape=(None, n_features_in),
    return_sequences=True))
model.add(LSTM(128, activation='relu'))
# Decoder
model.add(RepeatVector(n_pred))
model.add(LSTM(128, activation='relu',
    return_sequences=True))
model.add(LSTM(64, activation='relu',
    return_sequences=True))
# Output
model.add(TimeDistributed(Dense(n_features_out)))

```

decoder. In total, almost 300 000 weights and biases are to be tuned in the training process. The `None` notations in the output shapes of the different layers denote that the shape is variable. This variability in shape applies to the number of samples in the input sequence as well as the batch size, i.e. the number of stacks passed through the model at a time.

When training the model, the last six data points of each time series for each stack is reserved as labels, i.e. the targets for the output of the model during the training phase. The remaining preceding data points are used as input for the model. Hence the sample size of the input data is the number of samples for the given stack minus the six samples, that are used as labels.

However, not all of the stacks are used for training the models. A randomly selected portion of the stacks (approximately 20%) are reserved for testing the models after the training. Thus for cluster 1, 13 out of 67 stacks are reserved for testing and for cluster 2, 6 out of 33 stacks are reserved for testing. For the model that is trained without clustering the data, the combination of the test stacks from each of the clusters are used as the testing stacks, i.e. 19 stacks. However, the outlier stacks from the clustering analysis are still present in this dataset, therefore the total amount of stacks is 126, which is more than the combination of the stacks from the two clusters.

The models are trained with a mean-absolute-error (MAE) loss function and the Adam optimizer for 100 epochs. It is observed that the training loss for each of the models after the 100 epochs has reduced slightly below the validation

Table E.1: Recurrent neural network model summary.

Layer	**Output Shape**	**Parameters**
LSTM_1	(`None`, `None`, 64)	18 688
LSTM_2	(`None`, 128)	98 816
Repeat Vector	(`None`, 6, 128)	0
LSTM_3	(`None`, 6, 128)	131 584
LSTM_4	(`None`, 6, 64)	49 408
Time-Distributed Dense	(`None`, 6, 3)	195
Total parameters:		298 691

loss, indicating that the model has achieved a good fit without over-fitting the input data. The value of training loss after 100 epochs is approximately 0.03, whereas the validation loss is approximately 0.04, both have some variation between models.

E.5 PREDICTIONS

There is some randomness involved in training the RNN models, which leads to variability in the predictions. To mitigate this variability, 15 models are trained on each of the three datasets (cluster 1, cluster 2, and the complete dataset). Each of the 15 models are used to make the predictions on the testing data, and the average of the predictions is used as the prediction result of the ensemble of models.

The prediction results of a subset of the testing stacks are shown in Fig. E.7.a, Fig. E.7.c, and Fig. E.7.e for the cluster 1 models, cluster 2 models, and the models on all the data, respectively.

In the following paragraphs, the referred error is calculated by

$$e = y - \hat{y}$$

where y is the actual value and $\hat{y}$ is the predicted value.

The models trained on cluster 1 are, on average overestimating the voltage with the mean error across the test stacks going approximately linearly from -0.01 for the first prediction step to -0.05 for the sixth prediction step. Both the current and temperature are underestimated on average. The mean current error is steady between 0.02 and 0.03. The mean temperature error increases from 0.02 to 0.03 from the first to the sixth prediction. Fig. E.7.b shows boxplots of the prediction errors for each predicted variable across all the test systems for each prediction step. The middle line indicates the median, the box indicates the inter quartile range (IQR; from the 1^{st} quartile to the 3^{rd} quartile). The whiskers show the lowest data point larger than 1.5 IQR below

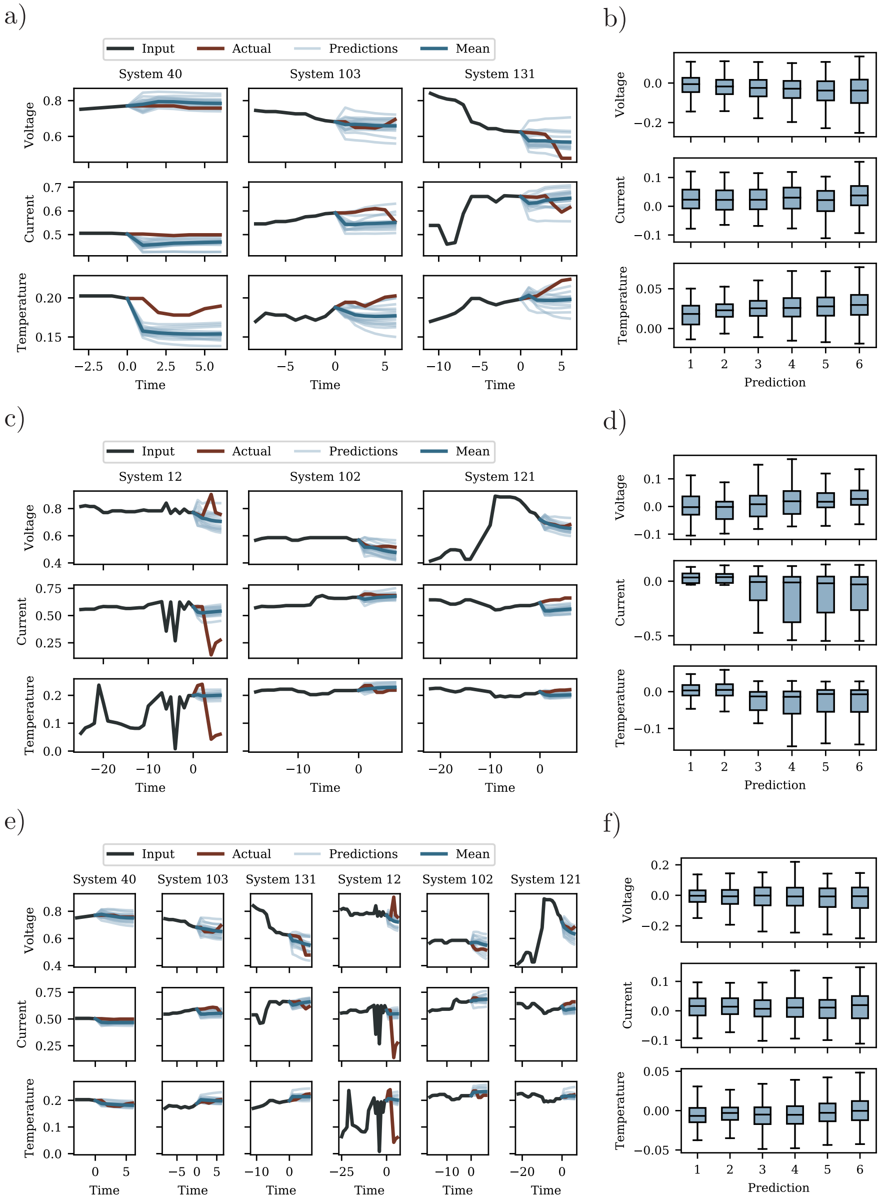


Fig. E.7: Subfigures a, c, e show prediction results for a selection of testing stacks of the models trained on cluster 1, cluster 2, and all the data, respectively. Subfigures b, d, and f show prediction error boxplots across all the stacks in the testing set of the models trained on cluster 1, cluster 2, and all the data, respectively.

the 1st quartile and the highest data point smaller than 1.5 IQR above the 3rd

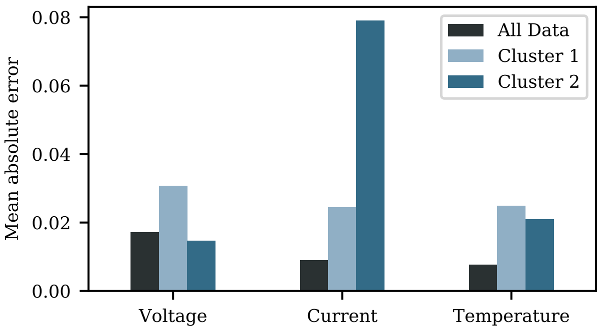


Fig. E.8: mean-absolute-error of the predictions for each of the trained models.

quartile.

The models trained on cluster 2 data are, on the other hand, underestimating the voltage with the mean error going from 0.00 to 0.03 from prediction step one to six. The mean current error starts off at -0.03 for prediction step one and two, but then falls to land at a median error at step six of -0.11. This sharp change in mean current error might be heavily influenced by the sudden drop in current seen in system 12, which the model is not able to predict. For the temperature, the mean error, starts off at 0.00 for the first two predictions but then falls to -0.03 for the last step. The same drop occurs in the temperature of stack 12, as for the current, but the magnitude is smaller, leading to lesser impact on the mean error. However, the drop in mean error is still present. boxplots of the prediction errors for each predicted variable across all the test systems are shown in Fig. E.7.d.

The models trained on the whole dataset seems to have the best performance, with a mean voltage error between -0.01 and -0.02, mean temperature error between 0.01 and -0.02, and mean temperature error between -0.005 and -0.01. Box plots of the prediction errors are shown in Fig. E.7.f.

From this look at the mean errors for each prediction step, it seems that the models that are trained on the whole dataset perform better than the model trained on subsets of the data, i.e. the detected clusters. This is more visible when looking at the mean-absolute-error for the models as depicted in Fig. E.8. This clearly shows the superiority of the all-data-models.

The all-data-models have a MAE of 0.017 on the voltage prediction, 0.009 on the current prediction, and 0.008 on the temperature prediction. The cluster 1 models have relatively low prediction MAE for both voltage, current, and temperature at around 0.03, 0.024, and 0.025, respectively. Cluster 2 models have comparable voltage MAE to the all-data-models of 0.014, but struggles with the current predictions, giving an MAE of almost 0.08. The temperature MAE is decent at 0.02, although this is almost three times that of the all-data-models.

E.5.1 Discussion

The all-data-models clearly outperforms each of the models that are specialized to a single cluster within the dataset. One explanation for this could be that the RNN model is capable of learning the different groupings within the dataset inherently, which would make the separate clustering obsolete. The all-data-model then has the obvious advantage of more training data, which is considered of key importance in this case, where the available data is limited to relatively few examples with high variance.

The high variance and the accuracy of the prediction results are expected to improve as more data is available for training. More data points from the existing systems as well as data from additional systems would benefit the model training.

E.6 Conclusion

Two approaches to predicting multi-time-step values of voltage, current, and temperature of field-operating fuel-cell-based backup-power systems six months into the future were studied. One approach used all training data to learn the evolution of the variables. The other approach sought to cluster the stacks based on similarities in their time-series evolutions and train separate models on each cluster. In both cases, the same encoder-decoder long short-term memory (LSTM) recurrent neural network (RNN) model architecture was used. To quantify similarities between the stacks, dynamic time warping (DTW) was used and principal component analysis (PCA) and density-based spatial clustering of applications with noise (DBSCAN) was used to detect clusters and eliminate outlier stacks. To mitigate variability in the prediction results, an ensemble of 15 RNNs were trained in each of the approaches.

The models that was trained on all of the training data was found to have better performance in predicting values of the testing stacks, than the models that were trained and tested on specific clusters. In conclusion, for the specific dataset in this work, it was not found that clustering and splitting the dataset prior to RNN model training improved performance.

An ensemble of LSTM RNN models, trained on the complete dataset, showed promising on-average performance in predicting future health indicating parameters. But struggles to make accurate predictions in irregular examples.

References

[1] M. J. Vasallo, J. M. Andújar, C. García, and J. J. Brey, “A Methodology for Sizing Backup Fuel-Cell / Battery Hybrid Power Systems”, *IEEE Transactions on Industrial Electronics*, vol. 57, no. 6, pp. 1964–1975, 2010.

[2] H. Paul, “Uninterruptible Power Supply (UPS)”, in *Hydrogen and Fuel - Technologies and Market Perspectives*, J. Töpler and J. Lehmann, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, ch. 7.

[3] M. F. Serincan, “Validation of hybridization methodologies of fuel cell backup power systems in real-world telecom applications”, *International Journal of Hydrogen Energy*, vol. 41, no. 42, pp. 19 129–19 140, Nov. 2016.

[4] Haimin Tao, J. Duarte, and M. Hendrix, “Line-Interactive UPS Using a Fuel Cell as the Primary Source”, *IEEE Transactions on Industrial Electronics*, vol. 55, no. 8, pp. 3012–3021, Aug. 2008.

[5] G. Bruni, S. Cordiner, V. Mulone, A. Giordani, M. Savino, G. Tomarchio, T. Malkow, G. Tsotridis, S. Bodker, J. Jensen, R. Bianchi, and G. Picciotti, “Fuel cell based power systems to supply power to Telecom Stations”, *International Journal of Hydrogen Energy*, vol. 39, no. 36, pp. 21 767–21 777, Dec. 2014.

[6] K. Goebel, A. Saxena, M. Daigle, J. Celaya, I. Roychoudhury, and S. Clements, *Introduction to Prognostics*, Dresden, Germany, 2012.

[7] M. Jouin, R. Gouriveau, D. Hissel, M. C. Péra, and N. Zerhouni, “Prognostics and Health Management of PEMFC - State of the art and remaining challenges”, *International Journal of Hydrogen Energy*, vol. 38, no. 35, pp. 15 307–15 317, 2013.

[8] R. E. Silva, R. Gouriveau, S. Jemeï, D. Hissel, L. Boulon, K. Agbossou, and N. Yousfi Steiner, “Proton exchange membrane fuel cell degradation prediction based on Adaptive Neuro-Fuzzy Inference Systems”, *International Journal of Hydrogen Energy*, vol. 39, no. 21, pp. 11 128–11 144, 2014.

[9] E. Lechartier, E. Laffly, M. C. Péra, R. Gouriveau, D. Hissel, and N. Zerhouni, “Proton exchange membrane fuel cell behavioral model suitable for prognostics”, *International Journal of Hydrogen Energy*, vol. 40, no. 26, pp. 8384–8397, 2015.

[10] G. Niu, B.-S. Yang, and M. Pecht, “Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance”, *Reliability Engineering & System Safety*, vol. 95, no. 7, pp. 786–796, Jul. 2010.

[11] W. Schmittinger and A. Vahidi, “A review of the main parameters influencing long-term performance and durability of PEM fuel cells”, *Journal of Power Sources*, vol. 180, no. 1, pp. 1–14, May 2008.

[12] M. Bressel, M. Hilairet, D. Hissel, and B. Ould Bouamama, “Remaining Useful Life Prediction and Uncertainty Quantification of Proton Exchange Membrane Fuel Cell Under Variable Load”, *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2569–2577, Apr. 2016.

[13] B. Dolenc, P. Boškoski, M. Stepančič, A. Pohjoranta, and Đ. Juričić, “State of health estimation and remaining useful life prediction of solid oxide fuel cell stack”, *Energy Conversion and Management*, vol. 148, pp. 993–1002, Sep. 2017.

[14] Z. Li, M. Ieee, Z. Zheng, M. Ieee, and R. Outbib, “Adaptive prognostic of fuel cells by implementing ensemble Echo State Networks in time varying model space”, *IEEE Transactions on Industrial Electronics*, vol. PP, no. c, p. 1, 2019.

[15] Z. Li and S. Jeme, “Remaining useful life estimation for PEMFC in dynamic operating conditions”, *2016 IEEE Vehicle Power and Propulsion Conference (VPPC)*, no. Cnrs 3539, pp. 1–6, 2016.

[16] K. Javed, R. Gouriveau, N. Zerhouni, and D. Hissel, “PEM fuel cell prognostics under variable load: A data-driven ensemble with new incremental learning”, *International Conference on Control, Decision and Information Technologies, CoDIT 2016*, pp. 252–257, 2016.

[17] R. Ma, E. Breaz, C. Liu, H. Bai, P. Briois, and F. Gao, “Data-driven Prognostics for PEM Fuel Cell Degradation by Long Short-term Memory Network”, in *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, IEEE, Jun. 2018, pp. 102–107.

[18] R. Ma, Z. Li, E. Breaz, and C. Liu, “Data - fusion Prognostics of Proton Exchange Membrane Fuel Cell Degradation”, *IEEE Transactions on Industry Applications*, vol. PP, no. c, p. 1, 2019.

[19] D. J. Berndt and J. Clifford, “Using dynamic time warping to find patterns in time series”, *KDD workshop.*, vol. 10, no. 16, pp. 359–370, 1994.

[20] S. Axelrod and B. Maison, “Combination of hidden Markov models with dynamic time warping for speech recognition”, in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, May 2004, pp. I–173.

[21] E. J. Keogh and M. J. Pazzani, “Scaling up Dynamic Time Warping for Datamining Applications”, pp. 285–289, 2000.

[22] A. Barr, G. Mathias, and D. Riu, “A Real-time Data-driven Method for Battery Health Prognostics in Electric Vehicle Use”, pp. 1–8, 2014.

[23] M. Ester, H.-p. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”, *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, 1996.

[24] F. Gers, “Learning to forget: continual prediction with LSTM”, in *9th International Conference on Artificial Neural Networks: ICANN ’99*, vol. 1999, IEE, 1999, pp. 850–855.

[25] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to Sequence Learning with Neural Networks”, pp. 1–9, 2014.

[26] F. Chollet *et al.*, *Keras*, \url{https://keras.io}, 2015.

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-466-9

AALBORG UNIVERSITY PRESS