TESIS DE LA UNIVERSIDAD DE ZARAGOZA

2021 364

Iván Sanz Gorrachategui

Hybrid and Intelligent Energy Storage Systems in Standalone Photovoltaic Applications.

Director/es
OYARBIDE USABIAGA, ESTANISLAO
BERNAL RUIZ, CARLOS

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

HYBRID AND INTELLIGENT ENERGY STORAGE SYSTEMS IN STANDALONE PHOTOVOLTAIC APPLICATIONS.

Autor

Iván Sanz Gorrachategui

Director/es

OYARBIDE USABIAGA, ESTANISLAO
BERNAL RUIZ, CARLOS

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

2021

Universidad Zaragoza
1542

PhD Dissertation

Hybrid and Intelligent Energy Storage Systems in Standalone Photovoltaic Applications

Author

Iván Sanz Gorrachategui

Advisors

Dr. Carlos Bernal Ruiz

Dr. Estanislao Oyarbide Usabiaga

Departamento de Ingeniería Electrónica y Comunicaciones
Escuela de Ingeniería y Arquitectura
University of Zaragoza, Spain

September 2021

Abstract

Remote systems such as communication relays or irrigation control installations cannot usually be powered by the electrical grid. One of the alternatives is to power these systems through solar panels, in what is known as standalone photovoltaic applications.

Most of these systems need a continuous operation, but a standalone photovoltaic installation cannot be powered during the night. For this reason, they use batteries to store excess energy during the day. These storage systems have been traditionally based on Valve Regulated Lead Acid (VRLA) batteries, but some effects can alter their performance in terms of reliability, operation cost and maintenance. One of the key issues that alter the energy behavior of the photovoltaic off-grid systems is the Partial State of Charge (PSoC) effect: Batteries cannot be completely charged as manufacturers indicate due to the day-night cycle. This gets the battery into an intermediate state of charge that effectively reduces its capacity, even halving it in some cases.

To mitigate the impact of these effects on the installation, batteries tend to be oversized with some security margins. These oversizing factors can be incredibly high and have a huge impact on the deployment and maintenance cost of the facility.

The first part of this thesis highlights some of these key concepts, analyzing which of them are critical in specific design cases, modeling them into a simulation tool, and as an outcome, establishing optimal sizing regions for the installations.

After the analysis, different ways of improving the performance of the installations are proposed. One idea to mitigate PSoC is to combine different storage technologies in a Hybrid Energy Storage Systems (HESS). HESSs have traditionally combined high energy density elements as batteries with high power density elements as ultracapacitors. An iteration of this idea is carried out throughout this thesis, where different types of batteries are combined. Each of them is best fitted to different power patterns in the application, such as daily cycles or emergency periods.

It is possible to further increase the performance by using intelligent algorithms to improve the functionalities of the Battery Management Systems embedded in these applications. To this end, failure prediction and health estimation algorithms are proposed as contributions of this work. These new algorithms endow the HESS with tools to predict possible energy disruption events and to anticipate aging, and thus, act accordingly.

Resumen

Debido a su localización, sistemas remotos como los repetidores de comunicaciones o los sistemas de control de riego no pueden ser alimentados desde la red eléctrica en muchas ocasiones. Una de las alternativas más comunes para alimentar este tipo de sistemas es el uso de paneles solares.

La mayoría de estos sistemas tienen un uso y un consumo constante, pero como es lógico, únicamente mediante el uso de paneles solares, no pueden ser alimentados durante la noche. Por este motivo, utilizan baterías para almacenar el exceso de energía durante el día. Tradicionalmente, estas baterías han estado basadas en químicas como el plomo regulado por válvula (VRLA). Existen distintos efectos como la temperatura, el envejecimiento, o el estado de carga parcial (PSoC) que afectan a estos sistemas de almacenamiento, reduciendo su fiabilidad y afectando a su coste de operación y a su mantenimiento. Para mitigar el impacto de estos efectos, las baterías suelen estar sobredimensionadas con ciertos factores de seguridad, que en ocasiones puede ser exageradamente grandes, y tienen un gran impacto en el coste de instalación, mantenimiento y reemplazo de las instalaciones.

La primera parte de esta tesis incide en la problemática de estas instalaciones, describiendo los distintos efectos que afectan a sus prestaciones, modelando los sistemas de almacenamiento y estableciendo como resultado las regiones de dimensionamiento óptimo de estos sistemas.

Después del análisis, se abordan distintas maneras de mejorar las prestaciones de dichas instalaciones. Para mitigar el estado de carga parcial, se propone la combinación de distintas tecnologías de almacenamiento, en lo que es conocido como sistemas de almacenamiento híbridos (HESS). Tradicionalmente, estos sistemas se han planteado como combinación de elementos de almacenamiento orientados a aplicaciones de energía (baterías) y elementos de almacenamiento orientados a aplicaciones de potencia (supercondensadores).

En esta tesis se propone una vuelta de tuerca sobre este concepto. Aquí, se propone la combinación de distintas químicas de baterías, cada una de ellas

enfocada a un tipo de uso en la aplicación, como los ciclados diarios, o el rol de emergencia.

Para ampliar las prestaciones de los sistemas de almacenamiento, también se propone el uso de algoritmia inteligente. En este trabajo se han desarrollado algoritmos de predicción de fallos y de estimación del envejecimiento de las baterías en estas instalaciones, de tal manera que se dota a los sistemas de almacenamiento híbridos con herramientas para anticipar distintos eventos potencialmente problemáticos y actuar en consecuencia.

Acknowledgment

First, I would like to acknowledge my advisors, Dr. Carlos Bernal Ruiz and Dr. Estanis Oyarbide Usabiaga, who made this research possible. Without their trust, work and support throughout these years, both in a professional and in a personal sense, this thesis would not have been existed. I would also like to thank Dr. Antonio Bono Nuez for his help in the last part of the thesis.

I would like to thank my laboratory colleagues, Alberto Sanz, Miguel Sanz, Eduardo Sebastián, Pablo Pastor, and very specially Alejandro Guillén, for sharing and discussing ideas, and their invaluable work and contributions. I would also like to acknowledge every student that has been in the laboratory throughout the years and has contributed in some manner to the thesis.

Similarly, I am very grateful towards the colleagues from Mondragon Unibersitatea, Dr. Érik Garayalde, Dr. Iosu Aizpuru and Dr. José María Canales, which whom we have worked very closely for several years now, and who have helped us adventure into the batteries world.

I would like to thank Epic Power Converters S.L. and all the staff for considering me one more, trusting in this work and supporting me both technically and financially when needed. I am also very grateful towards SICE, CHE and their staff for the support in the research.

I am very thankful to Mitsubishi Electric Research Laboratories in Boston for receiving me in my research stay and the support and excellence they showed towards me. Specifically, I would like to thank their current members Ye Wang, PhD, and Phil Orlik, PhD, and the former member Milutin Pajovic, PhD, for their help and support while being there and afterwards. Similarly, I would like to thank the Fulbright Commission for allowing me to live such an experience.

Last but not least, I would like to thank my family and my wife Clara for standing by me throughout these years, their love and support.

Iván Sanz Gorrachategui

Zaragoza, Spain

September 2021

Acronym list

ADAM – ADAptive Moment estimation

AN – Adversarial Network

BMS – Battery Management System

BMU – Best Matching Unit

C – Capacity

CC – Controlled Current

CHE – Confederación Hidrográfica del Ebro

CNN – Convolutional Neural Network

CSS – Cyclic Storage System

CV – Controlled Voltage

DAB – Dual Active Bridge

DAB-SRC – Dual Active Bridge Series Resonant Converter

DC – Direct Current

DoD – Depth of Discharge

ECM – Equivalent Circuit Model

EMS – Energy Management Strategy

EoD – End of Discharge

EoL – End of Life

ES – Early Stopping

ESR – Equivalent Series Resistance

ESS – Energy Storage System

EV – Electric Vehicle

GaN – Gallium Nitride

GPR – Gaussian Process Regression

GSM – Global System for Mobile communications

HESS – Hybrid Energy Storage System

HI – Health Indicator

IC – Incremental Capacity

IR – Internal Resistance

LAM – Loss of Active Material

LCO – Lithium Cobalt Oxide

LFP – Lithium Iron Phosphate

LLI – Loss of Lithium Inventory

LLP – Loss of Load Probability

LM – Levenberg-Marquardt
LSTM – Long-Short Term Memory network
LTO – Lithium Titanate Oxide
MaxRE – Maximum Relative Error
MES – Multiple Expert System
MLP – Multi-Layer Perceptron
MLR – Multivariable Linear Regression
MPPT – Maximum Power Point Tracking
MRE – Mean Relative Error
MSE – Mean Squared Error
NARX – Nonlinear AutoRegressive network with eXogenous inputs
NiCd – Nickel-Cadmium
NiMH – Nickel-Metal Hydride
OCV – Open Circuit Voltage
OTC – One-Time Constant
PbA – Lead-Acid
Pk – Peak
PSoC – Partial State of Charge
PV – Photovoltaic
PWM – Pulse-Width Modulation
ReLU – Rectified Linear Unit
RI – Resilience Improvement
RMS – Root Mean Square
RMSE – Root Mean Square Error
RNN – Recurrent Neural Network
RLR – Regularized Logistic Regression
RUL – Remaining Useful Life
SoC – State of Charge
SoD – Start of Discharge
SoH – State of Health
SOM – Self-Organizing Maps
SoP – State of Power
SSE – Sum of Squared Error
SSI – Steady SoC Improvement
SSS – Support Storage System

TIEDVD – Time Interval of Equal Discharging Voltage Difference

TTC – Two-Time Constant

VRLA – Valve-Regulated Lead Acid

Index

Figure Index

Table Index

Equation Index

Chapter 1.- Introduction

This chapter serves as an introduction to this dissertation. In it, the framework where it has been developed is presented. Going into further detail, standalone photovoltaic installations and their energy storage systems are described. The tools explored in the rest of the dissertation, Hybrid Energy Storage Systems, and Machine-Learning techniques, are introduced. Finally, the chapter presents the objectives of the thesis.

1.1 Background

1.1.1 Standalone photovoltaic applications

In recent years, communication technologies and infrastructures have advanced notably. Today, we have the possibility of accessing any kind of information regardless where we are. To make this possible, a large communication network is needed, with multiple nodes and intermediate systems. Mobile communications for example require numerous base stations and relays to provide services to a large wireless area. Other applications are also of vital value, such as the management of irrigation systems, which also need infrastructure for data collecting, communications, actuators...

These networks and systems fall under the umbrella of ambient intelligence. In the EU, the Horizon 2020 program during the past decade included the development of these systems as a strategic goal. This sector supposed an estimated yearly turnover of 1.780 M€ in the EU. Other international programs, such as NSF in the US or MOHURD in China also focused on these goals.

The deployment of these networks has an impact on multiple different services, such as Smart Cities, the management of natural resources as water and renewable energies, homeland security systems, transport services management, communication networks, or street lighting. Ultimately, these systems are needed in any scenario that requires sensing information and acting on a certain number of distributed stations.

However, and due to the remote location of some of these stations, it is not always possible to connect them to the distribution grid. Therefore, in these off-grid applications, it is necessary to use autonomous generation systems to power these facilities. Traditionally, fossil fuels have been used for this task, but in recent years, the use of renewable energy is becoming the go-to option. Among the renewable energy options, photovoltaic energy is probably the most common option, at least in small and medium-size installations, such as the one in Fig. 1.1. Other sources such as wind energy (micro-turbines) also have some presence in this niche.

Fig. 1.1. Standalone Photovoltaic Installation in Sigena, region of Aragón, Spain

However, renewable energies have a disadvantage regardless of their type, which is their availability. Usually, the application cannot depend on whether the renewable resource is available or not. For example, applications such as base stations have a mostly DC consumption throughout the day and night and cannot depend on the availability of solar or wind resources.

To make the energy generation problem independent from the application, these power systems usually include Energy Storage Systems (ESSs) that allow the system to be flexible. These systems serve a double purpose. On the one hand, they act as an energy buffer to power the system during the night, since there is no solar irradiation at that moment. On the other hand, they also serve as an emergency backup, powering the system during cloudy or foggy periods of several days.

1.1.2 Storage elements

Energy Storage Systems are also key elements in the development of modern societies since during the last decades the concern for sustainable development has increased on a worldwide scale. Currently, the Sustainable Development Observatory has established the Sustainable Development Goals (SDGs) as a means to achieve the objectives towards the 2030 Agenda. Among the 17 SDGs

that have been established, three of them are directly related to the ecologic transition and the energetic paradigm (SDG7, SDG12, and SDG13). Some of the strategic points common to these goals are the need for increasing the deployment of renewable energies and the concern on energy productivity, and this cannot be done without research and development on ESSs.

Research in the field of ESSs has risen in recent years. Multiple energy storage technologies have been developed for many different applications, such as those based on electrochemical reactions (batteries), electric field (capacitors), mechanical elements (flywheels)... The recent development in these technologies has been mainly driven by some applications such as electric vehicles (EV), DC micro-grids, and other industrial applications such as battery-powered elevators or peak-shaving systems. These applications use new storage technologies with high-power capabilities and deep charge and discharge cycles, such as lithium-based chemistries, as well as other high power-oriented technologies such as ultracapacitors (Lu et al. 2013; Thounthong, Raël, and Davat 2009; Carignano et al. 2017). Other conventional applications, such as energy storage in solar-powered micro-grids, use other bulk storage technologies, such as lead or nickel-related chemistries.

Regardless of their nature, the different storage technologies can be arranged according to their specific energy density and their specific power density. Ragone plots (Ragone 1968; Christen and Carlen 2000) are a well-known kind of diagram which represents these two magnitudes. As an example, Fig. 1.2 shows a Ragone plot with different storage technologies (Budde-Meiwes et al. 2013).

This diagram shows the tradeoff between power and energy for multiple storage technologies. On the one hand, elements such as capacitors or ultracapacitors are capable of handling high power rates but they are not capable of storing a large amount of energy. On the other hand, storage elements such as batteries, fuel cells, or diesel generators are capable of storing a large amount of energy, but deal worse with high power rates.

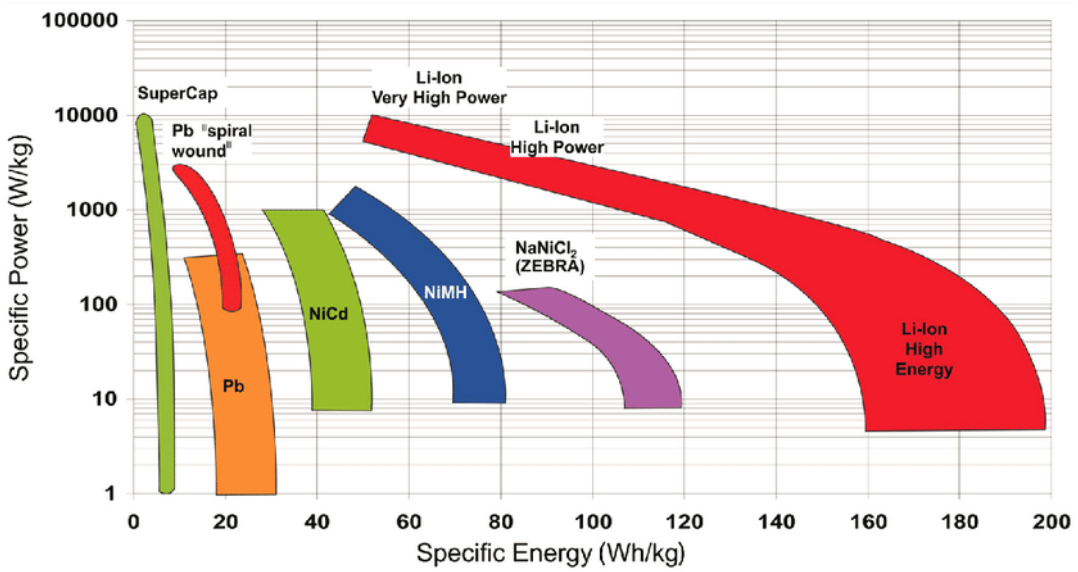

Fig. 1.2. Example of Ragone plot with different battery chemistries (Budde-Meiwes et al. 2013).

Even within batteries, some chemistries are best suited for energy density requirements whereas other chemistries are more oriented to power density applications, as can be seen in the plot. Ultimately, the selection of storage technology depends on the generation and consumption patterns of the specific application.

Different combinations of storage elements have been proposed in the literature to provide a flexible storage system, capable of providing energy at high power rates and of acting as a bulk energy reservoir. This has been typically done by combining ultracapacitors and batteries. These systems are referred to as Hybrid Energy Storage Systems (HESSs).

In standalone photovoltaic applications, batteries are used as energy storage elements in their ESS. Despite the increasing knowledge of the battery field and the development of new battery technologies (lithium-ion chemistries), the batteries used in large off-grid photovoltaic installations are mainly based on well-known chemistries such as Valve-Regulated Lead Acid (VRLA) or Nickel-Cadmium (NiCd). An example of one of these ESSs is depicted in Fig. 1.3.

Fig. 1.3. VRLA-based ESS in a photovoltaic standalone installation

However, these systems suffer from some issues in their daily operation, such as Partial-State-of-Charge (PSoC) or temperature-related issues, which worsen the performance of the batteries. To solve these problems, battery packs in these installations need to be oversized, and even then, they start having problems years before their expected End of Life (EoL). In future chapters, these issues will be expanded.

The BATT-Ex project [("BATT-Ex- Desarrollo de Extensores de Vida Útil de Baterías Para Sistemas Autónomos Alimentados Por Placas Fotovoltaicas | SICE" n.d.)], where this thesis started, aimed to solve, or at least mitigate these last issues. In the framework of said project, this thesis has counted with the support of the Government of Spain and the Confederación Hidrográfica del Ebro (CHE) state entity. This institution manages different elements regarding the river Ebro basin, which includes most of the northeastern region of Spain (see Fig. 1.4). They are in charge of tasks such as the control of irrigation channels, the control of water reservoirs and dams, rain and snow measurements... To do so, they count on a metering and communication network, with several installations distributed across the basin.

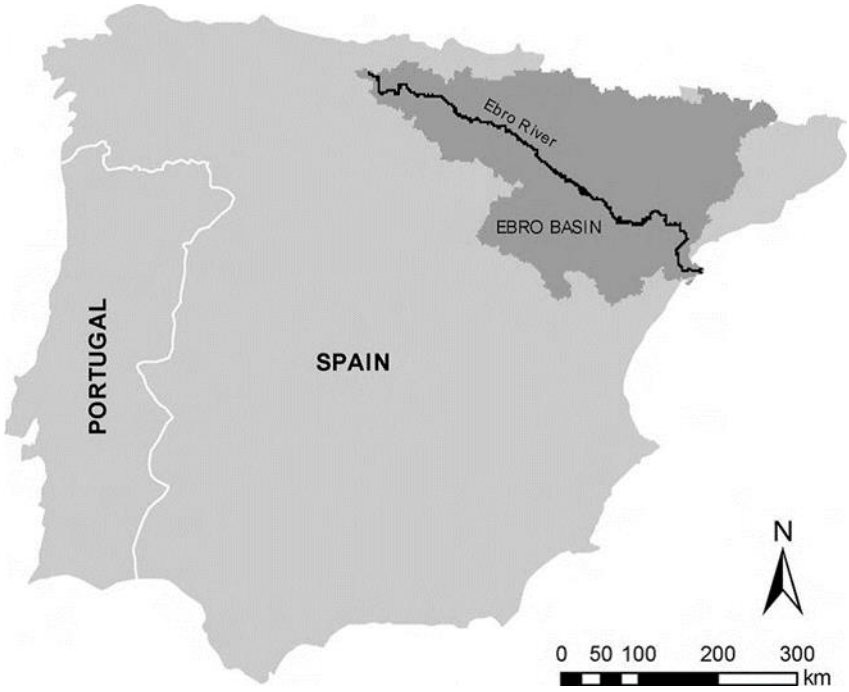

Fig. 1.4. Ebro basin map.

Among the installations in the basin, 333 of them are standalone and rely on solar panels and batteries to operate. Some of these installations have been used in the thesis as data-collecting stations. Others, to deploy hardware and software solutions, as will be developed in future chapters.

The ultimate goal of the project was to improve the performance of the ESSs in this kind of remote installations. To achieve this goal, two different and complementary approaches are explored within this thesis: the use of HESSs and intelligent algorithms.

1.1.3 Hybrid Energy Storage Systems

Conventional ESS rely only on one storage element. As explained above, depending on their generation and consumption patterns, certain applications may need different storage technologies. As a solution, HESSs, which combine different storage elements, have been developed in recent years.

The division of the storage element in multiple sub-systems allows flexibility in the charge and discharge processes, which can be controlled to a certain degree of freedom. HESSs have been typically addressed as the combination of power-oriented storage elements with energy-oriented storage elements, which

act as bulk energy storage. By using these different storage technologies, the system can optimize the energy flow so each storage subsystem is used with its best charge/discharge regime, improving the overall energy absorption processes and minimizing aging on the elements. Such storage technologies include batteries of different chemistries, ultracapacitors, fuel cells... For example, the combination of batteries and ultracapacitors has been studied previously in (Jian Cao and Emadi 2012; Khaligh and Zhihao Li 2010; Kroics 2015).

However, this case is not suitable in these standalone photovoltaic applications, since the consumption is mainly DC, and they are energy-oriented applications with a day-scale sparsity in power generation. Therefore, there is no need for a power-oriented element. For this reason, a different kind of HESS is going to be explored, by combining two energy-oriented elements to provide a solution for the problems in these installations.

1.1.4 Machine-learning techniques

The use of machine-learning techniques in battery research has a wide range of applications. Several studies use these tools as a means to model, diagnose or predict the behavior of batteries. Specifically, health-related problems have proven to be a challenging task for conventional estimation approaches, and are a promising field for data-driven algorithms.

Additionally, these techniques have also been applied to the field of standalone photovoltaic installations on multiple occasions with different approaches, such as predicting irradiation, predicting cloudy periods with or without weather forecast... Batteries in these installations have also been the subject of study with these tools. In this regard, State of Charge (SoC), State of Health (SoH), or Remaining Useful Life (RUL) are some of the most common battery-related parameters to be estimated that are studied in the literature.

In the last part of this thesis, different machine-learning techniques are going to be explored. Specifically, the goal is to develop algorithms to aid Battery Management Systems (BMS) to make better decisions in battery charging and management strategies, thus increasing their performance and reliability.

1.2 Objectives

The main hypothesis in this thesis is that Energy Storage Systems in standalone photovoltaic applications suffer from different effects that burden their performance, and that they can operate better by using different techniques. Among them, Hybrid Energy Storage Systems and machine-learning tools. Under this hypothesis, the following objectives are proposed:

- To study standalone photovoltaic installations and their Energy Storage Systems. This includes understanding their architecture, their generation and consumption patterns, how to size their storage systems, the storage technologies, and the key effects that burden their performance.
- To propose a flexible model capable of performing time-domain simulations of these systems and capturing these undesired effects.
- To study the viability of using Hybrid Energy Storage Systems as a means to mitigate the issues in the installations and to improve the performance. To this end, different storage technologies, architecture configurations, sizing, and control options must be studied. Finally, a simple but novel solution is going to be developed and implemented in a real system, and its performance through some months in operation is to be evaluated.
- To explore smart control techniques to endow Battery Management Systems with tools to diagnose the installations and address different battery-related problems. These tools are going to be machine learning based. Specifically, the problem of energy shortage prediction is going to be addressed. The goal is to predict a future blackout in a short period to act accordingly. Additionally, the aging estimation problem is going to be addressed as well, so battery replacement can be planned.

1.3 Outline

The general outline of this dissertation is shown in Fig. 1.5. It has been divided into three main parts and the conclusions chapter.

The first section, including Chapters 1 and 2, establishes the state of the art of standalone photovoltaic applications. A review of their storage elements, battery models, sizing methods, and most impactful issues is presented. As conclusions, two complementary ways to solve or mitigate these issues are proposed, namely a hardware approach that can be complemented by a smart management strategy. The hardware solution is based on Hybrid Energy Storage Systems with DC/DC converters, whereas the smart management strategy exploits data-driven algorithms to optimize the SoC and SoH of these installations.

The second part, which includes Chapter 3, studies the solution based on HESSs. The most suitable chemistries, architectures, and control strategies are explained. The sizing of these systems is also explored, and conclusions on their performance are obtained. The concept is validated through a hardware implementation on a real installation and the analysis of its performance over several months.

The third part, Chapters 4 and 5, explores machine-learning techniques that can be used to improve the performance of these HESSs. Specifically, Chapter 4 explores sequence-processing tools that process the voltage and current waveforms of the batteries and try to predict future failures. On the other hand, Chapter 5 addresses the aging estimation of lithium-ion batteries, developing some new techniques to diagnose cell degradation and to predict battery replacement.

Finally, Chapter 6 collects the main conclusions obtained throughout the thesis and establishes future research lines.

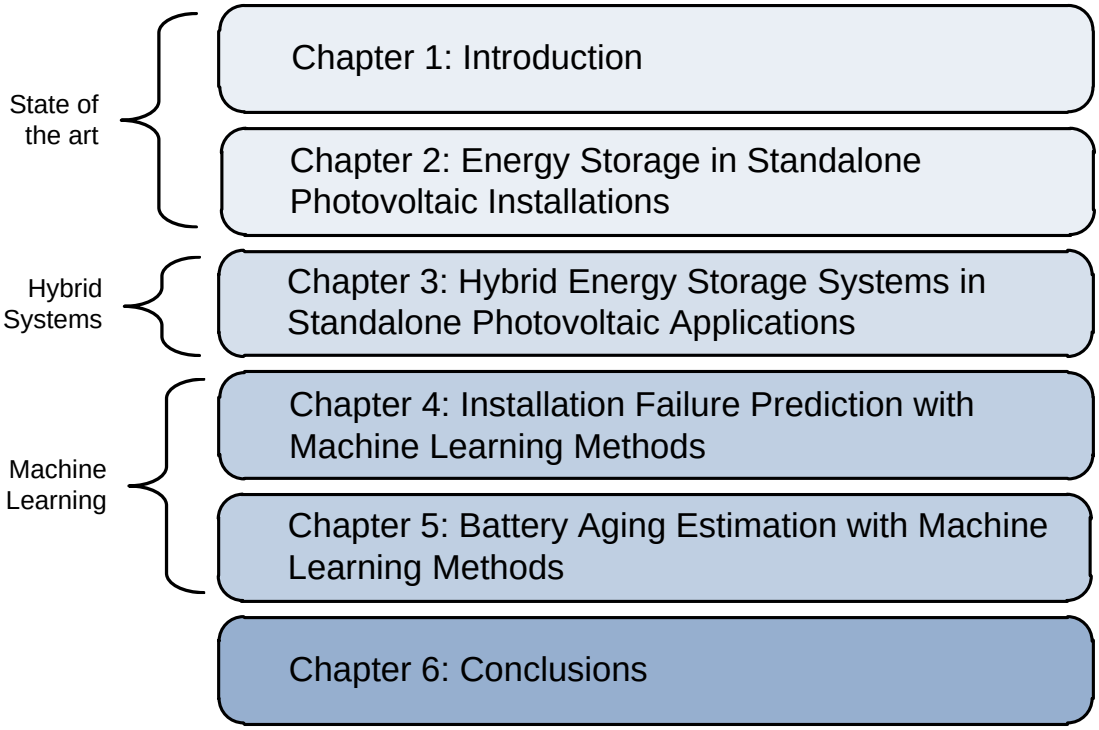

Fig. 1.5. Thesis structure

1.4 Publications

1.4.1 Journal papers

Sanz-Gorrachategui, I., Bernal, C., Oyarbide, E., Garayalde, E., Aizpuru, I., Canales, J. M., Bono-Nuez, A. (2018). "New battery model considering thermal transport and partial charge stationary effects in photovoltaic off-grid applications", Journal of Power Sources, 378, 311-321.

Sanz-Gorrachategui, I., Bernal-Ruiz, C., Oyarbide-Usabiaga, E., Bono-Nuez, A., Artal-Sevil, J. S., Garayalde-Pérez, E., Aizpuru-Larrañaga, I., Canales-Segade, J. M. (2019). "Partial State-of-Charge Mitigation in Standalone Photovoltaic Hybrid Storage Systems", Energies, 12, 4393.

Sanz-Gorrachategui, I., Pastor-Flores, P., Pajovic, M., Wang, Y., Orlik, P. V., Bernal-Ruiz, C., Bono-Nuez, A., Artal-Sevil, J. S (2021). "Remaining Useful Life Estimation for LFP Cells ins Second Life Applications", Transactions on Instrumentation and Measurement, Vol. 70.

Sanz-Gorrachategui, I., Pastor-Flores, P., Wang, Y., Orlik, P. V. (2021). "Remaining Useful Life Estimation for LFP Batteries Cells with Incremental Capacity Analysis", Transactions on Instrumentation and Measurement. **Under revision.**

Guillén-Asensio, A., Sanz-Gorrachategui, I., Bono-Nuez, A., Bernal-Ruiz, C., Sanz-Alcaine, J. M., Pérez-Cebolla, F. J. (2021). "Failure Prediction in Photovoltaic Standalone Installations by using Machine Learning Techniques", IEEE Access. **Under revision.**

1.4.2 Conference papers

Garayalde, E., Aizpuru, I., Canales, J. M., Sanz-Gorrachategui, I., Bernal, C., Oyarbide, E. (2017) "Análisis Experimental del Efecto de la Temperatura y la Tensión de Carga para la Optimización Energética de Sistemas de Almacenamiento en Instalaciones Fotovoltaicas Aisladas", SAAEI 2017.

Sanz-Gorrachategui, I., Gálvez, R., Artal, J. S., Pastor, P., Bernal, C., Oyarbide, E., Pérez, F. J., Garayalde, E., Anzola, J., Aizpuru, I. (2019) "Dimensionamiento de Sistemas Fotovoltaicos con Almacenamiento Híbrido", SAAEI 2019.

Pastor-Flores, P., Martín-del-Brío, B., Bono-Nuez, A., Bernal-Ruiz, C., Sanz-Gorrachategui, I., Pérez-Cebolla, F.J. (2019) "Clasificación de patrones de envejecimiento en baterías mediante SOM", SAAEI 2019.

Iosu Aizpuru Larrañaga, Jon Anzola García, Erik Garayalde Perez, José María Canales Segade, Iván Sanz Gorrachategui, Carlos Bernal Ruiz, Estanis Oyarbide Usabiaga y Jesus Sergio Artal Sevil (2019) "Análisis de Arquitecturas de Potencia Parcial para Aplicaciones de Corriente Continua", SAAEI 2019.

Pablo Pastor-Flores, Carlos Bernal-Ruiz, Iván Sanz-Gorrachategui, Antonio Bono-Nuez, Bonifacio Martín-del-Brío, Jesús Sergio Artal-Sevil, Fracisco J. Pérez-Cebolla. (2019) "Analysis of Li-Ion Battery Degradation using Self-Organizing Maps", IECON 2019.

A. Guillén-Asensio, I. Sanz-Gorrachategui, P. Pastor-Flores, J. S. Artal-Sevil, A. Bono-Nuez, B. Martín-del-Brío, C. Bernal-Ruiz. (2020) "Battery State Prediction in Photovoltaic Standalone Installations", EVER 2020.

I. Sanz-Gorrachategui, P. Pastor-Flores, A. Guillén-Asensio, J. S. Artal-Sevil, A. Bono-Nuez, B. Martín-del-Brío, C. Bernal-Ruiz. (2020) "Unsupervised Clustering of Battery Waveforms in Off-Grid PV Installations", EVER 2020.

Sanz-Alcaine, J.M., Sebastián, E., Sanz-Gorrachategui, I. et al. Online voltage prediction using gaussian process regression for fault-tolerant photovoltaic standalone applications. Neural Comput & Applic (2021).

1.4.3 Papers in the draft phase

"Aging exploration of LFP cells using UMAP and discrete Health Indicators"

Chapter 2.- Energy Storage in Standalone Photovoltaic Installations

This chapter delves into the problems of standalone photovoltaic installations, which rely on a solar-based energy generation system and a battery-based energy storage system. The focus is modeling and sizing these energy systems, the issues they face, and proposals on how to improve their performance and operation.

2.1 Conventional storage systems

Energy Storage Systems in photovoltaic standalone installations are designed to power the system when there is not enough sun irradiation. In these applications, different power generation patterns appear. Such patterns include deterministic variations as daily charge-discharge cycles and seasonal variation in weather and irradiation, but also other phenomena with a chaotic nature, such as sudden cloudy intervals and long periods with low solar irradiation due to fog or snow. The storage systems must store energy to power the installation during these events, and have to be designed taking into account the seasonal variation of irradiation. On the other side, consumption patterns may vary depending on the specific application. In the case being considered throughout this thesis, the consumption has been observed as mainly DC. An example of these current generation and consumption patterns in the installations is shown in Fig. 2.1.

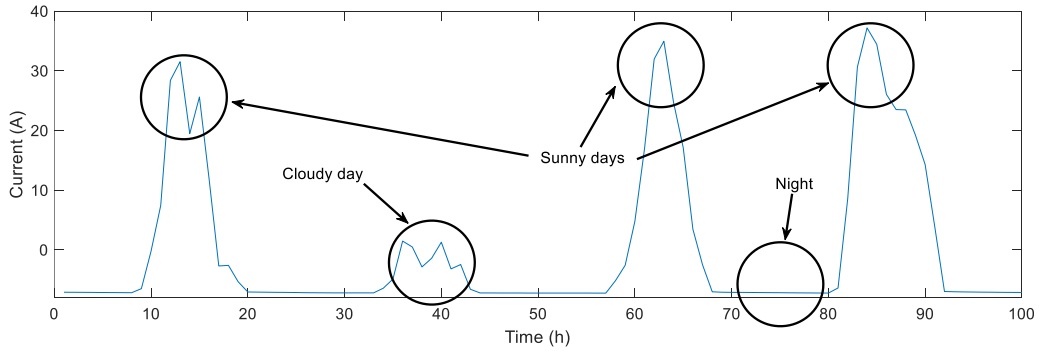

Fig. 2.1. Different generation / consumption patterns in the photovoltaic installation.

A general diagram of this type of installation is shown in Fig. 2.2. Here, the solar charger has been depicted as the Maximum Power Point Tracking (MPPT) block, since this is the most common type of charger nowadays (De Brito et al. 2013; Subudhi and Pradhan 2013; Esram and Chapman 2007). The output of the MPPT bus is connected to a DC bus, where the main battery pack and the load are connected as well.

Regarding the storage technology, this application does not require elements specialized in high energy densities, nor high power capabilities. As result, Lead-Acid batteries (gel or valve) or Nickel-Cadmium (NiCd) batteries have been

traditionally used for this purpose, since they are cheap and well-known technologies. In recent decades, Nickel-metal hydride (NiMH) has also been considered due to its better cyclic behavior (Zelinsky, Koch, and Young 2017).

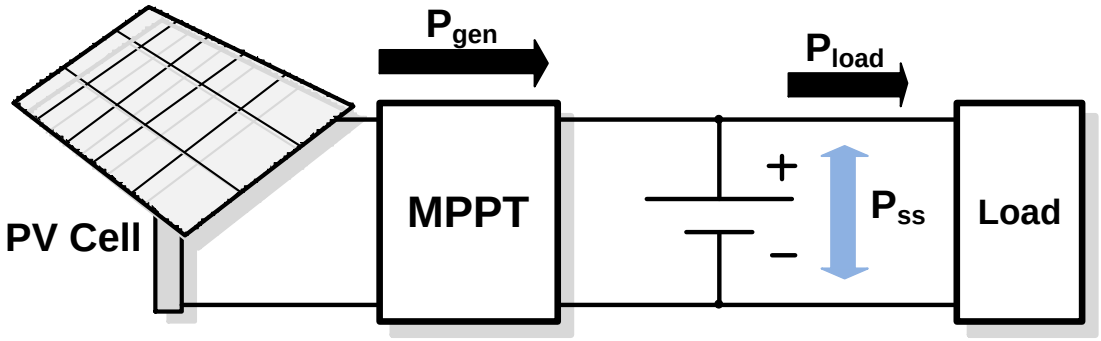

Fig. 2.2. Generic ESS architecture in a standalone photovoltaic installation.

Other new chemistries even better suited for cycling applications, such as Lithium-Ion batteries and their multiple variations, may be considered for this task too. They are designed to charge and discharge in a short period and will be able to fully charge during daily cycles. Nevertheless, these are still expensive technologies, and therefore, other solutions such as the combination of different elements may be considered.

Table 2.1 shows a brief review of the most common chemistries in the industry nowadays (Carrilero et al. 2018; EPEC 2021).

Chemistry	**Specific Energy (Wh kg^{-1})**	**Energy Density (Wh l^{-1})**	**Life cycles (100% DoD)**	**Max. discharge current (C-rate)**	**Cell voltage (V)**
Valve Regulated Lead-Acid (VRLA)	30-50	40-80	200-300	0.2C	2
NiCd	45-80	60-150	1000	1C	1.2
NiMH	60-120	120-280	300-500	0.5C	1.2
$LiCoO_2$ (LCO)	170-185	450-490	500-100	1C	3.6
$LiFePO_4$ (LFP)	90-125	130-300	1000-2000	5C DC 10C pulsed	3.2
LiNiMnCoO2 (NMC)	155-190	330-365	500-1000	1C DC 5C pulsed	3.7
$Li_4Ti_5O_{12}$ (LTO)	65-100	118-200	12000	10C DC 20C pulsed	2.5

Table 2.1. Features of different chemistries.

Lead-acid batteries are the most widely used batteries for different applications. They are the cheapest chemistry, but they do not age well under deep discharge cycles or high temperatures, whereas low temperatures reduce their effective capacity. In the niche of photovoltaic applications, their valve-regulated version (VRLA) is usually preferred due to lower maintenance costs.

Nickel-Cadmium (NiCd) batteries suppose an improvement on the performance of lead-acid batteries, with a better temperature behavior and increased service life. They also operate better under deep-cycle regimes. However, their use is deprecated due to Cadmium, which is a very toxic element with an expensive recycling process. Their cost is substantially higher than lead-acid batteries.

Nickel-Metal Hydride (NiMH) is an evolution over the NiCd technology, which does not include Cadmium and is much safer and easier to manage. However, this technology counts with a high self-discharge and a shorter service life.

Lithium Cobalt Oxide (LCO) is a technology with high energy density but low power density. It does not work well in high demanding applications in terms of C rate. Thus, it is the preferred technology in low power, portable applications. Its service life under high Depth of Discharge (DoD) cycles is low, and it is more expensive than other alternatives, thus it will not be considered for this application.

Lithium Nickel Manganese Cobalt's (NMC) main advantage is its great energy density, which allows for a more compact battery pack. Its high cell voltage also decreases the number of cells needed for a certain application. However, its service life is lower than other technologies such as LFP or LTO. On a side note, it contains cobalt, which makes it more expensive than LFP. It also contains hazardous materials, which can be a problem in some sectors as EV.

Lithium Titanate Oxide (LTO) technology promises a much larger number of life cycles and good thermal stability, but it is not competitive nowadays due to its cost. Besides, its low specific energy and low cell voltage make necessary

a larger number of cells, which increases additional manufacturing and installation costs (such as BMS or connectors) and undermines its development in a broad range of applications.

The Lithium Iron Phosphate (LFP) technology has some advantages over other Lithium-ion alternatives, such as a high number of life cycles and its competitive cost. On the other hand, the main disadvantages of this chemistry to others are a low energy density and a medium voltage level, which increases the size and weight of the storage pack. However, in the photovoltaic application, this is usually not a critical design parameter. The power capabilities of this technology are also not optimal, but more than enough in the application, since low current rates are expected.

2.2 Installation sizing

The sizing of standalone photovoltaic installations has been studied in the literature. Different sizing methods have been analyzed (Shrestha and Goel 1998; Shen 2009; Kazem, Khatib, and Sopian 2013; Kebaïli and Benalla 2015; Khatib, Ibrahim, and Mohamed 2016), and can be classified into averaged, analytical or numerical methods.

The sizing of the ESS in standalone photovoltaic installations is conditioned by the DC power consumption of the application. It can range from milliwatts in small harvesting sensors, to several kW in larger facilities such as a communication relay. It conditions the rest of the ESS since the sizing of the system must guarantee the energy supply needed by the installation.

Conventional ESSs in standalone photovoltaic installations usually rely on two main sizing parameters: the nominal energy of the storage elements and the nominal power of the solar panels. The sizing of these elements must fulfill the power consumption requirements, taking into account the nominal voltage of the load.

In the case of systems with a high peak-to-average power consumption ratio, the peak value would condition the selected storage technology and the sizing of the DC-DC converters. As introduced before, this is not going to be the case of study in this dissertation since the consumption is going to be considered mainly DC, although the study of these cases could be a future research line.

2.2.1 Battery sizing

Batteries count with two main sizing parameters: the Nominal Voltage, measured in Volts [V]; and the Nominal Capacity, measured in Coulombs [C] or Ampere-hours [Ah].

The voltage of battery cells is usually in the range of a few volts. For example, lead-acid cells have a nominal voltage of 2V. In a conventional ESS architecture, such as the one in Fig. 2.2, the battery pack is connected in parallel to the load (or loads) of the installation. Said loads usually count with their

voltage regulator (a DC-DC converter), with a specific input voltage, typically in the range of 12, 24, or 48 Volts. Therefore, multiple cells must be stacked in series to reach the desired load voltage.

The other parameter when sizing batteries for ESSs is capacity [Ah]. Only some fixed capacities are manufactured, so flexibility in the design of installations is limited. If needed, cells can also be stacked in parallel to increase the capacity of the battery pack. It is sometimes useful to express the battery size in terms of its Nominal Energy [Wh], which would be the product of its Nominal Voltage and Nominal Capacity.

Generally, the selection of the battery capacity will depend on the environment of the installation (geographical location, climate...), the load requirements, the target autonomy, and the security margins. Battery manufacturers and others such as IEEE provide guidelines to aid the designer in this task, with different correction factors due to aging, maximum cycling, or temperature. These factors oversize the batteries as a means to provide a safer operation. As an example, reference (IEEE 2019) includes recommendations on how to size these installations based on manufacturers' data and high oversizing factors.

2.2.2 Photovoltaic power sizing

The other sizing parameter is the nominal power of the solar panels. This value is generally given as the peak power generation. For a given panel technology, this parameter can also be specified as the effective area of the panels.

The focus of this thesis is on the storage system rather than the generation system, so the solar panels will only be addressed from a nominal power point of view, without looking further into panel technologies.

2.2.3 Sizing maps

The sizing problem can be studied graphically. The size of the panels and batteries are two degrees of freedom that define a space with infinite sizing

scenarios. This space can be represented as a surface (or map). Thus, each point on the map corresponds to a specific sizing scenario, with the third axis being the figure of merit considered (e.g. the availability of the system).

Different figures of merit have been defined in the literature for this purpose. Among them, the most common is probably the Loss of Load Probability, LLP (2.1) (Kazem, Khatib, and Sopian 2013; Khatib, Ibrahim, and Mohamed 2016). This figure represents the probability of a system running out of energy. It can be obtained as the portion of time that an ESS remains without energy (T_{ll}) compared to the total operation time ($T_{operation}$). Conversely, the system availability (2.2), which expresses the portion of time that the system is available, can be defined as the complementary figure of merit. An example of a map is shown in Fig. 2.3, taken from (Kebaïli and Benalla 2015). In this reference, these maps are used to size photovoltaic installations in different cities in Algeria.

$$LLP = \frac{T_{ll}}{T_{operation}} \tag{2.1}$$

$$availability = 1 - LLP \tag{2.2}$$

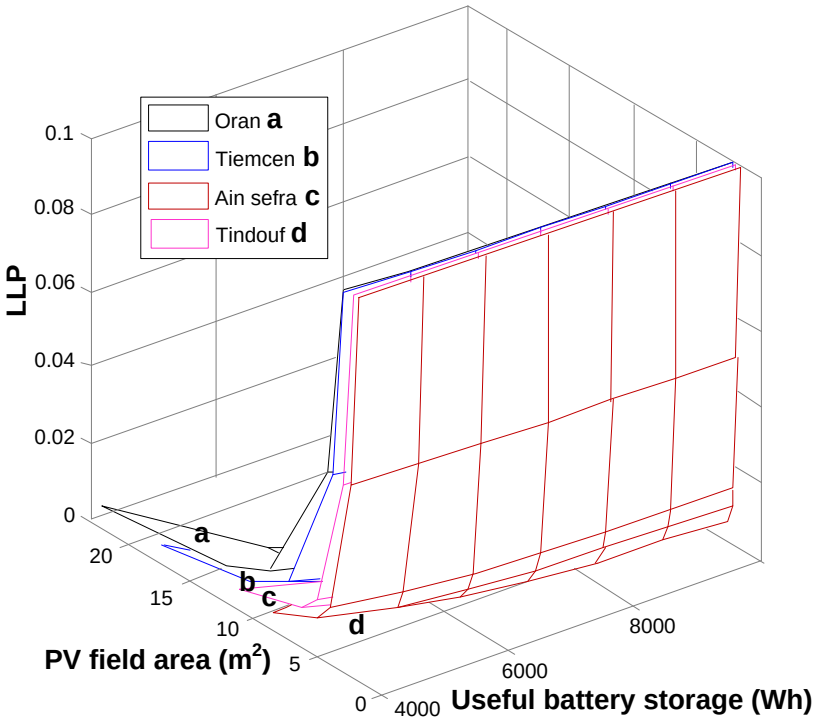

Fig. 2.3. Example of sizing maps (Kebaïli and Benalla 2015).

Other figures of merit can be included in these maps, such as installation cost, or performance compared to cost. Thus, by using these maps, the optimal design regions for each specific installation can be visually determined, as well as the cost and performance contour maps.

Sizing normalization

In Fig. 2.3, the magnitudes in the X and Y-axes are the PV field area [m^2] and the useful battery storage [Wh]. The specific values of these axes are closely related to the specific application under study, specifically to its consumption patterns. However, it is possible to normalize these parameters by the DC consumption of the application. By doing so, the sizing maps no longer depend on the specific values of the application, and general conclusions can be drawn from them.

The energy capacity of the battery can be normalized to the DC consumption of the installation, thus defining the *autonomy a* of the installation (2.3).

$$a[h] \triangleq \frac{E_{bat_{nom}}}{P_{load}} \tag{2.3}$$

Where $E_{bat_{nom}}$ is the nominal energy of the battery, and P_{load} is the DC consumption of the application.

The photovoltaic power can also be normalized to DC consumption. Therefore, the *generation ratio g* of the installation is defined (2.4).

$$g \triangleq \frac{P_{solar_{pk}}}{P_{load}} \tag{2.4}$$

Where $P_{solar_{pk}}$ is the peak value of the power of the solar panels, and P_{load} is again the DC consumption of the application.

The normalization of the design parameters allows comparing optimal design regions for different installations and obtaining conclusions on the sizing decisions being made. After the normalization, the sizing map is independent of the specific features of the application and only depends on the battery technology and the power generation / consumption patterns. From now on,

the design maps of the installations will be represented with these two normalized variables.

An example of a normalized sizing map is shown in Fig. 2.4. The total autonomy of the system is represented in the abscissa axis, and the generation ratio in the ordinate axis. The map shows the contour plots of the Loss of Load Probability in a color scale. It and has been obtained by simulating different installations, with different *generation ratios* and *autonomies.* The same irradiation and consumption patterns have been used as inputs. The simulation methods and models are described in forthcoming sections.

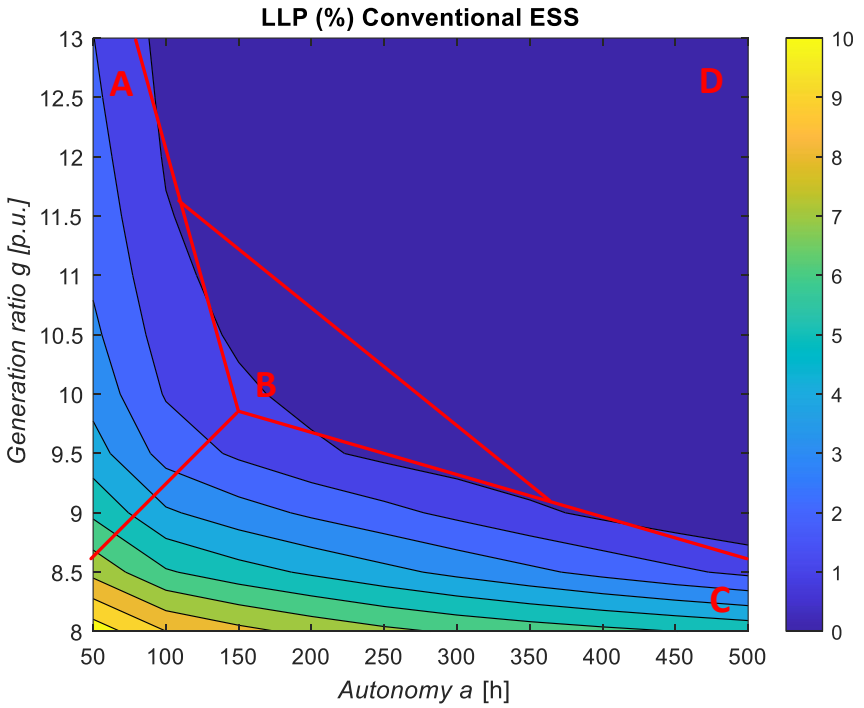

Fig. 2.4. Example of a normalized Loss of Load Probability sizing map.

Looking at the normalized sizing map, four main design regions can be appreciated. Maps with other battery technologies or generation patterns also share these same regions (although they may be located at different absolute values). These regions are described below.

- Region A: This design region is not desirable due to the low storage capacity (autonomy of the installation).

- Region B: Design region with a good tradeoff in performance (good LLP) without an oversized system.
- Region C: Undesirable design region due to low solar generation capabilities.
- Region D: Design region with good performance, but with an oversized system.

Other interesting conclusions can also be drawn from the maps to improve existing installations. In this case, it can be noted that below $g = 10$, it is more convenient to increase the number of solar panels rather than increasing the battery pack since LLP has an asymptotic behavior in the direction of autonomy. Reciprocally, with autonomies below 100 hours, it also makes more sense to increase the size of the battery pack rather than the solar panel power.

2.3 Degraded performance

Battery-based ESSs in standalone photovoltaic installations suffer from different effects that worsen their performance, either in the short term or in the long term (aging). Some of the most important effects are described below.

2.3.1 Ambient temperature

Temperature influences battery operation in different ways. The impact is greater depending on the specific chemistry of the cells, both in short-term operation and long-term degradation. These two effects are characterized independently by the manufacturers (Exide Technologies 2012; Saft Batteries 2007).

Short-term performance

The impact of temperature on battery capacity is shown in Fig. 2.5. The reference temperature is considered 20ºC (or 25ºC in some cases) in the tests. At this temperature, the battery shows 100% of its nominal capacity. Conventional lead-acid batteries increase their instantaneous capacity when used at temperatures above the nominal. On the other hand, below this temperature, their capacity decreases. This trend can also be observed in other chemistries such as NiCd and LFP batteries (Omar et al. 2013).

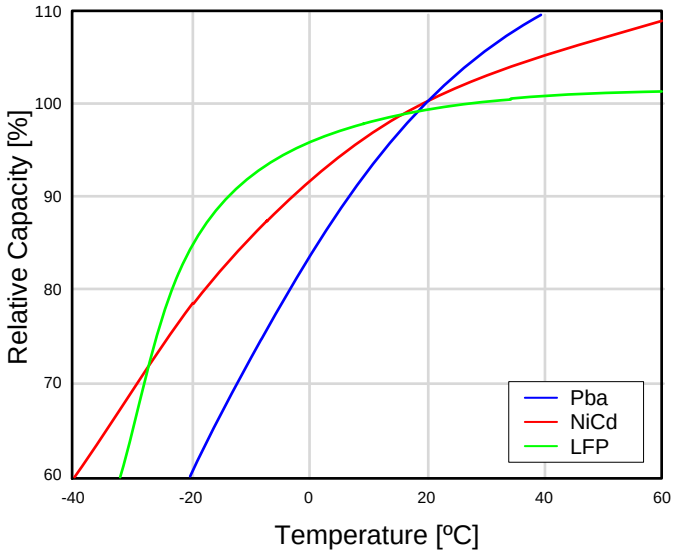

Fig. 2.5. Short-term temperature effect on capacity.

Generally, when capacity reduction is measured by manufacturers, the temperature is constant during the cycles. However, manufacturers do not provide data on capacity reduction when the temperature changes within the same cycle. In a standalone photovoltaic application, batteries are always charged during the day and discharged at night, and therefore they can operate at different ambient temperatures within the same cycle.

Cross-temperature tests have been carried out in the literature to determine the impact of this effect on the VRLA and the NiCd chemistries (Garayalde et al. 2017), and the results on the obtained charge are shown in Table 2.2. It can be appreciated that the results follow the trend described by the manufacturer. However, charging at the nominal temperature diminishes the impact of temperature on capacity by approximately 66% (from -34.7% to -12% when cold, from 16.4% to 5% when hot). The results obtained in said reference for the NiCd chemistry seem to contradict manufacturer's data, as they seem to work better in cold environments than in hot.

From these tests, a possible strategy for improving VRLA operation could be heating the cells. It would not be necessary to do so during the whole operation cycle, since heating during the charging process has a significant impact. In the case of NiCd, operation at temperatures above the nominal has proven to be detrimental, so in this case, cooling the cells would be a more appropriate strategy. These temperature-related strategies are not going to be explored in the thesis, but represent a plausible future research line.

Chemistry	**Charge temperature [ºC]**	**Discharge temperature [ºC]**	**Extracted charge (compared to 25ºC)**
PbA	25	5	- 12 %
	25	35	+ 5 %
	5	5	- 34,7 %
	35	35	+ 16,4 %
NiCd	25	5	+ 8,9 %
	25	35	-1 %
	5	5	+ 10,2 %
	35	35	- 10 %

Table 2.2. Experimental results from the cross-temperature experiments (Garayalde et al. 2017).

Temperature effect on aging

Battery aging is usually represented by the State of Health (SoH) of the battery, defined as in (2.5). Here C_{nom} represents the nominal capacity of the battery and C_{aged}, its aged capacity. SoH = 100% means a new battery, while SoH = 0% means that the battery has reached its End-of-Life (EoL). A common criteria for this is considering an aged battery when the aged capacity reaches 80% of its nominal capacity ($EoL = 0.8$ in the equation).

$$SoH(\%) \triangleq \frac{\left(\frac{C_{aged}}{C_{nom}} - EoL\right)}{1 - EoL} \quad x100 \tag{2.5}$$

Temperature also influences battery aging. At the reference temperature, the battery shows its nominal service life. Generally, cells last longer if used at a colder temperature, and age faster if used at warmer temperatures. Arrhenius' Law is used as a rule of thumb for lead-acid batteries, which estimates that the service life of the battery is halved for every 10ºC increase above the nominal temperature.

The impact of temperature on VRLA and NiCd aging is shown in Fig. 2.6. By comparing both temperature-related effects, it can be established that a battery above the nominal temperature improves its instantaneous performance, but ages faster. On the other hand, a battery below the nominal temperature ages slower, but has a lower storage capacity.

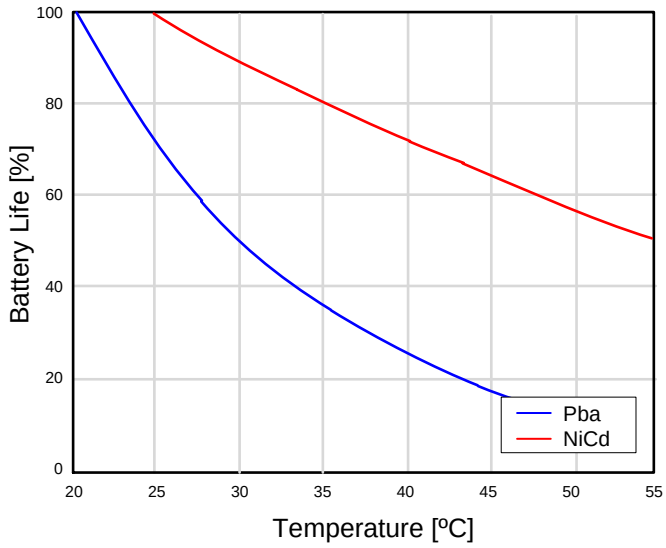

Fig. 2.6. Long-term temperature effect on battery life.

2.3.2 Continuous cycling

The short and medium-term performance

When included in ESS in standalone PV installations, batteries suffer from an additional effect that limits their operating capacity. This effect is related to the charging process of the batteries under daily cycles, and has been reported and defined in (Moseley and Rand 2012; Garayalde et al. 2017; Sanz-Gorrachategui et al. 2018; Swingler and Colgate 2018). This is known as Partial State of Charge (PSoC).

To understand the PSoC effect, the charging process of a battery needs to be explained. This process can be divided into two main stages, the current regulation phase (CC stage) and the voltage regulation phase (CV stage). Fig. 2.7 shows an example of a charging process with constant current.

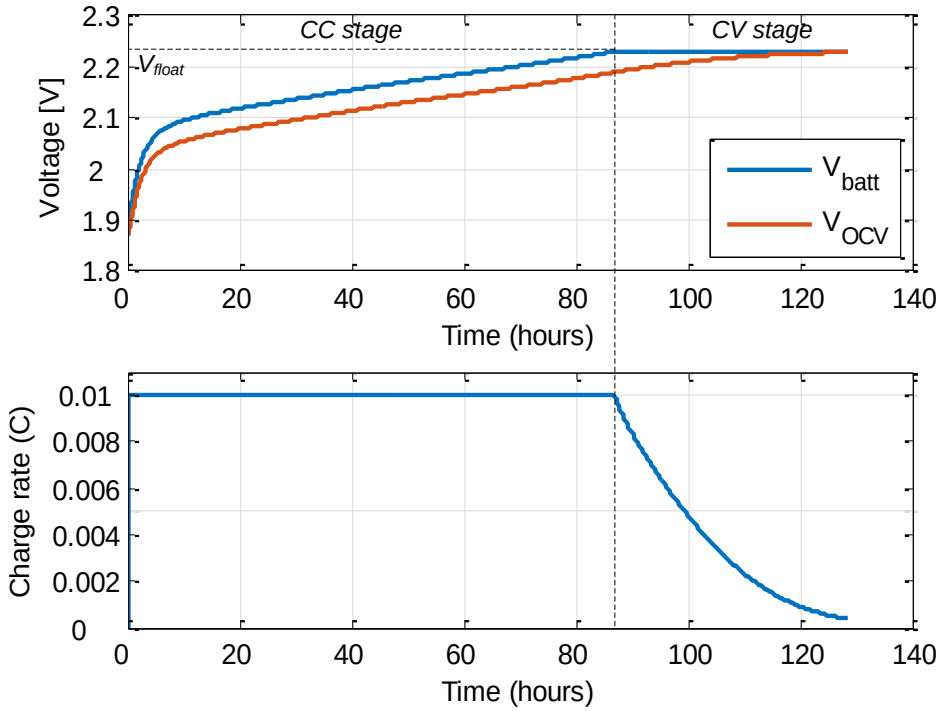

Fig. 2.7. Two-staged charge process with constant current.

As the battery charges during the CC stage, its internal, open-circuit voltage (v_{ocv}) begins to rise, as does its terminal voltage (v_{batt}). The difference between them depends on the series impedance of the battery. During this stage, the battery absorbs all the available generation power. Eventually, the battery

reaches the float voltage, and then the voltage regulation phase starts, where the battery charger controls v_{batt}. The current starts decreasing, and thus, the charging power decreases. Batteries must remain in this stage for a certain amount of hours (depending on the chemistry) before being completely charged. For example, the VRLA technology must remain in this stage for up to 72 hours (Exide Technologies 2012). The higher the series impedance, the sooner the battery reaches this stage. Additionally, chemistries with linear voltage curves reach this stage faster than chemistries with flat curves. These two elements are explained in the modeling section.

The charging process in a standalone photovoltaic system follows the same stages. Here, batteries are charged during the day and discharged at night. A typical daily charge waveform in one of these installations is shown in Fig. 2.8. At noon, the input power is maximum, thus generating a high current. This current rises battery voltage, which eventually reaches the float value and the battery enters the CV stage. Paradoxically, when irradiation is maximum, the absorbed power starts decreasing. Furthermore, sun irradiation hours are limited so batteries re-enter the CC stage as they start discharging when irradiation decreases. Thus, batteries cannot remain in float state for the number of hours required to reach the full-charge state, and they become partially charged.

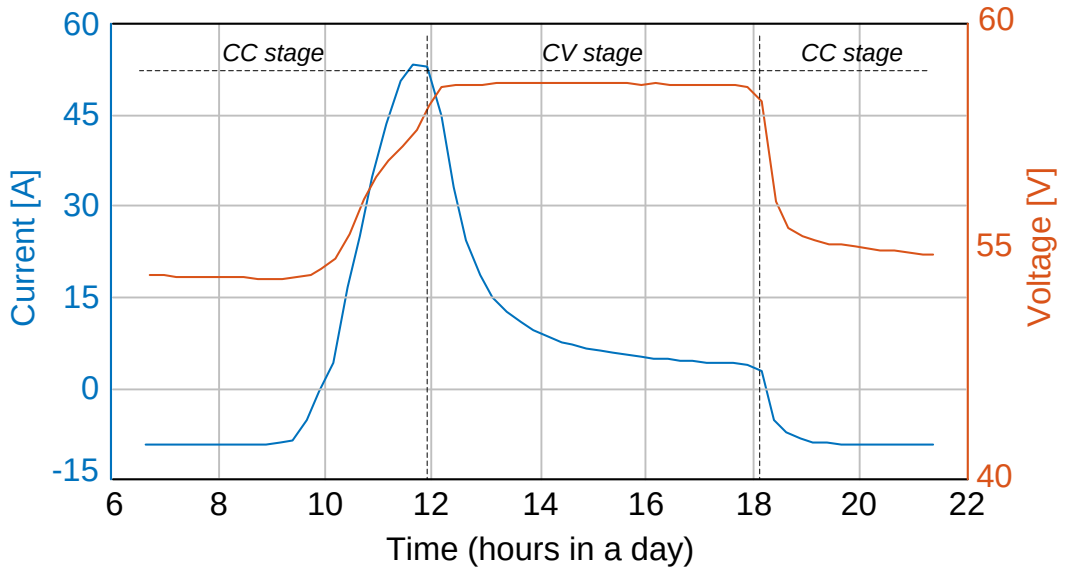

Fig. 2.8. Two-staged charge process in a standalone photovoltaic installation

This partial charge process is repeated daily in the standalone photovoltaic installations. In the medium term (months), it takes the battery into the Partial State of Charge (see Fig. 2.9). Even though batteries reach the float voltage every day, they are never fully charged, so they have less energy than expected and will be less ready if of a low-irradiation period happens.

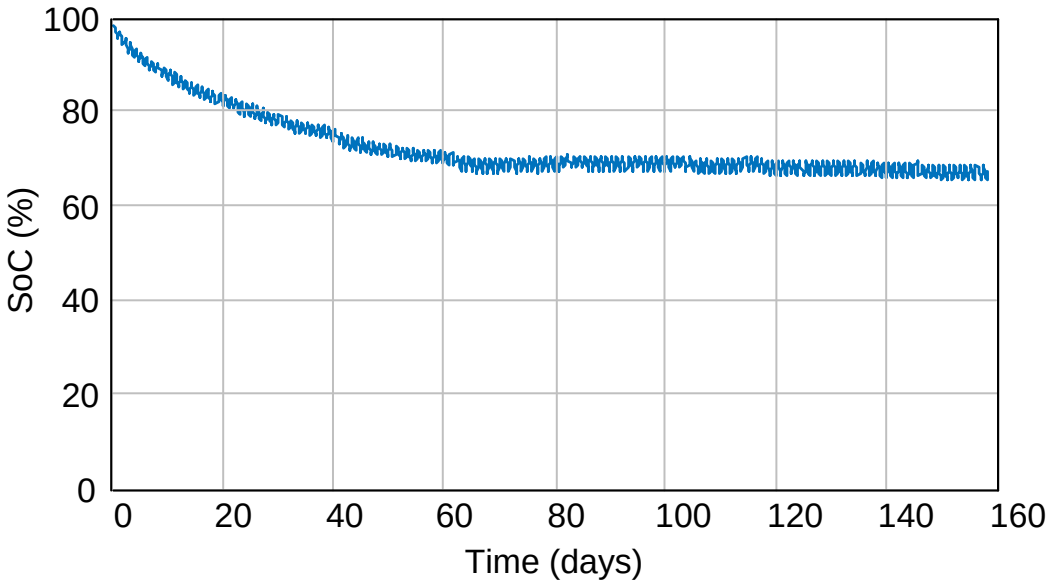

Fig. 2.9. SoC evolution under incomplete charges (PSoC).

Cycling impact on aging

Batteries in this application are under a continuous cycling regime, which accelerates aging. This effect is particularly significant in technologies such as VRLA when operating under deep cycles. This phenomenon is characterized by the Depth of Discharge (DoD), defined as in (2.6), where C_{nom} is the nominal capacity of the battery, and Q_{dis} is que discharged charge during the cycle.

$$DoD(\%) \triangleq \frac{Q_{dis}}{C_{nom}} \quad x100 \tag{2.6}$$

The impact of this effect depends on the specific battery technology since there are chemistries (lithium-ion variants) with better cycling performance and are not affected that much by it while other chemistries, better suited for stationary use, are greatly impacted.

Different standard tests are carried out by manufacturers to characterize this impact (such as IEC60896). During these tests, a continuous charge and

discharge cycle of the same depth (60%) is performed on the batteries along with a certain number of cycles, until the battery is considered to have aged. Afterward, aging is extrapolated in an inversely proportional curve for the remaining DoD values. An example of these tests is shown in Fig. 2.10.

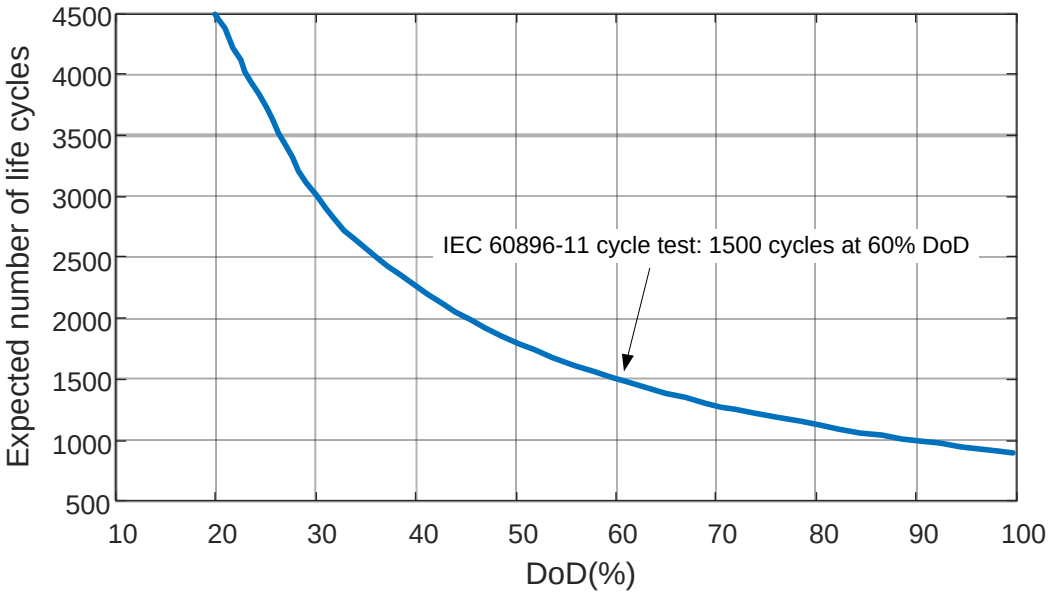

Fig. 2.10. DoD impact on aging for one example chemistry (OPzS solar by Exide).

2.3.3 High float voltage

As a means to mitigate the capacity reduction effect caused by temperature or PSoC, solar chargers use different charging strategies, usually based on increasing the float voltage of the battery.

Regarding the temperature, a common technique is to increase the float voltage when charging under low temperatures, and vice-versa. An example of this for VRLA batteries is shown in Fig. 2.11a.

However, this increased float voltage has a detrimental impact on battery aging, as shown in Fig. 2.11b. In general, using higher float voltages regularly, greatly increases aging for chemistries such as VRLA or NiCd. Other chemistries, such as some lithium-ion-based technologies, do not allow overcharging their cells due to permanent damage to the batteries.

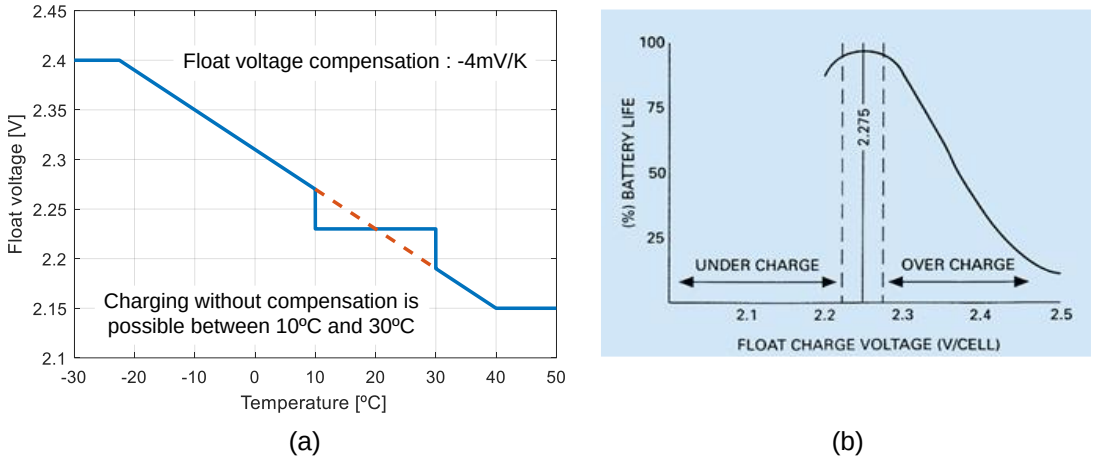

Fig. 2.11. Temperature-dependent float voltage compensation (Exide Technologies 2012) (a). Relationship between float voltage and aging for lead-acid technology (Yuasa Battery Corporation 1999) (b).

Regarding the Partial State of Charge, a common strategy is using additional absorption stages in the charging process (Fig. 2.12). These stages increase the voltage of the battery for a certain number of hours after the CV stage is reached. This allows some extra charge to flow in, overcharging the battery and reaching a higher SoC.

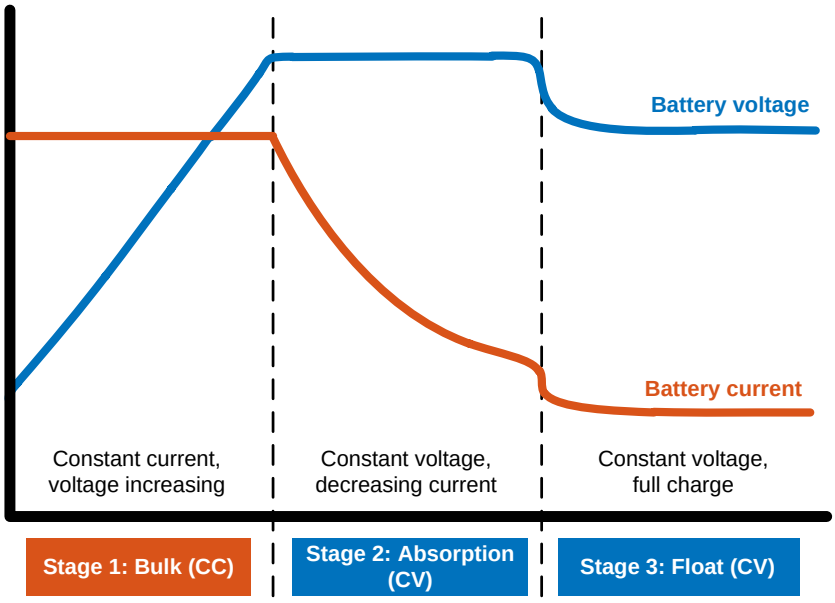

Fig. 2.12. Multiple-staged charge process

However, absorption stages do not completely solve the problem and accelerates aging. To avoid excessive degradation and other effects (e.g. gassing on VRLA batteries), some charger models include variations to this process, such as doing it monthly instead of daily. A possible improvement on these systems would be adding intelligent algorithms with predictive capabilities, only overcharging the battery if predicting low irradiation in forthcoming days.

2.3.4 Battery inefficiency

Batteries are not perfect energy accumulators i.e. not all the energy that is stored can be extracted. Different parameters, such as the energetic or the coulombimetric efficiency are defined in the literature to take into account this effect. Different battery chemistries have different efficiencies. For example, lead-acid batteries are generally considered to have an energetic efficiency of around 80%, while lithium-ion batteries are around 95%.

This effect is stressed by the C-rate of the application. Under high C-rates, batteries tend to provide less charge (Yuasa Battery Corporation 2006). However, in the standalone photovoltaic applications being considered, C-rates are not high (typically below C/10) so this additional effect is going to be disregarded.

2.4 Battery modeling

Battery models are needed for system simulation, sizing, and aging estimation. Additionally, they are the core of more complex online estimation algorithms implemented on in-situ applications, aimed to diagnose certain operation-related parameters. Such parameters include State of Charge (SoC), State of Health (SoH), State of Power (SoP), capacity or internal impedance (Tang, Liu, and Gao 2017; Tong, Klein, and Park 2015; Shahriari and Farrokhi 2013; Piller, Perrin, and Jossen 2001; Plett 2004). To this end, several techniques intending to model the battery behavior have been proposed in the literature. Considering the modeling techniques, they can be classified into different sub-classes such as electrical, physical, chemical or data-driven models.

Besides, simulation models are also interesting for developing future generations of ESS. These systems will include more than one storage element, becoming Hybrid Energy Storage Systems (HESS). In future years, research on developing Energy Management Strategies (EMS) will become relevant, and battery simulation models are key to develop, simulate and test these algorithms.

Future chapters in this dissertation will use simulation models of batteries. In this section, the modeling approaches for building these models are explained. An electrical model is going to be considered as a basic structure and is described in the following sub-sections.

2.4.1 Basic battery model

The main goal of this battery model is to mimic the electrical behavior of a real battery. This behavior is characterized by the voltage on battery terminals when a current circulates through it, or conversely, the current through the battery when a voltage is applied to its terminals. The basic electrical model structure is shown in Fig. 2.13. This model is also called the Equivalent Circuit Model of the battery, or ECM.

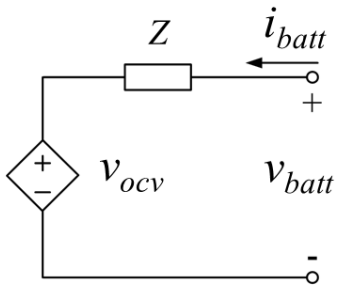

Fig. 2.13. Basic Equivalent Circuit Model of the battery.

The parameters in the figure are defined below:

- i_{batt}: represents the current through the battery terminals.
- v_{batt}: represents the battery voltage at its terminals, measured with or without current.
- Z is the output impedance of the battery. Different impedance models have been proposed in the literature, and are described in the following sections.
- v_{ocv}: is the internal, open-circuit voltage of the battery. That is the voltage at terminals minus the voltage difference in the output impedance Z.

2.4.2 Energy storage model

The ability to store energy is contained in the controlled voltage source v_{ocv}. The value of this voltage source depends on the energy of the battery, which is commonly expressed in terms of the State of Charge (SoC), as defined in (2.7).

$$SoC(\%) \triangleq \frac{q}{C_{nom}} \; x100 \tag{2.7}$$

Here, q represents the charge stored inside the battery, which is obtained by integrating i_{batt}. The parameter C_{nom} is its nominal capacity, as specified by the manufacturer. It is generally measured with a complete charge and discharge cycle at a nominal temperature (commonly, 20ºC or 25ºC). SoC can also be defined as a function of time, as in equation (2.8). This definition is used in the SoC estimation method known as Coulomb Counting, which can be used to determine the SoC of a battery in operation, by integrating the current (Ng et al. 2009).

$$SoC(t) = \frac{1}{C_{nom}} \cdot \int_{-\infty}^{t} i_{batt}(\tau)\, d\tau \;\; x100$$
$$= SoC(t_0) + \frac{1}{C_{nom}} \cdot \int_{t_0}^{t} i_{batt}(\tau)\, d\tau \;\; x100 \quad (2.8)$$

The relationship between SoC and v_{ocv} can be measured from a battery by charging and discharging it with a low current, obtaining an approximate curve. An example of these curves for an LFP cell is shown in Fig. 2.14.

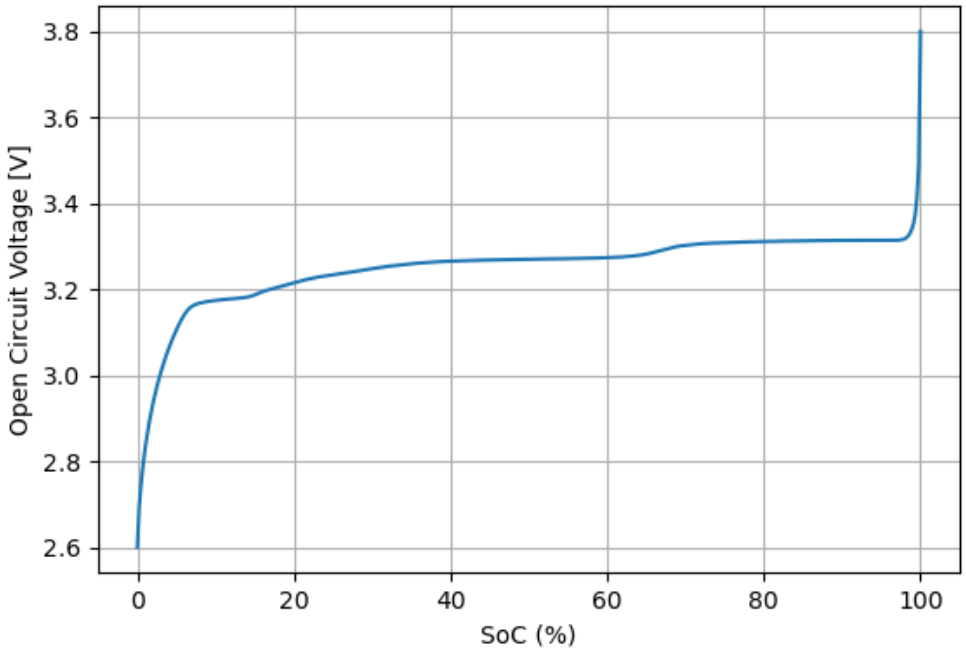

Fig. 2.14. Example of the SoC - v_{ocv} curve.

Equation (2.8) and the SoC – v_{ocv} curve build together the electrochemical model (M. Chen and Rincon-Mora 2006). This model estimates SoC by integrating the current and obtains v_{ocv} from the non-linear relationship. It is represented in Fig. 2.15.

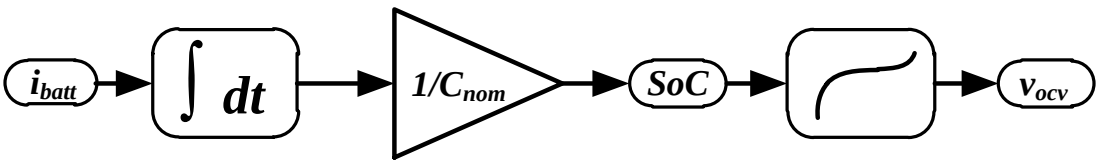

Fig. 2.15. Electrochemical model.

2.4.3 Impedance model

The battery model is completed with a circuit model of the output impedance. The impedance models are diverse and are selected mainly according to the battery chemistry and the most relevant dynamics in the application. Sometimes, a simple impedance model with an equivalent series resistance (ESR) may be enough.

However, diffusion effects or different dynamic characteristics are often taken into account. These effects are modeled by RC pairs added in series to the ESR. One Time Constant (OTC) or Two Time Constants (TTC) models (Fig. 2.16), including one or two RC pairs respectively, are commonly used for this purpose (Dubarry and Liaw 2007; Rahmoun, Biechl, and Rosin 2013; Plett 2004).

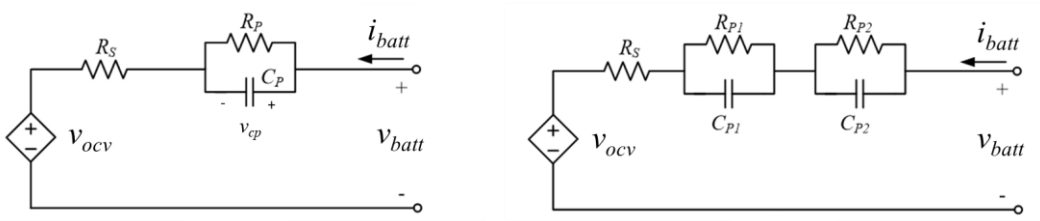

Fig. 2.16. OTC impedance model (left). TTC impedance model (right).

More complex impedance models are also proposed in the literature. The authors in (Coleman et al. 2007) use asymmetric impedance models (Fig. 2.17), to capture different behaviors when charging *(ch)* and discharging *(dis)* the batteries. These models are also known as hysteresis models.

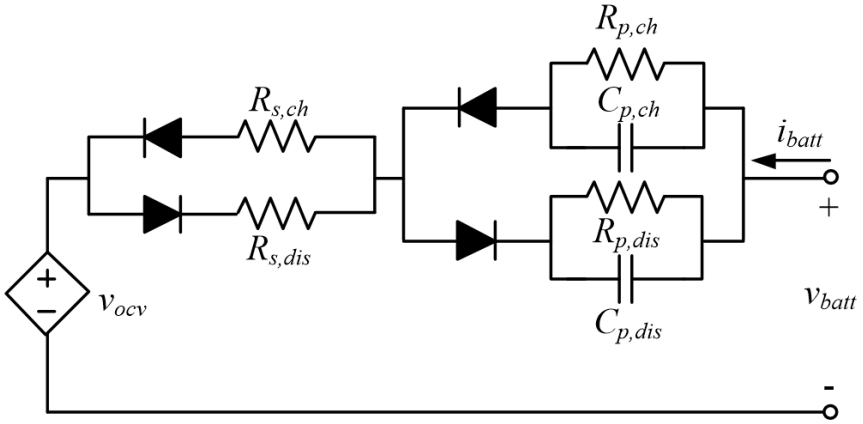

Fig. 2.17. Asymmetric impedance model.

Additionally, the circuit elements may also vary with external variables, such as temperature or aging. These considerations further complicate the modeling process. It has been decided that an OTC impedance model can replicate the dynamics of the application being studied, so it will be the option used for battery modeling purposes.

Regarding the different techniques developed in the state of the art for impedance modeling purposes, two main trends can be observed: large signal and small signal characterization.

Large signal characterization

Large signal methods analyze the transient and the steady behavior of cells when excited with pulse-shaped currents, of relatively high C-rates. The pulse response is analyzed, and the impedance model can be fitted from the voltage-current behavior.

These pulses are carried out at different SoC values, since the impedance model may vary depending on it. The SoC initial value ranges from a fully charged to a fully discharged cell, thus the impedance model is valid for the whole operation range of the battery. Examples of these large-signal methods are carried out in (Schweiger et al. 2010; Nital David 2016; Rodríguez et al. 2013).

Small signal characterization

Small signal methods are based on analyzing the frequency response of the cell. In this case, the cell is taken to a resting SoC and is then excited with a sinusoidal current, of small amplitude in terms of C-rate. The output voltage is analyzed, in amplitude and phase. By changing the frequency of the input, the whole transfer function of the system can be analyzed, thus obtaining the impedance. This technique is also known as Electrochemical Impedance Spectroscopy (Andre et al. 2011; Montaru and Pelissier 2010). It can be carried out at different rest SoC values, thus obtaining the impedance model for a wide range of operation points of the battery.

2.4.4 Defective performance modeling

As explained in 2.3.2, the Partial State of Charge is caused by the incomplete charging processes in the installations. To replicate this effect, the model of the battery must be able to work with the different charging phases of the battery: current regulation phase and voltage regulation phase. Since the transitions between these phases are conditioned by the voltage at the terminals of the battery, the impedance model is also important.

Other effects that worsen battery performance, such as temperature and aging are usually modeled as a reduction of the capacity of the batteries (Anbuky and Pascoe 2000; Pascoe and Anbuky 2004; Garayalde et al. 2017). However, as will be introduced further on, the scope of this work will be focused on mitigating the mid-term impact of PSoC in the installations, so these effects are not going to be considered in the battery model. Analyzing the temperature and the long-term impact on the solutions explored in the thesis will be a future research line, which has been initiated in (Sanz-Gorrachategui et al. 2018).

2.4.5 State-Space model

State-space modeling is a useful method for describing the equations of a time-domain model and is commonly used with other circuital models such as power converters. This approach defines memory variables, known as states, and combinational variables, which are a sub-product of the states and the inputs. The derivatives of the states are gathered in the "state equations", while the description of the outputs is gathered in the "measurement equations".

The Equivalent Circuit Model (with One Time Constant impedance model) and the electrochemical model of the battery, described in section 2.4.1 and shown in Fig. 2.15, can be written in the state-space description.

Current mode model

As a first approach, the input of the system is going to be i_{batt}, and the output, v_{batt}. This replicates the behavior of a battery working in current mode (a controlled current is applied through its terminals, and its voltage evolves freely). This is the case in the bulk stage in a solar charger. State of Charge,

which represents the energetic state of the battery, can be defined as an state variable under the state-space description. To model the voltage at the C_p capacitor in the impedance model, an additional state variable must be included (v_{cp}).

The state-space equations for this simple approach are contained in equations (2.9) and (2.10).

State equations:

$$\begin{cases} \dfrac{dSoC}{dt} = \dfrac{1}{C_{nom}}\, i_{batt} \quad x100 \\ \dfrac{dv_{cp}}{dt} = \dfrac{1}{C_p}\left(i_{batt} - \dfrac{v_{cp}}{R_p}\right) \end{cases} \tag{2.9}$$

Measurement equation:

$$\{v_{batt} = v_{ocv}(SoC) + i_{batt}\, R_s + v_{cp} \tag{2.10}$$

Here, the non-linear relationship between the SoC and the internal voltage is contained in the $v_{ocv}(SoC)$ term.

Voltage mode model

Conversely, v_{batt} can be used as an input, and i_{batt} as an output. This reflects the behavior of the battery when a constant voltage is applied to its terminals, and the current evolves freely. For example, on the float stage in a solar charger. The state-space equations for this case are (2.11) and (2.12).

State equations:

$$\begin{cases} \dfrac{dSoC}{dt} = \dfrac{1}{C_{nom}}\left(\dfrac{v_{batt} - v_{cp} - v_{ocv}(SoC)}{R_s}\right) x100 \\ \dfrac{dv_{cp}}{dt} = \dfrac{1}{C_p}\left(\dfrac{v_{batt} - v_{cp} - v_{ocv}(SoC)}{R_s} - \dfrac{v_{cp}}{R_p}\right) \end{cases} \tag{2.11}$$

Measurement equation:

$$\left\{i_{batt} = \frac{v_{batt} - v_{cp} - v_{ocv}(SoC)}{R_s}\right. \tag{2.12}$$

Full model

To simulate the complete behavior of a battery in the PV application and replicate the PSoC effect, both modes are complementarily needed. A Boolean parameter α can be defined for this purpose. This parameter takes the value "1" when in current mode, and the value "0" when in voltage mode. Thus, this model will use i_{batt} and v_{batt} as inputs and outputs. The complete state-space model is shown in (2.13) and (2.14).

State equations:

$$\begin{cases} \dfrac{dSoC}{dt} = \left(\alpha \dfrac{i_{batt}}{C_{nom}} + (1-\alpha)\left(\dfrac{v_{batt} - v_{cp} - v_{ocv}(SoC)}{C_{nom}\, R_s}\right)\right) x100 \\ \dfrac{dv_{cp}}{dt} = \dfrac{\alpha\left(i_{batt} - \dfrac{v_{cp}}{R_p}\right)}{C_p} + \dfrac{(1-\alpha)\left(\dfrac{v_{batt} - v_{cp} - v_{ocv}(SoC)}{R_s} - \dfrac{v_{cp}}{R_p}\right)}{C_p} \end{cases} \quad (2.13)$$

Measurement equation:

$$\begin{cases} i_{batt} = \alpha\, i_{batt} + (1-\alpha)\, \dfrac{v_{batt} - v_{cp} - v_{ocv}(SoC)}{R_s} \\ v_{batt} = \alpha\left(v_{ocv}(SoC) + i_{batt}\, R_s + v_{cp}\right) + (1-\alpha)\, v_{batt} \end{cases} \quad (2.14)$$

2.5 Potential improvements in these installations

After studying the behavior of conventional ESSs in standalone photovoltaic installations (operation, sizing, modeling, and issues), different alternatives may be proposed to enhance the operation of these systems. Specifically, this dissertation is going to delve in two different directions, which may be implemented complementarily to improve these ESSs. These alternatives are the use of different storage technologies (HESSs); and the use of smart algorithms to aid the Battery Management Systems (BMS).

2.5.1 Hybrid Energy Storage Systems

As introduced earlier, ESS in standalone photovoltaic systems have been based traditionally on chemistries such as VRLA. These technologies have higher series impedance and a more linear v_{ocv} curve than newer lithium-ion technologies, such as LFP. The latter can work better in these systems, mitigating the Partial State of Charge and the aging. However, lithium-ion technologies have a much higher cost than lead-acid chemistries, and therefore, HESSs can be considered as a solution with a good tradeoff between cost and performance. As introduced in Chapter 1, HESSs are based on dividing the storage system into multiple sub-systems (which can be of different storage technologies) and controlling them by using power conversion elements (such as DC/DC converters).

In this thesis, the hypothesis is that HESSs with two battery banks, based on different chemistries, can optimize the battery charging, discharging, and aging processes in both technologies. This effect can be achieved by adding a new auxiliary storage element to an existing ESS, complementing its operation. This new sub-system should be of different storage technology and specialized in different patterns of use. A DC/DC power electronics converter will control both systems, allowing for the assignation of complementary roles to each battery technology.

Including additional storage elements increases the complexity of the system, and adds new design and control challenges. Besides the different storage

technologies, the architecture of the system and the energy management strategies are also subjects of study. Chapter 3 looks into this approach, proposing different chemistries, control strategies, and implementing this concept in these installations.

2.5.2 Smart algorithms

Advanced charging strategies, such as multiple-stage charging with absorption phases (Fig. 2.12), have been previously introduced as a means to improve the partial charge effect on standalone PV installations. This particular strategy includes an additional absorption phase to improve the charge of the battery so it stores more energy. However, these overcharges come at expense of decreasing the useful life of the battery. For this reason, they are done periodically so the battery does not suffer that much.

This kind of overcharging strategy can be enhanced with predictive capabilities. The system can be overcharged only before a low irradiation period if it relies on a forecast of the future generation patterns. This way, the impact on battery aging is minimal, while preserving the improved operation. Different algorithms may be proposed for this predictive task. Chapter 4 establishes initial work on this research line, with some initial results and conclusions that show the promising features of these algorithms.

Additionally, when considering HESSs with more than one storage system, multiple strategies may be adopted to manage the energy and the batteries. To take better decisions, knowledge of the state and aging of the storage systems is required. Sometimes, aging estimation is problematic, thus developing new tools and aging estimation algorithms is desirable. Algorithms in this research line are also introduced in future chapters as a basis for these advanced control strategies. Chapter 5 explores some of these predictive tools to improve the performance and features of BMS in this specific application.

Chapter 3.- Hybrid Energy Storage Systems in Standalone Photovoltaic Applications

This chapter introduces Hybrid Energy Storage Systems as a solution for the problems of standalone photovoltaic installations. The storage technologies, sizing, and management of these systems are studied, and a specific solution is proposed for a case study. The system is implemented in a real installation, and its behavior is analyzed over some months of operation.

3.1 Introduction

The concept of HESSs has been already introduced in Chapters 1 and 2. These systems combine different energy storage technologies, resulting in a better overall performance. Typically, the HESSs that have been explored in the literature combine power-oriented storage elements with energy-oriented storage elements. This way, they can provide both high-power demands over a short period and low, and steady power demands over a long period.

However, not all HESSs combine power-oriented and energy-oriented elements. For example, they could combine different energy-oriented elements such as batteries of different chemistries. The specific storage technologies, their sizing, and operation modes depend on the application. Additionally, the architecture of the system becomes more complex as storage elements are added.

This chapter aims to overview these four key parameters (storage technologies, sizing, control strategies, and architectures) and apply them to the case of standalone photovoltaic installations. Finally, a simple HESS has been implemented in a real installation. Results of several months of operation are compared with the behavior of a conventional ESS of the same features in a nearby installation.

3.1.1 Hybrid Energy Storage Systems sizing

The process for sizing conventional standalone photovoltaic installations has been described in the previous chapter. Two sizing parameters have been defined as the main degrees of freedom: the photovoltaic nominal power and the total energy. They can be normalized to the power consumption, obtaining the expressions for generation ratio *(g)* and autonomy *(a)*. Both definitions have been rewritten in (3.1) and (3.2).

$$g \triangleq \frac{P_{panel_{pk}}}{P_{load}} \tag{3.1}$$

$$a \triangleq \frac{E_{ESS_{nom}}}{P_{load}} \tag{3.2}$$

Since g is not related to the storage element, its definition remains the same when considering HESSs. However, the autonomy problem changes slightly. As more storage elements are included, more degrees of freedom appear. Now, the nominal installed energy $E_{HESS_{nom}}$ is given by equation (3.3), where N is the number of storage sub-elements, and E_i is the nominal energy of the *i-th* storage element.

$$E_{HESS_{nom}} = \sum_{i=1}^{N} E_i \tag{3.3}$$

Since the power of the load remains constant, the overall autonomy of the HESS ($a_{HESS_{nom}}$) can be expressed as the sum of the contribution to the autonomy of each storage subsystem *i*, a_i (3.4).

$$a_{HESS_{nom}} = \sum_{i=1}^{N} a_i \,; \quad a_i \triangleq \frac{E_i}{P_{load}} \tag{3.4}$$

With a fixed value of the total autonomy $a_{HESS_{nom}}$, a new parameter can be defined for each storage sub-element. This parameter is related to the installed energy of the new element compared to the total. Thus, the Hybrid Factor f_i for each storage element can be defined as the ratio between the autonomy of the *i*-th element and the total autonomy of the HESS.

$$f_i \triangleq \frac{a_i}{a_{HESS_{nom}}}; \quad \sum_{i=1}^{N} f_i = 1 \tag{3.5}$$

Since the sum of all f_i values must be equal to one, N-1 f_i values are needed to fully define the sizing (besides *g* and *a).* Thus, for the case of two storage elements, only one f_i value is needed. In this case, it will simply be addressed as *f*.

When sizing a specific installation, the figures of merit introduced in chapter two (such as the Loss of Load Probability, LLP) must be analyzed. In a previous chapter, the contour plots of these figures of merit were presented and explained. However, when introducing the new degrees of freedom, these surfaces become hypercubes, whose coordinates are the variables *g*, *a,* and f_i.

Thus, the dimensionality of the space increases with each storage element, so it becomes difficult to generate visual representations of the problem.

Additionally, the figures of merit depicting the quality of the system depend largely on the specific chemistries and the roles assigned to each one of them, so the complexity of the problem grows enormously. The next section discusses different options for these energy management strategies in HESS.

3.1.2 HESS management strategies

In a conventional energy storage system with a single storage element, the strategies for controlling the power flows are limited. In these cases, the main battery is the only storage element operating as an energy buffer between the generation and the consumption nodes. The system must always absorb the most from the power source. At most, depending on the application and the storage technologies, different charging strategies such as the ones in Fig. 2.11a or Fig. 2.12 may be applied, but the controllability is very limited.

However, when considering multiple storage elements (and a flexible power conversion structure), the input and output power flow can be managed from or to different storage elements, or even between them. Multiple and new control options appear, with more degrees of freedom. In this thesis, the set of management techniques are referred to as Energy Management Strategies (EMSs).

EMSs depend on the technology of the storage elements and may have different goals, such as maximizing the energy absorption of the system, minimizing the aging of the storage elements, or trying to guarantee a minimum state of energy in the system. They establish how to use each storage element and therefore, have a critical impact on the performance and aging of the system.

Examples of such strategies have been proposed in the literature. One of the probably simplest strategies can be found in (Kim et al. 2014), where a priority-based strategy is proposed in a HESS with four storage elements: CTI capacitors, ultracapacitors, Li-ion batteries, and Lead-acid batteries. It always

gives priority to charging and discharging power-oriented elements (which store less energy), such as capacitors and ultracapacitors. This way, the last elements to be charged and discharged are the energy-oriented elements (batteries). This simple EMS has proven to be useful in the application. A flow chart that implements this EMS can be found in Fig. 3.1.

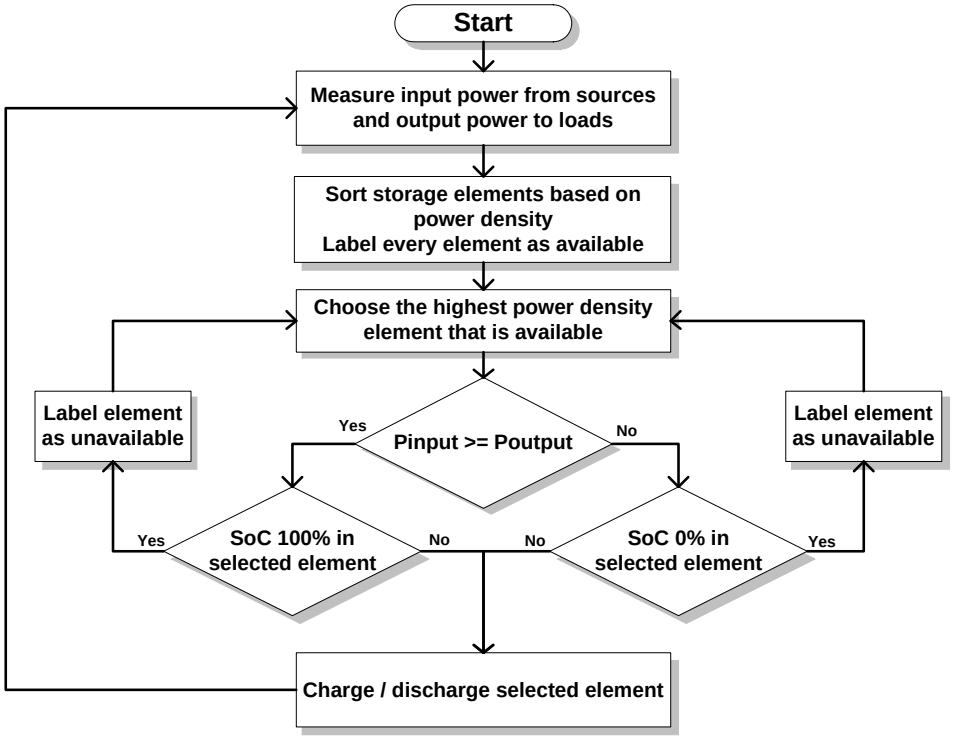

Fig. 3.1. Priority-based strategy flowchart.

In the flowchart, the storage elements are sorted depending on their power density, and then the highest power density element is chosen. If the element is fully charged and there is a positive net power input coming into the storage system, the following element is selected. This happens reciprocally when discharging.

However, multiple other criteria can be selected, depending on the specific elements and the goal of the system. As (Schiffer et al. 2007) point out, strategies could be based on determining a maximum depth of discharge on the elements, trying to minimize aging using health-monitoring parameters... To this end, multiple elements could be charged and discharged at the same time so they do not suffer deep cycles. It can be deduced that the control options are endless.

To implement these variated strategies, good estimation algorithms of different parameters related to the storage elements, such as SoC, SoH, or its equivalent impedance are necessary.

In the long term, deciding what to do when the storage elements have aged and the system starts failing is also important. Thus, replacement strategies may also be defined in addition to EMSs, with techno-economic criteria. However, over this chapter, the goal of the HESS is to improve the mid-term performance of the installations rather than optimize its long-term cost and amortization, so these strategies are not going to be explored.

3.1.3 HESS Architecture

Conventional Energy Storage Systems usually have a straightforward architecture, where the storage element is connected to the DC bus directly, or via a DC/DC converter. For example, in the case of the standalone photovoltaic installations under study, batteries are directly connected to de DC bus, in an architecture like the one shown in Fig. 3.2. Here, batteries act as an energy buffer between the solar converter (MPPT block in the figure) and the load.

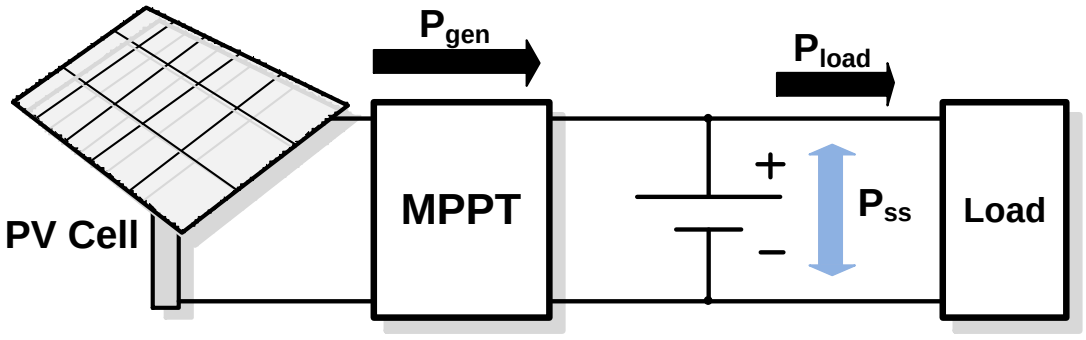

Fig. 3.2. Conventional ESS architecture in a standalone photovoltaic installation.

When incorporating multiple storage elements in a HESS, DC/DC converters become necessary to adapt the different voltage ranges of each storage technology and to control the power flow in the system. Different kinds of converter may be considered for this task. Generally, bidirectional converters are used since the storage elements need to be charged and discharged. However, unidirectional converters could also be considered. An example of these

architectures in the previous standalone photovoltaic installation is shown in Fig. 3.3. However, these are not the only architecture configurations, and some studies delve into this topic (Garayalde Perez 2020). Using bidirectional converters is usually more efficient, less complex, and better cost and size wise. For this reason, only bidirectional converters will be considered from now on.

In (Rahe 2016), the case of not using a converter is explored. In it, cells with different chemistries (lead-acid and LFP) are connected in parallel, directly to the DC bus of the installation. However, this is a passive hybrid system where the energy is balanced through non-controllable neither predictable parasitic impedances that, additionally, evolve in different ways as the batteries age. Because of that, this case is not going to be considered.

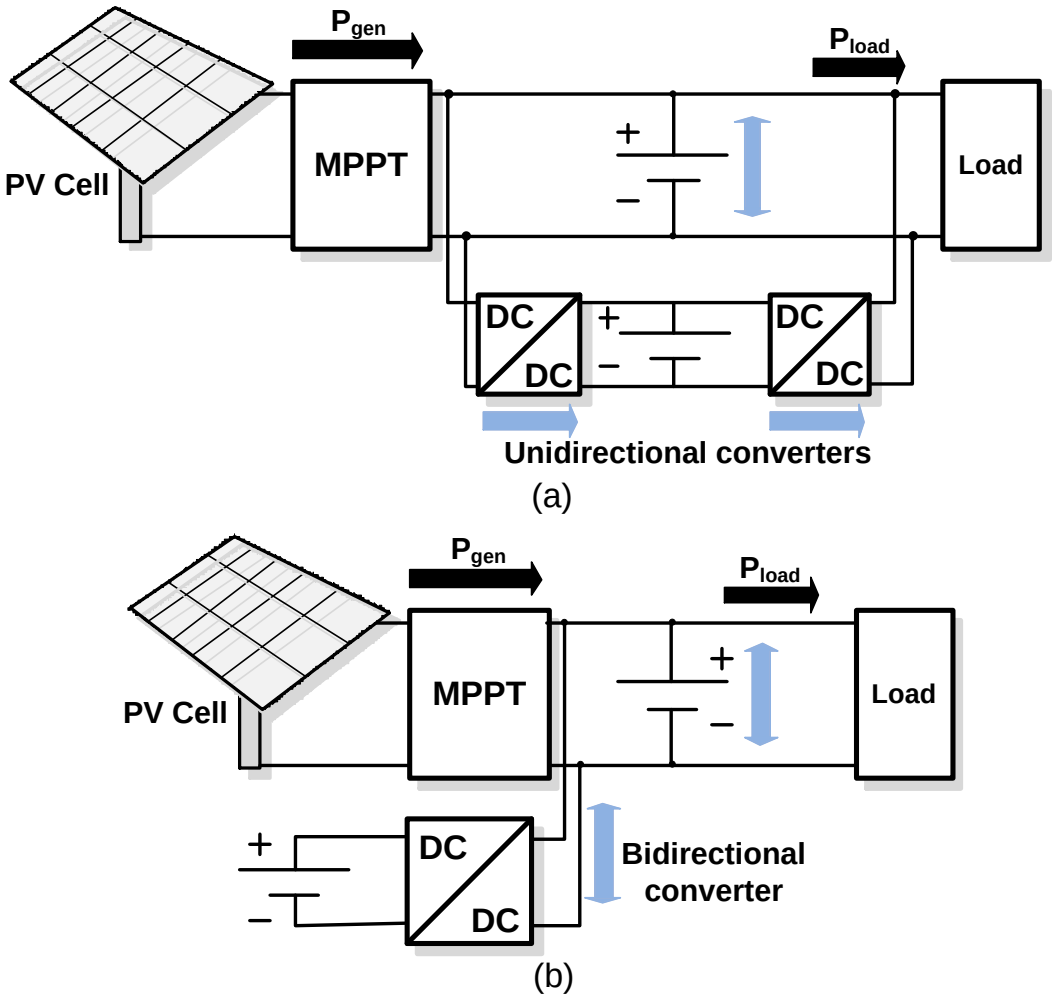

Fig. 3.3. HESS with unidirectional converters (a). HESS with a bidirectional converter (b).

It is necessary to differentiate the concept of HESS architecture from the topology of the converter. The topology of the converter is referred to as the physical electronic components (transistors, coils, capacitors, diodes ...) and their physical arrangement, connection, and mode of operation. Examples of topologies are conventional power electronics converters such as buck, boost, buck-boost, flyback, dual active bridge, forward, sepic, cuk...

From a power electronics point of view, several bidirectional conversion topologies may be used in these applications, both with and without isolation. Among the topologies without isolation, the most common option is the bidirectional synchronous buck-boost converter, see Fig. 3.4a. On the other side, if isolated topologies are considered, the most relevant bidirectional DC/DC converters for this application are the Dual Active Bridge (DAB) and the Series Resonant Dual Active Bridge (DAB-SRC), see Fig. 3.4b.

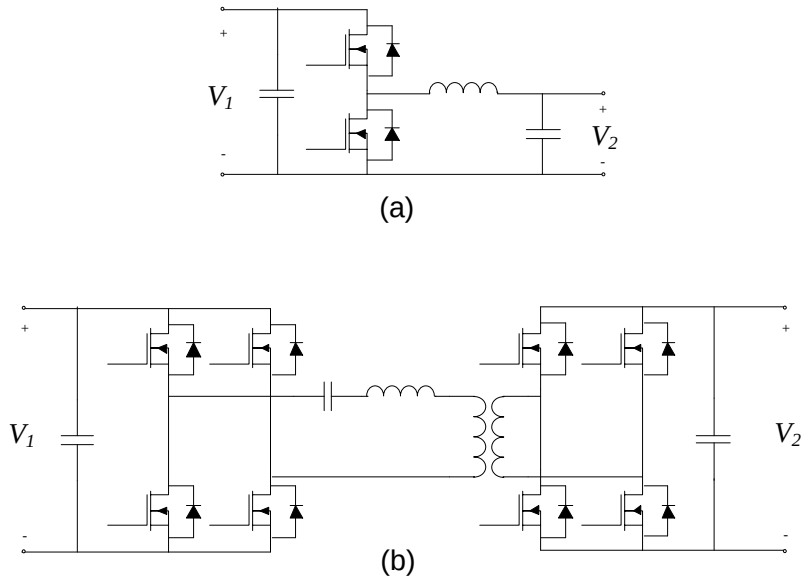

Fig. 3.4. Synchronous Buck-Boost converter (a). Dual Active Bridge-Series Resonant Converter (b).

On the other hand, HESS architecture is referred to as the electrical connections between the generation/consumption elements, the storage elements, and the converters. Multiple architecture configurations can be obtained from the combination of these elements. Fig. 3.5 shows the two most common architectures of a HESS comprising two storage elements: the cascaded architecture and the common DC bus architecture.

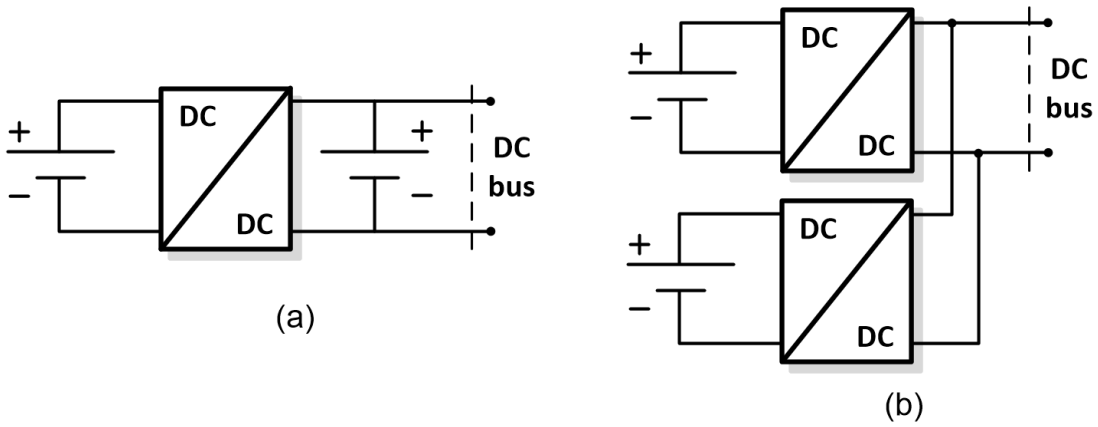

Fig. 3.5. Cascaded architecture (a). Common DC bus architecture (b).

New power electronics fields of research, such as Partial Power Processing, are developing new architectures of ESS by considering the storage element as an integral part of the converter (Artal-Sevil et al. 2020; Anzola et al. 2020). These new architectures are still in the research stage, but their application to HESS will be of interest in the forthcoming years.

3.2 Proposed HESS for PV installations

3.2.1 Overview

As introduced in section 2.5.1, HESS are proposed in this work to mitigate problems such as the Partial State of Charge in standalone photovoltaic installations. As a reminder, this problem is caused by daily cycles on a single, high-capacity battery. These partial cycles, when repeated in time, plunge the main battery into an intermediate SoC, never reaching a full charge state (Fig. 2.9). Due to this effect, when a period of several cloudy days starts, the battery is not fully charged and the system shuts down before its intended autonomy.

As a hypothesis, it is going to be assumed that a HESS with different storage technologies will be able to hold a higher SoC under this regime of repetitive charge and discharge patterns. As an initial approach to the problem, a very simple HESS, formed by two different storage sub-systems and controlled with a priority strategy (as described in section 3.1.2) is going to be selected. To help overcome the PSoC problem, the highest priority element should at least guarantee the energy of a daily cycle, so the other is not cycled with this daily rate. The control, technologies, and architecture of this initial system are described in the following sections.

3.2.2 Priority-based Energy Management Strategy

Compared to (Kim et al. 2014) and the priority-based strategy described in Fig. 3.1, the consumption in the application under study is steady and thus, the storage system does not need high power elements. There, the power density was the key factor when determining priority. On the contrary, this application does need high storage capacities so the HESS can last for days without irradiation.

In this case, the HESS will comprise two energy-oriented storage elements rather than power and energy-oriented elements. For the sake of simplicity, it has been decided that one of the storage sub-systems will always have a higher priority role whereas the other will always be secondary. The element with the

higher priority has been named the Cyclic Storage System (CSS) and the sub-system with the lowest priority, the Support Storage System (SSS). The CSS will fulfill a cycling role, whereas the SSS will act as bulk energy storage, providing the required energy if the CSS becomes completely discharged (see Fig. 3.6 where an example is depicted).

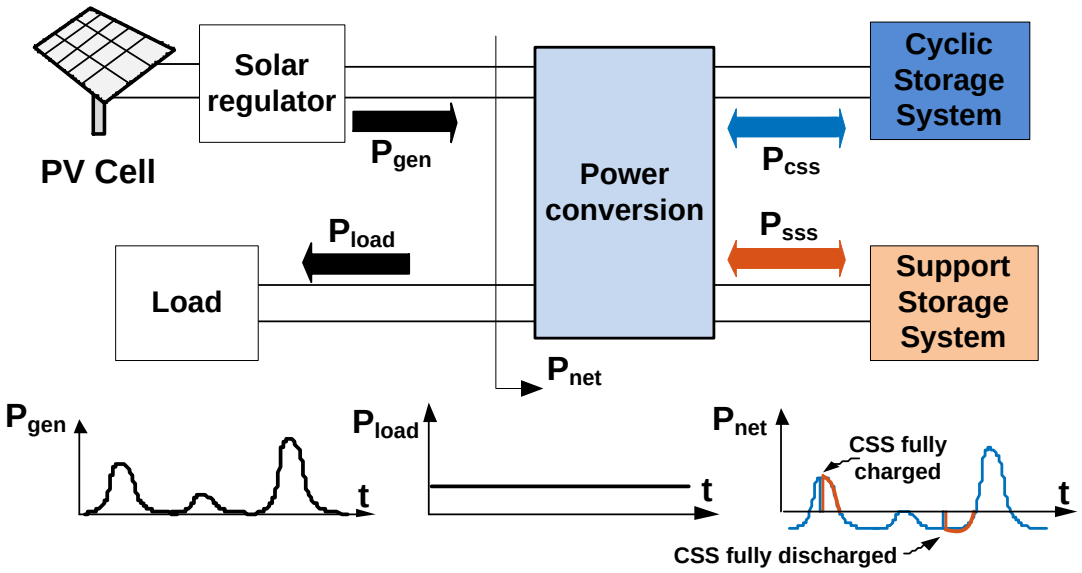

Fig. 3.6. HESS system in a standalone photovoltaic installation.

The net power input to the HESS in the figure is given by (3.6). Here P_{gen} is the power generated by the solar panels, and P_{load} is the power consumed by the load.

$$P_{net} = P_{gen} - P_{load} \tag{3.6}$$

The power conversion system implements the priority-based EMS by diverting the power flow to the sub-systems. As convention, the power flow towards the storage elements in the figure is considered positive when charging them. It can be appreciated how, when the CSS is fully charged, the SSS is charged with the extra energy. On the contrary, whenever the CSS becomes fully discharged, the SSS provides the power toward the load.

The multiple control modes and transitions of this specific strategy can be summarized in a state machine, see Fig. 3.7. Here, transitions are based on the net power input and the voltage of the batteries. The color code for the states in the figure is followed throughout the subsequent figures to indicate the states.

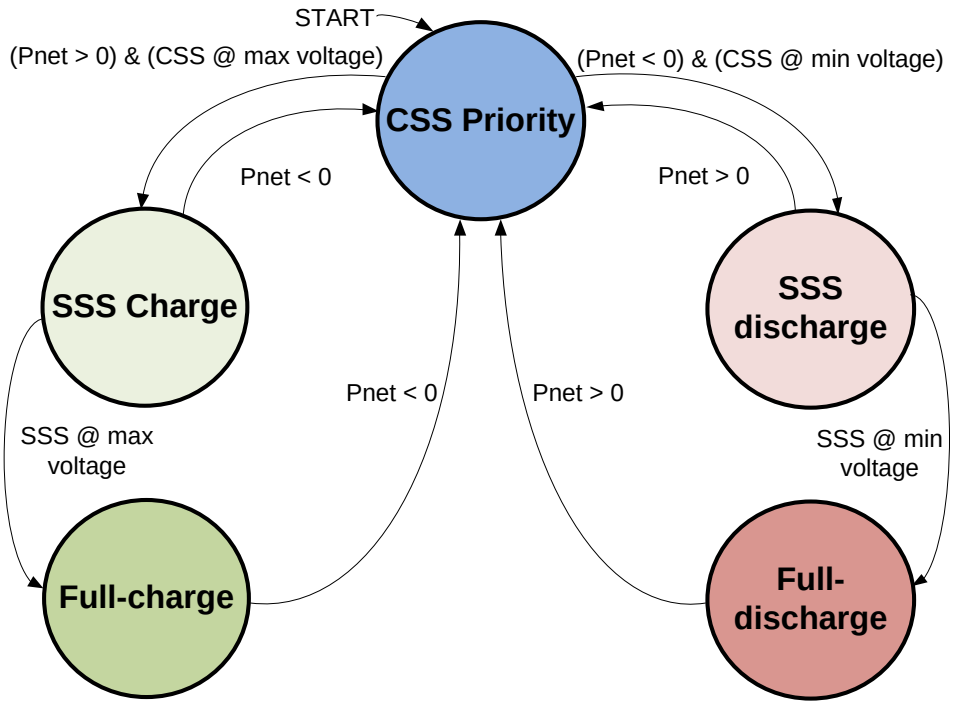

Fig. 3.7. Priority-based EMS state machine.

The different states are described below:

- **CSS priority**: This state is the regular operation mode. During the day, the CSS absorbs the input power, and during the night, it provides the power to the load. The SSS remains in a support role and is neither charged nor discharged. If the CSS is charged or discharged for a long period, it will eventually reach its maximum or minimum voltages. This will trigger a change in the state machine, leading to one of the following states.
- **SSS charge**: If the CSS reaches its maximum voltage and the net power input is positive, the HESS will start regulating the voltage of the CSS so it does not rise further. Thus, the current through it will start decreasing. There will be an excess of input power, which can be absorbed by the SSS. If this situation is held in time, the voltage of the SSS will eventually reach its maximum value, and the system will move into a Full-charge state. On the contrary, if the net power input becomes negative (for example, if sun irradiation decreases), the system will transit to the CSS priority state, so the CSS will start discharging.

- **Full-charge**: Both the CSS and the SSS are in a full-charge state, and there is a positive net-power generation. Voltages of both storage systems are regulated to their maximum value. Automatically, the solar regulator starts decreasing the generated power, generating only the power consumed by the load and the decreasing power absorbed by the batteries in the voltage-regulation phase. If the net power becomes negative, the system will transit to the CSS priority state, so and start a discharging phase.
- **SSS discharge**: If the CSS reaches its minimum voltage while discharging, the EMS will start regulating its voltage so it does not fall below that threshold. At this point, the CSS will provide less current than the one demanded by the load, and therefore, the SSS will inject the difference. Eventually, the CSS will be completely discharged and the SSS will be the only element powering the load, fulfilling its role as a support system. If during this state the net power input becomes positive, the system will transit to the CSS priority state, thus the SSS will stop discharging and the CSS will start charging. On the contrary, if the discharging mode is held in time, the SSS will eventually reach its minimum voltage and the system will transit to a Full-discharge state.
- **Full-discharge**: In this state, both the CSS and the SSS are at their minimum voltage levels, and the system can no longer provide energy to the application. The loads must be disconnected and the system will remain shut down until sun irradiation increases and the net power becomes positive. In this case, the system will transit to the CSS priority state, charging the cyclic system.

Simulations with these states and transitions are collected in the following figures:

- Fig. 3.8 shows the transition from the **CSS priority** state to the **SSS charge** state. In the beginning, the CSS is constantly cycling and increasing its state of charge. Eventually, the CSS becomes fully charged, and the SSS starts charging.

- Fig. 3.9 shows the transition from the **CSS priority** state to the **SSS charge** state and then to the **Full-charge state**. The CSS is constantly cycling. When it becomes completely charged, the system transits to the SSS charge state. But this time the SSS system is fully charged, so the HESS transits to the **Full-charge** state and stops absorbing energy,
- Fig. 3.10 shows the transition from the **CSS priority** state to the **SSS discharge** state. The CSS discharges completely along two days, and the SSS starts providing energy to the load (first red section). Afterwards, the system absorbs energy and transitions to the **CSS priority** state, but the CSS is then discharged again and the system enters again in the **SSS discharge** state (second red section).
- Fig. 3.11 shows a very low irradiation period, where the CSS is discharged completely and the system enters the **SSS discharge** state. The SSS system is also deeply discharged, and after some back and forth, the HESS finally reaches the **Full-discharge** state since the SSS becomes completely depleted.

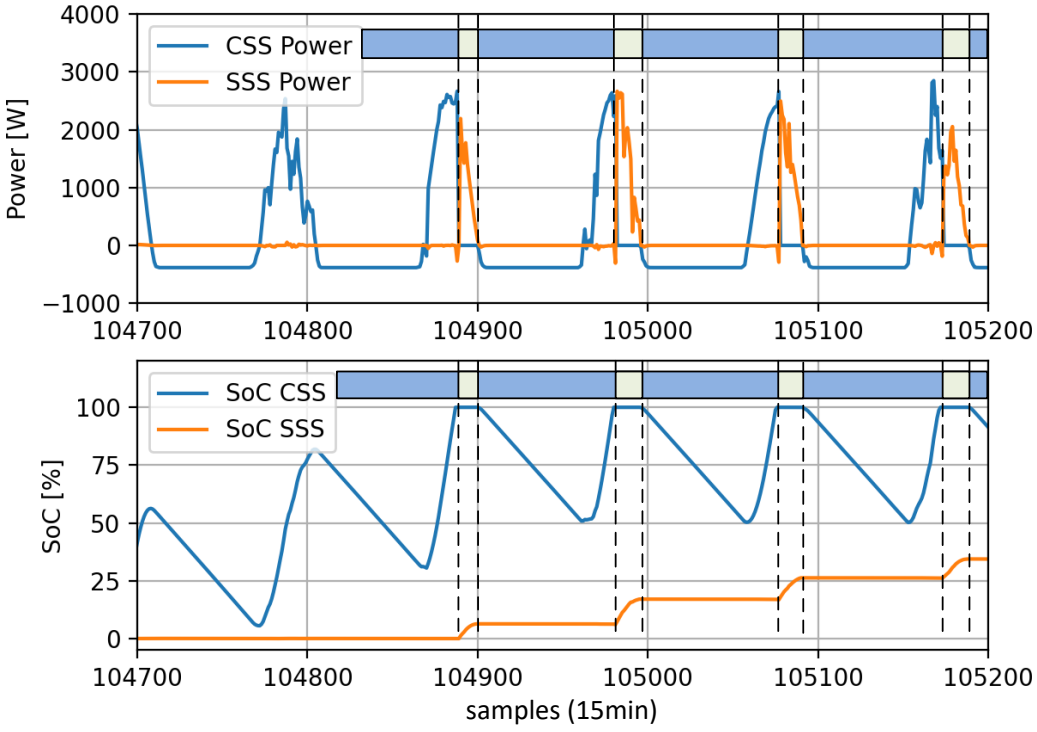

Fig. 3.8. The transition from CSS priority (blue) to SSS charge (light green).

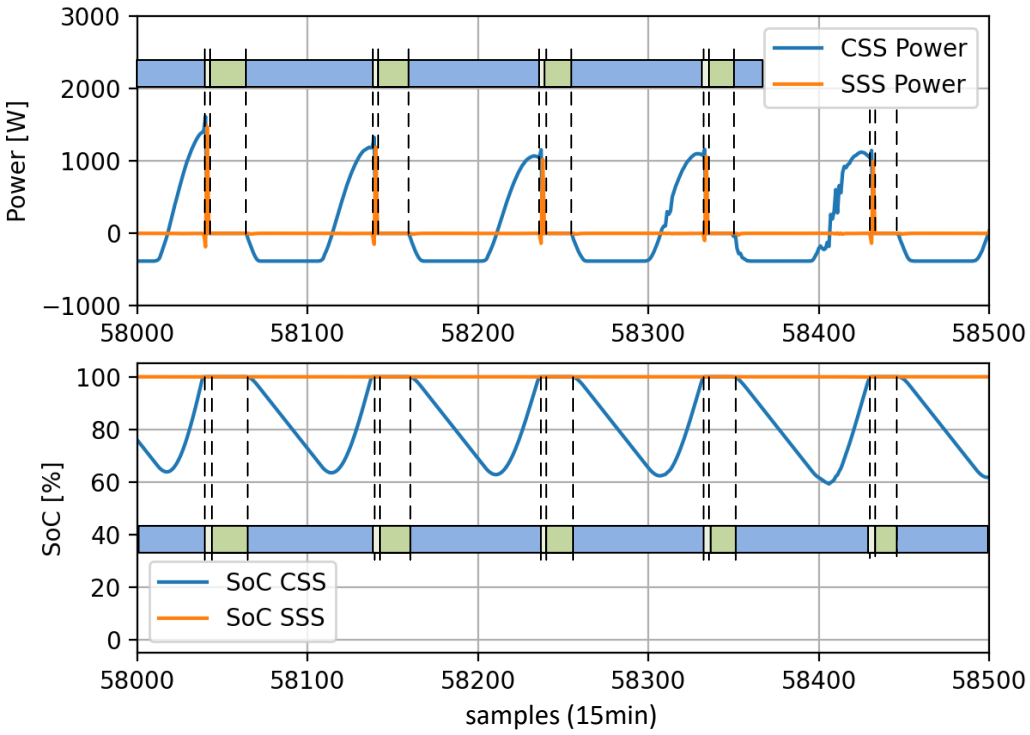

Fig. 3.9. The transition from CSS priority (blue) to SSS charge (light green) and Full-charge (green).

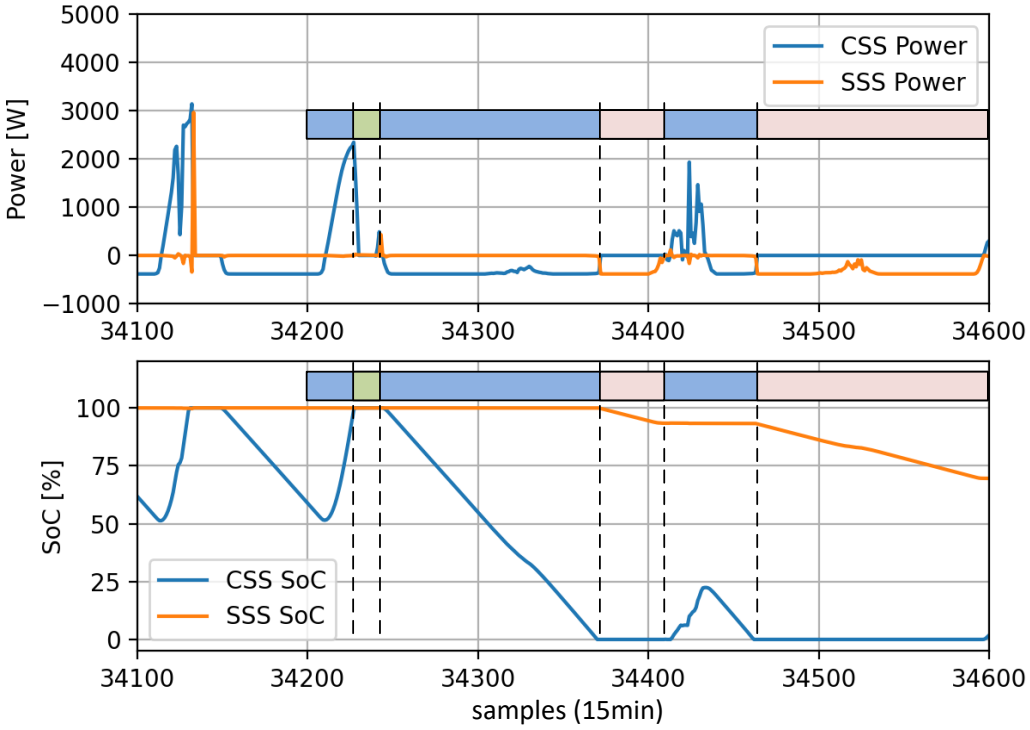

Fig. 3.10. The transition from CSS priority (blue) to SSS discharge (light red).

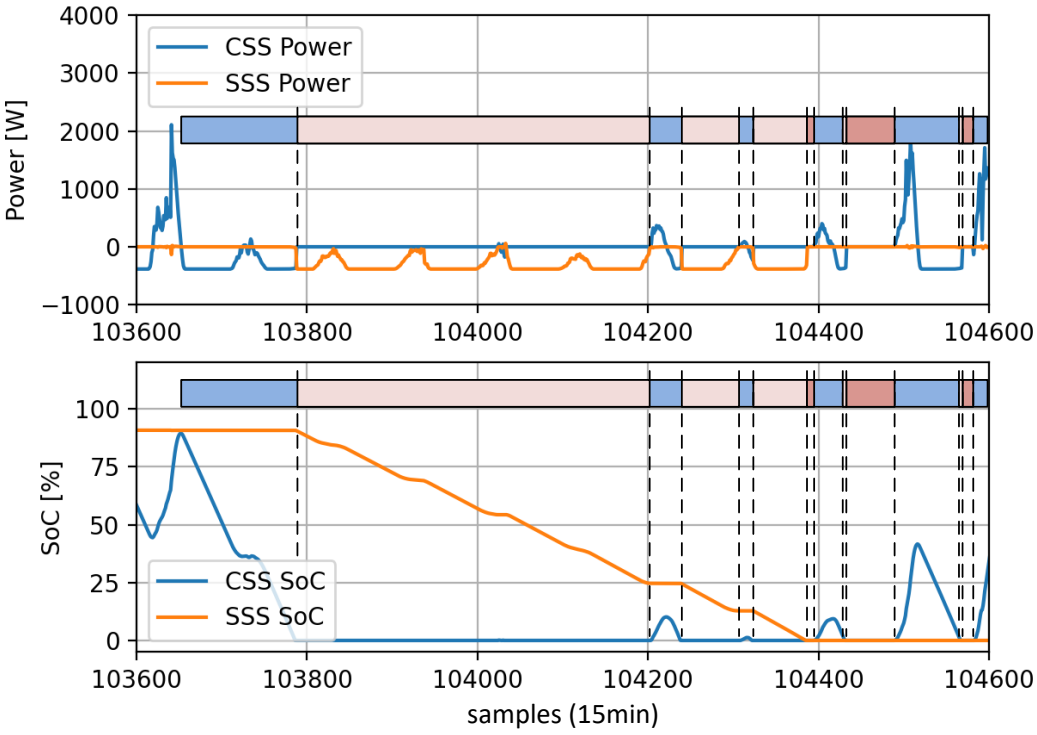

Fig. 3.11. The transition from SSS discharge (light red) to Full-discharge (red).

3.2.3 Storage technology

Based on these functional roles and requirements, the chosen technologies for the storage sub-systems are:

- LFP batteries for the CSS, due to their optimized characteristics for continuous cycling (Kim et al. 2014; Sepasi, Ghorbani, and Liaw 2015; Wang et al. 2016).
- VRLA batteries for the SSS. These batteries have good stationary characteristics (Kim et al. 2014; Burgess 2009; Exide Technologies 2012) but are much cheaper than lithium-ion packs, making a good technology for a large, bulk storage pack.

Thus, the selected EMS will give full priority to the charging and discharging processes in the LFP battery. Only when the LFP battery is discharged, the VRLA will start powering the load, and vice-versa. This role division fits with the features of each chemistry type (Kim et al. 2014), resulting in a higher overall SoC in the HESS and ensuring a higher resilience.

3.2.4 Architecture selection

As an architecture option for the HESS, the cascade architecture will be chosen (Fig. 3.5a). It is the simplest architecture from those introduced in the previous section, as it only features one power converter. This selection of architecture and chemistries allows installing a converter and an LFP battery pack in parallel to an existing VRLA battery in an installation, thus extending the useful life of already aged batteries. Furthermore, keeping the main lead-acid battery in parallel with the load makes the system more robust in case of failure of the DC/DC converter. This results in the architecture of Fig. 3.12.

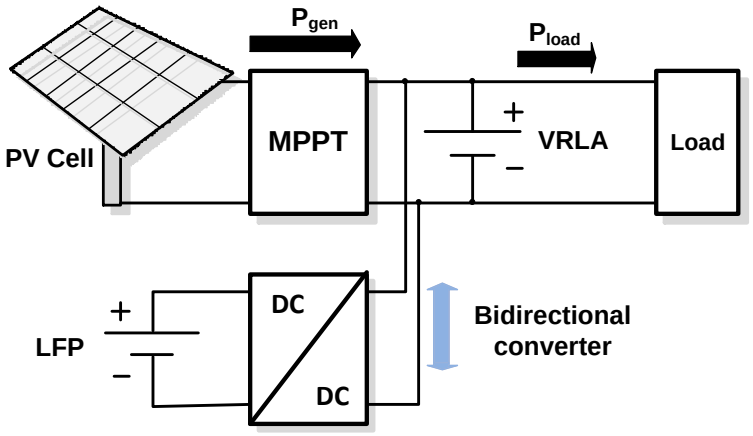

Fig. 3.12. Proposed HESS architecture.

As a first approximation, it is going to be considered that the solar charger (MPPT block) and the DC/DC converter are synchronized, so they act as a single control element. This is relevant in the case they both try to control the voltage in the VRLA battery with different voltage references, which may cause controllability issues.

3.2.5 Sizing

The sizing of a HESS is going to be compared with the sizing of a conventional ESS aimed for the same application. Sizing maps are going to be used for this purpose. To obtain these maps, both systems are going to be simulated for different values of their sizing parameters.

As introduced before, in an installation with a conventional ESS, the two coordinates of the sizing maps are the generation ratio g and the autonomy a

of the storage pack. As explained in section 3.1.1, when including one more storage system, an additional degree of freedom appears, the hybrid factor *f*. In this particular case, the definition in (3.7) is used, where *f* represents the fraction of the energy installed in LFP technology (E_{LFP}).

$$f = \frac{E_{LFP}}{E_{tot}} \tag{3.7}$$

Where E_{tot} is given by (3.8). Here, E_{VRLA} represents the nominal energy of the VRLA battery.

$$E_{tot} = E_{VRLA} + E_{LFP} \tag{3.8}$$

Therefore, the sizing maps depend now on three coordinates *(g, a, f)*. As before, multiple HESS will be simulated by varying these three parameters, obtaining the best design regions. These HESS are compared with a conventional ESS based on a single VRLA battery with the same E_{tot} as nominal energy.

The simulated hybrid system is shown in Fig. 3.13a, and the conventional system, in Fig. 3.13b.

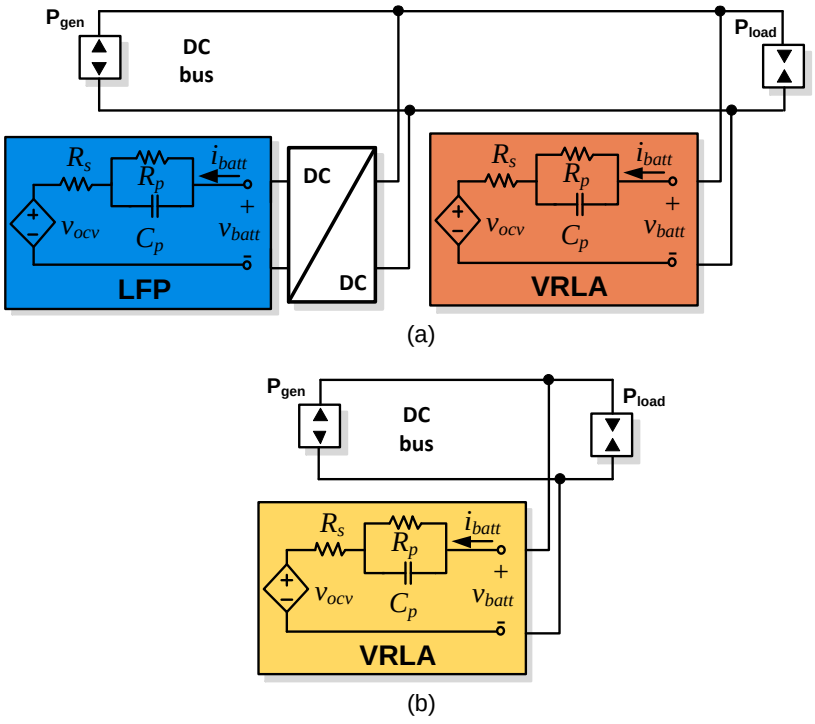

Fig. 3.13. Simulation scheme of the HESS (a). Simulation scheme of the conventional ESS (b).

Here, the load has been modeled as a DC power sink, and the PV source has been modeled by using solar irradiation data from an actual installation located in Aragón, Spain. The irradiation values have been sampled every 15 minutes during the years 2013, 2014, 2015, and 2016. The results of these simulations are considered valid for similar climates and latitudes.

The batteries are modeled as explained in Chapter 2. Each battery counts with many serialized cells (n_{cells}), and the SoC – v_{ocv} curves for one cell of each chemistry are shown in Fig. 3.14. The parameters of the Equivalent Circuit Model (R_s, R_p, and C_p) are shown in Table 3.1.

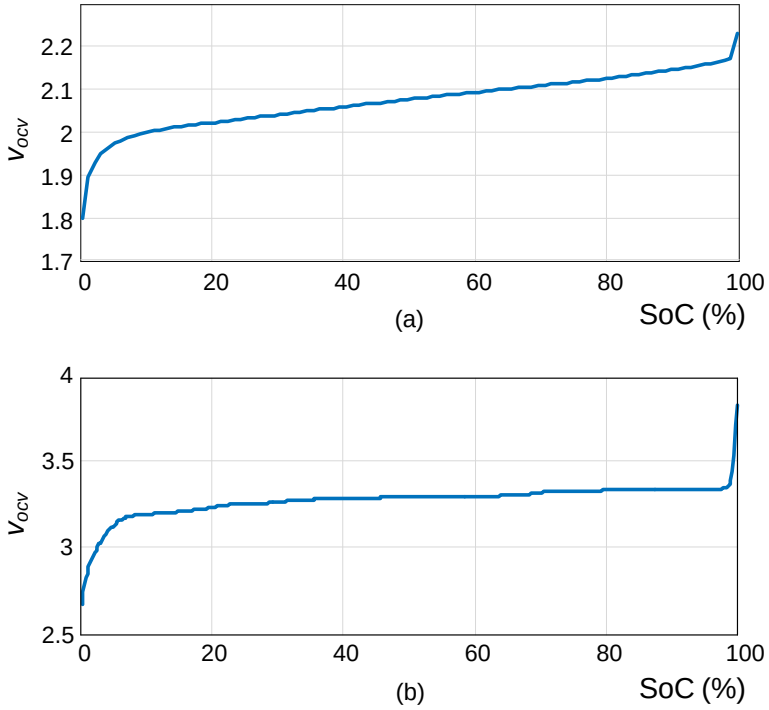

Fig. 3.14. SoC– v_{ocv} curve for one cell of VRLA technology (a). SoC– v_{ocv} curve for one cell of LFP technology (b).

Battery	Parameter	Value	Unit
VRLA	R_s	65	mΩ
	R_p	50	mΩ
	C_p	9.6e4	F
	n_{cells}	24	-
LFP	R_s	20	mΩ
	R_p	50	mΩ
	C_p	1.8e4	F
	n_{cells}	8	-

Table 3.1. Equivalent Circuit Model parameters for each chemistry.

Fig. 3.15 shows an example of the evolution of SoC in the batteries, during several months of simulation for one of the design points. It can be seen how the VRLA battery in the HESS (orange) only starts discharging when the LFP battery (blue) has a minimum SoC, hence fulfilling its role as the emergency battery. On the other side, the VRLA battery in the conventional ESS (yellow) is constantly charging and discharging due to the daily cycle, never achieving a full-charge state.

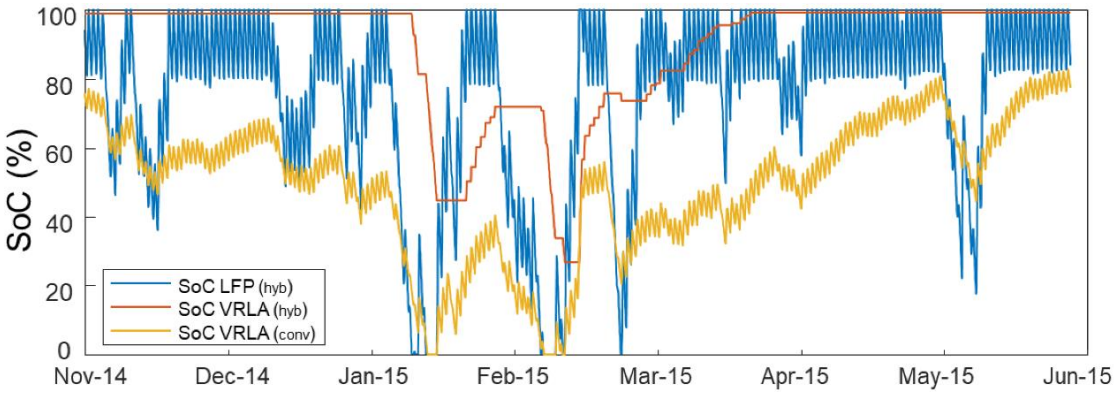

Fig. 3.15. Simulation example for several months in one design point.

The simulation results for the conventional ESS sizing are shown in Fig. 3.16. Since this system is only formed by VRLA batteries ($f = 0$), only one contour plot is obtained. The map represents in the abscissa axis the autonomy, and in the ordinate axis, the generation ratio. The Loss of Load Probability (LLP), as defined in Chapter 2, is shown through a color scale. This magnitude expresses the percentage of time that the system has been shut down along four simulation years.

The map presents the same asymptotic behavior as the one seen in Chapter 2. A label at $a = 150$h and $g = 9.5$ exhibits a design case that leads to an LLP of 1.827%, which would be quite high in a real application.

The simulation results for the HESS system contain one contour plot for each value of the hybrid factor *f*. Individual maps for different f values are shown in Fig. 3.17, representing in the abscissa axis the total autonomy of the system (autonomy of LFP + autonomy of VRLA) and the generation ratio g in the ordinate axis. As before, the contour plot for LLP is shown in color.

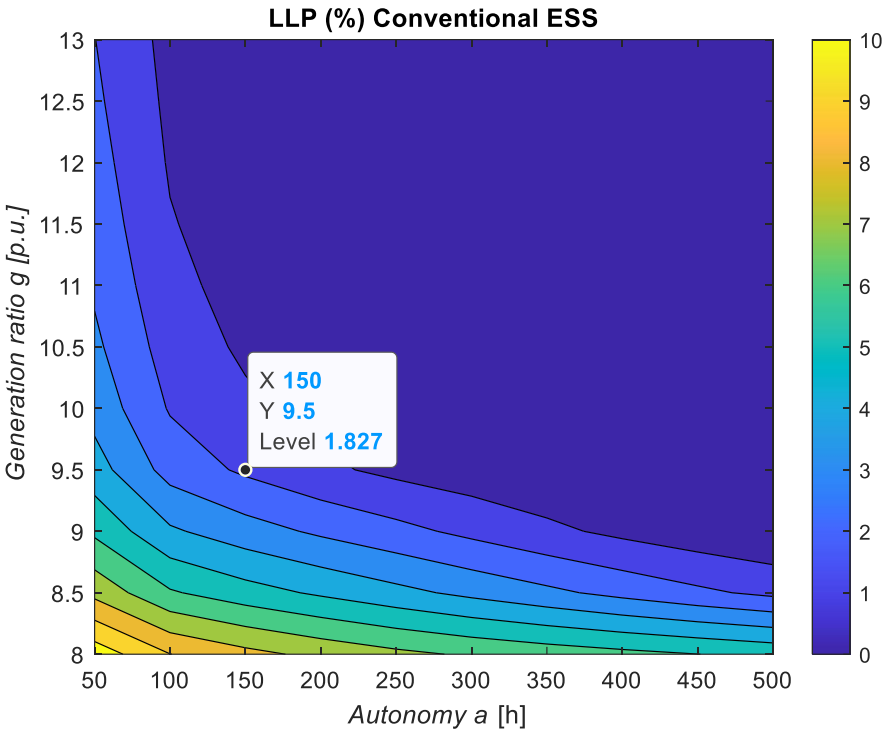

Fig. 3.16. Conventional ESS sizing map (f=0%).

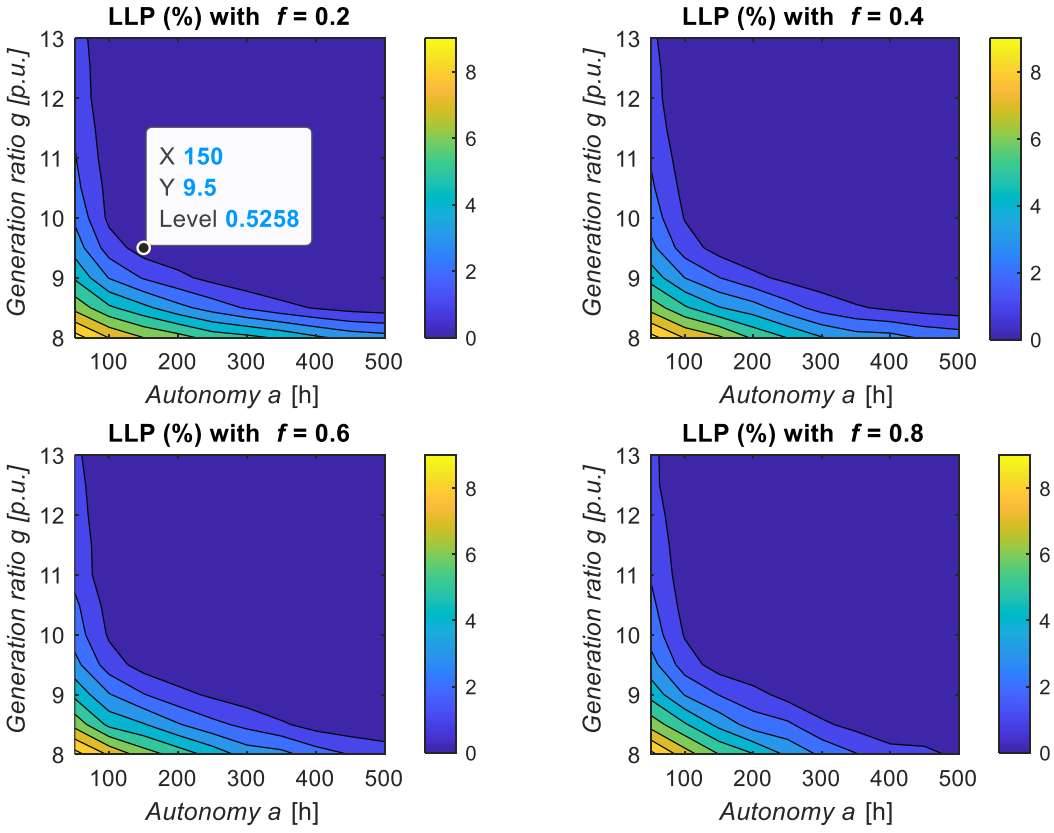

Fig. 3.17. Loss of Load Probability for HESS with different values of the hybrid factor f.

It can be seen how the different maps have similar behaviors, regardless of the value of the hybrid factor (the differences between 0.2, 0.4, and 0.6 are small). This will be analyzed further on. They also share the same asymptotical behavior as the contour plots for the conventional ESS, but with lower LLP values in general. As it can be observed at the map with $f = 0.2$, the design case $a = 150$h and $g = 9.5$ leads to an LLP value of 0.526%, three times better than the value obtained in a conventional ESS (LLP = 1.827%).

Due to the small differences in performance that appear in systems with different *f*, it is interesting to analyze their economic cost, so that the tradeoff between increasing the cost and the performance is visualized. The cost coefficients required for the comparative are shown in Table 3.2.

Element	Cost	Unit
Solar panel	2	*€/W*
VRLA Battery	0.2	*€/Wh*
LFP Battery	0.5	*€/Wh*

Table 3.2. Cost values used in the performance analysis.

The tradeoff *(Tr)* of the installation has been defined with the expression in (3.9). This variable compares the improvement in LLP versus the increase in cost. Here LLP_{HESS} and LLP_{ESS} are the LLP values obtained for the HESS and the conventional ESS respectively, and $Cost_{HESS}$ and $Cost_{ESS}$ are the cost values obtained for the HESS and the conventional ESS respectively.

$$Tr = \frac{\frac{1}{LLP_{HESS}/LLP_{ESS}}}{Cost_{HESS}/Cost_{ESS}} \tag{3.9}$$

For simplicity, to analyze the impact of the hybrid factor, a specific scenario has been selected. The performance of a conventional ESS in this scenario is going to be compared to the performance of a HESS with different hybrid factor. Specifically, the values of $g = 9.5$ and $a = 150$h have been selected as the simulation scenario. By varying *f*, the results in Fig. 3.18 are obtained.

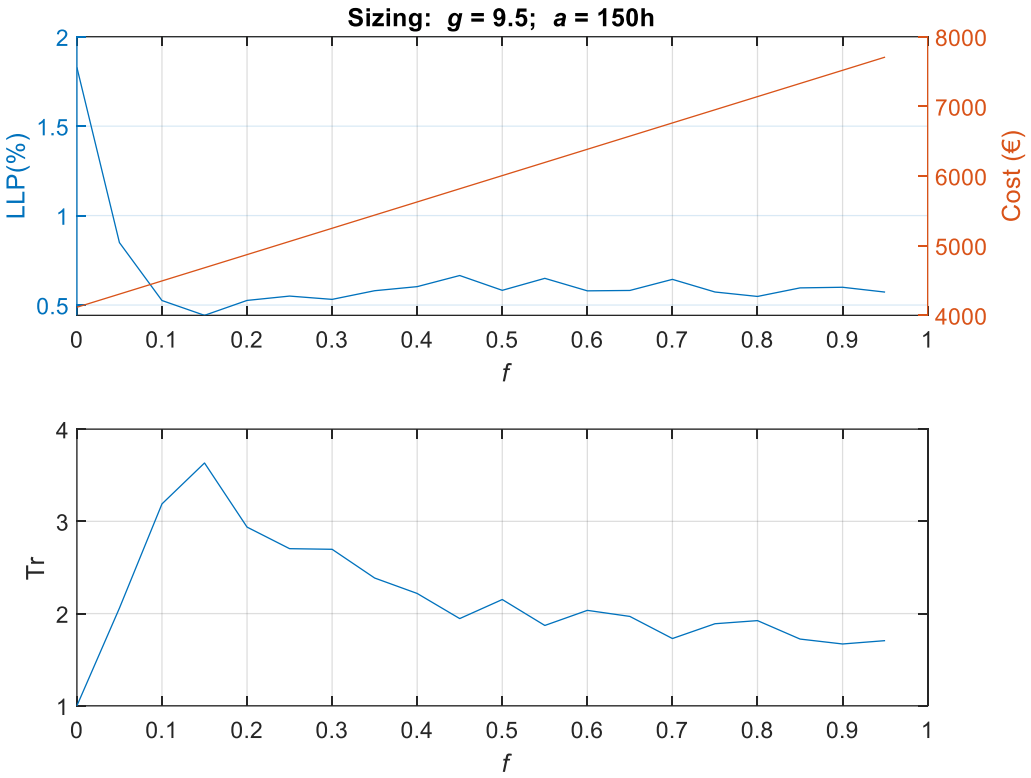

Fig. 3.18. Multiple f *values for a specific* g *and* a *scenario. LLP (%) and the cost of the system (top). Tradeoff (bottom).*

It can be seen how the benefits (understood as low Loss of Load Probability) are similar beyond $f = 0.1$, becoming stable past this value. With this fraction of LFP technology and the priority-based EMS, the LFP battery absorbs the full daily cycle, while the VRLA battery fills its role as an emergency battery, being ready for cloudy periods and thus, minimizing energy shortage. Further increasing f (thus, the LFP battery size) also increases the cost of the system, which deprecates its tradeoff, since the performance is already optimal. In this case, the best performance and tradeoff are found in a hybrid factor of $f = 0.15$, but this optimal value may be different in other scenarios.

This criterion for selecting the optimal hybrid factor can be applied to the whole design space, see Fig. 3.19. These results are interesting since they allow us to obtain conclusions about the optimal HESS configuration in different sizing regions. The results of the LLP with the optimal hybrid factor in each design case can be represented in the same contour plot, see Fig. 3.20, where four different regions are highlighted.

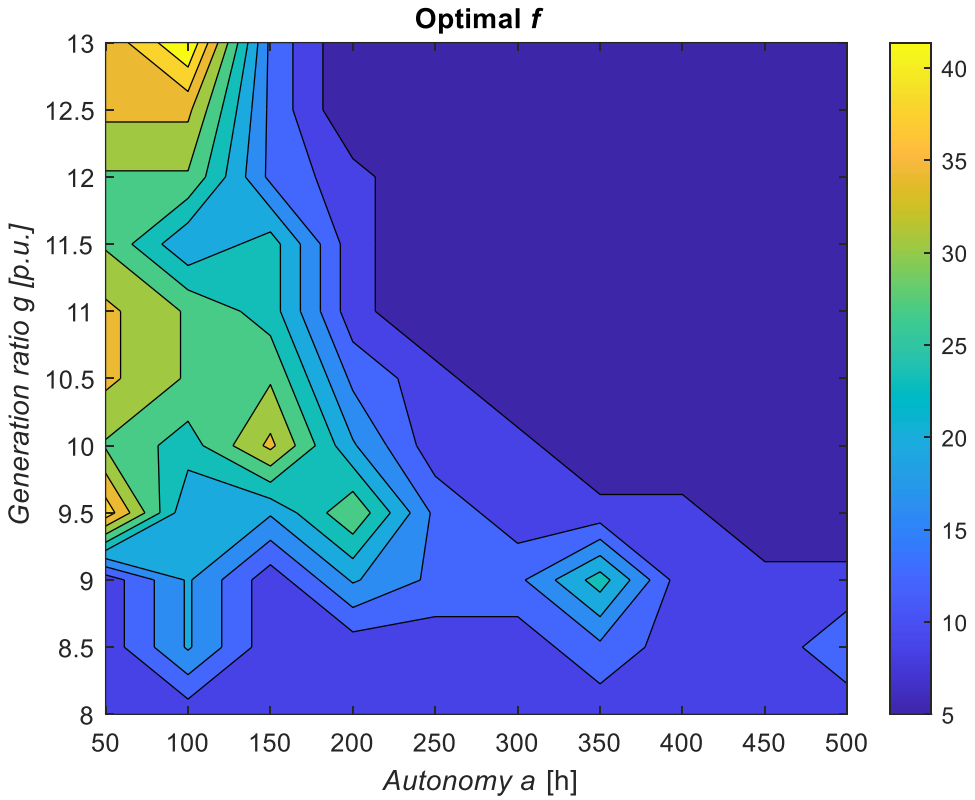

Fig. 3.19. Optimal f *sizing map.*

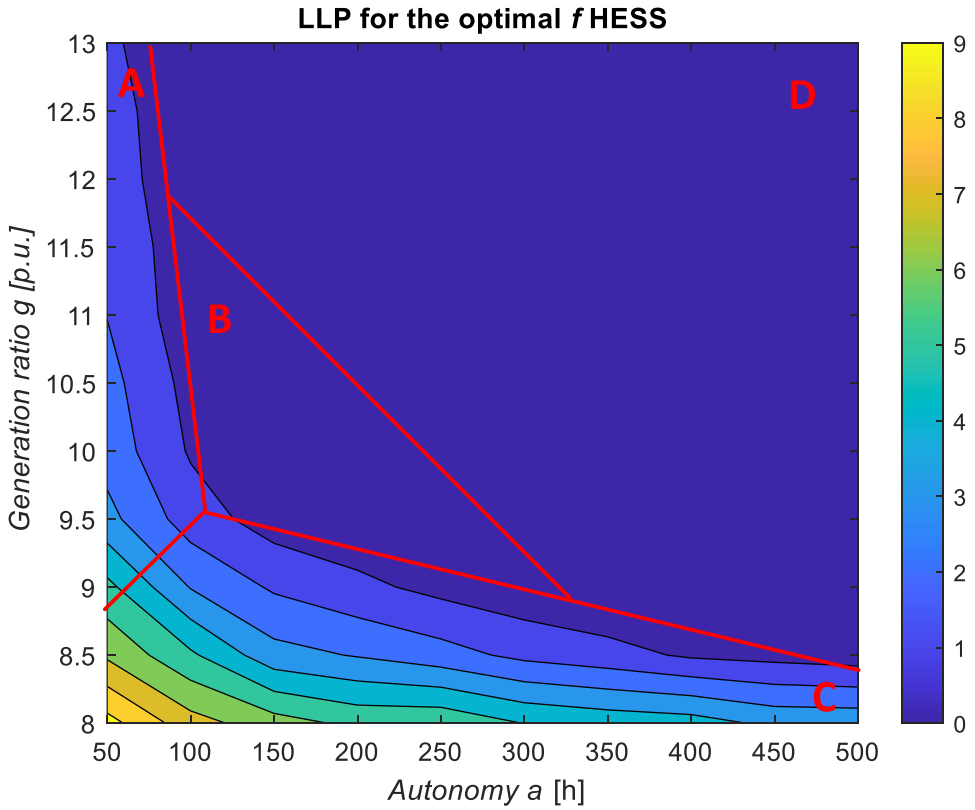

Fig. 3.20. Optimal LLP map.

The optimal f values shown in Fig. 3.19 can be overlayed with the A, B, C, and D sizing regions, presented in Chapter 2.

- **Region A** (below 100h of Autonomy), is a region with bad performance. Here, the optimal f tends to be higher ($0.3 \sim 0.4$), since, in this region, a better battery is required to increase the performance. This can be achieved in terms of quantity (adding more autonomy as seen in Chapter 2) or in terms of quality (better chemistry, more LFP), thus a higher hybrid factor is relevant.
- **Region B** (intermediate a and g) represents an interesting design region, where the system can operate properly without being oversized. Here, the optimal f value is between $0.15 \sim 0.3$. The example of the previous pages ($a = 150$h, $g = 9.5$) belongs to this region.
- **Region C** (low g) is also a region of bad performance. Here, the optimal f is more or less constant, around 0.1. In this region, the quality of the battery is not as relevant as the solar panels, so f is not as relevant and it remains low.
- **Region D** is a region of good performance, but a bad tradeoff, since the system is oversized. In this region, good battery quality is not needed, since the system already operates properly due to oversizing. For this reason, increasing f only increases the cost and worsens the tradeoff, thus, the optimal f is minimum.

In conclusion, it has been determined that the best performance-cost tradeoff in the sizing of these systems is found between 100 to 200 hours of autonomy, and with a g in the range of 9 to 10.5. In this cost-effective design region, the optimal hybrid factor f (the fraction of the battery pack that should be in LFP technology) is in the range of 0.15 to 0.3.

3.2.6 HESS impact on energy absorption and SoC

Until now, the proposed Hybrid Energy Storage System has been analyzed in terms of its sizing parameters and sizing figures of merit such as Loss of Load Probability. However, its performance in terms of energy can also be analyzed.

In this regard, two figures of merit are going to be obtained: the increment in the steady State of Charge of the HESS and its resilience.

The process is similar to the Loss of Load Probability simulations in previous sections: a basic HESS (Fig. 3.13a) is going to be compared to a conventional ESS, see Fig. 3.13b. Both systems will have the same overall storage capacity and will be connected to the same power generation and consumption patterns. In this case, the analysis will be focused on the impact of the HESS of the storage elements. Therefore, the generation power ratio g is fixed (specifically, $g = 9.5$ throughout the simulations). The hybrid factor f will vary in the range [0.05, 0.95] and the autonomy a, in the range [50, 500] hours.

State of Charge in a HESS

The State of Charge of a Hybrid Energy Storage System can be defined as the sum of the individual SoC of each storage element i, weighted by their hybrid factor f_i, as in equation (3.10).

$$SoC_{HESS}(t) = \sum_{i=1}^{N} SoC_i(t) \cdot f_i \tag{3.10}$$

To evaluate the benefits of the hybrid system, the Steady SoC Improvement (SSI) can be defined as the additional average SoC that the HESS achieves compared to the conventional ESS. The steady SoC improvement can be obtained as in equation (3.11), where $\overline{SoC_{ESS}}$ is the average SoC of the battery in the conventional installation and $\overline{SoC_{HESS}}$ is the average SoC of the HESS over the studied period.

$$SSI\ (\%) = \overline{SoC_{HESS}} - \overline{SoC_{ESS}} \tag{3.11}$$

This metric has been studied for multiple sizing scenarios, with different autonomies and hybrid factors in the HESS. These different cases have been simulated and the resulting contour plots of SSI are shown in Fig. 3.21. Please note that in the figure the y-axis represents the hybrid factor f rather than the generation ratio g.

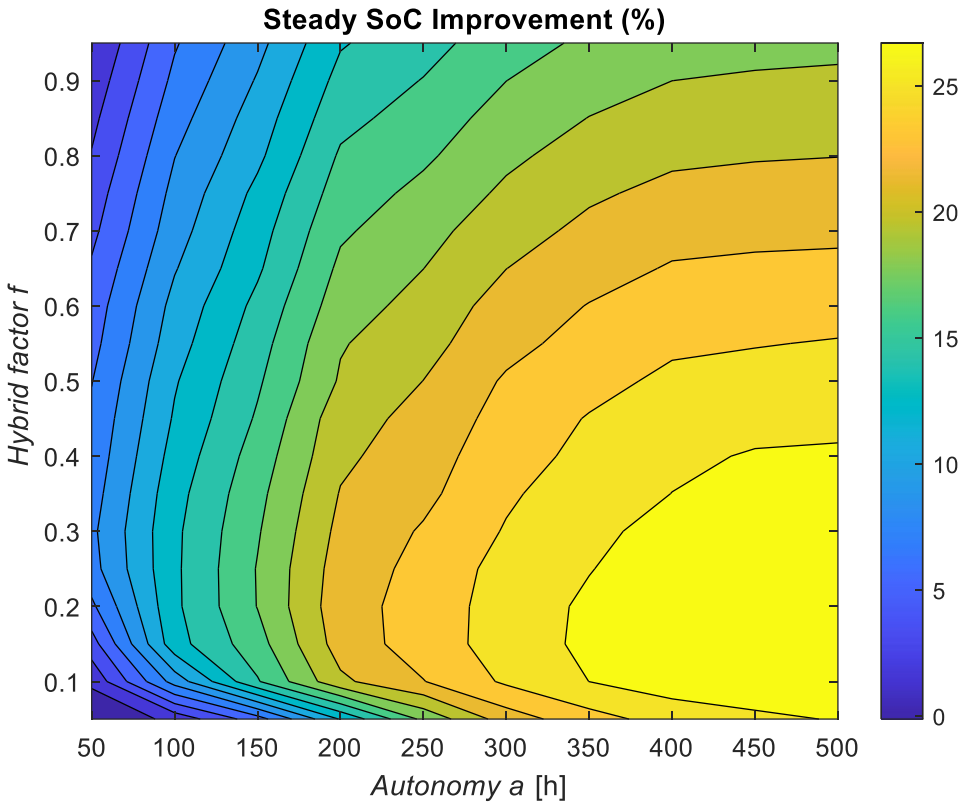

Fig. 3.21. Steady SoC Improvement map for different autonomies and hybrid factors.

For each autonomy value, the best hybrid factor f that ensures maximum energy absorption from the PV source is located from 0.15 to 0.30. Beyond this value, the SoC does no improve (even though the HESS has a larger LFP battery thus, becomes more expensive). This is a very interesting result, as it establishes that a hybrid VRLA + LFP system performs better than a full, single-pack LFP system (which has a much higher cost).

This can be explained as follows: in a single battery system, any of the studied chemistries (VRLA or LFP), suffers from the Partial State of Charge effect (PSoC), although it has less impact on the LFP battery due to its better cycling behavior. In a hybrid system with a priority-based strategy like the one implemented, only the LFP battery suffers the PSoC. With a medium-low hybrid factor, the LFP battery is small, so the impact of this effect on the overall State of Charge is small.

On the contrary, if the system is designed with a higher hybrid factor, the LFP battery has a larger capacity, and thus the overall impact of the PSoC

effect is higher. While there is still an important benefit over the conventional Energy Storage System, the benefits of the Hybrid Energy Storage System become less evident.

Resilience

In the framework of this work, resilience is defined as the speed at which the system recovers from a low-energy state. This parameter is relevant to characterize de behavior of the ESS (hybrid or not) when several periods of low irradiation are concatenated; after a first low irradiation period, the ESS must recover quickly a high SoC to be prepared for an eventual second low irradiation period.

To characterize this feature, the irradiation profile shown in Fig. 3.22 (recorded during the fall of 2015), is going to be used as irradiation input to test different HESS cases. This pattern shows a long period of cloudy days followed by a long period of sunny days. Thus, it is very convenient for analyzing to what degree the HESS is capable of recovering more energy after the low irradiation period. Note that the available input power is the same for both installations, but they will not absorb the same amount of energy.

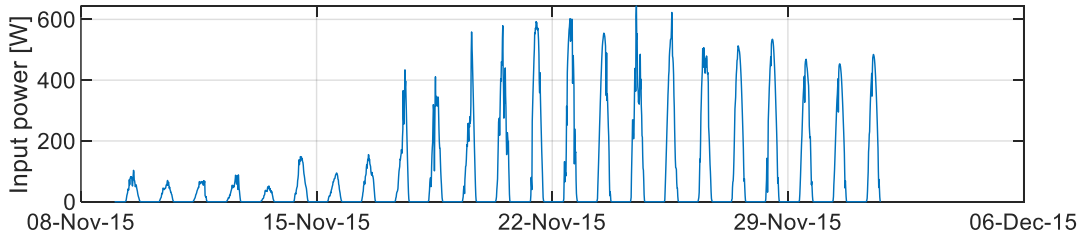

Fig. 3.22. Available input power (pattern used for obtaining Resilience).

The resilience of each installation has been obtained as the energy absorbed after the stressful event, during the following 15 days. The Resilience Improvement (RI) has been defined as in (3.12), where $E_{HESS15days}$ and $E_{ESS15days}$ are the aforementioned energy quantities integrated during the 15-day window.

$$RI\ (\%) = \left(\frac{E_{HESS_{15days}}}{E_{ESS_{15days}}} - 1\right)\ x100 \tag{3.12}$$

The map of the Resilience Improvement for different sizing parameters is shown in Fig. 3.23. In the low autonomy region, below 200h, the HESS can recover up to 45% more energy after the low irradiation period, increasing the reliability of the whole system (in the event of a second cloudy period, the HESS would have a higher SoC).

However, above the 200h threshold, the HESS improvement is not so evident, although is still better. For higher autonomy values, the improvement stabilizes around 5-10% with hybrid factor values above 0.25. For lower hybrid factor regions, the improvement decreases to the 0-5% range.

The general conclusion is that higher hybrid factors tend to higher resilience, although the impact is much noticeable in smaller HESS (with low autonomy) rather than in large systems.

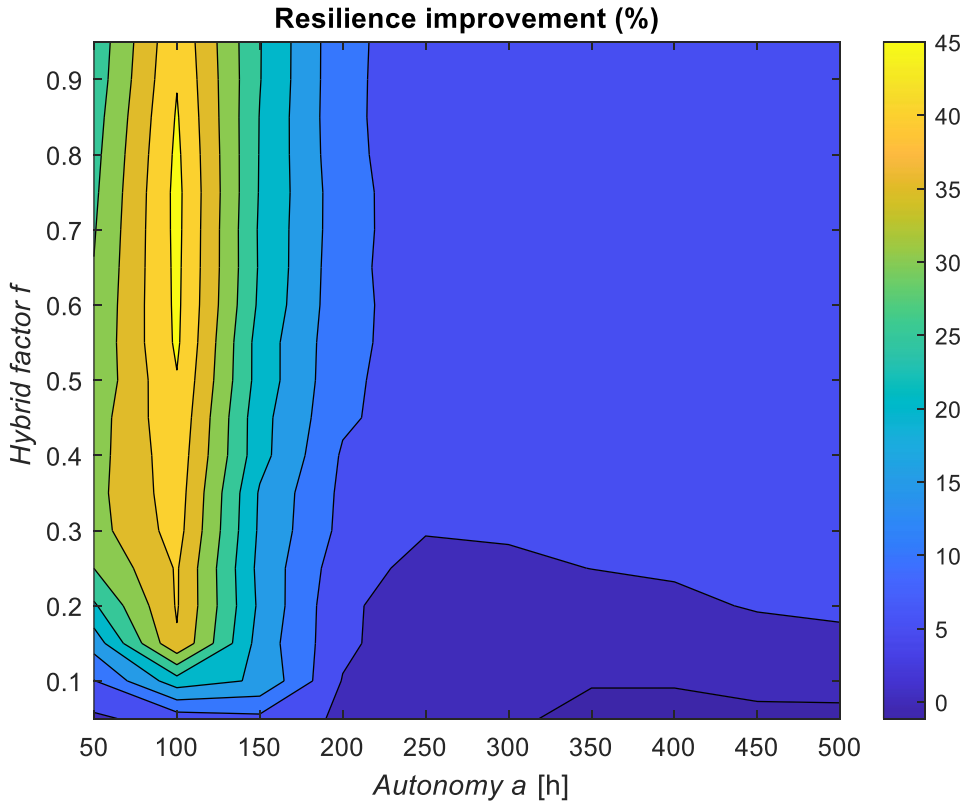

Fig. 3.23. Resilience improvement in the design region.

3.2.7 Hybridization power

The simulations in the previous sections have considered an ideal DC/DC converter for the HESS, with unit efficiency and no power limit. However, the power rating of this converter, if restrictive, may be relevant in the application since it will limit the power input towards the LFP battery. In general, the power rating will be very application-dependent, and will depend on the power rating of the panels, the power of the load, the Energy Management Strategy…

As an original contribution, this thesis is going to study a first approximation to the rated power of the DC/DC converter in this application. To obtain a first estimation of this design value, a hybrid system in the recommended design region ($a = 150$h, $g = 9.5$, $f = 0.2$) has been simulated, but with different values in power limit (P_{limit}) of the DC/DC converter. These values have been normalized to the power consumption of the site (P_{load}). The average SoC of the hybrid system ($\overline{SoC_{HESS}}$) throughout the simulations has been used as a figure of merit to determine the optimal power rating.

The results are shown in Fig. 3.24, where it may be appreciated how the average SoC reaches a plateau when the power limit of the converter rises above three times the load power.

Even though these results are scenario-dependent (since they also depend on the sizing of the rest of the elements), this ratio is going to be considered as a starting guideline for converter design in future sections.

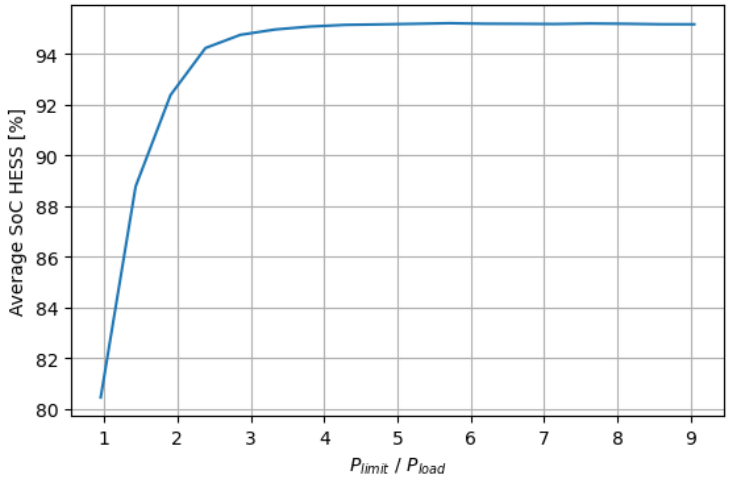

Fig. 3.24. Average SoC in the HESS vs Power limit of the DC/DC converter.

3.3 Experimental tests

3.3.1 Location of test sites

As introduced in Chapter 1, this work has been supported by the Confederación Hidrográfica del Ebro (CHE, Ebro river basin management entity), who has provided data from several remote ESSs facilities fed by photovoltaic sources. Some of these installations are going to be used to test the HESS concept.

Out of their 333 different sites, two facilities were selected in the so-called "Canal de Aragón and Cataluña", near the town of Binéfar, shown in Fig. 3.25a. Specifically, the sites have the numerals C092 (41°52'57.1"N 0°20'37.0"E) and C095 (41°52'02.0"N 0°22'15.1"E). Both installations count with the same equipment and functionalities, and were built at the same time. Besides, they are separated by only 2.85km, so their temperature and solar irradiation conditions are similar. Thus, a similar SoH can be expected in their batteries.

The HESS was installed on site C092. The DC/DC converter and the LFP battery pack were new, and it was decided to keep the existing VRLA battery as the emergency storage element. The installation also counted with communications and logging equipment to collect the data. The details of the converter are described in forthcoming sections. On the other hand, installation C095 was monitored as a reference ESS. Data for both installations was reported to a database system at the CHE facilities in Zaragoza. Fig. 3.25b shows an aerial view of the location of both installations.

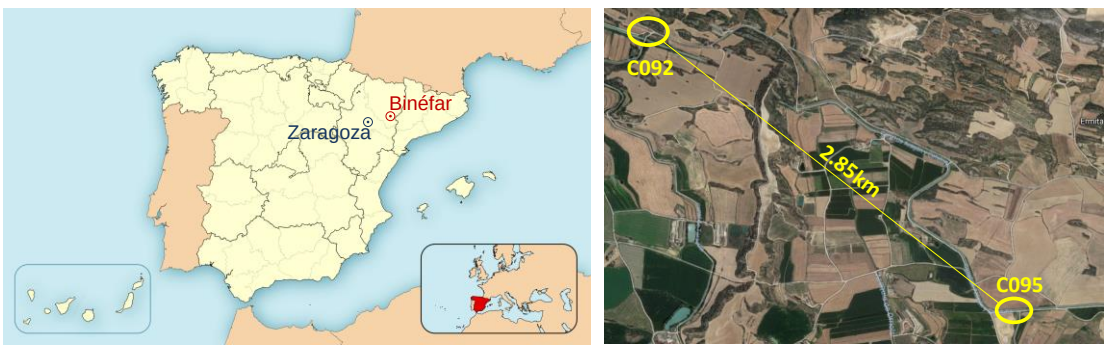

Fig. 3.25. Location of the town of Binéfar and the city of Zaragoza (top). Aerial view of the location of HESS and ESS installations to be compared (bottom).

3.3.2 DC/DC converter design

A bidirectional DC/DC power converter was designed to test the HESS. The specific details of the design of the converter are not a goal of this dissertation, since it has been developed as a means to implement the Energy Management Strategies and to test the HESS concept. However, to clarify the concept, some of the remarks of the design are gathered here.

The power consumption of the installation had been rated in 84W DC. Following the estimation in section 3.2.7, the DC/DC converter should have a nominal power of roughly 3 times this value (252W). To operate with some margin, it was decided to design the converter with a rated power of 300W. The converter did not require being isolated, so the bidirectional synchronous buck-boost topology was selected. To reduce the output ripple, an interleaved topology was implemented (Fig. 3.26). This is an evolution over the scheme in Fig. 3.4a with three parallel half-bridges, whose PWM carriers are phased 120º to reduce ripple at the output (Mohan, Undeland, and Robbins 2003).

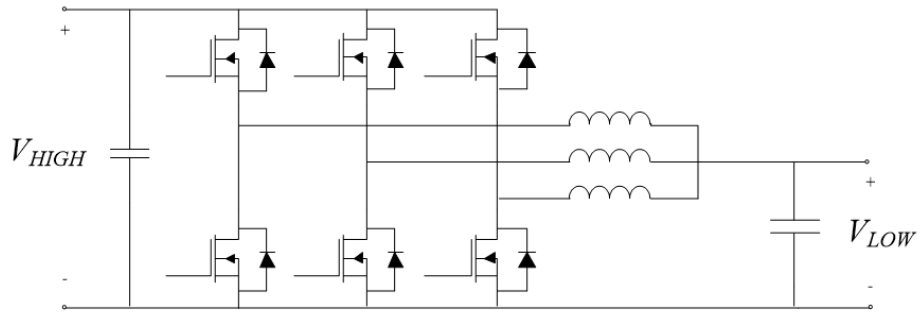

Fig. 3.26. Synchronous Interleaved Buck-boost converter.

The converter was initially designed using the EPC9033 GaN modules from EPC. Each of the modules consists of a GaN half-bridge and its drivers. The control algorithm for the converter and the EMS was programmed in a Cortex-M7 processor, using the digital platform Discovery STM32f746G by ST. Pictures of the converter and its measured efficiency are shown in Fig. 3.27. The connection scheme is shown in Fig. 3.28. The converter measures voltages (v_{LFP}, v_{VRLA}) and currents (i_{LFP}, i_{VRLA}) through both batteries as well as the overall current (i_{net}) exchanged with the installation. This data is required to implement the EMS.

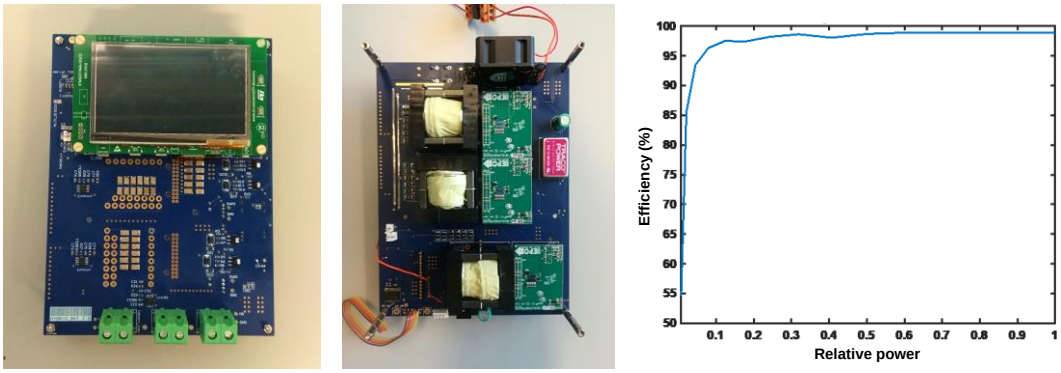

Fig. 3.27. Buck-boost converter front (left), Buck-boost converter rear (center), measured efficiency of the converter (right).

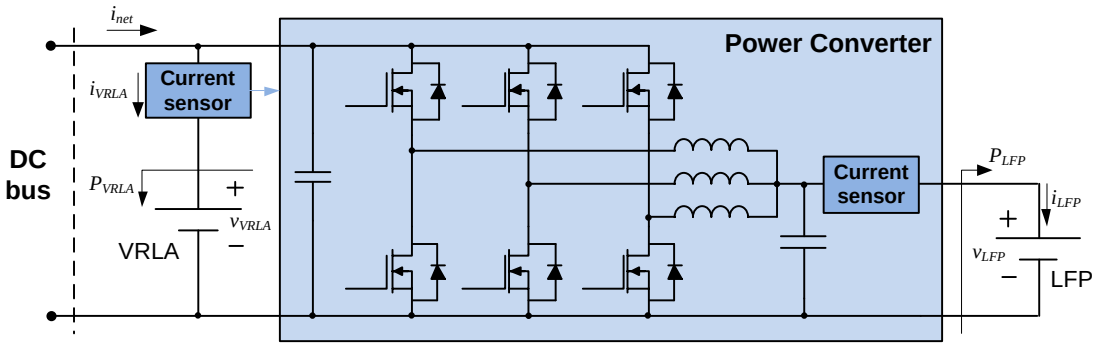

Fig. 3.28. HESS architecture and measured variables.

3.3.3 EMS implementation

The EMS of Fig. 3.7 has been implemented to control the energy flow in the system. Since the Cyclic Storage System and the Support Storage System now have an assigned storage technology, the names "CSS" and "SSS" in the state definition have been substituted by "LFP" and "VRLA" respectively.

However, the selected architecture has some small implications for the EMS. Considering the way the batteries are connected, the **LFP priority** state can theoretically be implemented by regulating the voltage of the VRLA battery to a fixed value. Since this voltage cannot change, all the charge and discharge current will automatically be provided to/by the LFP battery, and the VRLA battery will be "invisible" to the installation. However, by implementing this control in this architecture, both existing converters (the solar regulator and the DC/DC converter, which are not synchronized in this case) try to control

the voltage of the VRLA battery at the same time, leading to potential controllability issues. For this reason, the **LFP priority** state has been implemented by regulating the current through the VRLA battery to zero, rather than maintaining its voltage constant.

Additionally, a limit to the variable P_{LFP} (power to the LFP battery) was implemented in the DC/DC converter. While operating in the **LFP priority** state, if the charging power rises above the limit, the converter limits it at its maximum value and the VRLA battery is automatically charged with the excess. The power limit does not apply during discharging processes since the power consumed by the load is low and DC. If eventually P_{net} decreases and the VRLA battery starts providing energy to the LFP battery (P_{VRLA}<0), the state goes back to **LFP priority**. The implemented state machine is shown in Fig. 3.29.

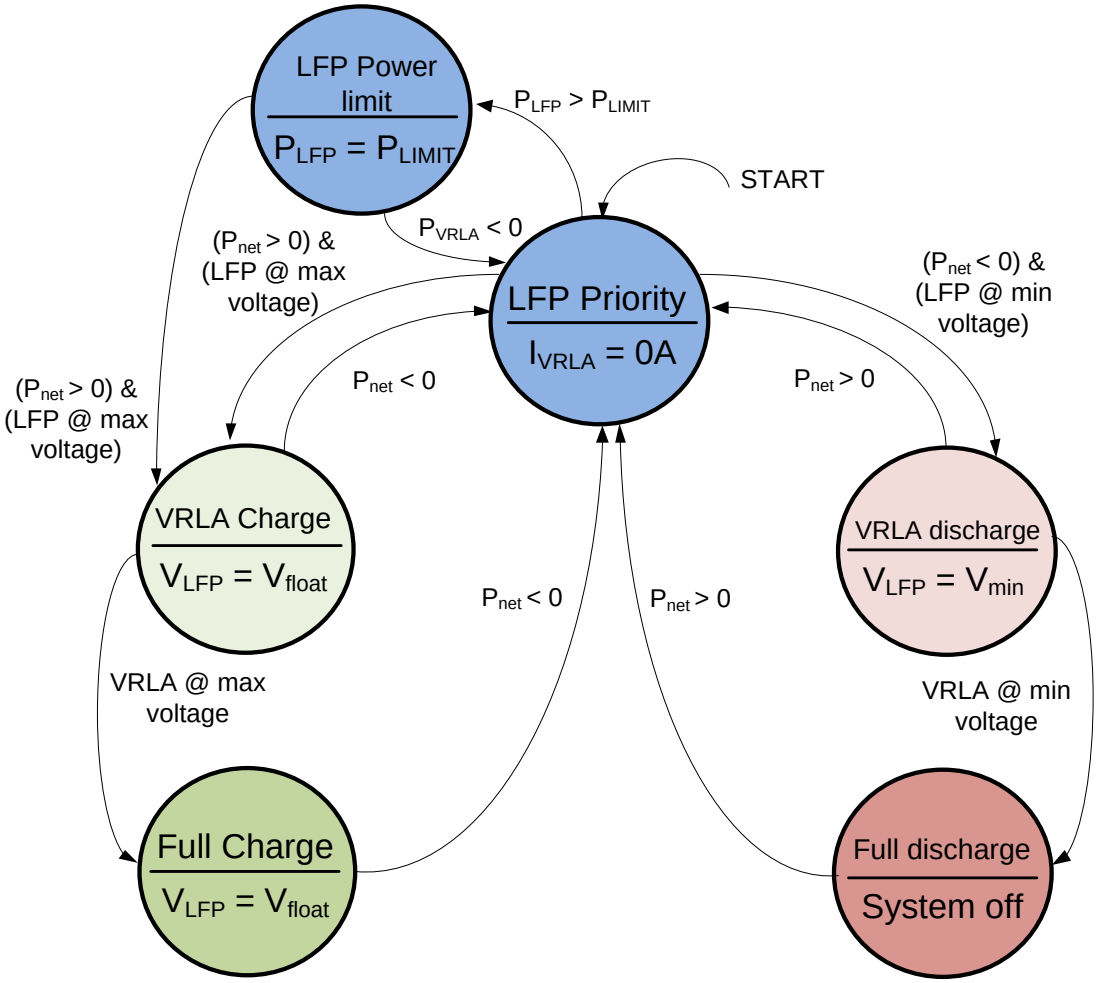

Fig. 3.29. The state machine of the implemented EMS.

3.3.4 Test bench

To test and debug the energy management strategy, a test bench was built at the laboratory. This bench emulates a real installation with all its components, such as solar generation, solar charger, DC load, storage systems, and the DC/DC converter, as well as the corresponding protections. A picture of this test bench is shown in Fig. 3.30, and Table 3.3 contains a description of the components.

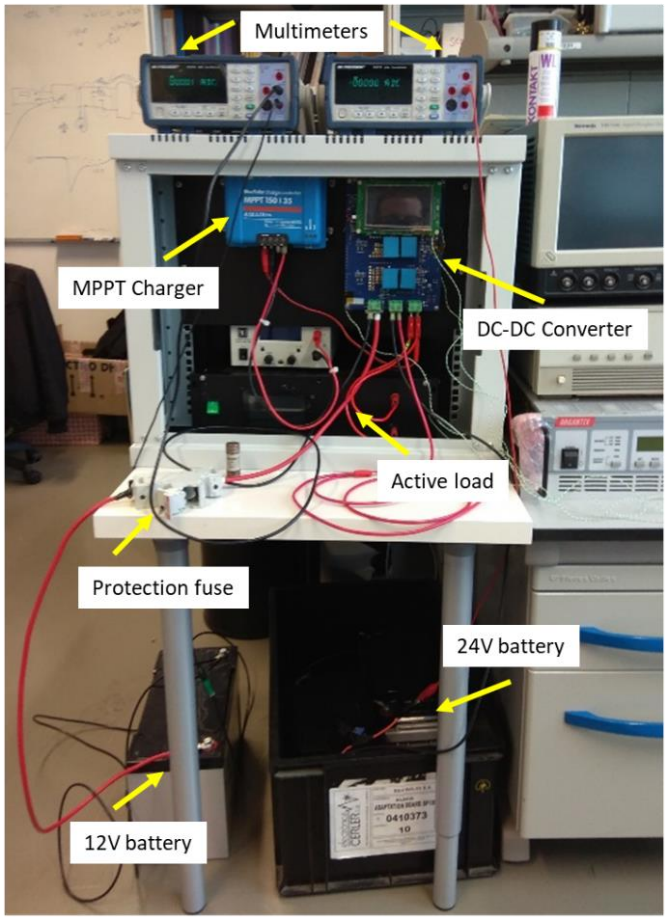

Fig. 3.30. Test-bench.

Element	Model
12V Battery	Yuasa NPC-12V-100Ah
24V Battery	2xYuasa NPC-12V-100Ah
DC/DC Converter	Ad-hoc
MPPT Charger	Victron BlueSolar MPPT 150/35
Multimeter	BK Precision 5491B
Protection Fuse	50A
Active load	Ad-hoc

Table 3.3. Test-bench elements.

3.3.5 Deployment

Hybrid installation

After the design and test of the DC/DC converter, the first installation of the hybrid system in C092 was carried out on July 3, 2018. As has already been mentioned, the existing VRLA batteries were kept as batteries for the support role. The converter and new LFP batteries were installed following the scheme of Fig. 3.28.

The C092 facility, where the installation took place, is shown in Fig. 3.31a. The final installation and its different components are shown in Fig. 3.31b. It includes the MPPT-Morning Star 60 solar charger, the DC/DC converter, the GSM module (used for data acquisition), and both battery packs. The VRLA battery consists of 12 Exide OPzS460Ah 2V cells, and the LFP battery consists of four ThunderSky 160AH-Wide type 3.2V cells.

The features of the setup are gathered in Table 3.4. The hybrid installation has a total nominal energy of 12.13kWh, an autonomy of 144h, a generation to consumption ratio g of 8.33, and a hybrid factor f of 0.17.

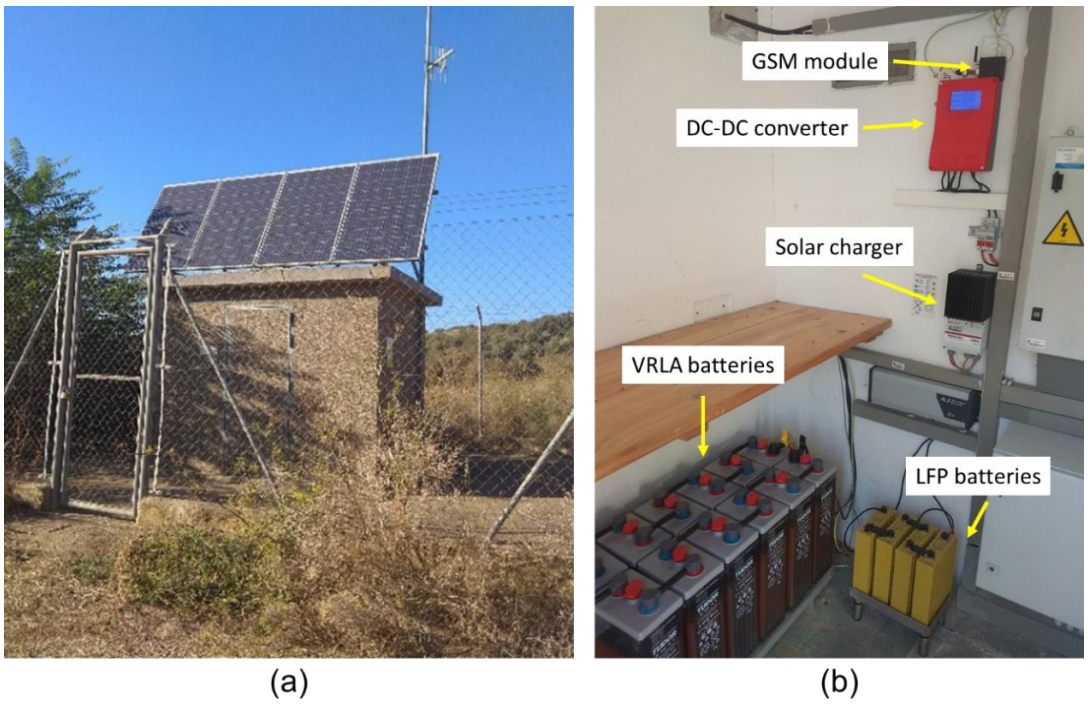

Fig. 3.31. C092 facility at the Canal de Aragón y Cataluña (a). Different elements in the HESS installation (b).

Element	Parameter	Value	Units
VRLA Battery	C_{NOM}	420	Ah
	V_{NOM}	24	V
	E_{NOM}	10.08	kWh
	Model	Exide 6 OPzS 420	-
LFP Battery	C_{NOM}	160	Ah
	V_{NOM}	12.8	V
	E_{NOM}	2.05	kWh
	Model	ThunderSky LYP160AHA(B)	-
Solar Charger	Model	MorningStar Tristar MPPT TS-45	-
Installation	DC Load	84	W
	Solar P_{PK}	700	W
	V_{NOM}	24	V
	g	8.33	-
	a	144	h
	f	0.17	-

Table 3.4. HESS Installation features.

The converter samples and reports data on a 15-minute basis from the following variables:

- VRLA battery voltage (v_{VRLA})
- LFP battery voltage (v_{LFP})
- Installation net current (i_{net})
- VRLA battery current (i_{VRLA})
- LFP battery current (i_{LFP})
- Converter temperature

The power converter sends the measured data to the Programmable Logic Controller (PLC) in the facility by using an ad-hoc protocol. The PLC sends the data to the database at CHE facilities, where it is stored. In parallel, the GSM modem sends the same data from the converter to a data storage device in the laboratory, where it is collected for redundancy and accessibility reasons.

At the time of the installation, a high ambient temperature (around 40ºC) was reported at the facility. Since the power modules used in the converter did not have good cooling capabilities, it was decided to limit of the converter to 120W (40W per module), to try to avoid problems caused by overheating. This limitation ultimately had consequences on the performance of the equipment, as described further on.

Reference ESS installation

The C095 facility was selected to be the reference site. This conventional, non-hybrid installation has been monitored and used as a reference to which compare the performance of the hybrid installation. For this purpose, a data logging device was installed on January 9, 2018. Some pictures of the reference site are shown in Fig. 3.32. The variables measured in the installation were the voltage and current of the VRLA battery, and the ambient temperature. The figures of absorbed energy, peak power, SoC, and others are post-processed from the collected data and compared against those obtained in the hybrid installation.

It is important to note that the battery pack of this conventional installation was not been expanded to match the total energy of both battery packs in the HESS. Thus, this reference installation counts with just the VRLA battery pack in Table 3.4, with a nominal energy of 10.08kWh, making a total autonomy of 120h under the same power consumption conditions.

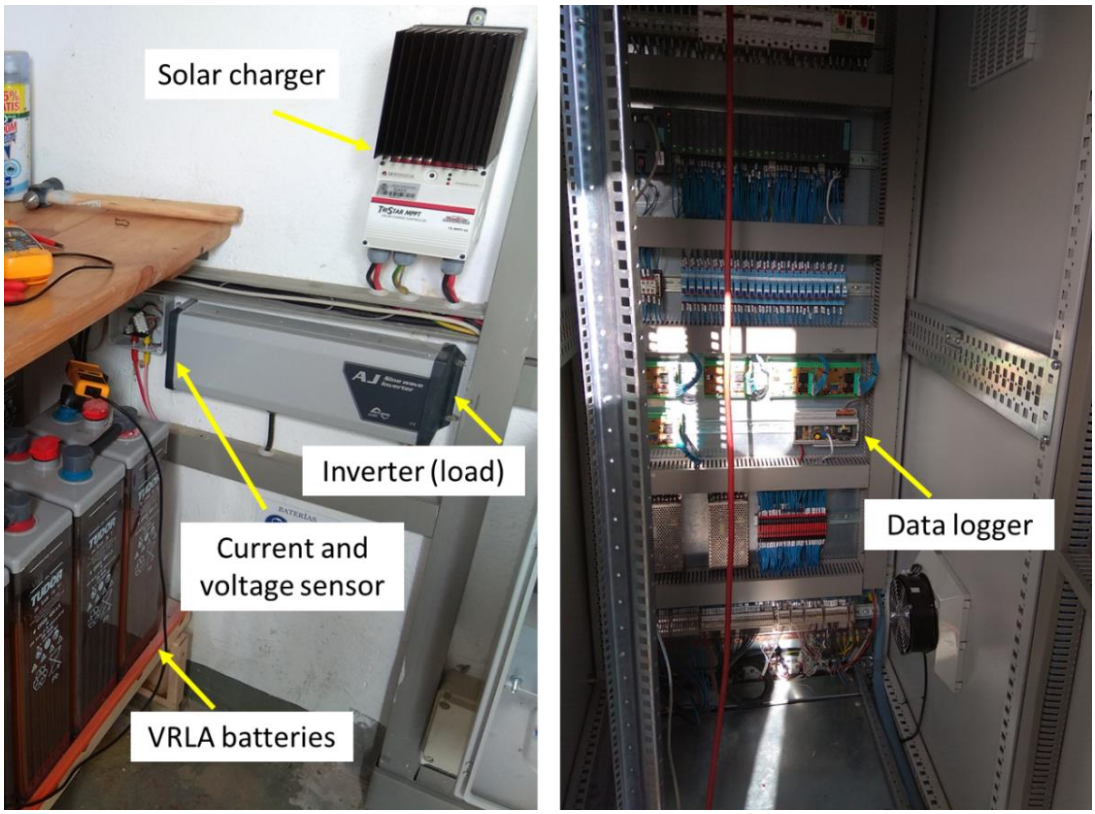

Fig. 3.32. Sensor and data logger in the control installation.

Simulation of this scenario

The specific scenario involving both installations has been simulated using the same procedure as in previous sections of this chapter. As explained above, the installations do not have the same installed energy, since the original VRLA batteries are equal size wise in both facilities, and the LFP pack has been added only to the HESS system. This means that the nominal energy of the HESS is 12.13kWh, whereas the nominal energy of the ESS is 10.08kWh. Thus, only the state of charge of the VRLA batteries in the systems (both of them with a nominal energy of 10.08kWh) is going to be compared.

Two different simulations have been carried out, with two different power limits in the DC/DC converter (the 300W limit originally conceived, and the 120W limit implemented on the first tests). The performance of the two different HESS has been compared to the performance of the conventional ESS. The results of these simulations are contained in Table 3.5.

According to the results, in both cases, the HESS operates with higher average energy than the conventional ESS. In the case of the 120W limit, the HESS stores on average 3.86kWh more than the conventional ESS. In the case of the 300W, this benefit rises to 4.25 kWh. Regardless of the power limit, the HESS holds more energy on average than the ESS even though the nominal energy of the additional LFP pack is roughly 2kWh. Fig. 3.33 and Fig. 3.34 show the evolution of the energy in the systems for both simulations.

In the comparison between power limits, the HESS with the 300W limit holds an average SoC of 61.4%, compared to 58.4% of the HESS with the 120W. In general, an operation with a higher power limit is preferred.

$\mathbf{P_{limit}}$	**120W**	**300W**
Average SoC HESS [%]	58.4	61.4
SoC LFP HESS [%]	17.7	51.7
SoC VRLA HESS [%]	66.7	63.4
SoC Conventional ESS [%]	35.9	
Overall Energy Increase [kWh]	3.86	4.25
VRLA Energy Increase [kWh]	3.50	3.15
VRLA SoC Increase [%]	30.8	27.5

Table 3.5. Simulation results.

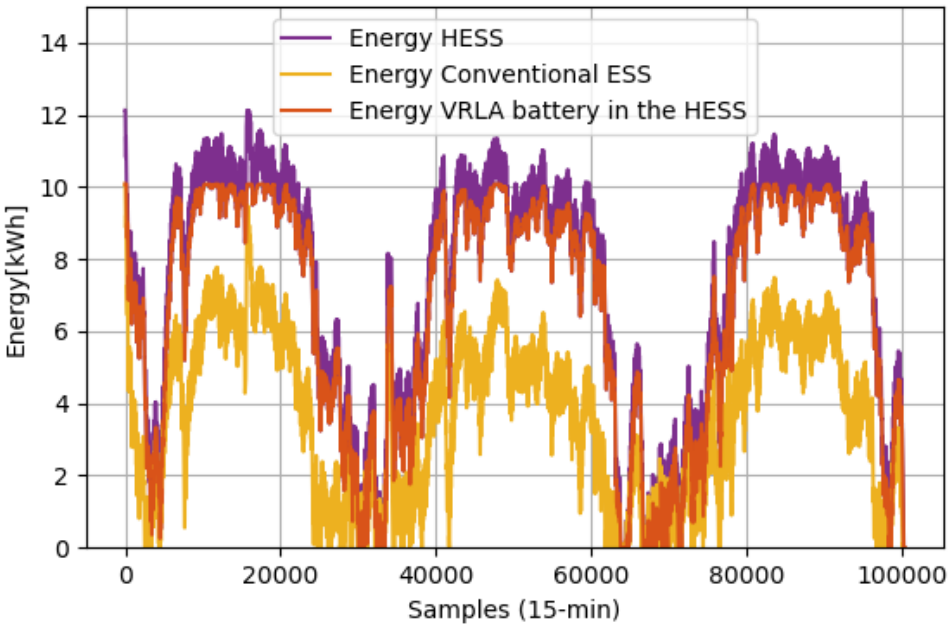

Fig. 3.33. Three-year simulation of the deployed systems (120W limit).

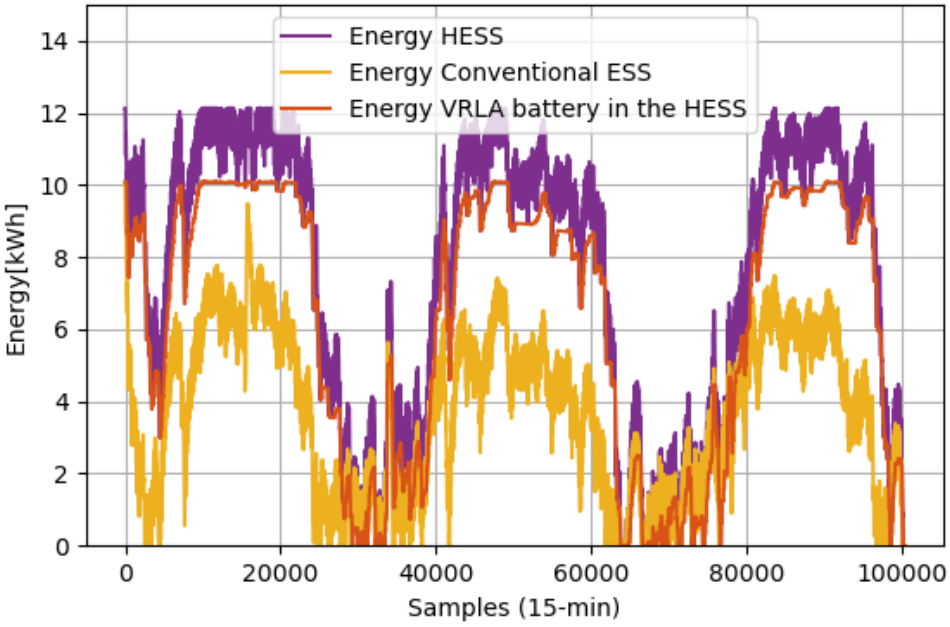

Fig. 3.34. Three-year simulation of the deployed systems (300W limit).

An interesting result may be appreciated if the VRLA energy metrics are compared. It may be appreciated how the system with a more restrictive power limit reaches a higher average SoC in the VRLA battery (at expense of the SoC in the LFP battery). In this case, the power limit constrains the power flow to the LFP battery, and the VRLA is charged more often with the excess. As a

future research line on EMS, it would be interesting to explore this constrained operation mode. An example of the operation with the 120W and the 300W where this simultaneous charge can be appreciated is shown in Fig. 3.35 and Fig. 3.36 respectively.

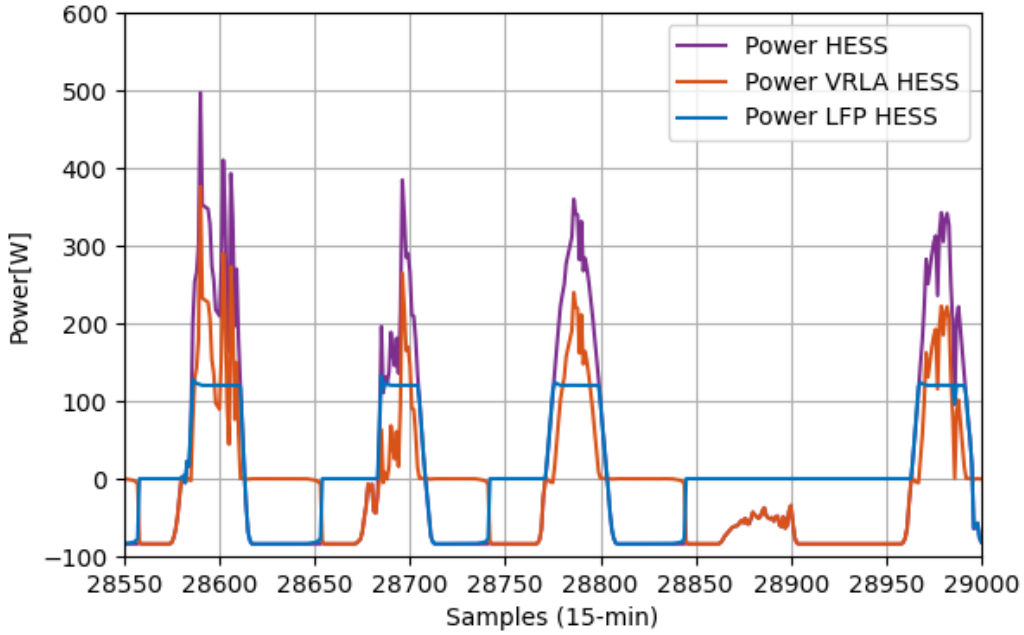

Fig. 3.35. Operation with the 120W limit.

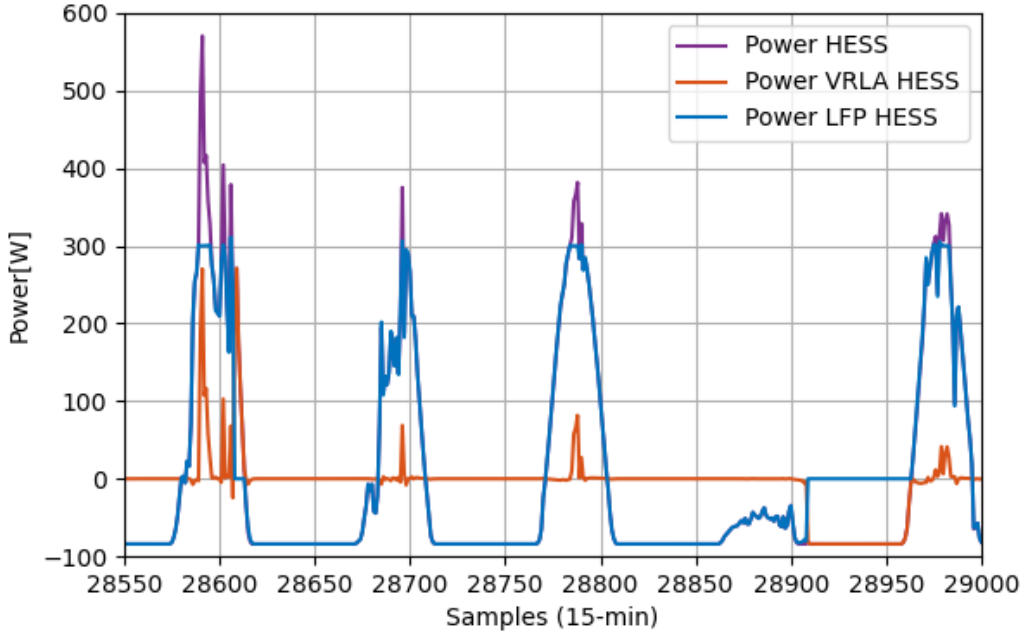

Fig. 3.36. Operation with the 300W limit.

3.3.6 Field tests results

Preliminary results

The first months after the deploy, during which the patterns of solar irradiation were favorable (without clouds), served as preliminary tests for the system. The power waveforms of the first 12 days are shown in Fig. 3.37a.

It can be seen how the converter limits the power to the LFP battery at 120W (blue plot) due to the safety limit. When the power generated by the PV system exceeds the limit, the excess goes into the VRLA battery (orange). The total absorbed power (yellow) is the sum of the power to the LFP battery and the VRLA battery. The energy stored by each battery during said period is shown in Fig. 3.37b. Here, it can be appreciated how the LFP battery is constantly charging and discharging, while the VRLA battery is charged with excess energy.

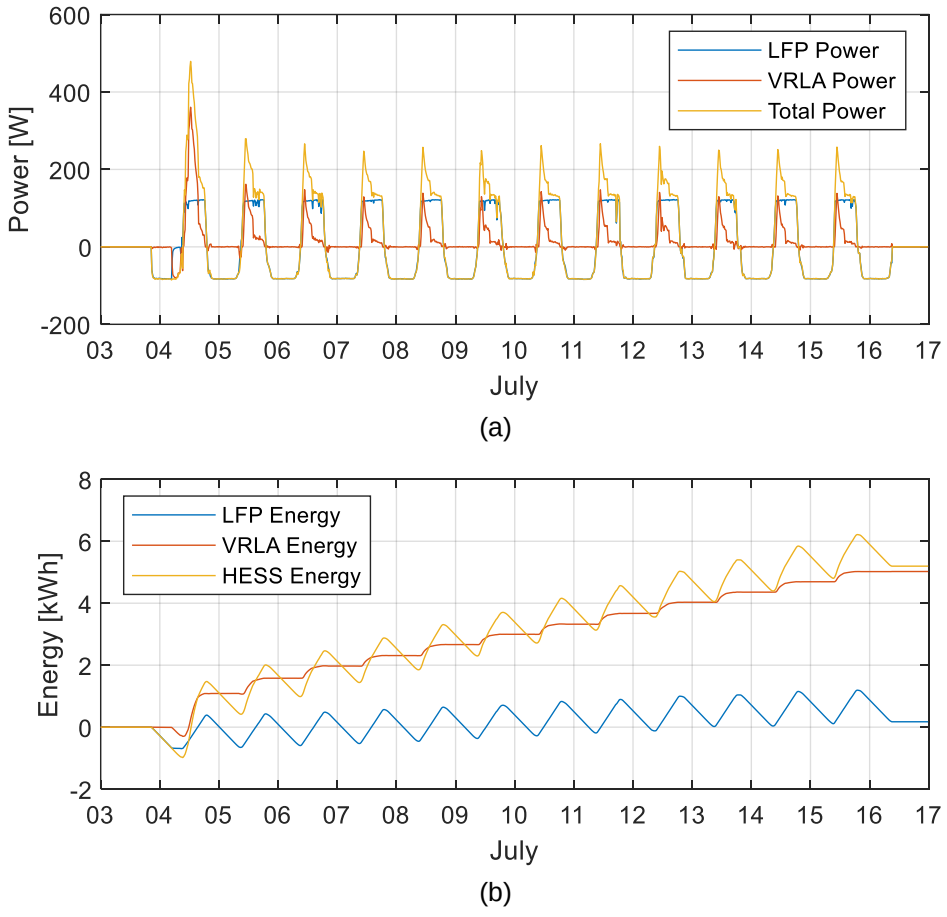

Fig. 3.37. Power waveforms in the HESS (a). Energy waveforms in the HESS (b).

During the first weeks of July, solar irradiation patterns were favorable, with high power generation. However, the power limit implemented did not allow the LFP battery to get to a full-charge state. The absorbed energy during the day (positive fraction of the blue waveform) and the discharged energy during the nights (negative fraction) by this battery are practically the same. At the end of the observed window, the energy in the LFP battery has barely risen.

Fig. 3.38 compares the net power waveforms in both installations (the HESS and the conventional ESS) during the same weeks of July 2018. It can be appreciated how the daily absolute energy of both facilities seems similar. To analyze this for a longer period, Fig. 3.39 shows the overall stored energy during the first two months. The HESS harvested 2kWh more than the conventional system, which is not a large margin, mainly due to the conservative power limit of the LFP charging process. This 2kWh amount will be referenced later.

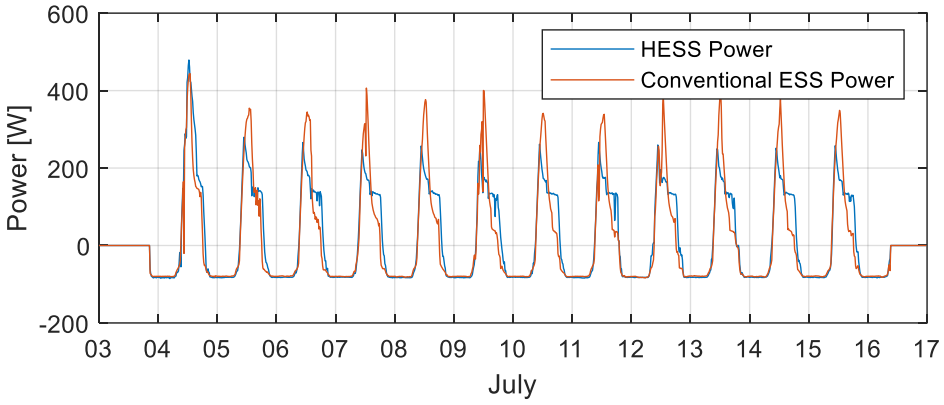

Fig. 3.38. Net power waveforms in the HESS vs the Conventional ESS.

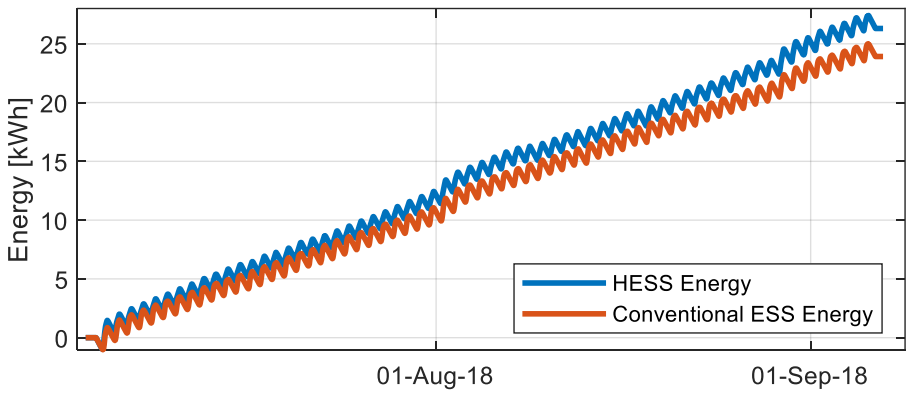

Fig. 3.39. Harvested energy in the first months in both systems.

The effect of the low power limit of the converter can also be seen in the voltage waveforms. As weeks pass and nights become longer, by the middle of August the LFP pack becomes completely discharged every night. When this happens, the VRLA battery in the HESS starts powering the load, and its voltage drops until sunrise. This can be appreciated in the voltage waveform of the VRLA batteries in the HESS, see Fig. 3.40.

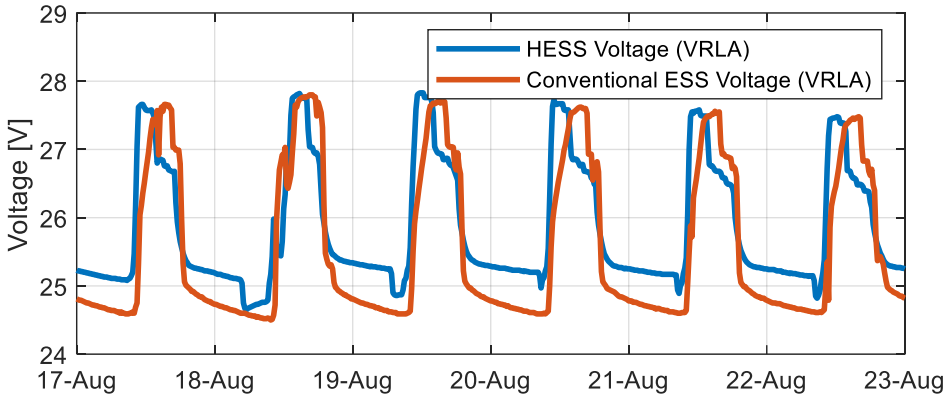

Fig. 3.40. VRLA voltage in both installations in mid-August.

A more detailed example of the same effect, in early September, is shown in Fig. 3.41. The figure shows the voltage of the VRLA batteries of both installations during one day. In blue, the hybrid installation, and in orange, the conventional.

In the A zone, it can be seen how the VRLA battery of the HESS reaches the float stage much faster than the VRLA battery of the conventional installation. This is due to its higher SoC.

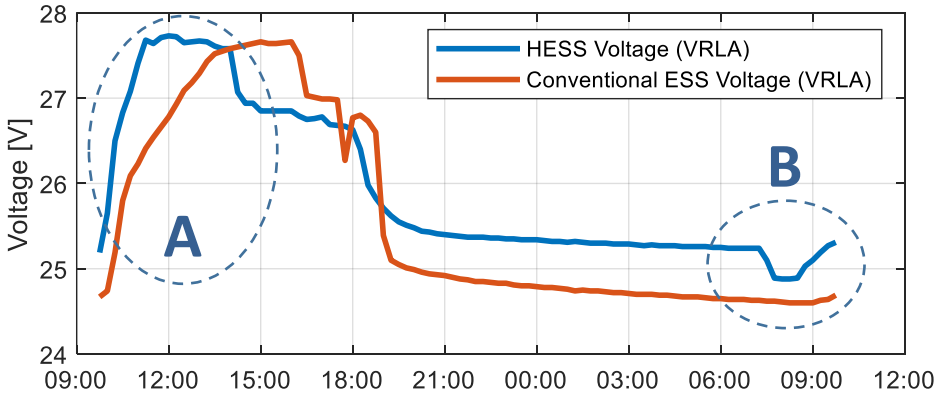

Fig. 3.41. VRLA voltage in both installations on September 2.

On the other hand, in the B zone, it can be appreciated the moment in which the LFP battery of the HESS (not shown) has been completely discharged and the VRLA battery starts providing power to the load (and thus, its voltage drops). Nevertheless, it remains above the voltage of the conventional system due to its higher SoC.

As a conclusion of these preliminary tests, it can be noted that the system is harvesting more than a conventional system, but it could work better. This has to be achieved by expanding the power limit of the converter.

Design iteration

To increase the power limit of the DC/DC converter, a cooling fan was incorporated. This way, the power modules were cooled and the power limit rose to 300W (100W per module). Further increasing the power limit was possible, but required a redesign of the DC/DC converter. The new converter was deployed on October 4, 2018. However, on the night of October 4 to October 5 a software problem disabled one of the three power modules. Therefore, the overall power was limited to 200W.

Fig. 3.42 shows the power flow in the HESS during the first week after the deployment. These waveforms are similar to the ones shown in Fig. 3.37a, but now the power limit of the LFP battery is reached at 200W. Fig. 3.43 shows the net power waveforms for both systems in the same period. As before, on a regular day without clouds, the HESS keeps generating energy after noon, whereas the production with the conventional ESS decreases.

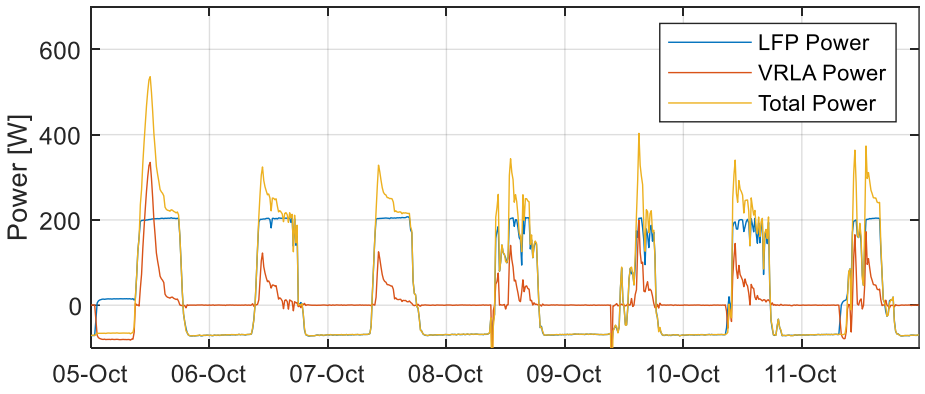

Fig. 3.42. Power waveforms in the HESS (October 2018).

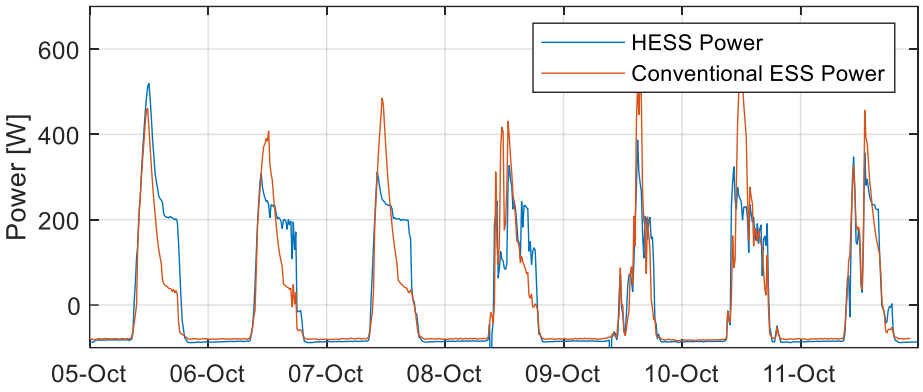

Fig. 3.43. Net power waveforms in the HESS vs the ESS (October 2018).

The benefit of the HESS becomes clearer when comparing the energy absorbed by both the HESS and the conventional ESS over a longer period. Fig. 3.44 shows the overall energy during 19 days of operation, where the HESS harvested substantially more energy than the conventional ESS. This means a considerable improvement in the operation of the installation. Specifically, an improvement ΔE=2kWh was achieved (from 1.6kWh to 3.6kWh) in the observed interval. This is the same amount of energy stored during the preliminary tests in the previous months, but it has been harvested in a much shorter period. Thus, the DC/DC converter is working substantially better, even with one of the three modules out of service.

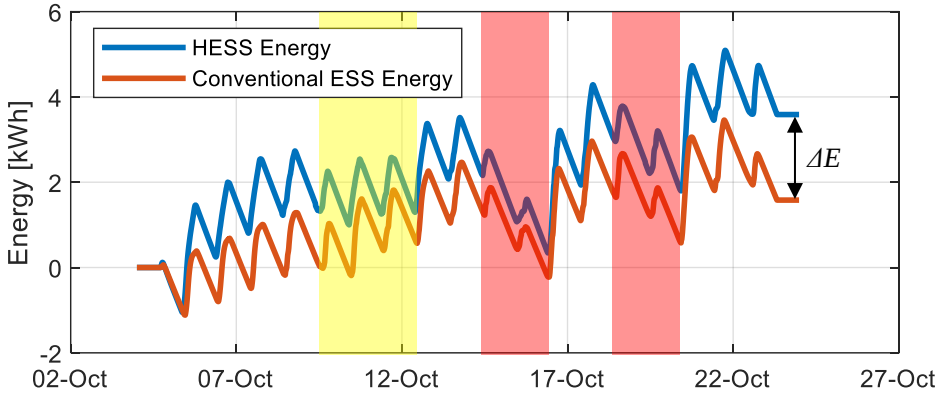

Fig. 3.44. Energy in both systems, with three low-irradiation periods highlighted.

The energy harvested right after a low irradiation event is going to be considered as an estimation of resilience. In the observed 19-day interval of Fig. 3.44, there are multiple low irradiation periods: two main low irradiation periods

between October 14-15 and October 18-19, highlighted in red; and a minor low irradiation period on October 9-11, highlighted in yellow. The energy harvested the day after these periods in both systems and the resulting Resilience Improvement is gathered in Table 3.6.

Period	Energy harvested by the HESS afterwards [kWh]	Energy harvested by the ESS afterwards [kWh]	Resilience Improvement
9-11th October	*2.08*	1.7	22%
14-15th October	*2.88*	2.55	13%
18-19th October	*2.94*	2.48	19%

Table 3.6. Energy harvested after the low irradiation periods.

At this point, the HESS has been operating for some months, and the VRLA battery in it should have reached a higher steady SoC than its counterpart in the conventional system. This can be appreciated in the VRLA battery voltage, see Fig. 3.45. In this case, as in Fig. 3.41b, the voltage of the VRLA battery drops during the night when the LFP battery (not plotted) becomes fully discharged, and the VRLA battery starts powering the load. If the voltages of both VRLA batteries are compared at this moment (when there is current flowing through them), a difference of 400mV (33.3mV/cell) may be noticed. According to the manufacturer, this would mean an approximate SoC increase of 16% (Exide Technologies 2012). It is expected that with more time, the system will evolve towards a higher improvement, closer to the theoretical benefit of 27.5% obtained in simulation with 300W as power limit (in simulation, it was during the summer months when the average SoC rose).

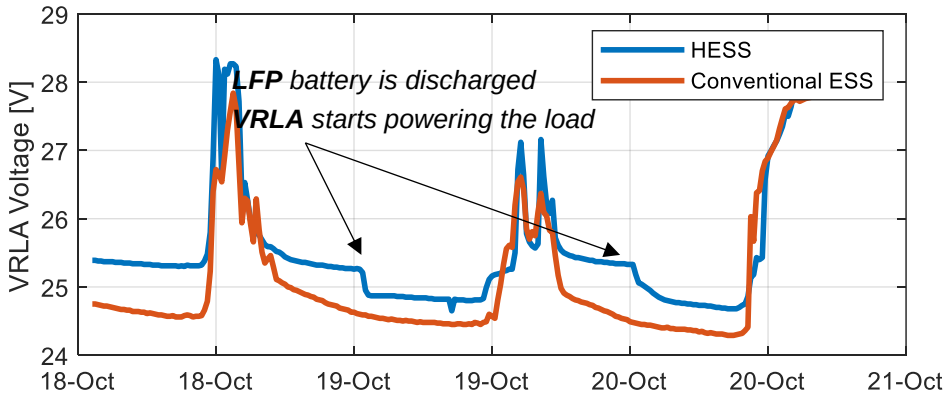

Fig. 3.45. The voltage of the VRLA battery (HESS and ESS).

Fogbank and system failure

The HESS was operating in this regime for two more months, until December 26, 2018. On December 22, a fog bank set in the area for 10 days. During these days, both systems ran out of energy. The current and voltage waveforms of both installations before the failure are shown in Fig. 3.46.

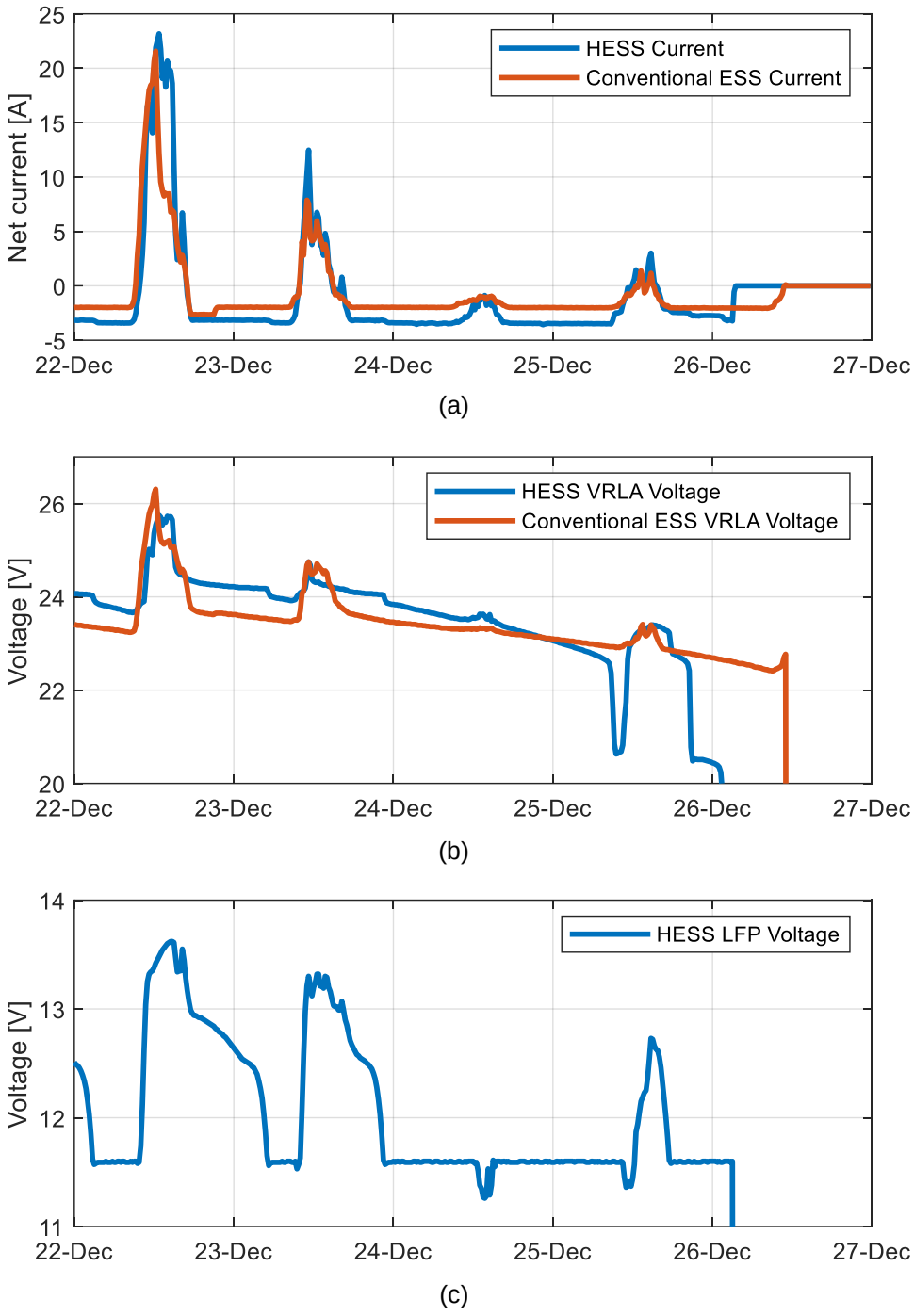

Fig. 3.46. Current and voltage waveforms during the fog bank.

Two main conclusions can be drawn from the waveforms:

- The hybrid system was subject to a higher power demand than the conventional system (Fig. 3.46a). This is caused by some loads that were disconnected only in the conventional installation and not in the hybrid installation. In the last weeks, the power demand at the HESS was 84W, almost twice compared to the 48W of consumption in the reference installation.
- The hybrid system had a higher SoC when the fog bank was set. This can be appreciated in the high voltage on Dec 22 (Fig. 3.46b). Therefore, the hybrid system had fulfilled its task. However, as of day 24, the higher consumption causes the voltage at the hybrid site to fall below the non-hybrid site. Two days later, on December 26, the HESS installation falls 12 hours earlier than the conventional one.

After the system was shut down on the 26th, the system remained off. On January 9, during the manual restart, its status was checked. Then it was verified that the batteries were completely discharged, and it was decided to disassemble the equipment and make a complete diagnose in the laboratory. During the analysis, it was observed that the power module that generated problems in October had very likely burdened the operation of the equipment, adding additional leak consumption to the system.

The energy balance during the last operation days of the system is shown in Fig. 3.47. It can be observed how the negative slope of energy (i.e. the power) is more pronounced in the HESS (blue) than in the conventional ESS (orange), which is explained by the extra consumption at the HESS. The yellow graph represents the predicted energy balance in the HESS under the same power demand as that of the conventional ESS. In this last case, two effects can be observed. First, the HESS exhibits better resilience, as has been explained in previous sections. Second, it can be seen that, considering the same power demand as the conventional ESS system, the SoC of the hybrid system at the end of the period (December 27) would have been much higher than the conventional system, guaranteeing three to four more days of autonomy.

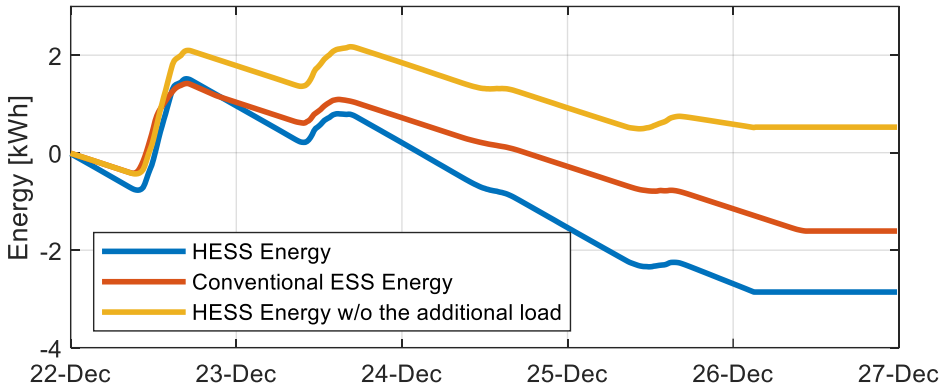

Fig. 3.47. Energy absorption and consumption during the fog bank.

3.3.7 Conclusions from the field tests

The following conclusions can be extracted from the field tests:

- The concept of the hybrid system works as expected. The hybrid system has an increased energy absorption capacity and takes the VRLA battery to a higher state of charge.
- After different low irradiation events during the month of October, the HESS has harvested around 13-22% more energy than the conventional ESS. Thus, it shows a higher resilience and is better prepared for consecutive low irradiation events.
- After four months in operation under the conditions described, the VRLA battery in the HESS has an estimated SoC 16% higher than the VRLA battery in the conventional ESS. It is expected that this value would evolve to a higher value after more time in operation and with a higher power limit in the converter.
- Both systems ran out of energy during the long fog bank. The State of Charge of the HESS at the beginning of the said event was higher, but it lasted shorter due to an increased power consumption.

In general, the HESS benefits have been validated and the system works better both in achieving a higher steady SoC and in increasing the resilience of the storage system.

3.4 Conclusions

PV Off-grid systems suffer from specific generation and consumption patterns that worsen their storage performance when using only one battery pack due to the Partial State of Charge phenomena. The study presented in this chapter proposes a simple HESS composed of two storage elements (VRLA and LFP batteries) as a strategy to overcome or at least minimize this issue, establishing the key parameters when sizing the storage elements.

An Energy Management Strategy has been implemented to coordinate the charge and discharge patterns of both batteries. This EMS is required to get simulation and experimental results, in such a way that the proposed HESS concept is validated. A simple strategy has been deduced and implemented in this work. However, there is a huge research field related to new EMSs, leading to strategies that are better suited to other applications.

The sizing of this HESS has been studied, and the simulation results show that the minimum autonomy of the systems should be around 150 hours, an optimal LFP size in the range of 15 to 30% of the overall installed energy (and the rest in the VRLA technology), obtaining the best performance tradeoff. The peak power of the solar panels should be an order of magnitude higher than the DC power of the load in the climate and latitude studied. The simulation results of different HESS sizing designs show an increment of the average SoC in the range from 5 to 25% and an improvement of the resilience of up to 45%, which is especially noticeable when considering small-size battery packs, with low overall autonomy.

The system has been experimentally tested on a real installation with a simple 200W DC/DC converter. The results show that the implemented HESS was able to harvest around 13-22% more energy than the conventional ESS after low irradiation periods, and obtained an increased SoC of around a 16% higher in the VRLA battery after some months in use, which could be improved if a longer period was analyzed.

In general, it can be stated that the hybrid system is noticeably better than the current system. The new system has different benefits:

- Better resilience capabilities. Due to the improvement in the energy absorption capabilities provided by the LFP battery, the system can harvest more energy after a long period of low irradiation.
- Better storage capabilities. The hybrid system reduces the PSoC issue, leading the emergency battery into a higher steady SoC than in the case of conventional ESS.
- Useful life extension. Since the VRLA battery is less cycled than in the ESS case, it ages slower. The specific impact of the HESS on this aspect still needs time to be checked.

The results of this study are subjected to several interpretations. In this application, new HESS do not need to be oversized to guarantee a minimum charge reserve when low irradiation days come. Even more, old ESS based on one monolithic VRLA battery can be retrofitted, by adding in parallel the DC/DC converter and a smaller LFP battery pack. The old battery can now operate in a support role and thus extend its lifetime.

As a future line, it is interesting to develop smart charge algorithms and battery maintenance strategies. Algorithms that predict system failures or energy shortages would be implemented, so batteries can be overcharged before these events, minimizing the probabilities of failure. Additionally, the EMS can be improved with the estimation of the state of health of the storage subsystems, so better suited decisions can be taken when selecting roles for the batteries.

An introduction to the above-mentioned predictive algorithms is carried out in the following chapters, as the first steps in this line of research.

Chapter 4.- Fault Prediction with Machine Learning Methods

This chapter addresses failure prediction in standalone photovoltaic installations, as a means to provide more information to energy management strategies and operators. This way, batteries can be intelligently charged or preemptive maintenance tasks can be scheduled.

4.1 Introduction

4.1.1 Motivation

In previous chapters, HESS have been proposed as a solution to mitigate most of the problems of conventional ESS in standalone PV installations. However, upgrading an existing ESS into a HESS requires a major modification in the energy system, since an additional battery pack and an additional power converter need to be installed. There are installations where this is not possible due to budget or room limitations.

Advanced battery charging strategies, such as 3-stage charging (Fig. 2.11), have been developed as a simple solution to improve the performance of existing ESS, without the need for additional hardware elements. These strategies do not provide a complete solution to the problems of the energy systems but are sometimes enough for the application. However, this speeds up the aging of the batteries due to higher float voltages or increased charging currents (Garayalde et al. 2017). Additionally, these overcharges are done periodically, without taking into account the state of the batteries. As an example, batteries are overcharged with the same periodicity on summer and on winter.

These overcharge strategies can be enhanced with predictive algorithms. If such algorithms foresee a period of low irradiation, with the risk of reaching low voltages, then the batteries can be overcharged. On the contrary, if a favorable scenario is predicted, the battery can be charged without the additional absorption phase, reducing the impact on aging.

Furthermore, even on systems without the capability of overcharging the batteries, it could be interesting to develop such techniques. An example is given: if future problems are predicted, maintenance tasks may be scheduled, so operators may be able to recharge the batteries using other sources such as diesel-powered generators. Currently, these tasks are carried out a-posteriori, after the system has failed, with the consequent service interruption and increased maintenance cost.

4.1.2 Background

The use of machine learning techniques in photovoltaic installations has been explored before. These techniques have been mainly applied to the estimation or prediction of the power generation patterns under uncertain scenarios like cloudy periods. (Anderson and Yakimenko 2018; Ospina, Newaz, and Omar Faruque 2019). Estimation of power consumption patterns is also a goal in applications such as micro-grids, where there is uncertainty in this regard (in the case being studied here, power consumption is mainly DC) (Brenna et al. 2018). The use of historic data from installations is critical to train and validate the prediction algorithms, even by using data from several years (Abdel-Nasser and Mahmoud 2019).

When addressing power generation prediction, these machine-learning techniques have been previously combined with weather forecasting services or tools to provide an estimate on future generation patterns, as in (Leva et al. 2017; Zhong et al. 2014). Cloud analysis has previously been used to make a short-term prediction on generation and irradiation (Anagnostos et al. 2019). Additionally, studies such as (Huang et al. 2015) propose a classification of different types of days, according to their daily irradiation patterns. These typologies are used as different future scenarios, to make different predictions that are then combined through fuzzy logic techniques.

Regarding the battery packs, most works are focused on estimating or predicting variables such as SoC or RUL, in a short-term window (Lipu et al. 2018; Caliwag and Lim 2019; Hong et al. 2020; Wu et al. 2020; Homan et al. 2019). Fewer studies are focused on predicting the voltage of the installation, even though it is the most critical variable when trying to predict a possible failure of the system (Yu et al. 2018). Since this is the approach taken throughout this chapter, voltage is going to be considered the prediction goal.

As for different data-driven techniques, most studies are centered in sequence-processing networks, such as Non-linear AutoRegressive eXogenous models (NARX), Long-Short Term Memory models (LSTM), or Convolutional Neural Networks (CNN) (Anderson and Yakimenko 2018; Lipu et al. 2018;

Anagnostos et al. 2019; Hong et al. 2020; Wu et al. 2020; NaitMalek et al. 2019). Specifically, LSTM seems to be the best-suited technique in terms of overall performance, according to the error metrics and the number of publications (Abbas, Nawaz, and Kamran 2019). Gaussian Process Regression (GPR) has also been considered in other studies, but its increased complexity makes it worse positioned when trying to implement these techniques in simple BMS (Pajovic et al. 2017; Sahinoglu et al. 2018; Sanz-Alcaine et al. 2021).

4.1.3 Objectives

This chapter aims to provide simple but effective algorithms to predict whether an installation is going to suffer from energy disruption in a certain time horizon. The hypothesis is that the future state of the energy system can be predicted by using a window of past observations of voltage, temperature, and current, and an input of the weather forecast for the future days. Two different algorithms, specifically designed for sequence processing, are going to be evaluated as options.

The outline of the chapter is as follows. Section 4.2 describes the methodology and database. Section 4.3 describes the algorithms and provides initial results for the algorithms and the prediction approach by using future data as an input, without addressing the weather forecast problem. Section 4.4 describes the algorithms used for modeling the future input data so the algorithm can be included in a real application with the weather forecast. Section 4.5 collects the results when incorporating this future estimation, and section 4.6 gathers the conclusions of the chapter.

As an observation, the work presented in the chapter has been developed by using real data from some installations in the Ebro basin, which belong to CHE. Each of them has different parameters such as power consumption, battery chemistry, photovoltaic power… The performance metrics presented throughout this chapter have been obtained with one of the installations (Morrón). Further, performance metrics for the rest of the installations are collected.

4.2 Methodology

4.2.1 Dataset description

The dataset used for training the algorithms is a proprietary dataset, which has been described in (Guillén Asensio et al. 2020; Sanz-Gorrachategui et al. 2020) and is an expanded version of the dataset used in previous chapters for simulation purposes. It comprises data from several standalone PV installations that belong to CHE. These installations are scattered along the Ebro Basin, and serve different purposes such as metering stations, irrigation channel control, and communication relays. Each installation has measured, stored, and reported different parameters and variables measured every 15 minutes, over the course of more than ten years. In this case, however, the database has been subsampled to a one-hour sampling time to reduce computational cost.

The stored variables related to the ESS are the voltage, current, and temperature of the batteries. As an example, Fig. 4.1 shows a couple of weeks of data belonging to the installation of Morrón (40°42'41.7"N, 0°20'50.0"W).

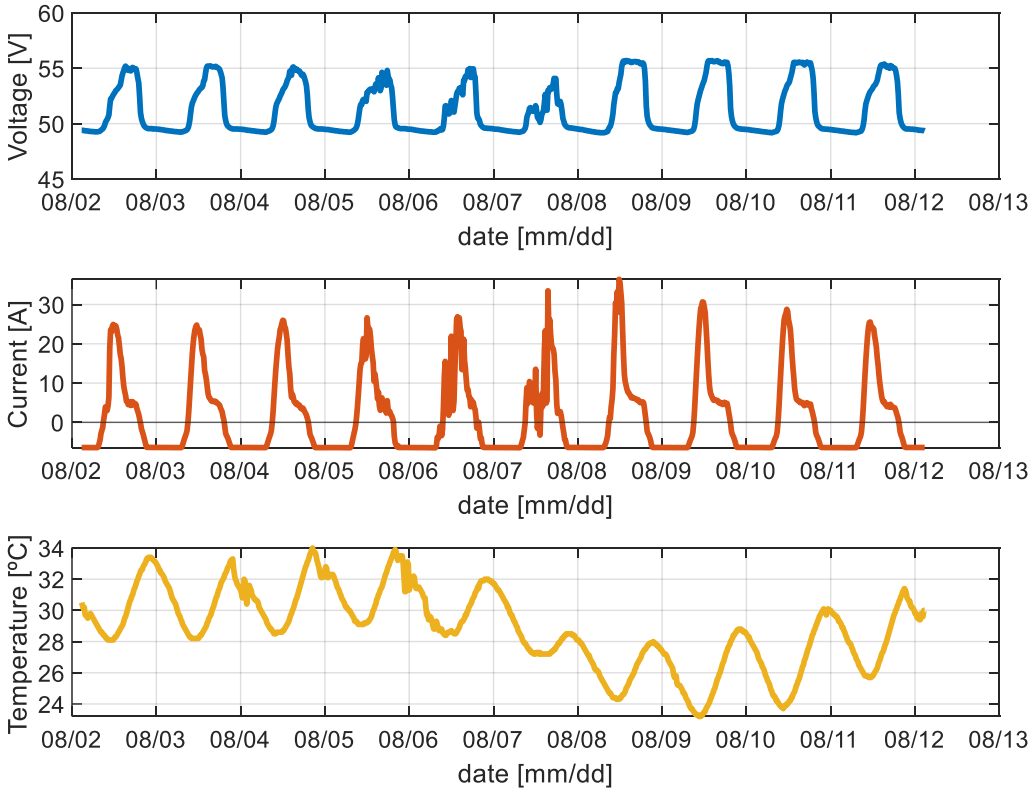

Fig. 4.1. Voltage, current, and temperature waveforms (Morrón).

The installation at the Morrón site, one of the installations contained in the dataset, has been selected for this work. The algorithms for performance prediction have been trained and validated on this installation, so they can predict the evolution and future behavior of the ESS at this location. The features of said installation are collected in Table 4.1, whereas Fig. 4.2 shows a photograph of the installation and the included battery pack.

Element	Parameter	Value	Units
Battery	Chemistry	NiCd	-
	V_{NOM}	48	*V*
	Capacity	920	*Ah*
	Nº Cells	40	
	Model	Sunica Plus Sun 920-1 SAFT	-
Installation	DC Load	250	*W*
	Solar P_{PK}	3120	*W*

Table 4.1. Morrón installation features.

Fig. 4.2. Morrón facility (left) and battery pack (right).

4.2.2 Estimation approach

Estimation goal

The main goal of this machine-learning algorithm is to determine whether the standalone PV installation may run out of energy in a short-term window. Thus, the output of the algorithms is going to be the voltage of the battery pack. When the voltage falls below a certain value (depending on the specific chemistry, conditions…), the installation shuts down, so it is critical to estimate

the voltage evolution and determine whether or not it falls below some emergency thresholds.

After consulting with the experts that run the installations in CHE, it has been decided that 48 hours is a reasonable estimation horizon for them to prevent failure if needed. To predict voltage at far time horizons, some studies (Sahinoglu et al. 2018) use a recursive approach, where the voltage output is predicted in a closer time-span and is then fed back as a new voltage measurement, to keep on predicting further in time. However, in this case, a direct estimation of the voltage measurement 48 hours in the future is going to be applied.

Inputs

To provide information on the past state of the battery, windows of 48 hours containing past waveforms of voltage, current, and temperature are going to be used as inputs for the algorithm since they are available at most installations (Sanz-Alcaine et al. 2021). Additionally, an estimation of the future current waveform during the following 48h is used to provide information on the evolution of the system. The future temperature was also initially considered but finally discarded after checking its low relevance in the voltage prediction. A summary of the input variables used in the prediction is shown in Fig. 4.3.

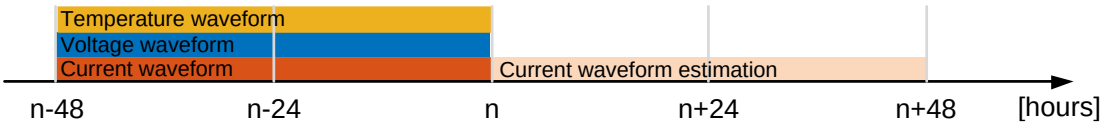

Fig. 4.3. Time windows of voltage, current, and temperature are used as inputs.

The general structure of the prediction algorithms will follow equation (4.1). Here, W is the size of the past observation window, $\hat{v}_{n+48}$ is the prediction of the voltage in 48 hours, v_n to v_{n-W+1} represent the past voltage samples, t_n to t_{n-W+1} represent the past temperature samples, i_n to i_{n-W+1} represent the past current samples and, i_{n+1} to i_{n+48}, the future current estimations.

$$\hat{v}_{n+48} = f(v_n, \dots v_{n-W+1}, t_n, \dots t_{n-W+1}, \hat{\imath}_{n+48}, \hat{\imath}_{n+1}, i_n \ \dots i_{n-W+1}) \qquad (4.1)$$

System overview

A scheme of the complete system is shown in Fig. 4.4. The core network, "Voltage prediction algorithm", receives the preprocessed inputs and obtains an estimate on the future voltage, 48 hours ahead. The past inputs are provided by the "Real-time inputs" block. The future current estimate is provided by the "Future current estimation" block. The "Preprocessing" block accommodates the different inputs so they can be processed by the prediction algorithm. Both the "Voltage prediction algorithm" and the "Future current estimation" blocks are trained with the dataset.

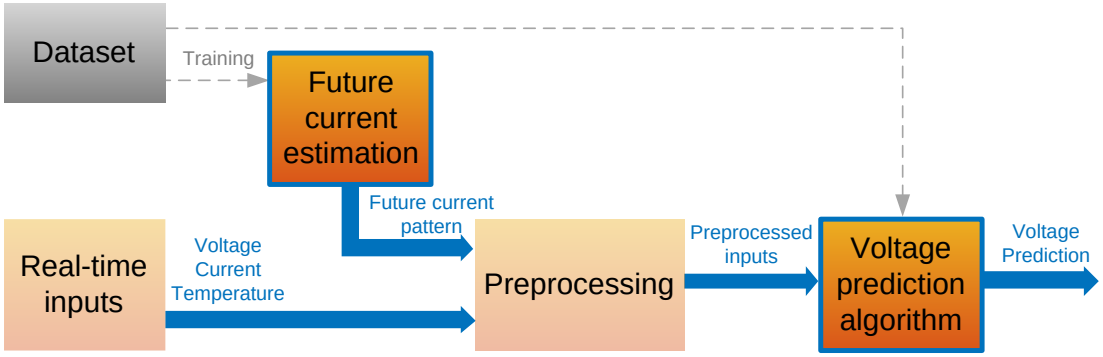

Fig. 4.4. Diagram of the voltage prediction process.

Both the "Voltage prediction algorithm" block and the "Future current estimation" block are developed in sections 4.3 and 4.4 respectively. The preprocessing block is described below.

Preprocessing

The ranges of voltage, temperature, and current are different and therefore, these variables need to be normalized before being processed by the networks. The normalization improves the convergence of the algorithm and ensures that some variables are not outweighed by others simply due to their absolute value. The voltage and current waveforms of each dataset have been normalized to the range [0, 1] by using the expression in (4.2), where each variable x is normalized by using its maximum and its minimum value in the whole set.

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{4.2}$$

4.3 Future voltage prediction

4.3.1 Neural network models

Different machine learning models have been considered for the prediction task. Specifically, these models need to be able to process time-domain data series, since the inputs to the algorithms are going to be data frames of different variables, measured during a given amount of hours. Additionally, the goal is to end up implementing these algorithms in a microcontroller embedded in an ESS/HESS, so networks with a lower complexity are preferred.

The algorithms NARX and LSTM have been chosen for this purpose. LSTM has shown good results in similar applications in the state of the art (NaitMalek et al. 2019; Ospina, Newaz, and Omar Faruque 2019; M. S. Hossain and Mahmood 2020), but is an algorithm with a certain degree of complexity, more suitable for cloud-like implementations. On the other side, NARX represents a simpler option with acceptable outcome, preferred for in-situ implementations.

Initially, to decouple the voltage prediction algorithm from the future current estimation, these algorithms are going to be designed using actual future current waveforms as inputs. How to estimate future current will be explored in further sections and the performance of the voltage prediction networks when using an estimate for the future current will be analyzed then.

Nonlinear AutoRegressive eXogenous models

NARX networks are recurrent algorithms, where part of the output is feedback into the input (Haykin 1998; Leontaritis and Billings 1985). They are best suited for modeling time-series, non-linear systems. Any NARX network follows one of the two architectures in Fig. 4.5, an open-loop architecture or a closed-loop architecture.

In the open-loop structure, true past measurements of the output variable are available at the time of the prediction and thus can be used as input. On the other hand, in the closed-loop architecture, these measurements are not available, and past estimates of the output are used as inputs instead.

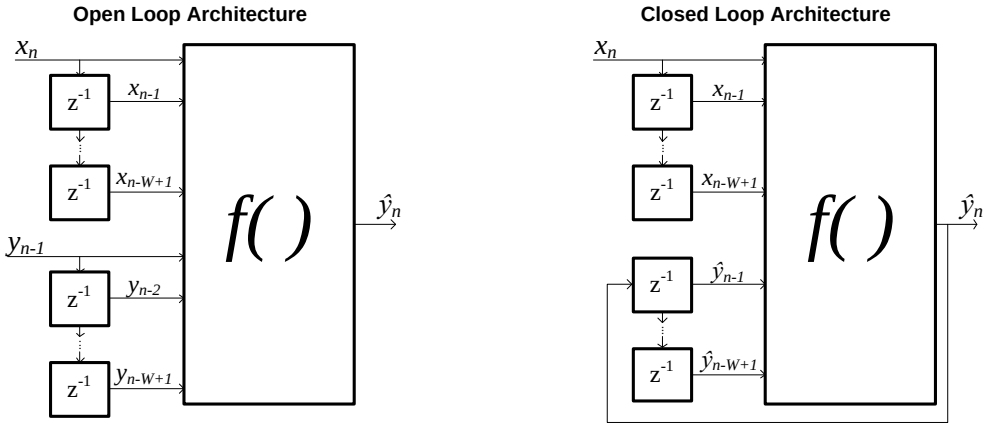

Fig. 4.5. Different NARX architectures.

In this application, as time progresses, new voltage measurements are available and therefore, there is no need for using past voltage estimates. For this reason, the open-loop architecture has been selected. The general definition of a NARX function with this architecture is described in (4.3), where f is a non-linear function, x_i is the input variable at the i-th sample, y_i is the output at the i-th sample and W is the size of the buffer or time window. Note that past samples of y are used to compute the next output.

$$\hat{y}_n = f(x_n, x_{n-1} \ldots x_{n-W+1}, y_{n-1} \ldots y_{n-W+1}) \tag{4.3}$$

Different options can be used to implement the nonlinear function f. In this case, a Multilayer Perceptron (MLP), also known as Feed-Forward Neural Network (Haykin 1998), has been selected. The structure of a generic Multi-Layer Perceptron is shown in Fig. 4.6. This network is usually made of three layers:

- An input layer, with as many input neurons as the dimensionality of the input array (m).
- Hidden layer(s). Here, the inputs are linearly combined (multiplied by certain weights $w_{l,i,k}$, and summed with a bias term) and then pass through a non-linear function (f_l) known as the activation function. Different functions can be used as activation functions, but the most common is the sigmoid. The main parameter to fix when designing an MLP is the number of neurons in this layer (p). A single or multiple hidden layers can be used.

- An output layer, with as many output neurons as desired outputs (n). On it, the outputs of each neuron in the hidden layer are linearly combined. An additional activation function (f_2) can be added after this layer, although this is usually done in classification problems rather than regression problems, which is the case here.

These structures may be used for classification or regression problems, depending on the activation function of the output layer. This is a regression problem since the output is a voltage estimate, so the output layer is linear.

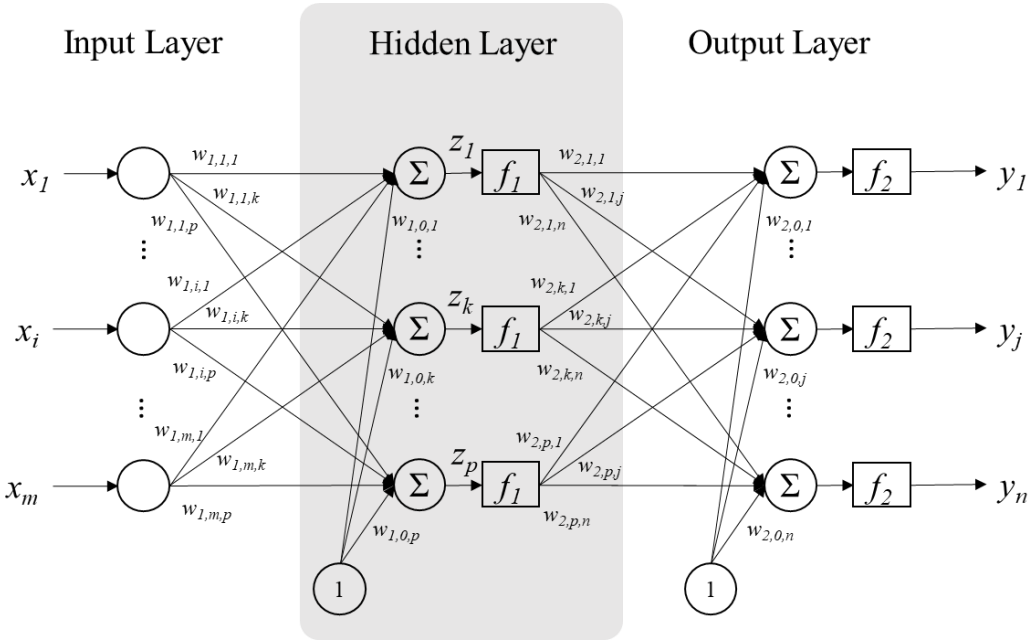

Fig. 4.6. MLP architecture.

Note that the inputs x_i of the MLP do not necessarily have a temporal relationship; they can be magnitudes of different nature. The NARX structure establishes the temporal relationship between them.

Long-Short Term Memory models

LSTM networks were initially proposed in 1997 as a means to deal with the vanishing gradient problem in training recurrent neural networks (RNN) (Hochreiter and Schmidhuber 1997). The structure of an LSTM layer is shown in Fig. 4.7. Each layer has an input $\mathbf{x}^n$, an output $\mathbf{h}^n$, and a memory term $\mathbf{c}^n$, which are vectors that change over time, being n the time variable. The vector $\mathbf{x}^n$ has dimensionality m, and both $\mathbf{h}^n$ and $\mathbf{c}^n$ have the same dimensionality p.

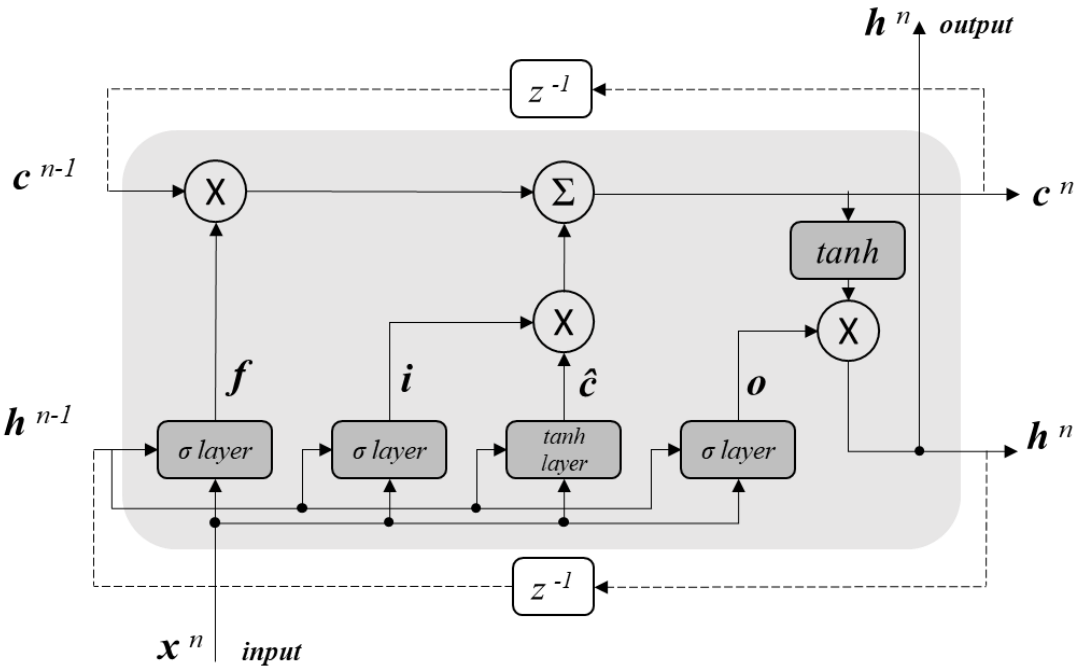

Fig. 4.7. LSTM layer architecture.

The LSTM layer controls information through three sub-networks, or "gates": The forget gate **f**, the input gate **i,** and the output gate **o**, which are also vectors of dimensionality p. The forget gate controls how much old information is removed from the memory term. The input gate controls how much new information in $\hat{\mathbf{c}}$ passes to the memory term. Finally, the output gate controls the impact the memory term has on the output h. In the figure, the σ layer and *tanh* layer blocks represent complete neural layers with sigmoid or hyperbolic tangent function as activation function respectively.

In each time instant n, the equations for each element k of the p elements in **f**, **i**, **o**, and $\hat{\mathbf{c}}$ are:

$$f_k = \sigma\left(\left[\mathbf{w}_k^f\right]^T \begin{bmatrix}\mathbf{h}^{n-1} \\ \mathbf{x}^n\end{bmatrix} + b_k^f\right) \tag{4.4}$$

$$i_k = \sigma\left(\left[\mathbf{w}_k^i\right]^T \begin{bmatrix}\mathbf{h}^{n-1} \\ \mathbf{x}^n\end{bmatrix} + b_k^i\right) \tag{4.5}$$

$$\hat{c}_k = tanh\left(\left[\mathbf{w}_k^{\hat{c}}\right]^T \begin{bmatrix}\mathbf{h}^{n-1} \\ \mathbf{x}^n\end{bmatrix} + b_k^{\hat{c}}\right) \tag{4.6}$$

$$o_k = \sigma\left(\left[\mathbf{w}_k^o\right]^T \begin{bmatrix}\mathbf{h}^{n-1} \\ \mathbf{x}^n\end{bmatrix} + b_k^o\right) \tag{4.7}$$

Where $\mathbf{w}_k^f$ is the array of $m+p$ weights associated with the k-th element in the $\mathbf{f}$ gate, and b_k^f is the bias term associated with the k-th element in the $\mathbf{f}$ gate. The same is applicable for the rest of the gates.

After the gates have been calculated, the updated memory term $\mathbf{c}^n$ and the output at the instant n, $\mathbf{h}^n$, can be obtained as in (4.8) and (4.9). Here, $\otimes$ represents the Hadamard product.

$$\mathbf{c}^n = \mathbf{c}^{n-1} \otimes \mathbf{f} + \hat{\mathbf{c}} \otimes \mathbf{i} \tag{4.8}$$

$$\mathbf{h}^n = \tanh(\mathbf{c}^n) \otimes \mathbf{o} \tag{4.9}$$

Structures with a higher complexity may be built by stacking multiple LSTM layers, so the $\mathbf{h}^n$ term of the previous layer acts as the input $\mathbf{x}^n$ to the following one. Additionally, post-processing functions can be added to obtain a scalar output from the term $\mathbf{h}^n$ of the final LSTM cell.

4.3.2 Network sizing

Both NARX and LSTM structures count with multiple parameters that need to be properly tuned while training the algorithm. These parameters include the number of hidden layers, the number of neurons in the layers, the activation functions of each layer, and others. To tune these parameters, structures with different sizes and functions have been trained and compared. The Root Mean Square Error (RMSE), defined as in (4.10), has been used to benchmark these different structures and parameters. The training process for these networks is described below.

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \tag{4.10}$$

Training process

A machine-learning algorithm should perform well with the training data and with new data. However, when training neural networks, overfitting may happen. Overfitting takes place if the network over-learns the training data,

achieving very low training errors but losing its generalization capabilities, and performing worse with new data. There are multiple techniques to avoid this effect. Generally, the dataset is divided into different subsets, which are used in different ways during the training process to minimize overfitting. In this case, the training process carried out has been early-stopping (Haykin 1998). To this purpose, the dataset has been divided into three subsets: the training, validation, and test subsets. The training set comprises 70% of the examples in the dataset. On the other side, the validation and the test subsets each made up 15% of the total examples in the set.

In early stopping, the training subset is used to fit the model on each iteration, and the validation subset is used to obtain an error metric after each iteration. When the validation error increases steadily, overfitting is starting to take place, so the training process stops and the iteration with the best validation performance (early-stopping point) is selected as the trained network, see Fig. 4.8. The test subset is used afterward to compare different algorithms with different structures or parameters. This way, none of the algorithms has interacted with the test data while training, thus obtaining a fair comparison.

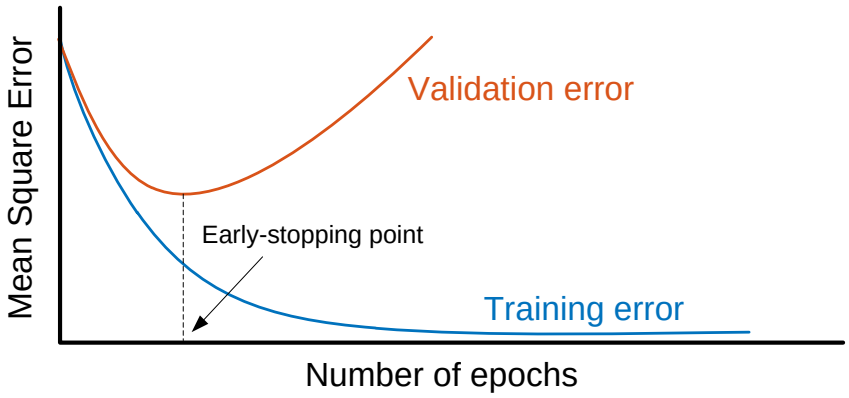

Fig. 4.8. Evolution of the error in the training process.

The ADAM (ADAptive Moment estimation) optimizer has been used for the training process. This algorithm is a stochastic optimizer with a variable learning rate, an evolution over the conventional Stochastic Gradient Descend algorithm (Kingma and Ba 2015), and it is probably one of the most common algorithms used for neural network training nowadays. As a cost function, the Mean Squared Error (MSE) has been used.

Network sizing

The NARX and the LSTM algorithms count with numerous parameters to be tuned while in training. This may be a difficult task, and usually, a parametric search is used to obtain the best-performing parameters.

In the NARX structure, the main parameters to be configured are the number of hidden layers, the number of neurons of the hidden layers, and the activation function of the layers. Since this is a regression problem, the activation function of the last layer will be linear. The number of neurons in the input layer equals the number of inputs (samples of the past voltage window, past temperature window, past current window, and future current window), and there is only one output, the future voltage estimation.

In the LSTM structure, the parameters to be configured are the number of LSTM cells and the size of memory variables $\mathbf{c}_n$ and $\mathbf{h}_n$ of each of them. As with the NARX structure, the last LSTM cell has a linear activation function since this is a regression problem.

For the NARX architecture, two hidden layers have led to better results than only one hidden layer. To find out the optimal number of neurons in these hidden layers, a parametric sweep has been carried out.

Similarly, in the LSTM network, two LSTM layers have been concatenated as they have led to better results than just one layer. The dimensionality of the memory variables for each layer has been swept to find the optimal size.

The results for the initial error metrics with both algorithms and different parameters have been benchmarked by using the RMSE, and are collected in Table 4.2.

Neurons 1st hidden layer	Neurons 2nd hidden layer	NARX Error	LSTM Error
25	12	0.667	0.310
50	25	0.595	0.269
100	50	0.430	0.222
200	100	0.568	0.246
300	150	0.632	0.240

Table 4.2. Error comparison for both networks and different hidden layer sizes.

As it can be observed, increasing the number of neurons of the structure above 100 and 50 in the first and second layers respectively does not have a meaningful impact on the error rates. In fact, it has proven to be detrimental. This has been reported for both algorithms. Thus, this has been the selected size for both algorithms.

To further reduce overfitting, dropout layers have been added after each hidden layer. Dropout layers are additional layers that modify the structure of the neural network. They randomly disconnect some neurons in these layers with a certain rate p when training. This makes the structure less dependent on specific neurons and reduces overfitting (Srivastava et al. 2014). In this case, a dropout rate p of 10% has been chosen.

Different activation functions for each structure have also been trained. The best results for the activation functions in the hidden layers have been obtained with the ReLU (Rectified Linear Unit) functions for the NARX structure and the hyperbolic tangent for the LSTM network (Hara, Saito, and Shouno 2015). As explained before, the output layer in both structures counts with a linear activation function since this is a regression problem and the output must be a real number.

The definitive parameters for both structures are collected in Table 4.3.

Algorithm	**Layer and neurons**	**Activation function**	**Optimizer**	**Drop-out**
NARX	100 – 50 – 1	ReLU	ADAM	10%
LSTM	100 – 50 – 1	Tanh	ADAM	10%

Table 4.3. Parameters of both networks.

4.3.3 Preliminary results

Different error metrics have been defined to evaluate the results of the different algorithms. Besides the RMSE, the Mean Absolute Error (MAE), and the Mean Relative Error (MRE) have also been used. The MAE and the MRE are defined by equations (4.11) and (4.12).

$$MAE = \frac{1}{m} \cdot \sum_{i=1}^{m} |y_i - \hat{y}_i| \tag{4.11}$$

$$MRE(\%) = \frac{1}{m} \cdot \sum_{i=1}^{m} \frac{|y_i - \hat{y}_i|}{y_i} \; x100 \tag{4.12}$$

Fig. 4.9 shows a sample time window with the results of the NARX and the LSTM predictions. In it, different cases have been included, such as multiple sunny (blue) and cloudy days (red). It can be appreciated how both algorithms follow the voltage waveform, both in sunny periods and on cloudy days. The LSTM seems to have higher accuracy, especially when predicting high voltage values.

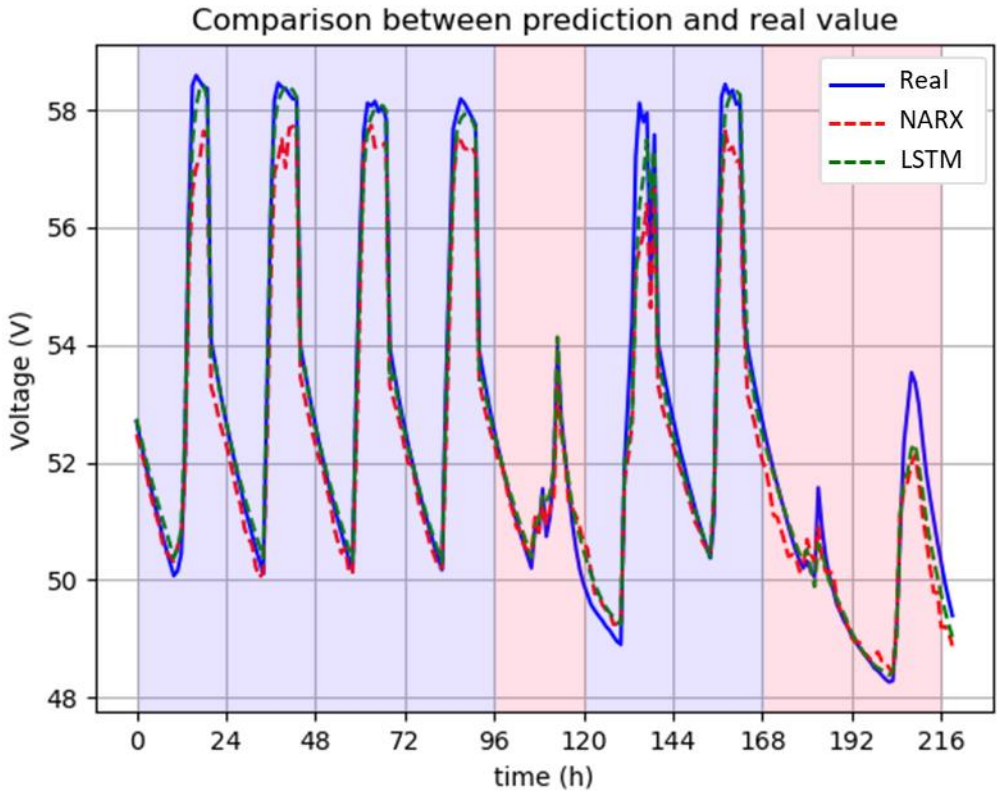

Fig. 4.9. Waveform prediction with both models and using future current as input.

To obtain comparable metrics, Table 4.4 includes the error figures obtained for both predictions. In general, results for both algorithms show that the LSTM prediction is better than the NARX prediction in all the error metrics. This can be especially appreciated in the high voltage part of the waveforms, which

belong to the float stage in the charging process. The LSTM algorithm is much more capable of following these values. However, these are not especially relevant intervals, since in these voltage ranges there is no failure probability. It can also be seen that the LSTM is a much slower algorithm, both in terms of training and inference time, which can be relevant since the ultimate goal of these algorithms is to be implemented in the Battery Management System of an ESS.

Algorithm	RMSE (V)	MAE (V)	MRE (%)	Training time	Inference time
NARX	0.670	0.464	0.863	1min 12s	432ms
LSTM	0.555	0.332	0.626	2h 23 min 11s	1min 38s

Table 4.4. Comparison of both network results.

A more exhaustive evaluation has been done in multiple PV standalone installations, to check if the algorithms are valid and expandable. The voltage ranges of the batteries in each installation are different, so the MRE (%) is the most adequate figure to compare the results (Table 4.5). It can be appreciated how the error metrics are in the table are in the range of 0.5-1%, which is comparable to the case of the Morrón installation previously considered.

Installation	NARX	LSTM
Acequia Ontiñena	0.811	1.048
Cartuja Monegros	0.549	0.481
Almenara Arba Luesia	0.879	0.956
Pluviómetro Candanchú	0.624	0.607
Repetidor Camero Nuevo	0.814	0.853
Repetidor Majalino	0.706	0.687

Table 4.5. MRE (%) results for other installations in the dataset.

As mentioned earlier, these predictions have been obtained by using future current measurements as inputs. To implement these algorithms in a real system, the future current pattern needs to be guessed. The following section goes deeper into this concept.

4.4 Prediction of the future current

So far, the algorithms have been using real future current waveforms as inputs to predict future voltage. This is possible in this stage of the study since future data is available in the dataset. However, when implementing these algorithms in a real system, future data will not be available as input. To solve this issue, a weather forecast could be used to predict the future current patterns, and these can then be used as inputs to the algorithm.

There are different approaches in the literature to daily photovoltaic generation clustering (M. S. Hossain and Mahmood 2020) by using irradiation patterns as input variables. An example of irradiation waveforms from one of the installations in the dataset, averaged by month, is shown in Fig. 4.10. However, in this case, the current is going to be used as an input variable instead of irradiation, due to its availability in the dataset.

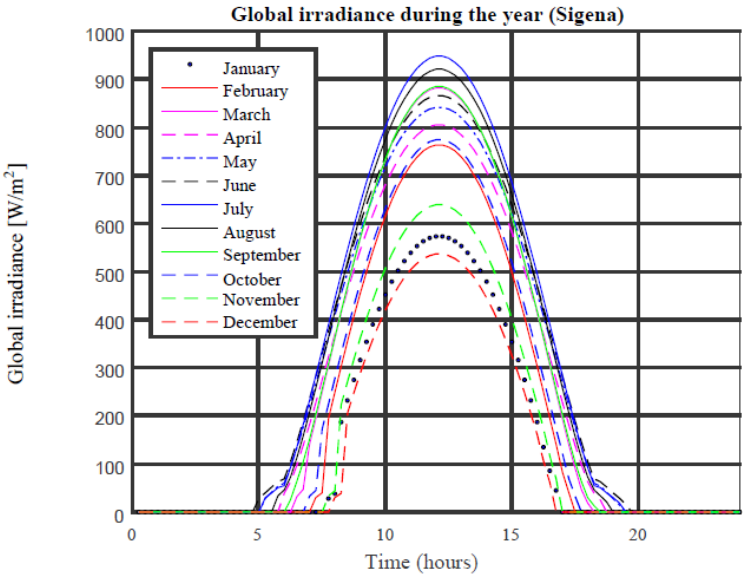

Fig. 4.10. Daily irradiation averaged by month.

The main idea is to obtain a classification of the different types of days that can appear in the installation under study (Sanz-Gorrachategui et al. 2020). Each of these types will have a representative current pattern. A weather forecast algorithm selects the type of day more likely to happen in the following days, and the voltage prediction algorithm uses the representative current pattern as the expected future current for voltage estimation purposes.

To obtain the classification of the types of current patterns, different features have been extracted from the waveforms. These features are going to be processed by non-supervised clustering techniques. The selected features are depicted in Fig. 4.11, and are the following variables:

- The voltage at the start of discharge (SoD): Voltage of the battery at the beginning of the night, when the solar panels stop producing energy.
- The voltage at the end of discharge (EoD): Voltage of the battery at the end of the night, when the panels start producing energy.
- Charge duration (hours): Time-lapse between the end of the discharge and the start of the following discharge.
- Discharge duration (hours): Time-lapse between the start of the discharge and the end of the discharge.
- Stored charge: Integral of the current during the charging process.
- The average temperature during the charge.
- The average temperature during the discharge.

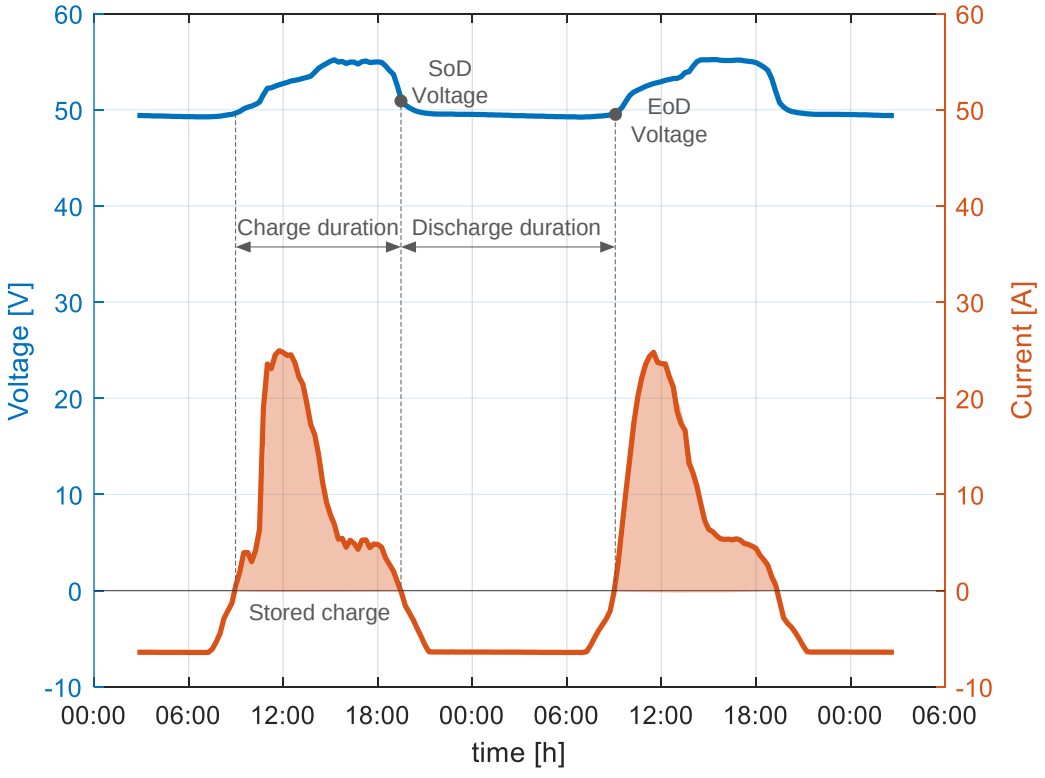

Fig. 4.11. Features for day classification.

4.4.1 Clustering algorithm

Self-Organizing Maps (SOM) and K-means have been used as clustering techniques to determine the day classification (Teuvo Kohonen 2001; T. Kohonen et al. 1996). SOM is an unsupervised neural network that is commonly used for pattern recognition, database visualization, and preliminary analysis of datasets. A SOM structure counts with a single input layer and a single output layer, the map itself (Fig. 4.12).

Given an input layer with K variables x_k $(1< k < K)$ and an output map with N x M neurons, each neuron *(i,j)* stores a vector of synaptic weights $\mathbf{w}_{ij}$, with K components *(1≤i≤N, 1≤j≤M)*. The map has two operation modes: training and inference. In training, the synaptic weights are initialized to a random value and are then adjusted. For each example in the dataset, the Best Matching Unit (BMU) is obtained as the neuron with the least distance d to that specific example, where d is defined as the Euclidean distance (4.13).

$$d_{ij}(\mathbf{w}_{ij}, x) = \sqrt{\sum_{k=1}^{K} \left(w_{ij_k} - x_k\right)^2} \tag{4.13}$$

Then, the synaptic weights w_{ij_k} of the BMU and the neurons on its neighborhood are adjusted to become closer to the input example, according to the expression in (4.14). Here, $\alpha(t)$ is the learning rate, which decreases as the training process progresses (4.15).

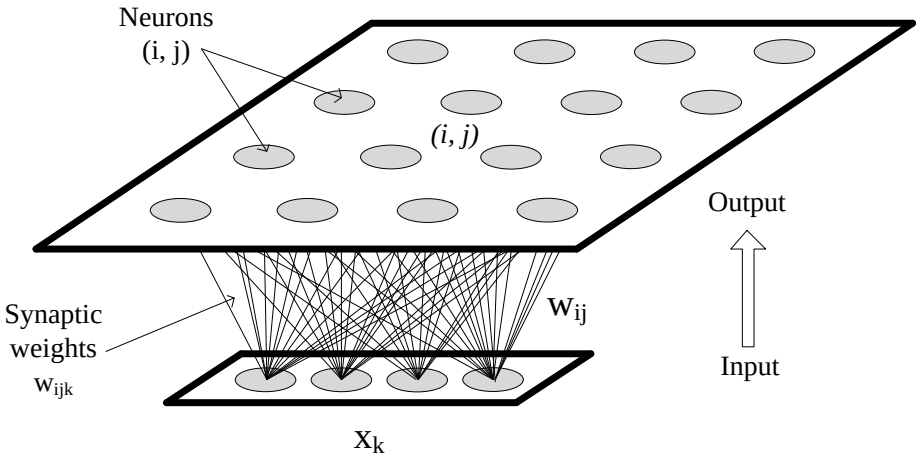

Fig. 4.12. Generic SOM.

$$w_{ijk}(t+1) = w_{ijk}(t) + \Delta w_{ijk}(t)$$
$$= w_{ijk}(t) + \alpha(t) \cdot \left(x_k(t) - w_{ijk}(t)\right) \quad (4.14)$$

$$\alpha(t) = \alpha_0 + \frac{t}{t_f}(\alpha_f - \alpha_0) \quad (4.15)$$

Where α_0 is the initial learning rate, α_f is the final learning rate and t_f is the number of iterations to get to the final learning rate.

After training the algorithm, the SOM has generated a non-linear, two-dimensional projection of the multi-dimensional dataset. Then, in the inference phase, the data is passed again through the algorithm, and the BMU is obtained for each example. By doing so, the map groups similar data patterns into closer regions.

4.4.2 SOM Results

The seven features described earlier have been extracted for each day in the database, then normalized in the range [0, 1], and finally used for training the SOM. A map of 10x6 neurons has been selected as a standard size.

The map can be visualized by using the U-matrix, which shows the distance between the weights of adjacent neurons. The component planes can be visualized as well, which are the map for each of the different weights associated with each of the variables, see Fig. 4.13.

In this case, the U-matrix does not show clear clusters. Looking at the component planes, the temperature maps are similar, thus these variables are redundant. They are correlated to the charge and discharge duration, since they mainly evolve in the vertical axis. As temperature and discharge duration are related to the seasons, it makes sense that some degree of correlation exists between them.

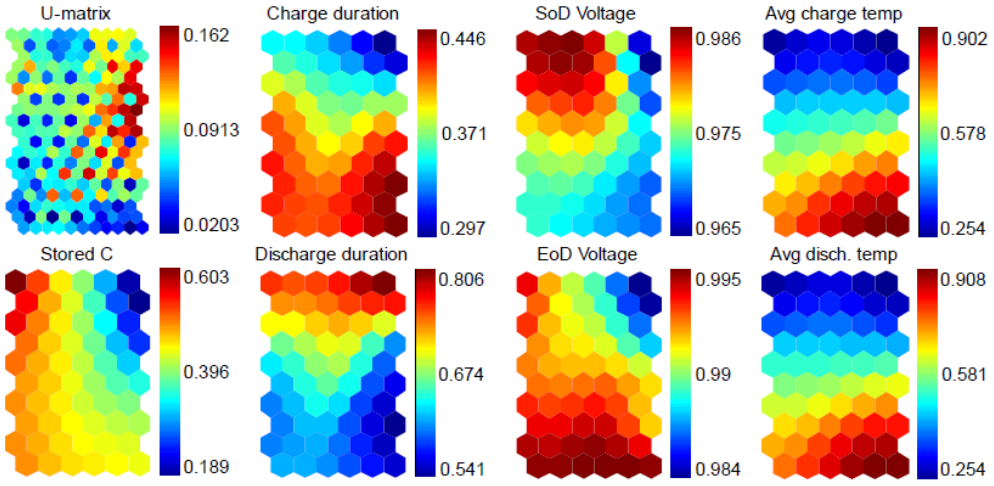

Fig. 4.13. SOM results with temperature input (U-matrix and component planes).

To avoid redundancy, it has been decided to remove the temperature-related variables, thus obtaining the map in Fig. 4.14. On it, the U-matrix shows a central area of average days, while clear different days are grouped in the upper-left and the bottom-right corners. This distribution is correlated with the variable Stored Charge, as it can be observed on its component plane (Stored Q). However, this behavior is not directly correlated with the component planes of other variables such as the SoD and EoD Voltages, and the charge and discharge durations, indicating that multiple sub-groups may appear. To obtain a clear division of the classes in the dataset, a clustering technique is going to be applied.

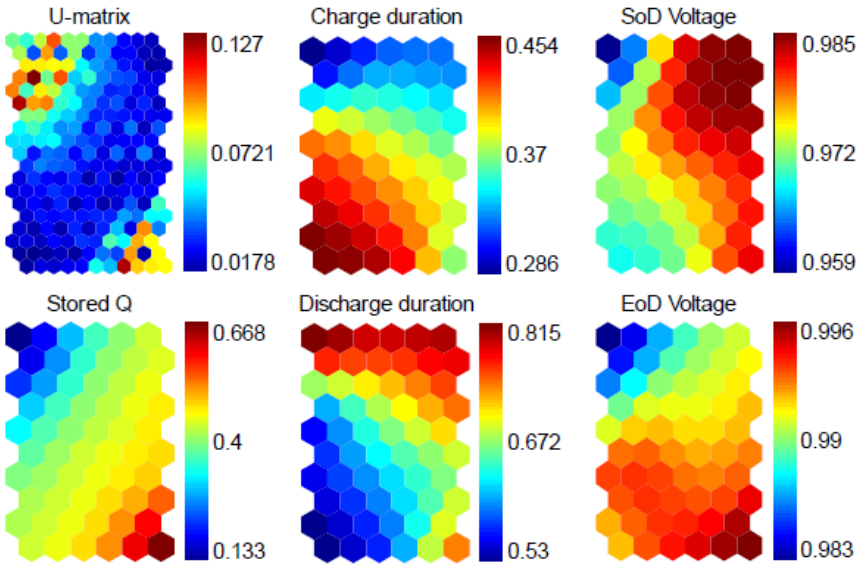

Fig. 4.14. SOM results without temperature input (U-matrix and component planes).

4.4.3 Clustering

To identify the different clusters (each of them representing a day type), the K-means algorithm has been applied to the outputs of the SOM. The K-means algorithm is a process that establishes k random centroids in the space of the data and assigns each example in the set to the closest centroid. Then, it updates the centroid to the mean value of each of its examples. This process is repeated iteratively until all the centroids are stable, meaning that the algorithm has converged. However, the K-means algorithm needs the number of clusters k as an input. To obtain the optimal number of clusters, different methods may be applied (Kodinariya and Makwana 2013). Among others, probably the simplest method is the "elbow" method, which is based on finding the elbow point in the Sum of Squared Errors (SSE) curve (4.16), where $\boldsymbol{k}$ is the number of clusters, S_i is the i-th cluster, $\boldsymbol{\mu_i}$ is the centroid of the i-th cluster, and $\boldsymbol{x_j}$ represents each of the examples assigned to the S_i cluster.

$$SSE(k) = \sum_{i=1}^{k} \sum_{x_j \in S_i} \left\| \boldsymbol{x_j} - \boldsymbol{\mu_i} \right\|^2 \tag{4.16}$$

The SSE value has been obtained for different numbers of k clusters, resulting in the curve in Fig. 4.15. Increasing the number of clusters decreases the value of the SSE, but worsens the clustering capabilities of the K-means algorithm. In this case, the optimal point (elbow of the curve) can be found between three and five clusters.

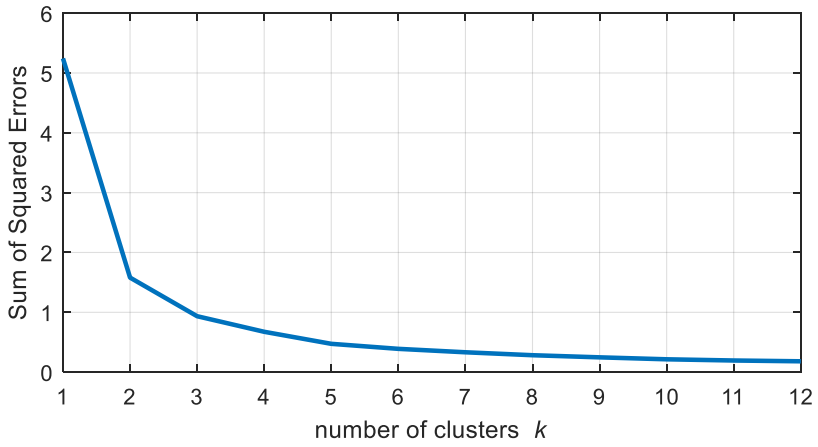

Fig. 4.15. SSE curve.

Fig. 4.16 collects the results of the SOM with three, four, and five clusters. After classifying each of the days into one of the clusters, the current pattern of each type of day has been obtained by averaging the current waveform of the days in each cluster. These average current patterns are shown in Fig. 4.17.

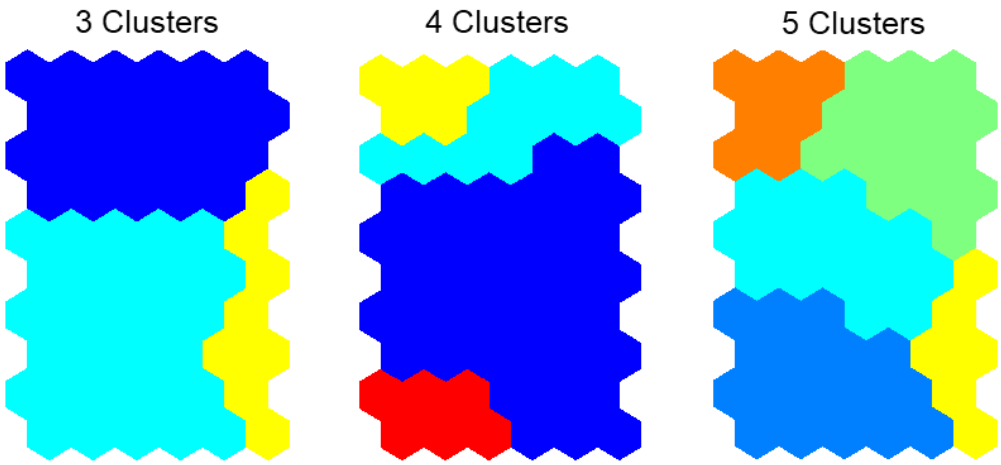

Fig. 4.16. SOMs with three, four, and five clusters.

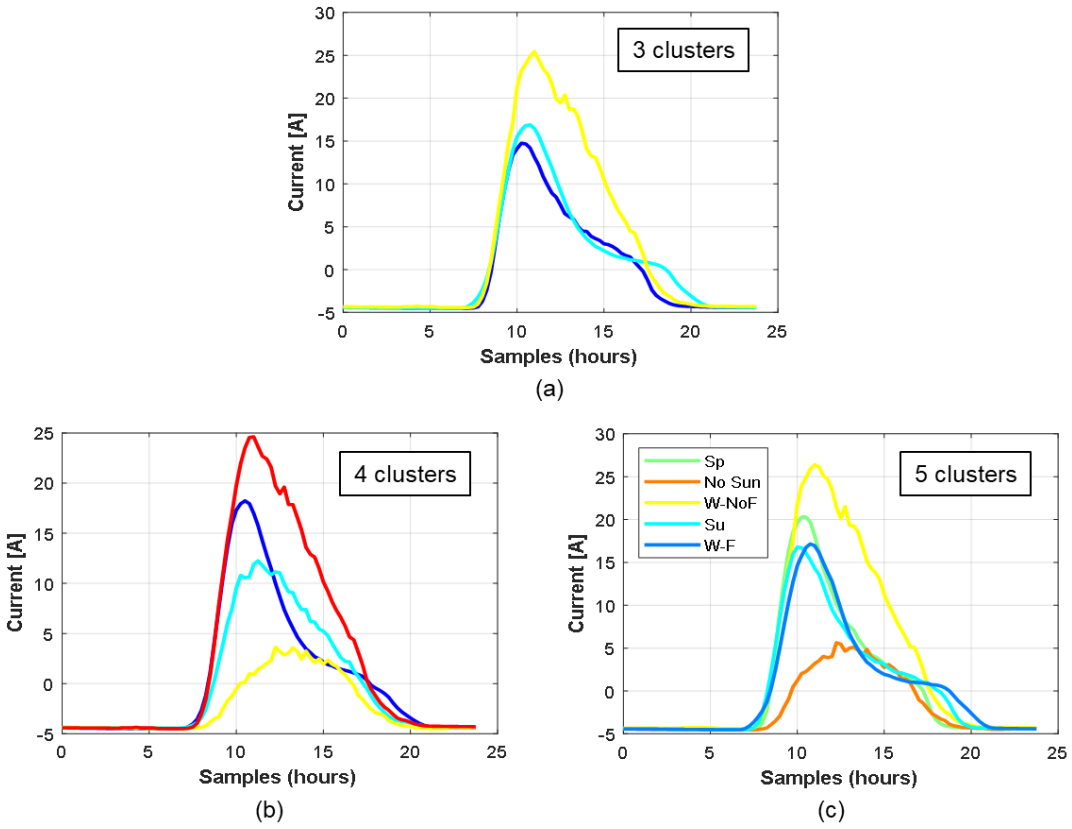

Fig. 4.17. Average current patterns with three (a), four (b), and five (c) clusters.

The case of three clusters does not separate sufficiently some of the day types, since there is not a clear low irradiation day, and there are two different medium irradiation days (cyan and blue cases). In the case of five clusters, there are multiple clusters for an average day (blue, cyan, and green). Here, a high seasonal correlation can be observed in the yearly distribution of these days, see Fig. 4.18.

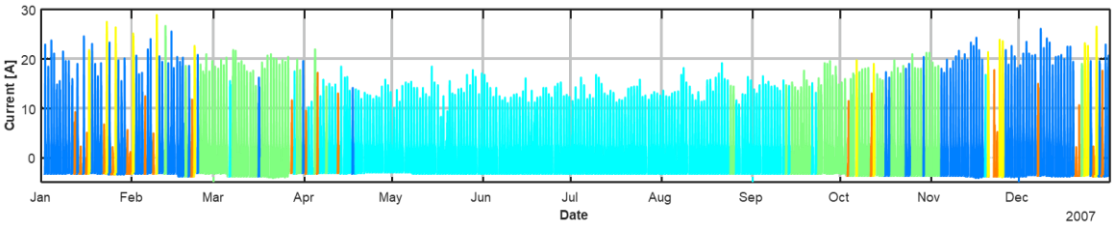

Fig. 4.18. Distribution of the 5 day types along a year.

According to the yearly distribution, each cluster can be labeled regarding the season and its irradiation in one of the following five categories:

- **No Sun** (orange): Days without much irradiation.
- **Summer** (blue): Average summer day.
- **Spring / fall** (cyan): Average spring or fall day.
- **Winter with float stage** (green): Average winter day.
- **Winter without float stage** (yellow): Sunny winter day after some low irradiation days, when the battery does not reach the float stage and absorbs all the generated energy.

As it can be observed, the **Summer**, **Spring/fall** and **Winter with float stage** classes have a very similar current pattern. These patterns are not distinct enough to deserve a separate class. For this reason, the case of five clusters has been discarded.

Among these three cases of study, using four clusters provides the most distinct types of day. Here, the seasonal component does not appear, and the different current patterns are correctly separated in terms of generation. The results of this map and the different average current patterns are shown in Fig. 4.19.

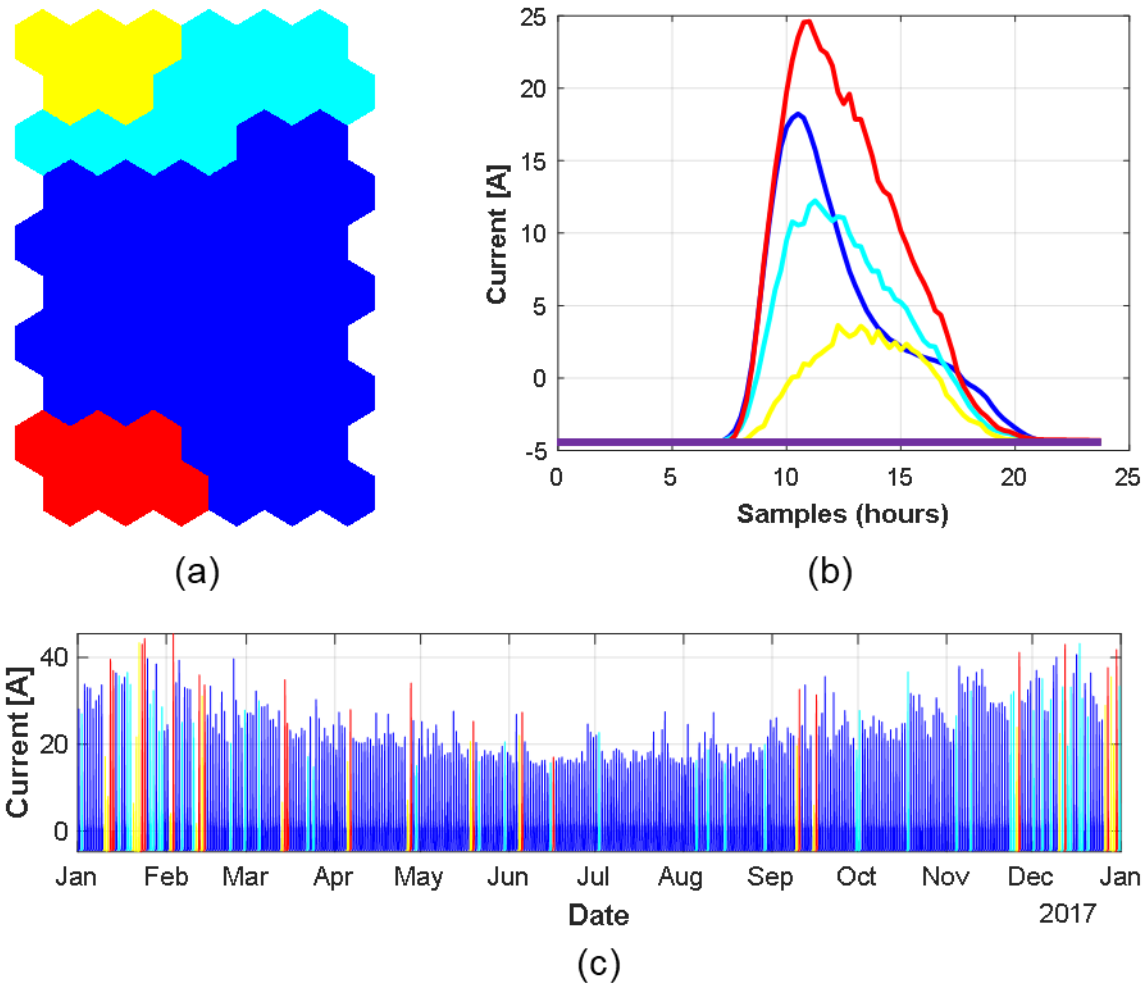

Fig. 4.19. Final SOM and k-means results. SOM (a). Current pattern from each cluster (b). Distribution of each pattern along a year (c).

The classes for the case of four clusters have been labeled as:

- **Low irradiation day** (yellow): These are rainy or foggy days when solar irradiation is at its lowest.
- **Medium day** (blue): This is usually a cloudy day, with some sunny periods. Absorption is not regular along with the whole day and rarely reaches the maximum available power.
- **Good irradiation day** (cyan): This is a regular sunny day, with the battery already in a high SoC. The battery absorbs the maximum available power in the morning, and around 11am, it reaches the float stage. From here, the current decreases until the evening, when the battery starts discharging.

- **High absorption day** (red): This is a high irradiation day, with the battery in a low SoC at the beginning. On it, the absorption follows the same trend as the good irradiation day, but it takes longer for the battery to reach the float stage, since its starting SoC was lower, and therefore it is capable of absorbing an increased amount of charge.

An additional class has been added to the taxonomy:

- **Day without absorption** (purple): This waveform emulates zero energy absorption. This situation occurs on days with either snow blocking the solar panels, a blizzard, or technical failures.

In the dataset, there are very few events like this, so the algorithm does not create a separate cluster for them. However, it shows a clear different current pattern, so it has been added manually.

When implementing the whole system, the weather forecasting algorithm will select two hypotheses for the future current of the forthcoming two days (48 hours). Then, the voltage prediction algorithm will use these hypotheses as the future current input. This automatic weather forecast algorithm is not a goal of this thesis, but could be developed as a future research line. Current versions of the system use an operator to select the future current patterns.

4.5 Results

4.5.1 Prediction results

The same time window of Fig. 4.9 is included in Fig. 4.20, but here the prediction has been estimated by using the current pattern from the day types determined with the SOM. The predictions are obviously not as good as when using real future data, but they still provide an approximate idea of the behavior of the storage system in the forthcoming days.

The numerical error metrics have been obtained for the whole test subset, and are collected in Table 4.6. The new metrics are larger than those obtained previously, with real future current, but are still acceptable for the application.

Algorithm	**RMSE** [V]	**MAE** [V]	**MRE** [%]
NARX	0.986 (0.670)	0.656 (0.464)	1.214 (0.863)
LSTM	0.883 (0.555)	0.550 (0.332)	1.029 (0.626)

Table 4.6. Final prediction error metrics.

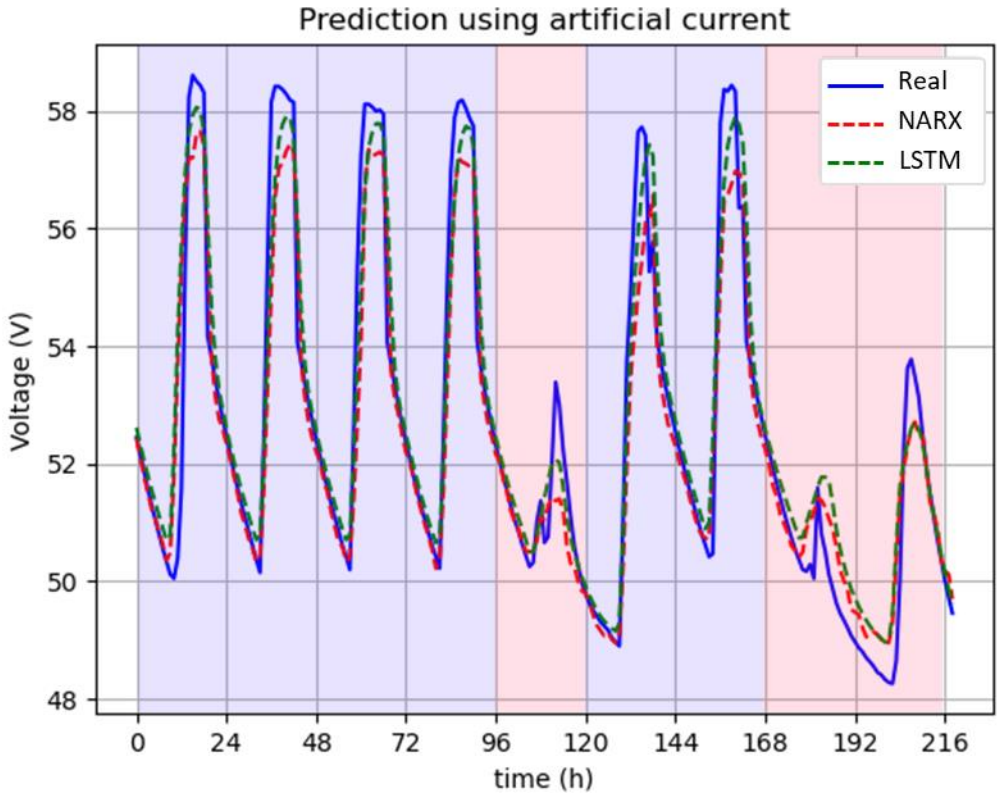

Fig. 4.20. Waveform predictions using artificial current as input.

A longer waveform has been selected and depicted in Fig. 4.21. This specific window has been selected since it includes multiple different irradiation cases and an eventual failure of the system. In the figure, it can be appreciated how the prediction during regular days is accurate, specifically during the discharges.

The prediction in medium or low irradiation days is also accurate, up until the 432h mark in the graph, when voltage drops below 48V. Here, four consecutive low irradiation days take place, and the battery suffers from a deep discharge. In this case, the algorithms do not predict accurately the whole voltage evolution. This effect is reasonable. In the database, there are few events with four or more consecutive low irradiation days, so the neural networks are not trained with this type of situation. Thus, are not capable of replicate properly these extreme cases. However, the decreasing trend of the voltage is indeed captured. This trend can be used as an alert indicator, so the operators can anticipate a low energy state within this 48h interval. To obtain an estimate of the error in these low-energy events, the error figures for voltages below 50V have been obtained in Table 4.7.

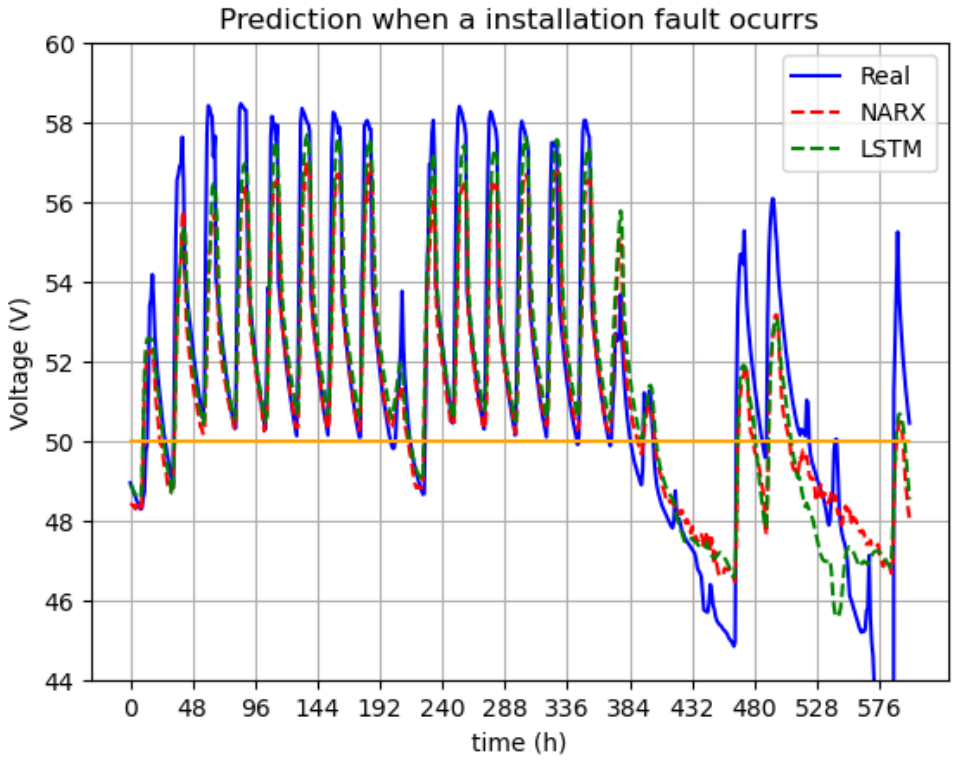

Fig. 4.21. Prediction in a system failure situation.

Voltage range	Algorithm	RMSE [V]	MAE [V]	MRE [%]
Full voltage range	NARX	0.986	0.656	1.214
	LSTM	0.883	0.550	1.029
< 50V	NARX	1.128	0.898	1.898
	LSTM	1.298	1.060	2.231

Table 4.7. Prediction error metrics for voltage ranges below 50V.

It is interesting to note that the LSTM algorithm makes more pessimistic estimations than the NARX, around the 528h mark. It could be considered as a more adequate algorithm if a more conservative approach is needed.

4.5.2 Preliminary implementation

A preliminary version of the whole system has been implemented in CHE servers. In this implementation, both an estimate for the following 24 and 48 hours have been included. A capture of the software tool is shown in Fig. 4.22. In it, the red plot is the actual measured voltage value in the installation. The green plot is the estimate for the voltage 24 hours ahead, and the orange plot, 48 hours ahead. In this version, the future current patterns are selected automatically by a weather forecast service that runs in the same server. The results are promising but the final error evaluation needs a larger amount of data, of some years, to be evaluated.

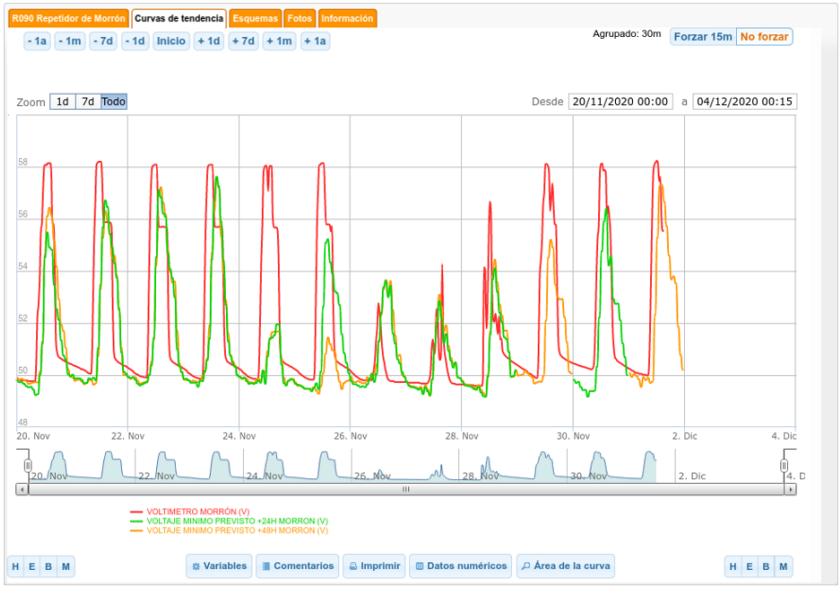

Fig. 4.22. Capture of the implemented system in CHE servers.

4.6 Conclusions

The prediction of future voltage in standalone photovoltaic installations has been addressed in this chapter. By processing temperature, voltage, and current waveforms, already measured in most installations, the proposed algorithms are capable of estimating future voltage, thus detecting possible system blackouts with a wide enough margin to respond accordingly. These algorithms can be used by BMSs as a means to improve existing charging methods, which accelerate aging by overcharging batteries periodically. By only overcharging them when a failure is predicted, the impact on aging is minimized.

To this end, two different sequence-processing algorithms have been proposed for this task: NARX and LSTM. The initial error metrics obtained are accurate, with an RMSE of 0.670V for the NARX algorithm and 0.555V for the LSTM algorithm, on battery packs with a nominal voltage of 48V. However, these error metrics have been obtained when using the real future current as an input to the algorithms.

To address the problem of using future current waveforms, the current generation patterns have been studied and classified by using a non-supervised machine-learning technique, Self-Organized Maps. It has been observed that, in terms of irradiation, four main different types of day appear. A fifth type (without any irradiation) has been added to consider panel failure or very low irradiation periods (snow). A weather forecast algorithm selects these waveforms as hypotheses for future current scenarios, to be used as future inputs to the voltage prediction networks.

After using the predicted current patterns as the obtained error metrics are 0.986V RMS in the case of the NARX algorithm, and 0.883V RMS in the case of the LSTM algorithm. The error for low voltage ranges, below 50V in the studied installation, are 1.128V RMS in the case of the NARX algorithm and 1.298V RMS in the LSTM. These results are still considered valid for the application, since the decreasing voltage trend is still captured, as seen in Fig. 4.21.

The conclusions of the comparison between both algorithms show that LSTM provides better error metrics than NARX in general, but worse results when analyzing voltages below 50V. LSTM is a more complex algorithm to train and implement, obtaining higher training and inference times, which could be detrimental when implementing the algorithm in a simple BMS in an application, but adequate if a cloud-like implementation is considered.

On the other side, the NARX network obtains worse averaged error metrics, but better results when observing voltages below 50V. These results are accurate for the goal of the application. Its training and inference times are much lower, and thus it would be preferable for its implementation in a simple BMS or DC/DC converter.

After the results showed in this chapter, both algorithms have been applied to multiple installations, obtaining comparable results. However, these networks have been re-trained with data from each different installation. As a future research line, it would be of interest to develop a universal algorithm, capable of been deployed on any installation with similar features, without the need for retraining it specifically. This algorithm should also be capable of learning from the installation while in operation, adapting its predictions to the specific behavior of that battery pack.

Chapter 5.- Aging Estimation with Machine Learning Methods

This chapter addresses aging estimation for lithium-ion batteries like the ones proposed for HESS in previous chapters. This is necessary to develop models that are more robust and control algorithms. The goal is to determine the remaining useful life of batteries embedded in the application, by using simple test cycles, extracting key features from them, and training machine learning algorithms.

5.1 Introduction

5.1.1 Motivation

In the previous chapters of this dissertation, VRLA and LFP batteries have been used in HESS for standalone photovoltaic installations. The aging of lead-acid related technologies (such as VRLA), has been studied in depth in the past decades (Pascoe and Anbuky 2004; Sun, Jou, and Wu 2011; Marchildon, Doumbia, and Agbossou 2015) and is well known. On the other side, the aging of lithium-ion batteries has been extensively studied and has proven to be a challenging problem (Wright et al. 2003; Broussely et al. 2001; Bloom et al. 2001; Schuster et al. 2015; Chaoui and Ibe-Ekeocha 2017; Sepasi, Ghorbani, and Liaw 2015), and has gained attention in the recent years, probably motivated by the EV application.

For years, a lot of effort has been put into obtaining different aging models for different battery chemistries, with different goals. Concepts such as State of Health (SoH), End of Life (EoL), and Remaining Useful Life (RUL) have been explored as measurements for battery aging. They have different meanings and purposes depending on the specific application. On some applications, they report the health of the battery, so it can be replaced accordingly. In other applications (such as Energy Management Strategies in HESS), the aging estimation algorithm may play a fundamental role in the decisions taken by the system.

Second life of batteries is another useful application of health estimators (Martinez-Laserna et al. 2018). A common conclusion is that a suitable second life application is a low-demanding application in terms of power requirements and depth of discharge (DoD), such as energy storage in on-grid systems (Jiang et al. 2018; Martinez-Laserna et al. 2018; E. Hossain et al. 2019). Its role in off-grid systems still needs to be evaluated, since it tends to be a more intensive application in terms of DoD, but it is a promising application as well.

Remaining Useful Life estimation studies have traditionally been focused on in-situ applications i.e. estimating the RUL of batteries that are already embedded in an application and monitored. Here the main goal is to predict the failure threshold (D. Liu et al. 2015; Lyu Li et al. 2019; Khelif et al. 2017; D. Liu et al. 2013; C.

Chen and Pecht 2012; J. Liu and Chen 2019). These approaches usually require knowledge of the past use of the battery (Martinez-Laserna et al. 2018).

In this dissertation, Chapter 3 has been focused on improving the performance of ESS in terms of energy absorption and the State of Charge. Long-term performance has not been addressed, since there is still uncertainty in the aging of LFP batteries. The whole picture in lithium-ion aging is much wider, and the goal of future work. This chapter aims to provide some novel steps in RUL estimation for this kind of battery, as well as some additional tools to be used by BMS on in-situ applications. Additionally, these algorithms should remain simple enough to be used in low processing-power microprocessors, such as those used in these applications.

Although these techniques have been developed with data from LFP batteries, their portability to other chemistries is the subject of study, since in the HESS application there is interest in developing enhanced life estimation algorithms for the VRLA batteries as well. However, one of the singularities of this HESS application is the fact that the battery packs may already be in place, with non-monitored, used batteries. This is usually not desired when using health estimation methods, since they usually require past knowledge of the use of the battery to estimate its state of health and to predict its RUL. Thus, it is interesting to develop tools able to obtain health-related parameters without knowing the past of the batteries. Additionally, these algorithms would be needed in the case of using second life batteries in PV applications, which is a promising niche (since in this case, past information is not available).

5.1.2 Background

Many studies have been focused on determining the aging mechanisms of lithium-ion batteries. Multiple aging factors have an impact on cells' life and performance. These include:

- High and low temperatures (Jaguemont, Boulon, and Dubé 2016; Amine, Liu, and Belharouak 2005).
- High current cycles (Chandrasekaran 2014).

- Overcharges and overdischarges (Ohsaki et al. 2005).
- Mechanical stress (Christensen and Newman 2006).
- Time, or calendar life (Grolleau et al. 2014).

These different aging factors have a different impact on the degradation of the cell. A widely accepted approach is that there are three main degradation types: the increase of the batteries' equivalent series resistance, the Loss of Active Material (LAM), and the Loss of Lithium Inventory (LLI) (Barré et al. 2013).

Regardless of the aging factors or mechanisms, many different techniques and algorithms have been developed for SoH and other health-related parameters in the literature such as State of Power (SoP). Traditionally, two main approaches have been used: model-based and data-driven techniques. Model-based algorithms are suitable for battery behavioral models (such as State of Charge estimators) but have also been used for aging estimation. Model-based used for battery aging estimation include Kalman or Extended Kalman Filters (Plett 2004; Zou et al. 2015), stochastic techniques (He et al. 2011; Saha et al. 2009), or Particle Filters (C. Chen and Pecht 2012; Z. Liu et al. 2017; Miao et al. 2013; Wei, Dong, and Chen 2018). These algorithms are adaptive and are usually meant for tracking cells' life while it is operating.

Due to the large number of agents involved in battery aging, a data-driven approach is also suitable for this task (Xiong, Li, and Tian 2018; Y. Li et al. 2019). However, data-driven approaches need large and reliable datasets to be properly trained, which can be difficult to obtain. Different techniques have been applied to this field, such as Multi-Layer Perceptron, Support Vector Machines, Gaussian Process Regression, Long-Short Term Memory networks, or Echo networks (Yang et al. 2018; D. Liu et al. 2013; 2015; Pajovic, Orlik, and Wada 2018; Jia et al. 2020).

Until very recently, publicly available datasets were small (D. Liu et al. 2015; Ren et al. 2018). Practically the entirety of the studies referenced above uses the same dataset (Goebel and Saha 2010), which only counts with a handful of cells. The dataset introduced in (Severson et al. 2019) was published recently (March 2019) and is the largest currently available for the LFP chemistry. It contains information on 124 cells, which are cycled until their EoL. In the original paper,

the authors use this dataset for RUL prediction purposes, introducing some novel health indicators (HI). Their focus is the early RUL estimation on a monitored cell before it shows capacity-fade-related effects.

5.1.3 Objectives

In this chapter, the work in (Severson et al. 2019) will be taken as a starting point. The same dataset is going to be used, but the goal of the algorithms will be RUL and capacity estimation. Through few test cycles, features from the measured voltage and current waveforms are extracted and processed by the data-driven algorithms. As a simplification, this work is going to consider that the conclusions reached for this database (which is based on individual cells) are valid for complete ESS systems based on the same technology.

Different estimation approaches are going to be addressed, with both classification and regression tools. Additionally, different HI are going to be used as features. Some of them have already been widely used in the literature, while others are less explored. The goal is to develop simple machine-learning techniques, capable of providing estimations on different health-related variables of the battery packs in the HESS application.

The chapter is organized as follows. Section 5.2 describes the methodology that has been followed through the rest of the chapter, including the different estimation algorithms, a description of the dataset, and the selected features. Sections 5.3, 5.4, and 5.5 collect the development of the machine-learning algorithms and their results. Section 5.6 gathers the conclusions of the chapter.

5.2 Methodology

5.2.1 Estimation approach

When estimating battery aging, most studies in the state of the art track the evolution of capacity and other health-related parameters from the beginning of the life of the cell. They use past information to predict the evolution of the cell, and determine its End-of-Life (EoL). On many occasions, HESS are built by upgrading the existing ESS systems with an additional battery pack. Thus, there is no information available on the past use of the existing batteries. For this reason, there is interest in developing life-estimation tools that do not rely on knowledge of the past of the storage packs.

On a different axis, two main approaches can be taken when studying RUL estimation: "direct" or "indirect" estimation (D. Liu et al. 2015; Lianbing Li et al. 2019). The direct estimation approach uses cell capacity as a feature to predict RUL, whereas indirect estimation does not. Direct estimation is usually possible if the cell is in a controlled environment (e.g. a laboratory) and it is possible to cycle the battery to extract features. On the other side, indirect estimation does not rely on capacity as a feature to determine cell aging and remaining useful life. It is aimed at industrial applications where the battery is not cycled completely, and therefore it is not possible to disconnect cells to measure their features.

Conventional ESS with only one storage device would need the indirect approach since it is not possible to freely discharge the battery to obtain its capacity (doing so would imply the system to shut down). In the case of HESS, since the system has multiple storage devices, they can be discharged in a controlled manner by charging other elements, obtaining capacity, and other related features. This allows for applying direct estimation approaches on applications with HESS.

Thus, this chapter is mainly going to focus on the direct estimation of life-related variables of the battery, not knowing the past usage of the cells. Three main different approaches are going to be followed:

- **Life stage classification**. With this approach, cells are classified into different categories, depending on whether they can be used in a long term, or should be replaced soon.
- **Remaining Useful Life estimation**. This approach aims to give an estimation of the number of remaining useful life of a certain cell. Here, RUL [cycles] is defined as the remaining number of cycles until the battery reaches its end of life.
- **Straightforward capacity estimation and capacity curve regression**. These approaches use the features and algorithms explored in previous sections to perform capacity estimation. On the one hand, a straightforward approach is proposed to estimate capacity from other related features, aimed for applications where it is not possible to measure it directly. On the other hand, the capacity evolution for future cycles is predicted from a past window of capacity measurements.

5.2.2 Dataset description

The dataset that is going to be used to train the algorithm in this chapter has been introduced in (Severson et al. 2019). It is currently the largest publicly available database, containing information from 124 commercial LFP/graphite cells. The original study uses this dataset to estimate RUL after the 100^{th} cycle of a cell that has been monitored from the beginning of its life.

In the dataset, the cells are cycled until their EoL, defined as the moment when their capacity falls below 80%. Cell life ranges from 170 cycles to 2237 cycles in extreme cases, averaging around 750 cycles. This provides more than 90000 full discharge cycles in the whole dataset, with several measured features on each cycle.

The specific model of the cells is the APR18650M1A from A123 Systems, with 1.1Ah of nominal capacity. The cells have different high-current charge patterns, but they have all been discharged with the same discharge current (4C) along with all their life. During all the cycles, the ambient temperature was controlled to be 30ºC. Waveforms for voltage, temperature, charge, and time were measured for each cycle of each cell. The dataset also provides the observed capacity and an estimate of the internal resistance of the cell at each test cycle.

The dataset is provided by the original authors in three data batches. In the paper, they use certain batches for training purposes and others for testing, but there are clear differences in the behavior of the cells in each batch (Fermín-Cueto et al. 2020). For example, batch #1 contains cells with long life, above 500 cycles, while in batch #2 none of the cells lasts longer than this. For this reason, data across all the batches will be used for training and testing here.

As a note, the case studied in this dissertation is the case of LFP batteries in HESS for PV applications. The Energy Management Strategy described in the experimental setup in Chapter 3 gives priority to the charge and discharge of LFP batteries. Under this regime, they suffer from deep daily cycles. In this sense, the dataset fits the singularity of the application. However, the author is aware that the dataset contains cells cycled with high current profiles, and this is not the case in the studied HESS application with a conventional sizing. Nevertheless, in the absence of a better-suited dataset, this is going to be considered as a starting point.

5.2.3 Feature selection

The algorithms described below use several features obtained from the test cycles. Some of these features, such as Capacity or Internal Resistance, are well known and are usually used for SoH or RUL estimation purposes. On the other side, other features such as Time Interval of Equal Discharging Voltage Difference (TIEDVD) or Incremental Capacity have been less explored. Nevertheless, they have proven to be good Health Indicators (HI) in other battery aging studies (D. Liu et al. 2015; 2013; Ansean et al. 2019).

The original study in (Severson et al. 2019), where the dataset was introduced, considered each cell as a separate example, since the algorithms were simple enough to be trained on just 124 examples, and tracked the evolution of each cell. On the contrary, this approach will consider each cycle in the dataset, rather than each cell, as a separate example. The number of examples increases from 124 cells to more than 90000 cycles and so, complex data-driven algorithms, which require more data examples, can be trained. This approach will have relevant implications, as discussed later.

With this approach, each example has its own set of features, and will have as output the number of RUL cycles of the cell at that specific moment of its life. Some of these features, such as the measured capacity and internal resistance, are already provided by the dataset for each cycle of each cell. Other features are computed from processing the waveforms of each cycle. Each feature has been analyzed separately, and has proven to contain useful information towards health estimation. The main features that are going to be considered are described below.

Capacity

Capacity, C [Ah] is defined as the amount of charge extracted in each full-discharge cycle. Fig. 5.1 shows an example of the evolution of this feature for one of the cells in the dataset. Degradation can be seen along with the lifespan of the cell, with an accelerated decay near the EoL. As explained in section 5.2.1, when considering direct RUL estimation, capacity is available as an input since the cells can be cycled as needed, which will be the case throughout most of the chapter. By contrast, section 5.5 will have capacity as the estimation goal, for applications where it is not possible to measure it.

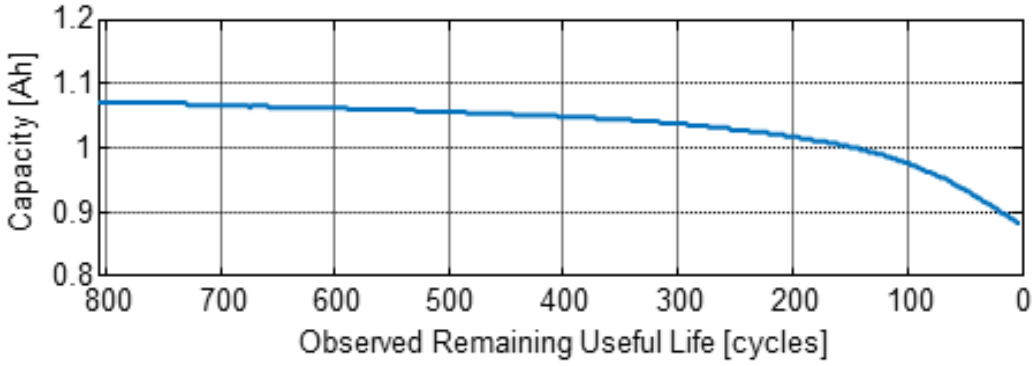

Fig. 5.1. Evolution of Capacity for one cell in the database.

Internal Resistance

Internal Resistance, IR [Ω] is defined as the equivalent series resistance of the cell. One measurement per cycle is provided in the dataset, obtained by averaging ten current pulses at 80% of State of Charge (SoC). An example of its evolution is shown in Fig. 5.2, where it can be observed how the internal resistance increases abruptly when the cell is near its EoL.

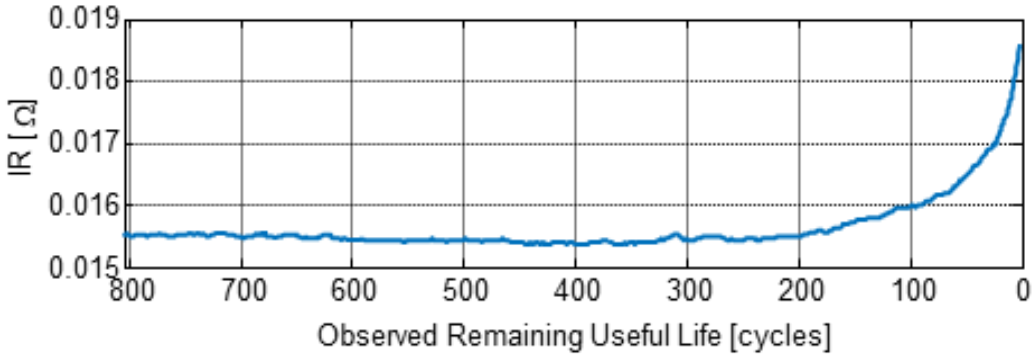

Fig. 5.2. Evolution of Internal Resistance for one cell in the database.

Time Interval of Equal Discharging Voltage Difference

Time Interval of Equal Discharging Voltage Difference, TIEDVD [min] is the time lapse between two voltage thresholds (v_{max}, v_{min}) as defined in (5.1), while discharging a constant current. An example of this is shown in Fig. 5.3. It has been used in the literature for indirect estimation purposes in some industrial applications (D. Liu et al. 2015; 2013) when it is not possible to completely discharge the cell to measure its capacity (e.g. an Uninterruptible Power Supply).

$$TIEDVD\ [min] = t_{v_{min}} - t_{v_{max}} \tag{5.1}$$

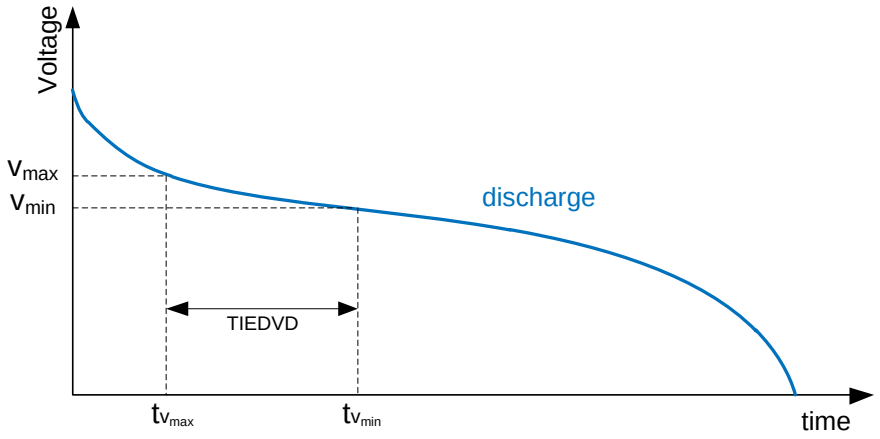

Fig. 5.3. TIEDVD extraction from a discharge cycle.

This feature has been obtained from the discharge voltage waveforms of each cycle in the dataset. Different threshold voltages have been tried for estimation purposes, and it has been found that the tuple ($v_{max} = 3.3$V, $v_{min} = 3.15$V) shows good estimation capabilities for this dataset. It can be appreciated in Fig. 5.4, where it decreases linearly with the remaining useful life.

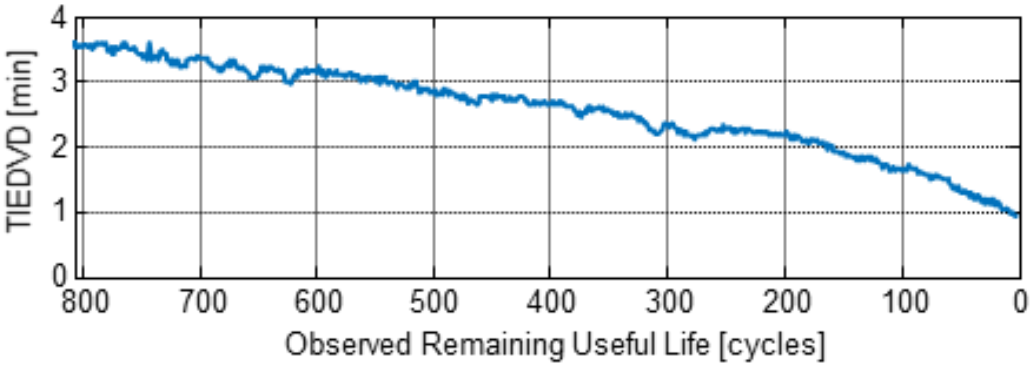

Fig. 5.4. Evolution of TIEDVD for one cell in the database.

Incremental Capacity

The previously introduced features have been used in the literature as health indicators for SoH estimation purposes. Additionally, this dissertation will include the Incremental Capacity (IC) as a feature to be considered in the estimation algorithms.

Incremental Capacity, IC [Ah V^{-1}] is the derivative curve of the charge vs the voltage of the cell (5.2) and can be obtained from the time-domain discharge waveforms (Fig. 5.5a). If the battery is seen as a nonlinear capacitor with a complex output impedance (Günther, Feldmann, and ter Maten 2005; Macdonald and Brachman 1955; Wyatt 1978; Sanz-Gorrachategui et al. 2019), it is the equivalent to the nonlinear capacitance curve of the cell, which indicates the energy storage capabilities across the voltage range of the cell. It has been described before as a relevant Health Indicator in different studies (Ansean et al. 2019).

$$IC = \frac{\delta q(v)}{\delta v} \tag{5.2}$$

The Incremental Capacity curves show different capacitance peaks at different voltage values, depending on each chemistry and aging process. These peaks show the regions of voltage where the cell stores energy (Fig. 5.5b), and change slightly as the cell ages. Similar capacitance peaks obtained with lower discharge currents have been studied for LFP batteries, and have been used as health indicators (Ansean et al. 2019; Severson et al. 2019). The peaks and their variation have been proven useful for battery life estimation, and for determining different degradation mechanisms (Ansean et al. 2019; Dubarry, Baure, and Anseán 2020).

In those studies, IC waveforms are obtained with low current rates on the cells. However, these approaches have not been extensively applied to high current waveforms. In the database used here, the discharge cycles were done at 4C, so these curves will be considered as an approximation of IC curves. When processing genuine IC curves, better results would be expected. The results of this study will be useful to determine whether these high current profiles are useful for this same task. If this is the case, it will allow the application of these techniques to real in-situ installations, where the cells are subject to real high-current discharges.

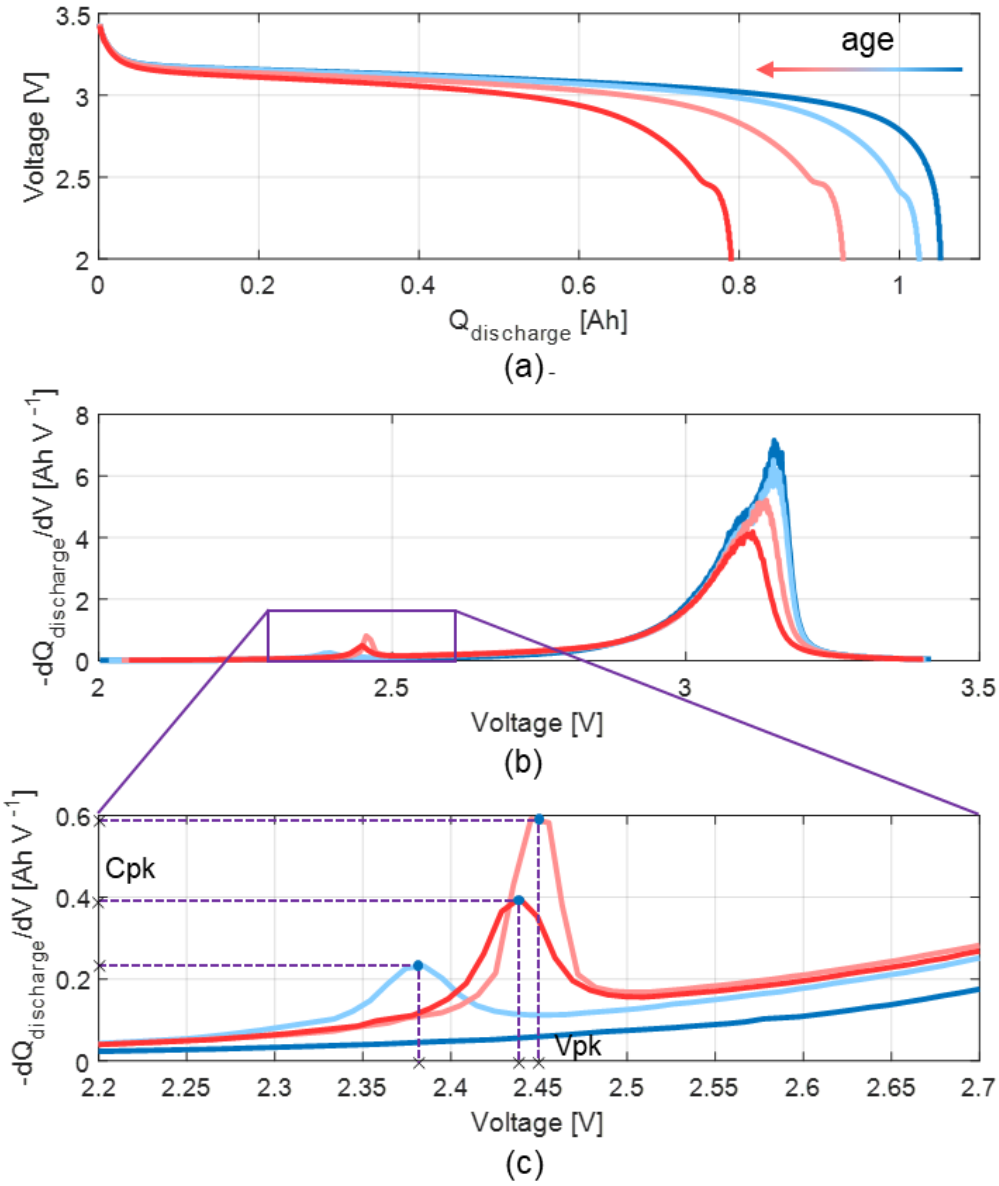

Fig. 5.5. Capacitance peak advent. Voltage vs. Charge waveform and its variation as the cell ages (a) IC curve, and its variation as the cell ages (b). Zoom in the lower voltage part, where the secondary Capacitance peak appears (c).

Each discharge cycle in the database can be minimally pre-processed to obtain its IC waveform. Since a cycle-level processing approach is being considered, each example in the dataset will have associated one IC waveform and the features that can be extracted from it.

Two approaches are going to be considered to process the information of the IC curves. The simpler algorithms that are going to be developed will use pre-processed data from the IC curves. On the other side, the more complex algorithms will be able to process whole IC curves to use as much information as possible.

In particular, the pre-processed features that are going to be used in the simpler algorithms are the following:

- **Capacitance peak, C_{pk} [AhV^{-1}]:** As the cell ages, a small disturbance appears in the low voltage part of the discharge waveform. It grows until a certain moment in its life and then becomes smoother when the cell is near its EoL. This generates a peak in the cell equivalent capacitance, as seen in Fig. 5.5c. This capacitance peak can be used as another feature. Fig. 5.6 shows how this peak evolves as the cell ages, increasing in the first hundreds of cycles, and then decreasing in the last 150 cycles approximately. However, it has a higher uncertainty when the peak is near its maximum. Additionally, when the peak is near its minimum (the first 80 cycles of life), the peak cannot be detected correctly.

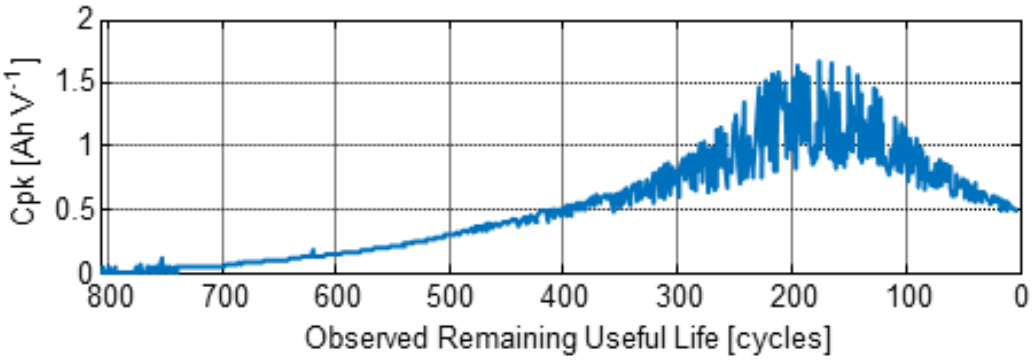

Fig. 5.6. Evolution of Capacitance Peak for one cell in the database.

- **Voltage at Capacitance Peak, V_{pk} [V]:** Additionally, the voltage at which the capacitance peak is observed (Fig. 5.5c) is used as an additional

feature, since it also varies with the cells life. The evolution of this feature is shown in Fig. 5.7. Similarly, it can be seen how in the first 80 cycles of life it cannot be captured correctly.

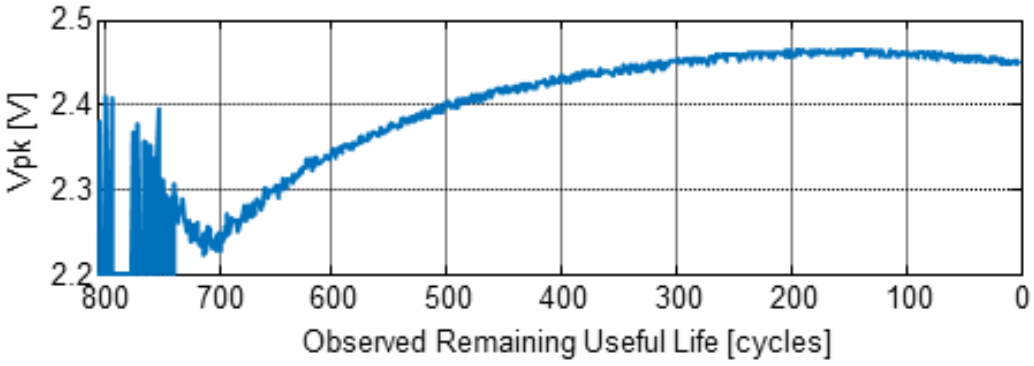

Fig. 5.7. Evolution of Voltage at Capacitance peak for one cell in the database.

It has been determined empirically that these variables are related and, when used independently as inputs for estimation algorithms, provide worse results than when being combined in a single variable due to redundancy in the information.

Particularly, it has been observed that the feature V_{pk} is highly correlated with $\log_{10}(C_{pk})$. For this reason, both variables have been combined into a single feature, to be used as input in the algorithms. The feature Pk is thus defined as in (5.3).

$$Pk = \frac{V_{pk}}{\log_{10}(C_{pk})} \tag{5.3}$$

5.2.4 Multiple measurements per variable

The cell under test can be cycled δ times to obtain multiple measurements of the previously introduced features. This way, each input example could have up to δ measurements for capacity, internal resistance, TIEDVD... Therefore, the network is provided with more information and better RUL estimation results can be expected. A greater value for δ increases the number of measurements and could improve the prediction capabilities, at expense of cycling more (and thus, aging) the cell for test purposes. It has been found that $\delta = 10$ provides the best tradeoff in accuracy. Later on, this aspect will be further developed.

The variation between consecutive cycles of some of these features can also be considered as an additional feature. Specifically, after an extensive search, the difference in capacity between cycles has proven to be useful. It has been renamed Capacity Fade.

Capacity Fade

Capacity Fade (ΔC_δ [C]) has been defined as the variation in capacity between two cycles separated δ cycles apart. An example of its evolution is shown in Fig. 5.8. As happens with Capacity and Internal Resistance, in the last hundreds of cycles, its behavior changes drastically.

By conducting δ test cycles to a battery, δ-1 measurements for capacity fade can be obtained, in addition to the δ measurements from each of the features previously introduced.

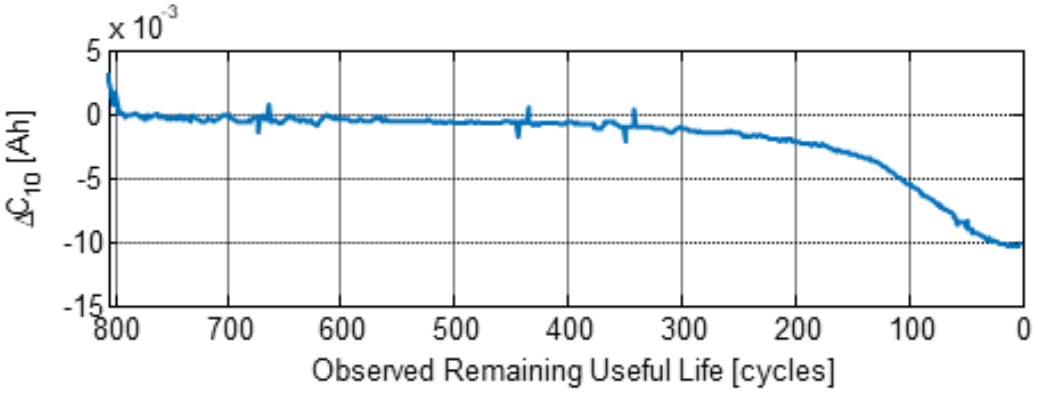

Fig. 5.8. Evolution of Capacity Fade for one cell in the database.

Summing up, each example in the dataset will have its own set of features, collected from the last δ cycles of the battery. These features will include:

- δ measurements for the features C, IR, TIEDVD, and Pk, belonging to the last δ cycles.
- δ -1 measurements for the feature ΔC, obtained by comparing the current Capacity with the δ-1 previous cycles.
- A single value for the RUL at that cycle, which will be the output to estimate.

5.3 Life stage estimation

The evolution of most of the features depicted in the previous section shares a distinctive pattern among most of the features: all of them except for TIEDVD and V_{pk} show a change in their trend as the battery approaches the last few hundreds of cycles of its life.

This effect can be appreciated in features such as Capacity, which decays faster in the last 200 cycles approximately. The threshold has been described as the "aging knee" of the capacity curve (Martinez-Laserna et al. 2018), and according to (Martinez-Laserna et al. 2016), batteries in this life region should not be considered for further use or a second life application. The effect does not only affect Capacity since features such as Internal Resistance or Capacity Fade also have a change in their behavior from this threshold on.

The effect has been reported across the whole dataset, in all the cells, to a certain degree. This suggests the possibility of training a classification algorithm to determine whether a cell is in this EoL region or not. This way, cells will fall into two different life stages:

- **Short RUL cells**: The cell has passed the "aging knee" and is in its last cycles of useful life. Cells in this stage should be replaced.
- **Long RUL cells**: The cell has not passed the "aging knee", and thus it is still suitable for use in an application.

5.3.1 Classification algorithm

The Regularized Logistic Regression (RLR) algorithm is proposed for the classification task (Haggstrom 1983), which is based in the conventional logistic regression algorithm. This network is probably the simplest data-driven classification algorithm, and has been chosen for this reason. In contrast with the conventional logistic regression algorithm, RLR includes a regularization term in its cost function to minimize overfitting, as explained below.

The model that this algorithm tries to fit is given by equation (5.4), where the predicted output $\hat{y}$ is the probability for the cell to belong to **Short RUL cell** or

Long RUL cell classes, **x** is an n-dimensional feature array, and **w** is an n-dimensional weight array.

$$\hat{y}(\mathbf{x}, \mathbf{w}) = \frac{1}{1 + e^{-\mathbf{w}^T \mathbf{x}}} \tag{5.4}$$

From here, the output class is obtained by thresholding $\hat{y}$ with 0.5, with $\hat{y} = 0$ meaning **Short RUL** and $\hat{y} = 1$ meaning **Long RUL**.

The key of the algorithm lies in the **w** vector, which needs to fit to provide an accurate classification. To obtain the best values for **w**, a training algorithm such as Gradient Descent fits **w** by minimizing the cost function in the expression (5.5).

$$J(\mathbf{w}) = \frac{1}{m} \sum_{i=1}^{m} \big(E(\mathbf{x}_i, \mathbf{w}, y_i)\big) + \lambda \sum_{j=1}^{n} {w_j}^2 \tag{5.5}$$

Here, the first summand represents the error in the prediction, with m being the number of examples in the training dataset and the error function $E(\mathbf{x}_i, \mathbf{w}, y_i)$ being the traditional cost function used for conventional logistic regression for a single example $\mathbf{x}_i$. Its full expression is given by (5.6), where y_i is the actual output associated with $\mathbf{x}_i$.

$$E(\mathbf{x}_i, \mathbf{w}, y_i) = -y_i \log\big(\hat{y}(\mathbf{x}_i, \mathbf{w})\big) - (1 - y_i) \log\big(1 - \hat{y}(\mathbf{x}_i, \mathbf{w})\big) \tag{5.6}$$

When working with data-learning tools, high variance (i.e. complex) algorithms may overfit the training data if trained for too long, over-learning the specific training examples. This results on very low errors with the training data, but poor generalization capabilities and thus, high errors with the test data.

Regularization techniques are a common tool used to avoid this issue. In the expression (5.5), the second summand represents the Ridge regularization term (Hoerl and Kennard 1970) that makes the algorithm less prone to overfitting by penalizing large weights in the **w** vector (with n being the number of weights of the RLR network).

5.3.2 Threshold selection

To establish the threshold that divides the **Short RUL cell** and **Long RUL cell** classes, an initial RLR algorithm has been tested. Different thresholds have been selected, and the algorithms have been compared to obtain the best-performing threshold. When comparing classifiers, different metrics have been traditionally used in the literature. Among them, *precision* (5.7) and *recall* (5.8).

$$precision = \frac{tp}{tp + fp} \tag{5.7}$$

$$recall = \frac{tp}{tp + fn} \tag{5.8}$$

In the previous expressions, *tp* is the true-positive rate, *fp* is the false-positive rate and *fn* is the false-negative rate.

To compare different classifiers, the F_1 score (5.9), which combines both metrics, is conventionally used (Sokolova, Japkowicz, and Szpakowicz 2006). It has been obtained for different threshold values as a means to establish the best.

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{5.9}$$

Fig. 5.9 collects the F_1 scores of different training regularized logistic regression algorithms for different threshold values in the [110, 200] range, where the "aging knee" takes place. The best threshold for dividing the "Short RUL cell" and "Long RUL cell" classes has been found in 150 cycles.

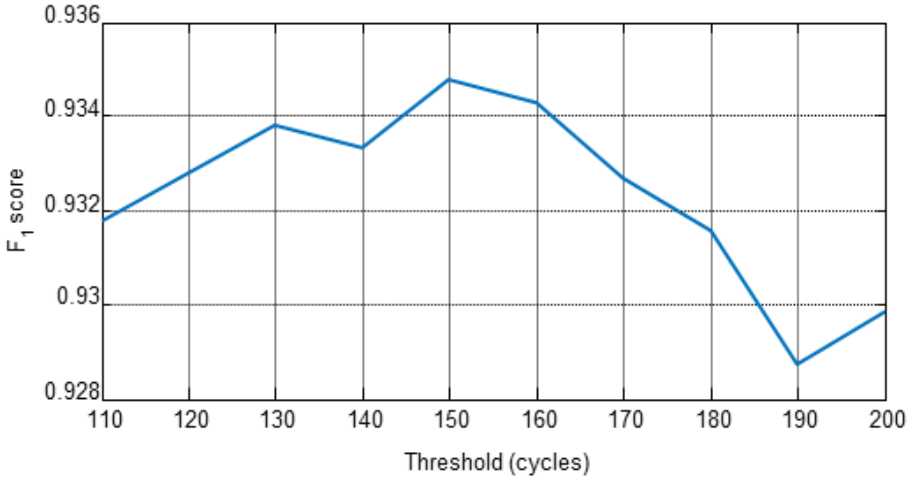

Fig. 5.9. F_1 score for classification with different threshold values.

5.3.3 Determining the optimal number of test cycles

The number of test cycles δ has been introduced before as a means of obtaining more measurements for the features of the cell under test. This way, each input example in the dataset is comprised of δ cycles and their respective measurements. A specific number of cycles δ provides δ measurements for each feature except ΔC, and δ-1 measurements for ΔC.

A larger δ allows gathering more information about the cell. This comes at expense of reducing the RUL for test purposes (due to the δ test cycles). Thus, there is a tradeoff between δ and the success rate achieved.

To determine the optimal value for δ, the success rate of the RLR algorithm has been used. Different sets of variables have been tested, with a different number of measurements per feature, and applying certain transformations. The best combination of features found is gathered in Table 5.1, which makes a total of 4δ measurements.

Feature	Samples used as input	Number of measurements
C	C_δ	1
ΔC	$[\Delta C_2, \dots \Delta C_\delta]$	δ-1
Pk	$[Pk_1, \dots Pk_\delta]$	δ
IR	$[IR_1, \dots IR_\delta]$	δ
TIEDVD	$[TIEDVD_1, \dots TIEDVD_\delta]$	δ

Table 5.1. Features used in the algorithms.

The Regularized Logistic Regression algorithm has been evaluated with different values of δ, using 60% of the dataset for training, 20% for validation, and 20% for test purposes. Fig. 5.10 shows the success rate (accuracy), defined as the number of correct classifications over the number of total classifications, as a function of δ.

It can be seen that values of δ above 10 do not improve the accuracy of the method, while additional measurements further reduce the RUL of the cell. Thus, the case of $\delta = 10$ will be the only considered in the rest of the analysis.

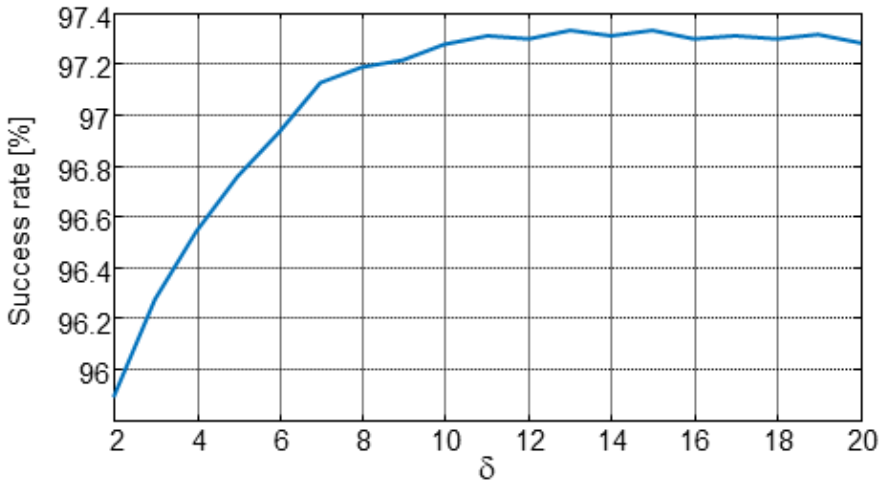

Fig. 5.10. Accuracy for each δvalue.

5.3.4 Results

The confusion matrix for the case of $\delta = 10$ is shown in Fig. 5.11. The overall accuracy in the classification obtained with this method is around 97.27%, with 93.7% accuracy for detecting "Short RUL cells", and 98.1% accuracy for "Long RUL cells".

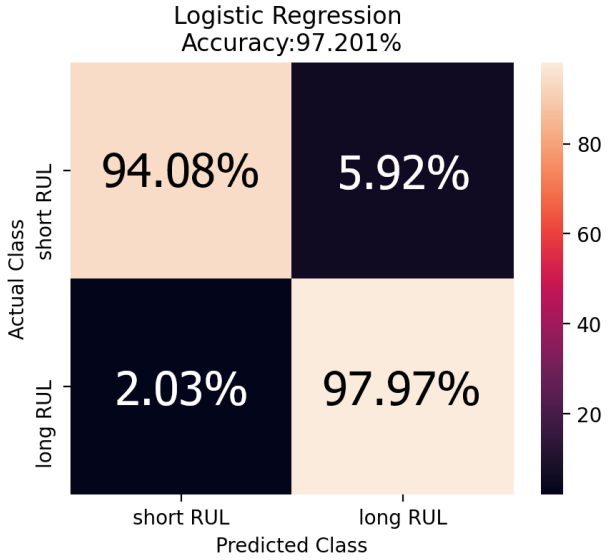

Fig. 5.11. Confusion matrix for the case of δ=10.

It should be emphasized that this high accuracy is achieved for cells in an unknown moment of their life, i.e., there is no information on the past aging of the cells, and the number of cycles they have suffered before the measurements is unknown.

5.4 Remaining Useful Life estimation

After observing the classification capabilities of the features, a different approach is taken, where these features are processed to perform actual RUL estimation, i.e., predicting the specific number of remaining cycles for a given cell. Unless made explicit, the regression algorithms will use the same 4δ inputs as the classification algorithm described in the previous section. This means that Capacity will be available as a feature, thus performing what has been previously defined as *Direct Estimation*.

In the following sub-sections, the regression algorithms that have been implemented and tested are described.

5.4.1 Multivariable linear regression

The first RUL regression method considered is Regularized Multivariable Linear Regression (MLR), as it is probably the simplest. The output of this algorithm for a given input is given by the expression (5.10).

$$\hat{y}(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T \mathbf{x} \tag{5.10}$$

Where $\hat{y}$ is the predicted RUL, $\mathbf{x}$ is a $(4\delta + 1)$ dimensional (due to the bias term) input array and $\mathbf{w}$ is a $(4\delta + 1)$ dimensional weight array.

To obtain the best configuration for the weight vector $\mathbf{w}$, a training algorithm such as Gradient Descent, fits $\mathbf{w}$ by minimizing the cost function (5.11).

$$J(\mathbf{w}) = \frac{1}{m}\sum_{i=1}^{m}(\mathbf{w}^T\mathbf{x}_i - y_i)^2 + \lambda\sum_{j=1}^{n} {w_j}^2 \tag{5.11}$$

Where the first summand is the traditional least-squares term, and the second summand is the Ridge regularization term, introduced to avoid overfitting as in the case of the classification problem. Again, the parameter m is the number of training examples.

5.4.2 Multi-Layer Perceptron and Multiple-Expert System

As an alternative to Multivariable Linear Regression, some more sophisticated algorithms have been considered. One of these algorithms is the Multilayer Perceptron (MLP), which has been introduced previously in section 4.3.1.

The MLP has been trained with the whole dataset (maintaining 60% for training, 20% for validation, and 20% for test) for RUL regression. The inputs to this algorithm are the same as for the case of Multivariable Linear Regression. A single hidden layer has been used, and the optimal size was found at 32 neurons by conducting a parametric search.

In section 5.3, the classification approach has proven that the "Short RUL" and the "Long RUL" classes have different behavior. As an improvement to a regular MLP, "Expert" MLPs have been proposed. These experts have been trained and tested with two different subsets: one for short RUL cells (those below 150 RUL cycles) and the other for long RUL cells (those above 150 RUL cycles).

The expert algorithms have been combined in a Multiple-Expert System (MES), shown in Fig. 5.12. Here, the "RUL Classifier" block (the RLR algorithm described in section 5.3) chooses which expert to use after classifying the input in one of the classes.

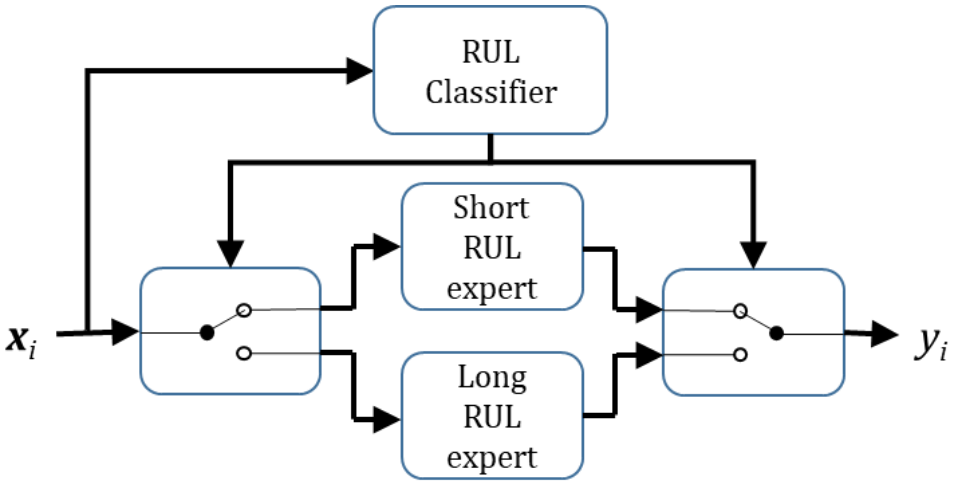

Fig. 5.12. MES algorithm scheme.

5.4.3 Convolutional Neural Networks

The algorithms introduced above use features that are pre-processed from the waveforms of each cycle in the database. However, in this other approach, the idea is to process the raw waveforms instead. The goal is to use as much information available from the dataset, thus using minimal preprocessing. An example of these raw waveforms is shown in Fig. 5.13.

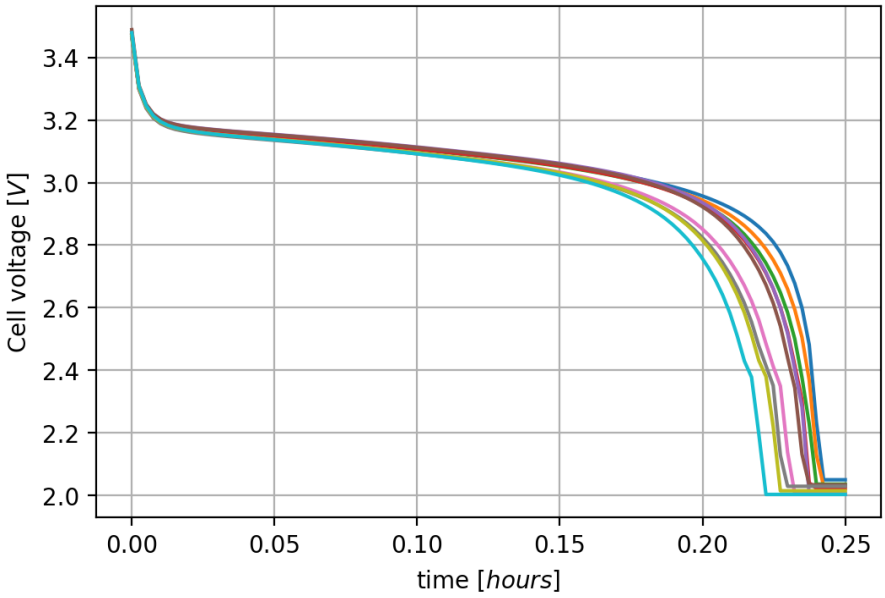

Fig. 5.13. Discharge waveforms of a single cell along with its life.

Algorithms capable of handling variable-length sequences, such as Long-Short Term Memory (LSTM) networks, have been considered for this task. Other considered algorithms include Convolutional Neural Networks (CNN), which are capable of handling multi-dimensional signals. The initial study with LSTM networks did not achieve meaningful results, so the CNN development will be described.

CNNs are a subtype of machine-learning algorithms, generally used to process time-domain sequences or spatially related sequences such as images (Gu et al. 2018). They are based on using 1D or 2D convolutional filters on the input sequence, reducing its dimensionality. After the convolutional layer, the dimensionality of the signal/image is reduced into a subset of more relevant features. The weights of the filter are trained and adjusted with the same methods as conventional neural networks.

Convolutional layers are often combined with pooling layers, further reducing the dimensionality of the input data. The most common pooling techniques are Maximum Pooling or Average Pooling (Cireşan, Meier, and Schmidhuber 2012; Mittal n.d.), and are applied to the output of the convolutional layer. Additionally, the outputs of the convolutional and pooling layers are often post-processed by an additional network, such as a linear layer or an MLP, to obtain the final output, which would be the RUL estimate in this case. In the context of CNN, these prost-processing layers are also named "fully connected layers". An example of this process is shown in Fig. 5.14.

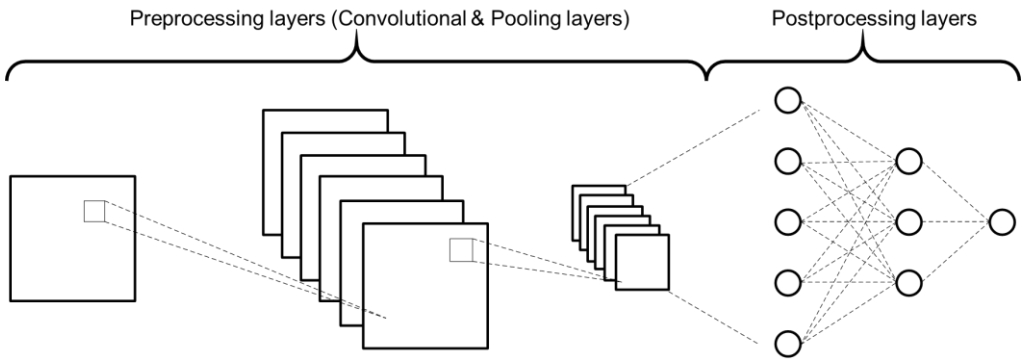

Fig. 5.14. Example of a CNN with a post-processing network.

After several attempts with the time-domain waveforms, the performance of the algorithms did not match the expectations, with high error results when trying to predict the RUL cycles. For this reason, the approach was shifted to IC raw waveforms. As introduced before, IC is defined as the derivative curve of the charge against the voltage of the cell for each charge (or discharge in this case) cycle (5.2). Thus, it can be obtained with simple pre-processing from the voltage waveforms contained in the dataset.

An example of IC waveforms is shown in Fig. 5.15, which corresponds to the discharge waveforms shown in Fig. 5.13. As introduced in section 5.2.3, IC curves show different capacitance peaks at different voltage values, depending on each chemistry and aging process. The peaks and their variations along cell life have been proven to contain information about the aging process, thus they can be used as a HI (Ansean et al. 2019). Both raw and normalized versions of these waveforms have been considered for the algorithms.

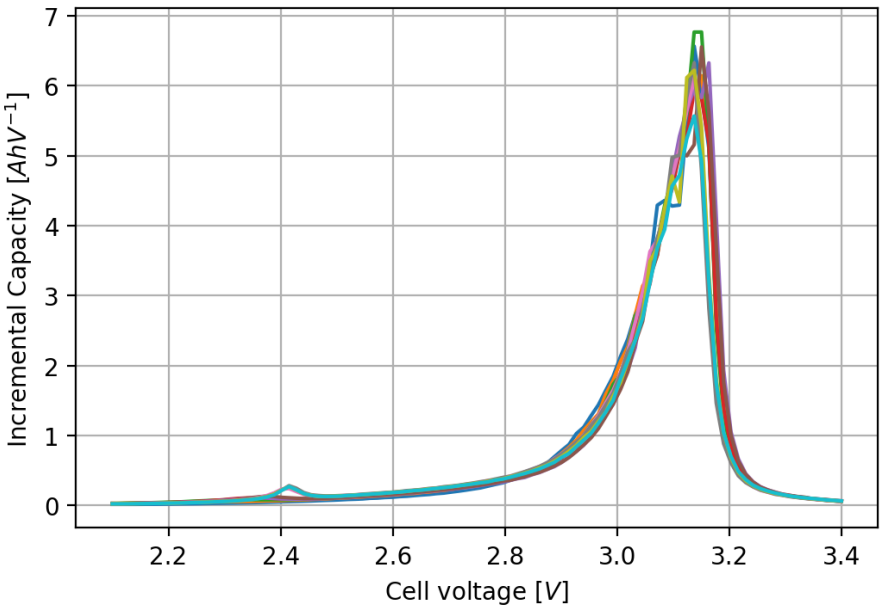

Fig. 5.15. Examples of Incremental Capacity Waveforms.

Studies in the field of medical signal processing (Lawhern et al. 2016) have grouped time-domain signals from the same nature into matrixes, forming virtual images, where one dimension is time and the other may be distance or of other nature. This allows applying 2-dimensional convolutional layers to the waveforms, which have proven to be effective in fields such as image recognition or object identification and classification.

Using these approaches as inspiration, multiple IC waveforms are grouped into virtual images, which will be processed by 2-dimensional CNN and pooling layers. In this case, the virtual images have voltage in one of the axis and time (over several discharge cycles) in the other.

The number of waveforms considered has been maintained as $\delta = 10$ from the previous sections, since provides a good tradeoff between accuracy in the prediction and using a low number of cycles. Therefore, ten Incremental Capacity waveforms have been used to create each single virtual image.

The cycles in the dataset have a different number of samples (and so do the IC curves). As a starting point, these curves have been resampled to 100 samples each in the range [2.1, 3.4V]. A virtual image is formed by arranging ten 100-sample IC waveforms into a 10x100 matrix (Fig. 5.16).

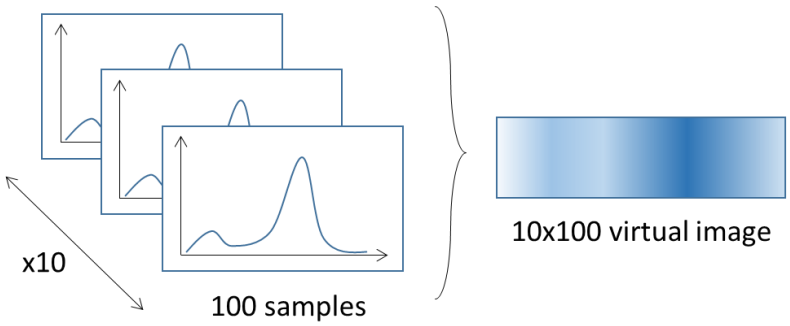

Fig. 5.16. Incremental Capacity waveforms grouped in a virtual image.

Tuning the CNN structure may be a difficult task since there is a great number of hyperparameters to be adjusted (Schmidhuber 2015). Some of the more common parameters are the number of input and output channels of each convolutional layer, the shape of the convolutional kernel (or filter) on each convolutional layer, the stride parameter, or the shape of the pooling kernel after each convolutional layer. Parametric searches have been conducted to obtain the best values of these hyperparameters, and the optimal configuration found is shown in Table 5.2.

Layer	Type	Parameters
I	Convolutional	Channels in:1 out:2 Kernel 5x1 Stride 2x1
II	Average Pooling	Kernel 2x2
III	Convolutional	Channels in:2 out:4 Kernel 1x5 Stride 1x2
IV	Average Pooling	Kernel 1x2

Table 5.2. CNN configuration.

After the convolutional and pooling layers (pre-processing network from now on), the algorithm reduces the input from 10x100 = 1000 features into four channels with shape 1x12. These pre-processed features are then unwrapped into a single array of 48 features. To increase the performance of the algorithm, it was decided to include additional features after the pre-processing layer, to be post-processed with the 48 features that output the convolutional layers. Specifically, C, IR, and TIEDVD gave the best results. Two cases have been considered: the case of including just one measurement of each feature and including all 10 measurements.

For the post-processing network, simpler machine learning algorithms have been tested. The best results were obtained using a two-hidden-layer MLP and the same number of neurons as inputs in the hidden layers. As an example, Fig. 5.17 shows the structure of the network for the case of one measurement for each additional feature, up to 48+3=51 inputs to the post-processing layer.

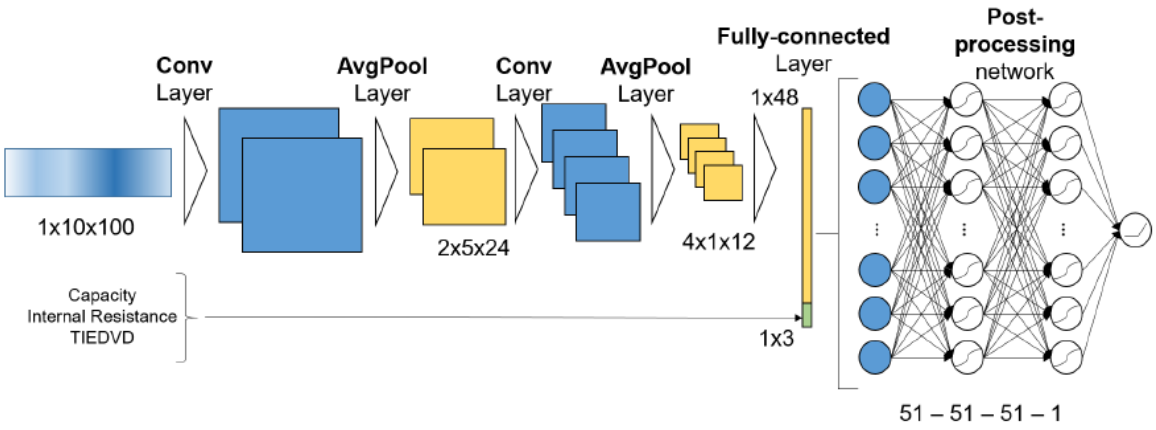

Fig. 5.17. CNN structure.

5.4.4 Preliminary results

The algorithms introduced in the previous sections have been trained with different approaches. In the first design stage of the algorithms, the hypothesis for the training process was to consider each cycle in the database as an individual example, uncorrelated with the rest of the examples.

The training method selected for the initial tests was Early Stopping (ES) (Girosi, Jones, and Poggio 1995). It has been briefly described in Chapter 4, but it will be explained in detail down below. With this technique, the cycles of the dataset are randomly assigned to three subsets: the training, validation, and test subsets:

- Training subset: on each iteration of the training process, the output for the training subset is obtained, and the parameters of the network are modified to reduce the error between the obtained output and the real output.
- Validation subset: on each iteration of the training process, the output for the validation subset is obtained, and its error is computed. The

network related to the best validation error is selected as the best network at the end of the training process.

- Test subset: This subset is not used during the training process. It is used afterward to benchmark different algorithms by checking the error obtained with this subset.

This subset division is the go-to option in the literature since it is the most straightforward option that avoids overfitting. In this case, 60% of the dataset has been used for training, while 20% has been used for validation and the remaining 20% for test purposes. Further, this subset division will be revisited.

To compare the algorithms, different error metrics are going to be used. Specifically, the Root Mean Square Error (RMSE), as defined in (5.12), and the Mean Relative Error (MRE), as defined in (5.13).

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \tag{5.12}$$

$$MRE(\%) = \frac{1}{m}\sum_{i=1}^{m}\frac{|y_i - \hat{y}_i|}{y_i}\ x100 \tag{5.13}$$

Where m is the number of examples in the test subset, y_i is the actual output for the *i-th* example in the test subset, and $\hat{y}_i$ is the output of the algorithm for the *i-th* example in the test subset.

The preliminary error metrics, obtained for the simpler algorithms and trained with the aforementioned approach, are contained in Table 5.3. As it may be observed, algorithms that are more complex and capable of capturing non-linearities, such as Multi-Layer Perceptron and the Multiple Expert System (MES), perform much better.

Algorithm	RMSE (Cycles)	MRE (%)
MLR	90	53.81
MLP	52	23.03
Multiple Expert System	49	15.2

Table 5.3. Preliminary error metrics.

Different subset division

In the preliminary training stage, each of the cycles of the cells in the dataset has been considered as a single, uncorrelated example. The subset division for the training process has been carried out by randomly selecting these cycles for the training, validation and subset groups. This implies that different cycles derived from the same cell could end up in the training, validation and/or test subsets. By using this division, the algorithm learns some information from every cell when training. The result is a high performance in the error metrics, which is ultimately caused by something similar to overfitting, not in the conventional sense, but at a "cell level". Furthermore, this training process would not be realistic from an application point of view, since when estimating RUL for a new cell, the algorithm has never seen other cycles from that cell in training.

Different options for the subset division are depicted in Fig. 5.18. The approach followed in the preliminary training is the one in Fig. 5.18a. In it, the subsets are created by randomly taking 60% of all the cycles for training, 20% for validation, and 20% for test. To try to obtain error metrics closer to a realistic implementation, a different approach is going to be taken for the subset division, as shown in Fig. 5.18b. This approach randomly divides the cells into training cells, validation cells, and test cells (with the same 60%-20%-20% ratio). The training, validation and test subsets contain the cycles of their corresponding cells. This way, there are not multiple cycles from the same cell in different subsets, and the results are closer to the expected error metrics in an implementation stage.

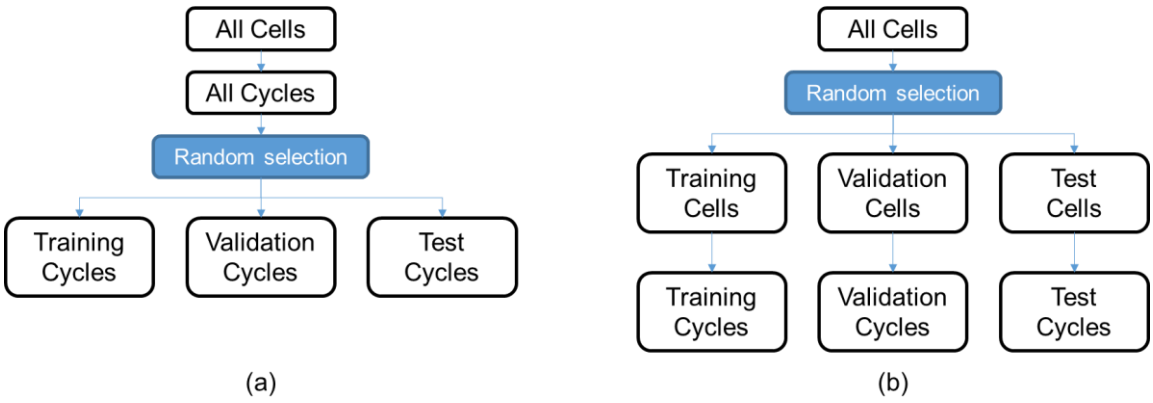

Fig. 5.18. Different subset divisions. Prone to overfitting at the cell level (a). Immune to overfitting at the cell level (b).

NOTE: The results for the life stage estimator of section 5.3 are updated and have been obtained with this subset division.

Regularization techniques

Additionally, to avoid overfitting to a higher degree, different regularization methods have been applied. These regularization methods include Ridge regularization (Hoerl and Kennard 1970), Lasso regularization (Tibshirani 1996), and other techniques such as dropout layers (Goodfellow et al. 2020; Ganin, Lempitsky, and Ru 2015; Srivastava et al. 2014).

Ridge (5.14) and Lasso (5.15) regularization include the weights of the network into the loss function, penalizing high weights and achieving smoother output functions.

$$J(\mathbf{w}) = \frac{1}{m}\sum_{i=1}^{m}(f(\mathbf{x}_i, \mathbf{w}) - y_i)^2 + \lambda \sum_{j=1}^{n} {w_j}^2 \tag{5.14}$$

$$J(\mathbf{w}) = \frac{1}{m}\sum_{i=1}^{m}(f(\mathbf{x}_\mathbf{i}, \mathbf{w}) - y_i)^2 + \lambda \sum_{j=1}^{n} |w_j| \tag{5.15}$$

In the previous equations, J is the cost function value, m is the number of examples in the database, $\mathbf{x}_i$ is the i-th example in the dataset, y_i is the output for said example, n is the number of weights in the network, w_j is the j-th weight and λ is a parameter set by the user.

Dropout layers, as explained in the previous chapter, have also been used as a different method to avoid overfitting. Different drop-out rates (p) have been considered for this task.

5.4.5 Results

The test results for the algorithms are collected in Table 5.4. It may be observed how the error rates are, in general, worse than those obtained in the previous section are. This reaffirms the fact that overfitting at the cell level was taking place, and that the new subset division gives results closer to the expected in a real implementation of the algorithms.

Similarly, the general trend is that algorithms that are more complex tend to obtain better error metrics, since the worst error rates, both for RMSE and MRE, are obtained with the simpler algorithm (MLR). However, the results obtained with this simple algorithm are slightly better than the results that were obtained by the authors in the original study, where the dataset was introduced. The algorithms used there were similar, although the approach to RUL estimation was not strictly the same.

Algorithm	**RMSE (Cycles)**	**MRE (%)**
MLR	109.9	60.57
MLP	86.3	36.29
Multiple Expert System	86.2	19.06
CNN IC waveforms only	100.81	21.15
CNN Normalized IC waveforms	86.69	21.14
CNN IC + Additional features (1 sample)	84.21	17.31
CNN IC + Additional features (10 samples)	82.67	18.17
CNN Ridge ($\lambda = 0.003$)	**79.09**	16.06
CNN Lasso ($\lambda = 0.003$)	83.49	16.8
CNN Dropout layers ($p = 5\%$)	91.26	16.19

Table 5.4. Error metrics for MLR, MLP, MES, and variants of CNN.

In the case of MLP-based algorithms (regular MLP and Multiple Expert System), the results are interesting. The RMSE with both are similar, but the MRE metrics are much higher for the regular MLP. This means that the MLP estimates worse than the MES for the case of cycles with a low actual RUL since the MRE error is higher when trying to estimate lower RUL values (error of only a few cycles for cells with low RUL leads to a large MRE). To clarify this, Fig. 5.19 provides information on the RMSE and MRE obtained for each value of actual RUL. Note that MRE is 10-25% for practically the whole range of actual RUL, except for the first 100 cycles. Thus, it can be concluded that MRE does not provide a good representation of the error when predicting low RUL values.

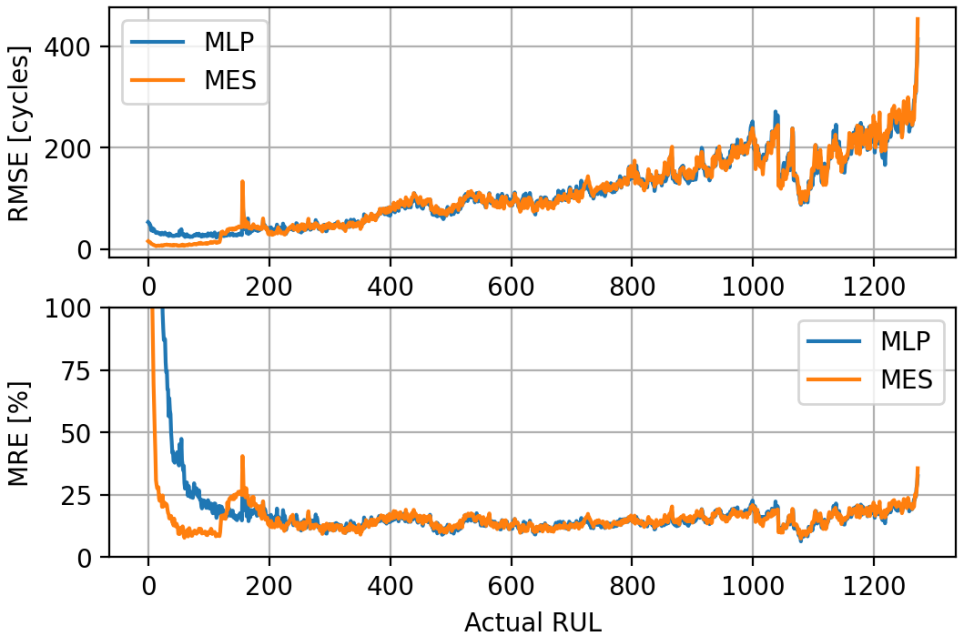

Fig. 5.19. RMSE and MRE rates for the MLP and MES algorithms for different actual RUL values.

Regarding the CNN approach, some variations to the basic structure, regularization method, and the number of additional features have been tried. The details for each of these variations are collected in Table 5.4 as well. It is worth noting how the version of the algorithm that normalizes IC waveforms, performs much better than the non-normalized version.

By normalizing, the original shape of the waveforms is not preserved, and therefore some information is lost (information of capacity is contained as the integral of the IC curve, and thus is lost when normalizing these curves). However, it seems that the training process is easier for the algorithm due to the normalized data, and makes it capable of performing better even without the capacity information. When adding the additional features (as capacity or internal resistance), the algorithm recovers the information and performs even better.

Among the different regularization techniques implemented, the best results were obtained when using Ridge regularization (with $\lambda = 0{,}003$). The other different techniques gave similar results than the non-regularized case in the best scenario (the reported error metrics), and far worse results when using extreme values for their tuning parameters.

As a means of comparison, Table 5.5 shows the results of some recent studies (Severson et al. 2019; Ma et al. 2020) that aim to perform RUL prediction with the same dataset. Their approach is however a little bit different from the approach taken here since they track information from all the past life of the battery, whereas here, only ten test cycles are taken as inputs. It can be appreciated how the CNN algorithm with Ridge regularization is on par with the best performing algorithm in Table 5.5, but only requiring ten test cycles rather than information on the whole past life of the battery.

Algorithm	**RMSE (Cycles)**
Elastic Net	118
Broad Learning	115.97
Extreme Learning Machine	105.99
Broad Learning-Extreme Learning Machine	75.78

Table 5.5. Error metrics of algorithms found in the literature with the same dataset.

As a final example of RUL estimation, Fig. 5.20 shows the output of the best-performing algorithm for all of the test cycles of one specific cell in the dataset (#99). It can be appreciated how the RUL prediction curve follows accurately the RUL line, which decreases linearly as the cell ages.

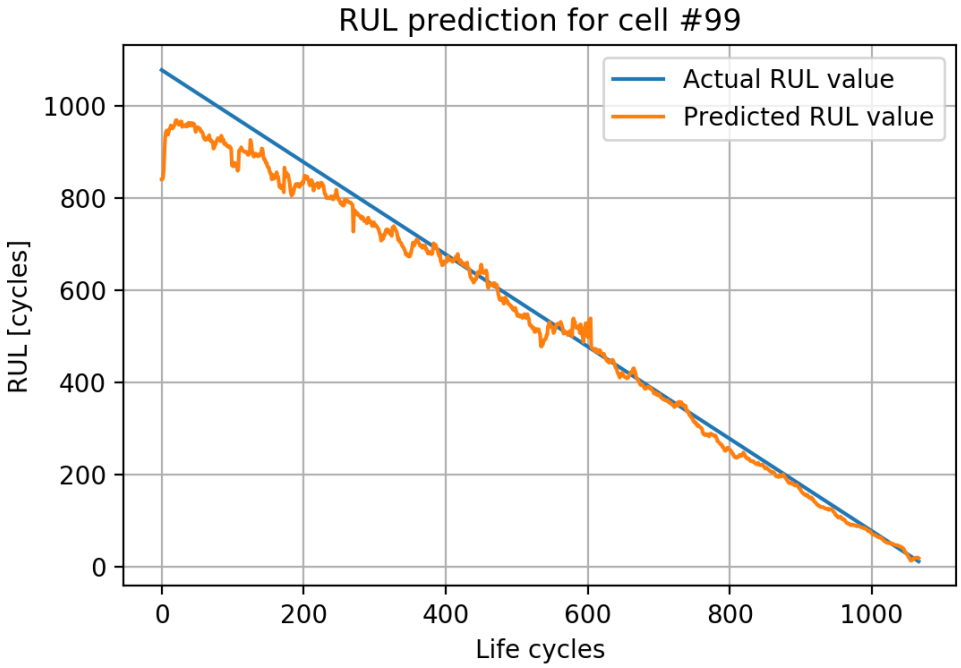

Fig. 5.20. Multiple Expert outputs for all the test cycles belonging to cell #100.

5.5 Capacity estimation

Traditional approaches to battery-health estimation focus on capacity instead of RUL as the main feature to be predicted. However, both variables are closely related. Capacity has been used as one of the features used for RUL estimation in the previous section, whereas, in this section, the capacity estimation problem is addressed.

Here, the problem of capacity estimation is addressed with two different approaches: straightforward capacity estimation and capacity curve regression.

5.5.1 Straightforward capacity estimation

Batteries in a conventional ESS application such as standalone photovoltaic installations cannot be completely discharged to measure their capacity and estimate their health, since this would mean an energy disruption event. This also happens in other applications such as Uninterruptible Power Supplies (UPS). For this reason, there is interest in developing algorithms capable of directly estimating capacity from features available without discharging the battery pack. Additionally, this estimate is very relevant for a BMS to compute other battery figures such as SoC, which, to be accurate, need feedback on how the capacity is changing as the cell ages.

Not all the features introduced in the previous sections are available under this hypothesis. Besides capacity and capacity fade, it must be noted that IC curves and the C_{pk} and V_{pk} features are obtained when completely cycling the cell, or at very low SoC. For this reason, only IR and TIEDVD are going to be used for capacity estimation purposes.

Machine-learning techniques have been trained to output directly the capacity of the cell. In this case, the CNN approach is not valid, since the IC waveforms have been discarded as a feature. Additionally, the accuracy of the Life Stage Classifier is reduced to 93.6% when using only IR and TIEDVD as features, which worsens the error rate of the Multiple Expert System. Hence, MLP will be used as the algorithm for the capacity estimation task. In this case, each input pattern will

comprise ten measurements for TIEDVD and ten measurements for IR, preserving the optimal value for δ found in section 5.3. The optimal number of neurons in the hidden layer has been found at 25 by conducting a parametric search. The output of the algorithm is going to be capacity at that very same cycle.

As an example, Fig. 5.21 shows the actual capacity values and the output of the algorithm for one cell in the dataset.

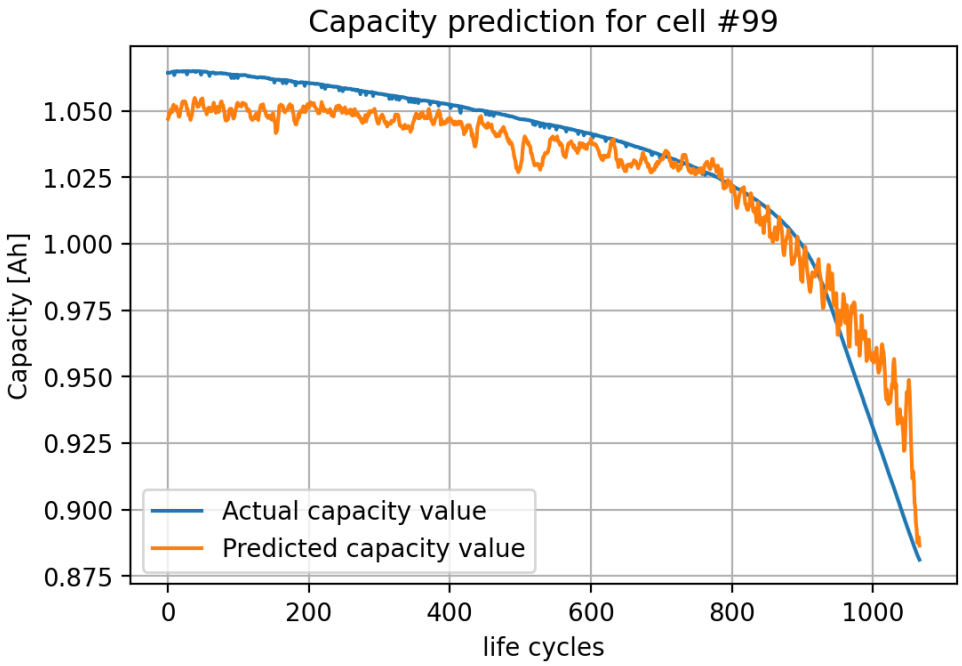

Fig. 5.21. Capacity estimation for cell #99.

The results of the algorithm applied to the whole dataset are contained in Table 5.6. The accuracy in the prediction of capacity is high according to the error metrics obtained.

Algorithm	**RMSE [Ah]**	**MRE (%)**
Capacity Estimation	0.02	1.49

Table 5.6. Error results for Capacity Prediction.

5.5.2 Capacity curve regression

This approach aims to estimate the evolution of the capacity curve (Fig. 5.1) from a set of past measurements. In this case, the hypothesis is that the cell has

been monitored from the beginning of its life, and therefore, the increasing window of the past evolution of capacity is known.

A simple capacity degradation model, which has been explored previously in the literature (He et al. 2011), is considered for capacity modeling. It assumes that the function of capacity aging in the cells, $C(t)$, can be modeled as a sum of two exponential functions, as given by (5.16).

$$C(t) = a\, e^{bt} + c\, e^{dt} \tag{5.16}$$

This aging model is parametrized by four parameters, $\{a,\ b,\ c,\ d\}$. A regression algorithm, such as Levenberg-Marquardt (LM) (Levenberg 1944), can be used to determine the evolution of the capacity curve by fitting these parameters.

As cycles are carried out, the window of past capacity measurements increases and the regression is better. However, to provide good accuracy in the prediction, the algorithms need a large window size. This is in part due to the abrupt decay of the capacity in the last 150 cycles of life. With a window size of a few hundred cycles, the algorithm does not see this decay, and cannot fit this part of the curve accurately (Fig. 5.22).

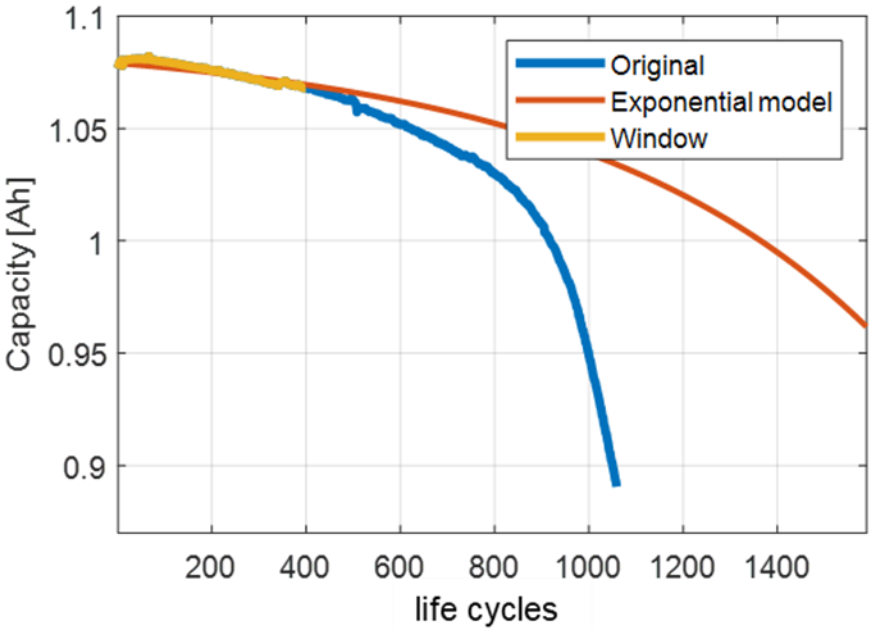

Fig. 5.22. Capacity curve regression with the LM algorithm.

The RUL estimation algorithms developed earlier can be used to enhance this regression approach. These data-driven techniques can provide a future estimate for the End of Life (EoL) point of the battery. Since the hypothesis is that the cell has been monitored from the beginning of its life, the EoL point is given by (5.17).

$$\widehat{EoL}_k = \widehat{RUL}_k + n_k \tag{5.17}$$

Where $\widehat{EoL}_k$ is the End of Life estimate at time k, $\widehat{RUL}_k$ is the estimate for RUL given by the data-driven algorithm at time k, and n_k is the number of cycles the cell has already lived at time k.

The new EoL point can be added to the past window of capacity measurements, so the values for the $\{a,\ b,\ c,\ d\}$ parameters can be fitted with higher accuracy. Now, the EoL capacity value acts as an "anchor point", ensuring that the fitting algorithm converges to a reasonable solution.

As seen in Fig. 5.23, the results obtained using the EoL estimation as an anchor point are much accurate than those obtained without it (Fig. 5.22). Now, the estimate for the EoL point allows the algorithm to detect the decay in the capacity that takes place around cycle #800. Previously, the algorithm did not count with the EoL estimate, so it made a much more optimistic prediction of cell life.

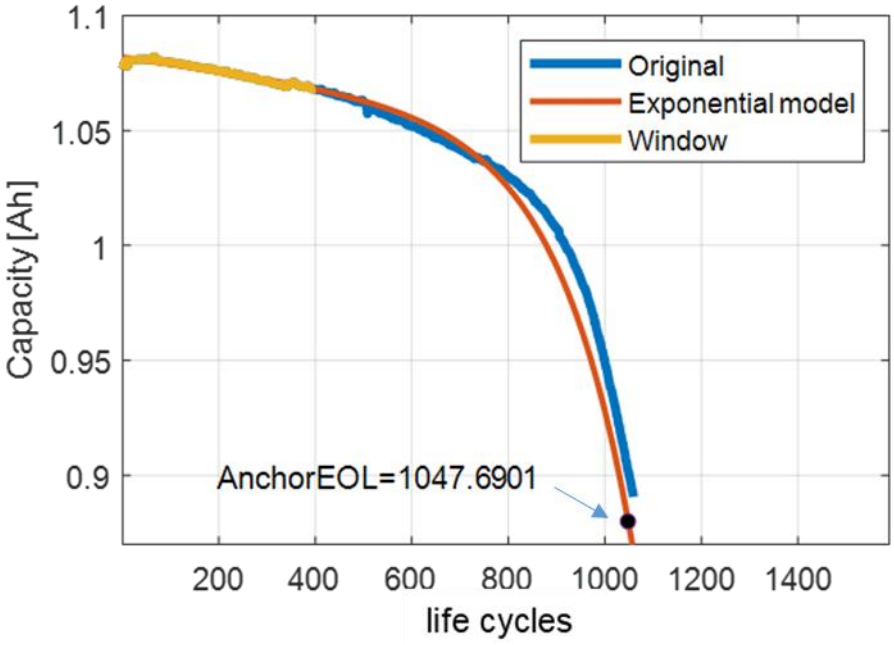

Fig. 5.23. Capacity curve regression with LM and the EoL estimate.

Fig. 5.24 shows the average Maximum Relative Error (MaxRE), with and without the EoL estimation, and an increasing window size, as defined in (5.18).

$$MaxRE_k(\%) = \frac{1}{n}\sum_{i=1}^{n} \max\left(\frac{\left|\hat{C}_{i,k} - C_i\right|}{C_i}\right) \quad x100 \tag{5.18}$$

Where $\hat{C}_{i,k}$ is the prediction of capacity for cell i with a window size of k, C_i is the actual capacity of cell i, and n is the number of cells in the test subset. As the cell lives, the window size increases, and better results are expected. It is interesting to note that past 100 cycles of life, the estimations are consistently accurate when using the EoL estimate. The regression algorithm without anchor point takes double (around 200 cycles) to start making reasonable predictions for the evolution of the capacity, and even then, using the EoL estimate is still better.

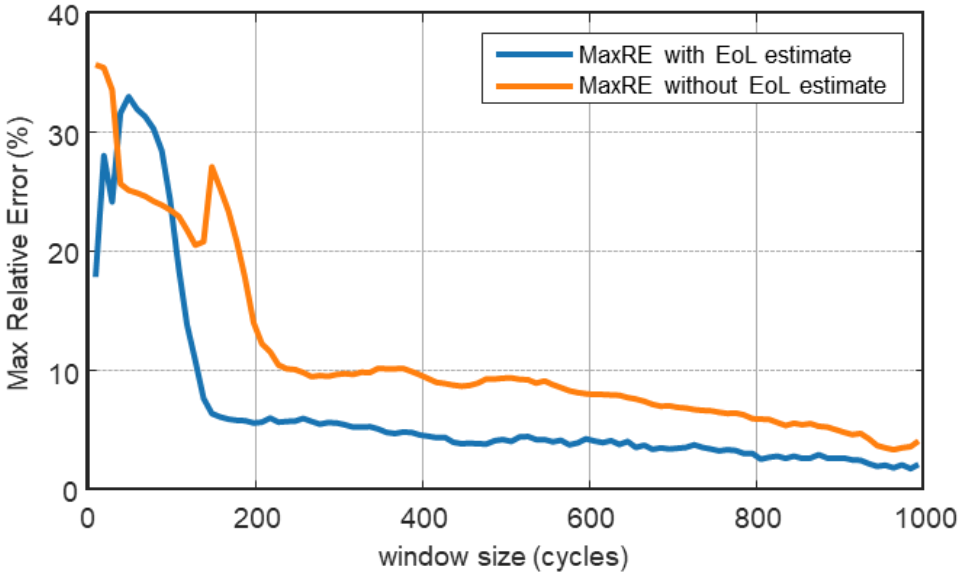

Fig. 5.24. Maximum Relative Error (MaxRE) for different window sizes.

The average RMSE and MaxRE for both methods are collected in Table 5.7. It can be observed how the regression with the EoL estimate performs consistently better. This is especially significant in the case of MaxRE, where the algorithm without an EoL estimate performs around 35% worse.

Regression type	RMSE (Ah)	MaxRE (%)
With EoL estimate	0.0445	10.05
Without EoL estimate	0.0545	13.54

Table 5.7. Capacity Estimation Error Metrics.

5.6 Conclusions

Estimation of health-related parameters of batteries is a popular topic in battery-related research. Such parameters include RUL, capacity, or SoH among others. There is interest in developing tools in this field applied to lithium-ion batteries. Due to the complexity of the problem, data-driven approaches are appropriate for the task, although until very recently, publicly available datasets were small and simple.

In this chapter, the problem of Remaining Useful Life estimation has been addressed. The main hypothesis assumes that past information on cell use may not available. For this reason, most approaches in the literature are not suitable here, since they tend to track capacity from the beginning of cell life.

By conducting simple charge and discharge cycles, key features are obtained from the voltage waveforms (such as IC waveforms), which have proven to be good health indicators in other applications. These controlled cycles can be done in a controlled environment, such as a laboratory, or an in-situ application, if the storage system is a HESS with more than one storage element. Using said features, the algorithm can determine different health-related features of the cells.

Both traditional and recent machine learning techniques have been proposed for the task. Simple classification algorithms such as RLR are capable of classifying used cells into "Short Remaining Useful Life" and "High Remaining Useful Life" categories with an overall accuracy higher than 97%.

Additionally, different methods and features are explored, aiming to estimate the exact number of remaining cycles for cells in an unknown moment of their lives. Such methods include conventional linear regression techniques and MLP, or newer algorithms such as CNN. The best results are obtained with Ridge regularization applied on the CNN algorithm, obtaining 79.09 cycles of RMSE. These results are on-par with other recent algorithms applied to this same database, but with the benefit of not needing all the past information of the cell, but only a few recent cycles.

One of the most singular features explored has been IC waveforms, which have been described in the literature as a convenient HI. To extract this feature, the cells need to be cycled with a low-current profile, which is not the case in this database. One additional goal was to determine if these high-current IC waveforms were sufficiently good to be used as a HI, even though they are not genuine IC curves. The results of the analysis certify that these high-current IC waveforms still contain useful information on battery degradation, and can be used as a HI as well. As a future line, it would be interesting to obtain a database with low-current IC waveforms and use the same algorithms.

A simple machine-learning tool has been proposed for the problem of capacity estimation in applications where the cell cannot be controllably discharged. Using a small subset of the features introduced in this chapter, the algorithm provides estimates for capacity with a relative error of around 1.5%. This is useful for BMSs, which need to provide SoC and SoH estimates, since capacity is an integral part of the calculation of such variables, and it is not always available in an application.

As the last contribution, a discussion on the different levels of overfitting when working with this kind of dataset has been included. Here, overfitting takes place due to the data being correlated on a cycle-level basis and a cell-level basis. Realizing this is level key to obtain reliable results from the data-driven algorithms, especially approaching implementation stages, where completely new cells need to be used in the test subsets.

Chapter 6.- Conclusions and research lines

This chapter collects the main conclusions of the previous chapters and establishes plausible future research lines.

6.1 Conclusions

Throughout this thesis, the Energy Storage Systems of small standalone photovoltaic installations intended for metering and communication purposes have been studied. These storage systems are usually based on conventional battery chemistries such as lead-acid, NiCd, or NiMH. Their architecture and sizing methods have been explained. These ESSs suffer from different key problems, which include temperature-related issues and incomplete charging cycles. Both effects worsen the performance, both in the short-term due to Partial State-of-Charge and effective capacity reduction, and in the long-term due to accelerated aging.

The industry proposes partial solutions for these problems, such as multiple-stage charging processes. However, these solutions do not mitigate completely the partial charging problem and accelerate aging. On the other hand, this thesis explores different solutions to deal with them, both modifying the architecture of the system and by using intelligent algorithms.

As a first solution, Hybrid Energy Storage Systems are proposed. This solution comprises two storage subsystems: a Cyclic Storage System (CSS) and a Support Storage System (SSS). The CSS provides energy during the daily cycle, whereas the SSS acts as a bulk energy reservoir. Storage systems with different roles require chemistries with different features, therefore, specific storage technologies have been proposed for each storage system. Specifically, conventional battery technologies such as VRLA are proposed for the SSS role, due to their lower price and stationary features. On the other side, newer battery chemistries, such as LFP, are proposed for the CSS.

The sizing of these specific storage systems has been studied with simulation data from real installations, and a novel sizing procedure has been defined. It has been found that the best energy ratios are between 15-30% of the total storage system for the cyclic subsystem, and conversely, around 70-85% of the total storage system for the support subsystem. The recommended autonomy of these systems in the studied locations has been estimated at around 150 to

200 hours, and the power-peak-to-consumption ratio of the solar panels has been estimated at around 10.

The mid-term performance of these hybrid systems has been analyzed. The role division and the optimized sizing provide a theoretical improvement of the partial SoC of up to 25%. The resilience of these systems has also been studied, obtaining an increment up to a 45% of the absorbed energy after a blackout period. The long-term benefits of these hybrid systems are yet to be proven, although an extension on the battery life is expected, since each storage subsystem is better suited for its specific role, and the lead-acid batteries in the installation are not under deep daily cycles.

An ad-hoc DC/DC converter has been designed and implemented to test this concept in a real installation, and its performance has been compared to another conventional, non-hybrid system. After some months in operation, the lead-acid battery of the HESS had its SoC increased by a 16%.

Industrial solar chargers in these applications tend to overcharge the batteries to mitigate the problem of incomplete charging processes. This practice, however, comes at expense of accelerating the aging of the batteries. Machine-learning solutions have been also proposed as a complement to these chargers. The studied approach aims to predict the voltage of the battery pack 48 hours ahead. This way, there is enough time margin to address a potential failure in the installation. Additionally, if the solar charger does not have the capacity of overcharging the batteries, this is a wide enough margin so that the installation operators can intervene. Two different prediction techniques have been studied for this task, NARX, and LSTM. The results obtained with both techniques are positive, predicting the trend of the voltage waveform and whether is going to be a failure situation. However, the exact waveform when de voltage is low is difficult to predict, since the datasets do not count with many examples of this situation.

The application of machine-learning techniques to the estimation of health-related parameters in lithium-ion batteries has also been studied as a complementary diagnose tool for HESS. Counting with multiple storage packs,

these systems can cycle batteries at little energy expense. This allows for test cycles, which can be then post-processed with smart algorithms. Different techniques have been proposed to this end, such as linear regression, logistic regression, MLP, and CNN. Additionally, multiple health indicators have been studied as features. The results show that the exact remaining useful life of these cells can be predicted with reasonable accuracy. There is also a high precision in classifying cells in the last stage of their life. This can be useful for operators, so the cells can be replaced accordingly.

6.2 Summary of contributions

The main scientific contributions of this thesis towards HESS in standalone photovoltaic installations are listed below.

- A review on standalone photovoltaic installations, their sizing parameters, storage technologies, and underperforming causes (Chapter 2).
- A battery runtime-simulation model that combines the different operation modes of the battery and allows for replicating Partial State of Charge (Chapter 2).
- Sizing of the storage packs of HESS in this application. A sizing method based on sizing maps has been developed, which determines the best ratios in the size of the solar panels and the size of the different battery packs for an installation with given power consumption. The sizing of the DC-DC converter in this application has been addressed as well. (Chapter 3).
- An improvement of the behavior of the installation with the HESS over an installation with a conventional ESS. Models of these installations have been developed to find the best-case scenarios and to determine how much does a hybrid system improves the performance over a conventional system. The results have been backed up with experimental validation (Chapter 3).
- Failure prediction algorithms for photovoltaic installations. Voltage prediction tools have been developed to determine future energy-disruption events. (Chapter 4).
- Development of a daily solar generation taxonomy and classification based on daily features obtainable in the installations. This classification allows for making hypotheses on the future current generated in the installations (Chapter 4).
- Proposal of different Health Indicators in LFP batteries towards health estimation techniques. Among them, the variables V_{pk} and C_{pk} are introduced as novel features (Chapter 5).

- Development of a life-stage estimation tool based on Regularized Logistic Regression, to determine whether a used battery is suitable or not of being re-used in another application (Chapter 5).
- Proposal of non-intrusive, Remaining Useful Life estimation techniques to estimate the number of cycles a battery can still withstand in a deep cycle application, such as Cyclic Storage Systems in HESS, without the need of knowing all the past usage of the battery (Chapter 5).
- Proposal of a capacity estimation technique to anticipate the decay of capacity in LFP batteries in a deep cycle application (Chapter 5).

6.3 Future research lines

Multiple research lines have been proposed along the text to continue the work developed in this thesis. This section collects them and ponders on the future steps to take.

- As it has been introduced, standalone photovoltaic installations in this application have a mostly DC consumption pattern, without high power spikes. Thus, the simulations throughout the thesis and the sizing conclusions have been obtained under the DC consumption hypothesis. It would be of interest to do the same simulations with other consumption hypotheses, from a different application, to check how it influences the sizing of the hybrid system.
- The HESS proposed in Chapter 3 has been focused on mitigating the impact of the Partial State of Charge in the application. However, it would be interesting to develop aging models of the batteries to obtain conclusions of the long-term benefits of the hybrid system. The HESS can help reducing aging on VRLA since the number and depth of the cycles that these batteries withstand are reduced. However, analyzing empirically the benefits in aging may take some years to complete. To obtain verification of the aging conclusions, it would be interesting to have the HESS working for several years.
- Additionally, when analyzing the limited hybrid power in Chapter 3, it could be seen how in some scenarios, working with a power limit in the DC-DC hybrid converter is beneficial for the VRLA battery, since it preserves a higher SoC. This is an interesting observation since it implies that a smaller (and probably cheaper) DC-DC converter may be better for some scenarios than a high-power converter. Studying these limited operation modes is an interesting future research line.
- Regarding system failure prediction, the algorithms developed in Chapter 4 have been trained and applied to a certain number of installations that belong to the CHE network. However, for each installation, a separate database has been created and a separate

network has been trained. This results in a different network for each one of the installations. Before the deployment of the algorithm in a new installation, data in that installation needs to be monitored to create the database and train the algorithm. For this reason, it could be of interest to study the universality and adaptability of these algorithms. The goal would be to develop a universal algorithm, which would be deployed on any installation and learn from the data measured there, thus adapting to that specific installation.

- Regarding the aging estimation algorithms in Chapter 5, it could be of interest to apply such techniques, developed for LFP batteries, to other chemistries. This is conditioned on the availability of datasets, but would improve the generalization of such algorithms. Specifically, applying this to VRLA batteries could help to analyze the long-term impact of the HESS.
- Furthermore, these algorithms could be then used as a tool to develop new Energy Management Strategies, taking into account the aging of both battery packs and optimizing the long-term behavior of such HESS. These EMS could process aging to include the replacement of the battery packs, optimizing the use of the HESS with techno-economic criteria.

Conclusiones y líneas futuras

Los sistemas de almacenamiento de energía en instalaciones fotovoltaicas aisladas de la red han sido estudiados a lo largo de la tesis. Estos sistemas han estado basados tradicionalmente en químicas como el plomo-ácido, el níquel-cadmio o el níquel-manganeso. También se han estudiado sus métodos de dimensionamiento, arquitecturas y distintos problemas a los que se ven sujetos. Concretamente, se ha destacado el impacto de la temperatura y de las cargas incompletas, que reducen las prestaciones de las instalaciones reduciendo la capacidad de las instalaciones y acelerando el envejecimiento de las mismas.

Desde la industria se han propuesto distintas soluciones para paliar estos problemas, como los procesos de carga con etapas de sobrecarga. Sin embargo, estas soluciones no mitigan el problema completamente, y por el contrario aumentan el envejecimiento. Esta tesis explora otras soluciones para lidiar con estos sistemas. Por un lado, modificando la arquitectura del sistema y por otro, utilizando algoritmia inteligente.

Los sistemas de almacenamiento híbridos (HESS) se han propuesto como una solución inicial. Se propone dividir el elemento de almacenamiento principal en dos sub-sistemas: el sistema de almacenamiento cíclico (CSS) y el sistema de almacenamiento de apoyo (SSS). El CSS se encarga de proveer energía durante el ciclo diario, mientras que el SSS actúa como almacenamiento de emergencia. Se han propuesto las tecnologías VRLA para la batería de soporte y LFP para la batería cíclica.

El dimensionamiento de estos sistemas híbridos se ha simulado usando perfiles de simulación de instalaciones reales. Se ha planteado un método de dimensionamiento basado en superficies, y se han determinado las mejores ratios de hibridación entre el 15-30% para el CSS. Se ha determinado que la autonomía objetivo de estos sistemas está de 150 a 200 horas, y la potencia de paneles recomendable se ha establecido en unas 10 veces la potencia DC de consumo de la instalación.

También se ha analizado el funcionamiento a medio plazo de estos sistemas híbridos. La división de roles y la estrategia simple propuesta consigue en

simulación una mejora teórica de hasta un 25% del SoC. También se ha definido y analizado la resiliencia de estos sistemas, y se ha obtenido una mejora de hasta un 45% en la energía absorbida en un sistema de almacenamiento híbrido frente a uno convencional tras un estado sin energía.

Por otro lado, las ventajas a largo plazo de estos sistemas tienen que ser comprobadas todavía. A priori, es de esperar que el sistema híbrido propuesto prolongue la vida de las baterías de plomo actuales, debido a que no están sujetas a ciclados profundos diariamente. Son las baterías de litio las que están sujetas a este ciclado diario, más acorde con sus características.

Para comprobar experimentalmente los resultados obtenidos en simulación, se ha implementado un sistema híbrido en una instalación de la Confederación Hidrográfica del Ebro. El sistema híbrido ha sido comparado con uno convencional de las mismas características. Tras varios meses en operación, el estado de carga de la batería de plomo en el sistema híbrido ha aumentado un 16%.

Para mitigar el problema de la carga parcial, los cargadores industriales actuales realizan sobrecargas periódicas a las baterías. Estas técnicas mitigan parcialmente el problema a costa de envejecer las baterías por operar a tensiones elevadas. Por este motivo, resulta interesante el uso de algoritmia inteligente para ver cuándo es necesario realizar estas sobrecargas, en vez de hacerlas periódicamente. Para ello, se han propuesto técnicas de aprendizaje automático basadas en datos, que predicen si el sistema se va a caer con una antelación de 48h, situación en la cual se puede efectuar una sobrecarga. Se han aplicado los algoritmos NARX y LSTM a esta tarea, obteniendo una detección de la tendencia de caída cuando se va a producir un fallo. Sin embargo, la detección de la forma de onda exacta no es exacta, debido a los pocos ejemplos de estas situaciones en las bases de datos.

También se ha estudiado la viabilidad de aplicar técnicas de aprendizaje basado en datos al problema de estimación de la salud de los sistemas de almacenamiento basados en litio. Esto permitirá dotar de herramientas a los sistemas de gestión de las baterías para planificar reemplazos programados

(mantenimiento preventivo) o gestionar los ciclos diarios para minimizar el envejecimiento. Algoritmos como la regresión lineal, logística, MLP y CNN han sido entrenados con este propósito, obteniendo un sistema capaz de clasificar la etapa de la vida de la batería con una alta precisión (97.2%), y algoritmos capaces de estimar el número exacto de ciclos de vida restante que le quedan a una batería, con un error de 79 ciclos en baterías de en torno a 1000 ciclos de vida. También se han propuesto estos algoritmos para hacer un seguimiento y predicción de la capacidad futura de la celda, con errores de 0.05 Ah en baterías de 1.1Ah.

A continuación, se listan las principales líneas de investigación a futuro que se desprenden de los contenidos de la tesis.

- En las instalaciones fotovoltaicas estudiadas, el consumo es principalmente en corriente continua. Esto ha servido como hipótesis para la normalización de las variables de dimensionamiento. Como línea de futuro, resultaría interesante estudiar otros patrones de consumo, y su impacto en las métricas de dimensionamiento estudiadas.
- El sistema de almacenamiento híbrido propuesto en el capítulo 3 ha sido utilizado para mitigar los efectos del estado de carga parcial (PSoC) en la operación a medio plazo del sistema. Sin embargo, también resultaría interesante estudiar las ventajas a largo plazo. Con dicho sistema, las baterías basadas en plomo sufren un ciclado menos intenso, y es de esperar que su vida útil se extienda. Como línea de futuro, sería conveniente desarrollar modelos de envejecimiento para las múltiples químicas involucradas, y determinar cuantitativamente esta extensión de la vida útil. Además, bajo estos nuevos modelos, nuevas estrategias de gestión de la energía pueden ser propuestas y validadas.
- Adicionalmente, en el capítulo 3 se ha podido ver como en determinados escenarios, tener un límite de potencia en el convertidor de hibridación resultaba ventajoso para la batería de plomo, ya que hacía que conservase un SoC promedio más elevado. Este es un

resultado interesante, ya que supone que un convertidor pequeño (y presumiblemente más barato) puede operar mejor que un convertidor de elevada potencia. Esta operación limitada resulta un campo de estudio interesante para optimizar el diseño de dichos convertidores.

- En lo referente a la predicción de fallos en instalaciones desarrollada en el capítulo 4, para cada instalación, se ha necesitado una base de datos propia y entrenar un algoritmo distinto. Esto limita en cierta manera la implementación práctica de estos algoritmos para instalaciones nuevas, ya que primero deben grabarse datos en dichas instalaciones para entrenar los algoritmos. Por lo tanto, resulta interesante estudiar la viabilidad de hacer un algoritmo universal y adaptable, capaz de ser desplegado en cualquier instalación, y aprender en vivo de los datos que va midiendo, de tal manera que adapte su comportamiento a cada instalación específica.
- EN lo referente a la estimación de vida realizada en el capítulo 5, resulta interesante expandir el estudio, realizado con química LFP, a otras químicas. Esto depende en gran medida de la disponibilidad de bases de datos, pero mejoraría la capacidad de generalización de las técnicas desarrolladas. Concretamente, aplicar dichos algoritmos a baterías VRLA ayudaría a analizar el impacto a largo plazo de los HESS.
- Adicionalmente, estos algoritmos pueden ser usados para desarrollar nuevas estrategias de gestión de la energía, que tengan en cuenta el envejecimiento de las baterías, y que optimicen el funcionamiento a largo plazo del pack. De este modo, se podrían tener en cuenta criterios de remplazo de las baterías para optimizar los HESS con criterios tecno-económicos.

References

Abbas, Ghulam, Menaa Nawaz, and Farrukh Kamran. 2019. “Performance Comparison of NARX & RNN-LSTM Neural Networks for LiFePO4 Battery State of Charge Estimation.” In *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, 463–68. IEEE. https://doi.org/10.1109/IBCAST.2019.8667172.

Abdel-Nasser, Mohamed, and Karar Mahmoud. 2019. “Accurate Photovoltaic Power Forecasting Models Using Deep LSTM-RNN.” *Neural Computing and Applications* 31 (7): 2727–40. https://doi.org/10.1007/s00521-017-3225-z.

Amine, K., J. Liu, and I. Belharouak. 2005. “High-Temperature Storage and Cycling of C-LiFePO4/Graphite Li-Ion Cells.” *Electrochemistry Communications* 7 (7): 669–73. https://doi.org/10.1016/j.elecom.2005.04.018.

Anagnostos, D., T. Schmidt, S. Cavadias, D. Soudris, J. Poortmans, and F. Catthoor. 2019. “A Method for Detailed, Short-Term Energy Yield Forecasting of Photovoltaic Installations.” *Renewable Energy* 130: 122–29. https://doi.org/10.1016/j.renene.2018.06.058.

Anbuky, A.H., and P.E. Pascoe. 2000. “VRLA Battery State-of-Charge Estimation in Telecommunication Power Systems.” *IEEE Transactions on Industrial Electronics* 47 (3): 565–73. https://doi.org/10.1109/41.847897.

Anderson, William W., and Oleg A. Yakimenko. 2018. “Using Neural Networks to Model and Forecast Solar PV Power Generation at Isle of Eigg.” *Proceedings - 2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering, CPE-POWERENG 2018*, 1–8. https://doi.org/10.1109/CPE.2018.8372522.

Andre, D., M. Meiler, K. Steiner, Ch Wimmer, T. Soczka-Guth, and D. U. Sauer. 2011. “Characterization of High-Power Lithium-Ion Batteries by Electrochemical Impedance Spectroscopy. I. Experimental Investigation.” *Journal of Power Sources* 196 (12): 5334–41. https://doi.org/10.1016/j.jpowsour.2010.12.102.

Ansean, David, Victor Manuel Garcia, Manuela Gonzalez, Cecilio Blanco-Viejo, Juan Carlos Viera, Yoana Fernandez Pulido, and Luciano Sanchez. 2019. “Lithium-Ion Battery Degradation Indicators Via Incremental Capacity Analysis.” *IEEE Transactions on Industry Applications* 55 (3): 2992–3002. https://doi.org/10.1109/TIA.2019.2891213.

Anzola, Jon, Iosu Aizpuru, Asier Arruti, Argine Alacano, Ramon Lopez, Jesus Sergio Artal-Sevil, and Carlos Bernal-Ruiz. 2020. "Partial Power Processing Based Charging Unit for Electric Vehicle Extreme Fast Charging Stations." In *2020 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 1–6. IEEE. https://doi.org/10.1109/VPPC49601.2020.9330881.

Artal-Sevil, Jesus Sergio, Carlos Bernal-Ruiz, Jon Anzola, Iosu Aizpuru, Antonio Bono-Nuez, and Jose Miguel Sanz-Alcaine. 2020. "Partial Power Processing Architecture Applied to a Battery Energy Storage System." In *2020 IEEE Vehicle Power and Propulsion Conference, VPPC 2020 - Proceedings*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/VPPC49601.2020.9330966.

Barré, Anthony, Benjamin Deguilhem, Sébastien Grolleau, Mathias Gérard, Frédéric Suard, and Delphine Riu. 2013. "A Review on Lithium-Ion Battery Ageing Mechanisms and Estimations for Automotive Applications." *Journal of Power Sources* 241 (November): 680–89. https://doi.org/10.1016/J.JPOWSOUR.2013.05.040.

"BATT-Ex- Desarrollo de Extensores de Vida Útil de Baterías Para Sistemas Autónomos Alimentados Por Placas Fotovoltaicas | SICE." n.d. Accessed April 26, 2021. https://www.sice.com/idi/batt-ex-desarrollo-de-extensores-de-vida-util-de-baterias-para-sistemas.

Bloom, I, B.W Cole, J.J Sohn, S.A Jones, E.G Polzin, V.S Battaglia, G.L Henriksen, et al. 2001. "An Accelerated Calendar and Cycle Life Study of Li-Ion Cells." *Journal of Power Sources* 101 (2): 238–47. https://doi.org/10.1016/S0378-7753(01)00783-2.

Brenna, Morris, Federica Foiadelli, Michela Longo, and Dario Zaninelli. 2018. "Energy Storage Control for Dispatching Photovoltaic Power." *IEEE Transactions on Smart Grid* 9 (4): 2419–28. https://doi.org/10.1109/TSG.2016.2611999.

Brito, Moacyr Aureliano Gomes De, Luigi Galotto, Leonardo Poltronieri Sampaio, Guilherme De Azevedo Melo, and Carlos Alberto Canesin. 2013. "Evaluation of the Main MPPT Techniques for Photovoltaic Applications." *IEEE Transactions on Industrial Electronics* 60 (3): 1156–67. https://doi.org/10.1109/TIE.2012.2198036.

Broussely, M, S Herreyre, P Biensan, P Kasztejna, K Nechev, and R.J Staniewicz. 2001. "Aging Mechanism in Li Ion Cells and Calendar Life Predictions." *Journal of Power Sources* 97–98 (July): 13–21.

https://doi.org/10.1016/S0378-7753(01)00722-4.

Budde-Meiwes, Heide, Julia Drillkens, Benedikt Lunz, Jens Muennix, Susanne Rothgang, Julia Kowal, and Dirk Uwe Sauer. 2013. "A Review of Current Automotive Battery Technology and Future Prospects." *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*. https://doi.org/10.1177/0954407013485567.

Burgess, Webb L. 2009. "Valve Regulated Lead Acid Battery Float Service Life Estimation Using a Kalman Filter." *Journal of Power Sources* 191 (1): 16–21. https://doi.org/10.1016/j.jpowsour.2008.12.123.

Caliwag, Angela C., and Wansu Lim. 2019. "Hybrid VARMA and LSTM Method for Lithium-Ion Battery State-of-Charge and Output Voltage Forecasting in Electric Motorcycle Applications." *IEEE Access* 7: 59680–89. https://doi.org/10.1109/ACCESS.2019.2914188.

Carignano, Mauro G, Ramon Costa-Castell O, Vicente Roda, Norberto M Nigro, Sergio Junco, and Diego Feroldi. 2017. "Energy Management Strategy for Fuel Cell-Supercapacitor Hybrid Vehicles Based on Prediction of Energy Demand." *Journal of Power Sources*. https://doi.org/10.1016/j.jpowsour.2017.06.016.

Carrilero, Isabel, Manuela González, David Anseán, Juan C. Viera, Joaquín Chacón, and Paulo G. Pereirinha. 2018. "Redesigning European Public Transport: Impact of New Battery Technologies in the Design of Electric Bus Fleets." In *Transportation Research Procedia*, 33:195–202. Elsevier B.V. https://doi.org/10.1016/j.trpro.2018.10.092.

Chandrasekaran, Rajeswari. 2014. "Quantification of Bottlenecks to Fast Charging of Lithium-Ion-Insertion Cells for Electric Vehicles." *Journal of Power Sources* 271 (December): 622–32. https://doi.org/10.1016/j.jpowsour.2014.07.106.

Chaoui, Hicham, and Chinemerem Christopher Ibe-Ekeocha. 2017. "State of Charge and State of Health Estimation for Lithium Batteries Using Recurrent Neural Networks." *IEEE Transactions on Vehicular Technology* 66 (10): 8773–83. https://doi.org/10.1109/TVT.2017.2715333.

Chen, Chaochao, and Michael Pecht. 2012. "Prognostics of Lithium-Ion Batteries Using Model-Based and Data-Driven Methods." *Proceedings of IEEE 2012 Prognostics and System Health Management Conference, PHM-2012*, 1–6. https://doi.org/10.1109/PHM.2012.6228850.

Chen, M., and G.A. Rincon-Mora. 2006. "Accurate Electrical Battery Model

Capable of Predicting Runtime and I–V Performance." *IEEE Transactions on Energy Conversion* 21 (2): 504–11. https://doi.org/10.1109/TEC.2006.874229.

Christen, Thomas, and Martin W. Carlen. 2000. "Theory of Ragone Plots." *Journal of Power Sources* 91 (2): 210–16. https://doi.org/10.1016/S0378-7753(00)00474-2.

Christensen, John, and John Newman. 2006. "Stress Generation and Fracture in Lithium Insertion Materials." *Journal of Solid State Electrochemistry* 10 (5): 293–319. https://doi.org/10.1007/s10008-006-0095-1.

Cireşan, Dan, Ueli Meier, and Juergen Schmidhuber. 2012. "Multi-Column Deep Neural Networks for Image Classification." *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, February, 3642–49. http://arxiv.org/abs/1202.2745.

Coleman, Martin, Chi Kwan Lee, Chunbo Zhu, W.G. William Gerard Hurley, Chi Kwan Chi Kwan Lee, Chunbo Chunbo Zhu, W.G. William Gerard Hurley, Chi Kwan Lee, Chunbo Zhu, and W.G. William Gerard Hurley. 2007. "State-of-Charge Determination from EMF Voltage Estimation: Using Impedance, Terminal Voltage, and Current for Lead-Acid and Lithium-Ion Batteries." *IEEE Transactions on Industrial Electronics* 54 (5): 2550–57. https://doi.org/10.1109/TIE.2007.899926.

Dubarry, Matthieu, George Baure, and David Anseán. 2020. "Perspective on State-of-Health Determination in Lithium-Ion Batteries." *Journal of Electrochemical Energy Conversion and Storage* 17 (4). https://doi.org/10.1115/1.4045008.

Dubarry, Matthieu, and Yann Liaw. 2007. "Development of a Universal Modeling Tool for Rechargeable Lithium Batteries." *Journal of Power Sources* 174: 856–60. https://doi.org/10.1016/j.jpowsour.2007.06.157.

EPEC. 2021. "Battery Comparison of Energy Density - Cylindrical and Prismatic Cells." 2021. https://www.epectec.com/batteries/cell-comparison.html.

Esram, Trishan, and Partrick L. Chapman. 2007. "Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques." *IEEE Transactions on Energy Conversion* 22 (2): 439–49. https://doi.org/10.1109/TEC.2006.874230.

Exide Technologies. 2012. "Handbook for Stationary Vented Lead-Acid Batteries Part 2 : Installation , Commissioning and Operation" 2: 1–68.

Fermín-Cueto, Paula, Euan McTurk, Michael Allerhand, Encarni Medina-Lopez, Miguel F Anjos, Joel Sylvester, and Gonçalo dos Reis. 2020. "Identification and Machine Learning Prediction of Knee-Point and Knee-Onset in Capacity Degradation Curves of Lithium-Ion Cells." *Energy and AI* 1: 100006. https://doi.org/https://doi.org/10.1016/j.egyai.2020.100006.

Ganin, Yaroslav, Victor Lempitsky, and Lempitsky@skoltech Ru. 2015. "Unsupervised Domain Adaptation by Backpropagation." PMLR. http://proceedings.mlr.press/v37/ganin15.html.

Garayalde, E., I. Aizpuru, J.M. Canales, I. Sanz, C. Bernal, and E. Oyarbide. 2017. "Análisis Experimental Del Efecto de La Temperatura y La Tensión de Carga Para La Optimización Energética de Sistemas de Almacenamiento de Instalaciones Fotovoltaicas Aisladas." *SAAEI 2017*.

Garayalde Perez, Erik. 2020. "Hybrid Energy Storage Systems via Power Electronic Converters," February. https://hdl.handle.net/20.500.11984/1783.

Girosi, Federico, Michael Jones, and Tomaso Poggio. 1995. "Regularization Theory and Neural Networks Architectures." *Neural Computation* 7 (2): 219–69. https://doi.org/10.1162/neco.1995.7.2.219.

Goebel, Kai, and Bhaskar Saha. 2010. "DASHlink - Li-Ion Battery Aging Datasets." 2010. https://c3.nasa.gov/dashlink/resources/133/.

Goodfellow, Ian J, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. "Generative Adversarial Nets."

Grolleau, Sébastien, Arnaud Delaille, Hamid Gualous, Philippe Gyan, Renaud Revel, Julien Bernard, Eduardo Redondo-Iglesias, and Jérémy Peter. 2014. "Calendar Aging of Commercial Graphite/LiFePO4 Cell - Predicting Capacity Fade under Time Dependent Storage Conditions." *Journal of Power Sources* 255 (June): 450–58. https://doi.org/10.1016/j.jpowsour.2013.11.098.

Gu, Jiuxiang, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, et al. 2018. "Recent Advances in Convolutional Neural Networks." *Pattern Recognition* 77 (May): 354–77. https://doi.org/10.1016/j.patcog.2017.10.013.

Guillén Asensio, Alejandro, Iván Sanz-Gorrachategui, Pablo Pastor-Flores, J. Sergio Artal-Sevil, Antonio Bono-Nuez, B. Martin-del-Brio, and C. Bernal-Ruiz. 2020. "Battery State Prediction in Photovoltaic Standalone

Installations." In *2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER)*. https://doi.org/78-1-7281-5641-5/20.

Günther, Michael, Uwe Feldmann, and Jan ter Maten. 2005. "Modelling and Discretization of Circuit Problems." *Handbook of Numerical Analysis* 13: 523–659. https://doi.org/10.1016/S1570-8659(04)13006-8.

Haggstrom, Gus W. 1983. "Logistic Regression and Discriminant Analysis by Ordinary Least Squares." *Journal of Business & Economic Statistics* 1 (3): 229. https://doi.org/10.2307/1391344.

Hara, Kazuyuki, Daisuke Saito, and Hayaru Shouno. 2015. "Analysis of Function of Rectified Linear Unit Used in Deep Learning." In *Proceedings of the International Joint Conference on Neural Networks*. Vol. 2015-September. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/IJCNN.2015.7280578.

Haykin, Simon O. 1998. *Neural Networks: A Comprehensive Foundation*. Prentice/Hall International.

He, Wei, Nicholas Williard, Michael Osterman, and Michael Pecht. 2011. "Prognostics of Lithium-Ion Batteries Based on Dempster-Shafer Theory and the Bayesian Monte Carlo Method." *Journal of Power Sources* 196 (23): 10314–21. https://doi.org/10.1016/j.jpowsour.2011.08.040.

Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9 (8): 1735–80. https://doi.org/10.1162/neco.1997.9.8.1735.

Hoerl, Arthur E, and Robert W Kennard. 1970. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." Vol. 12.

Homan, Bart, Marnix V. ten Kortenaar, Johann L. Hurink, and Gerard J.M. Smit. 2019. "A Realistic Model for Battery State of Charge Prediction in Energy Management Simulation Tools." *Energy* 171: 205–17. https://doi.org/10.1016/j.energy.2018.12.134.

Hong, Jichao, Zhenpo Wang, Wen Chen, Le Yi Wang, and Changhui Qu. 2020. "Online Joint-Prediction of Multi-Forward-Step Battery SOC Using LSTM Neural Networks and Multiple Linear Regression for Real-World Electric Vehicles." *Journal of Energy Storage*. https://doi.org/10.1016/j.est.2020.101459.

Hossain, Eklas, Darren Murtaugh, Jaisen Mody, Hossain Mansur Resalat Faruque, Md. Samiul Haque Sunny, and Naeem Mohammad. 2019. "A

Comprehensive Review on Second-Life Batteries: Current State, Manufacturing Considerations, Applications, Impacts, Barriers & Potential Solutions, Business Strategies, and Policies." *IEEE Access* 7: 73215–52. https://doi.org/10.1109/ACCESS.2019.2917859.

Hossain, Mohammad Safayet, and Hisham Mahmood. 2020. "Short-Term Photovoltaic Power Forecasting Using an LSTM Neural Network." In *2020 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2020*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ISGT45199.2020.9087786.

Huang, Chao Ming, Shin Ju Chen, Sung Pei Yang, and Chung Jen Kuo. 2015. "One-Day-Ahead Hourly Forecasting for Photovoltaic Power Generation Using an Intelligent Method with Weather-Based Forecasting Models." *IET Generation, Transmission and Distribution* 9 (14): 1874–82. https://doi.org/10.1049/iet-gtd.2015.0175.

IEEE. 2019. "IEEE Recommended Practice for Sizing Lead-Acid Batteries for Stand-Alone Photovoltaic (PV) Systems." IEEE Std 1013-2019. 2019. https://doi.org/10.1109/IEEESTD.2019.8845030.

Jaguemont, J., L. Boulon, and Y. Dubé. 2016. "A Comprehensive Review of Lithium-Ion Batteries Used in Hybrid and Electric Vehicles at Cold Temperatures." *Applied Energy*. Elsevier Ltd. https://doi.org/10.1016/j.apenergy.2015.11.034.

Jia, Jianfang, Jianyu Liang, Yuanhao Shi, Jie Wen, Xiaoqiong Pang, and Jianchao Zeng. 2020. "SOH and RUL Prediction of Lithium-Ion Batteries Based on Gaussian Process Regression with Indirect Health Indicators." *Energies* 13 (2): 375. https://doi.org/10.3390/en13020375.

Jian Cao, and A. Emadi. 2012. "A New Battery/UltraCapacitor Hybrid Energy Storage System for Electric, Hybrid, and Plug-In Hybrid Electric Vehicles." *IEEE Transactions on Power Electronics* 27 (1): 122–32. https://doi.org/10.1109/TPEL.2011.2151206.

Jiang, Yan, Jiuchun Jiang, Caiping Zhang, Weige Zhang, Yang Gao, and Na Li. 2018. "State of Health Estimation of Second-Life LiFePO4 Batteries for Energy Storage Applications." *Journal of Cleaner Production* 205 (December): 754–62. https://doi.org/10.1016/J.JCLEPRO.2018.09.149.

Kazem, Hussein A., Tamer Khatib, and K. Sopian. 2013. "Sizing of a Standalone Photovoltaic/Battery System at Minimum Cost for Remote Housing Electrification in Sohar, Oman." *Energy and Buildings* 61 (June): 108–15. https://doi.org/10.1016/J.ENBUILD.2013.02.011.

Kebaïli, S., and H. Benalla. 2015. "Optimal Sizing of a Stand-Alone Photovoltaic Systems under Various Weather Conditions in Algeria." *Revue Des Energies Renouvelables* 18 (2): 179–191. http://www.cder.dz/download/Art18-2_1.pdf.

Khaligh, Alireza, and Zhihao Li. 2010. "Battery, Ultracapacitor, Fuel Cell, and Hybrid Energy Storage Systems for Electric, Hybrid Electric, Fuel Cell, and Plug-In Hybrid Electric Vehicles: State of the Art." *IEEE Transactions on Vehicular Technology* 59 (6): 2806–14. https://doi.org/10.1109/TVT.2010.2047877.

Khatib, Tamer, Ibrahim A. Ibrahim, and Azah Mohamed. 2016. "A Review on Sizing Methodologies of Photovoltaic Array and Storage Battery in a Standalone Photovoltaic System." *Energy Conversion and Management* 120 (July): 430–48. https://doi.org/10.1016/J.ENCONMAN.2016.05.011.

Khelif, Racha, Brigitte Chebel-Morello, Simon Malinowski, Emna Laajili, Farhat Fnaiech, and Noureddine Zerhouni. 2017. "Direct Remaining Useful Life Estimation Based on Support Vector Regression." *IEEE Transactions on Industrial Electronics* 64 (3): 2276–85. https://doi.org/10.1109/TIE.2016.2623260.

Kim, Younghyun, Jason Koh, Qing Xie, Yanzhi Wang, Naehyuck Chang, and Massoud Pedram. 2014. "A Scalable and Flexible Hybrid Energy Storage System Design and Implementation." *Journal of Power Sources* 255: 410–22. https://doi.org/10.1016/j.jpowsour.2013.12.102.

Kingma, Diederik P., and Jimmy Lei Ba. 2015. "Adam: A Method for Stochastic Optimization." In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. International Conference on Learning Representations, ICLR. https://arxiv.org/abs/1412.6980v9.

Kodinariya, T., and Prashant R. Makwana. 2013. "Review on Determining Number of Cluster in K-Means Clustering." *Undefined*.

Kohonen, T., E. Oja, O. Simula, A. Visa, and J. Kangas. 1996. "Engineering Applications of the Self-Organizing Map." *Proceedings of the IEEE* 84 (10): 1358–84. https://doi.org/10.1109/5.537105.

Kohonen, Teuvo. 2001. *Self-Organizing Maps*. 3rd ed. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-56927-2.

Kroics, Kaspars. 2015. "System for Start of Internal Combustion Engine with Hybrid Battery-Supercapacitor Source." In *2015 56th International*

Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), 1–6. IEEE. https://doi.org/10.1109/RTUCON.2015.7343167.

Lawhern, Vernon J., Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. 2016. "EEGNet: A Compact Convolutional Network for EEG-Based Brain-Computer Interfaces." *Journal of Neural Engineering* 15 (5). https://doi.org/10.1088/1741-2552/aace8c.

Leontaritis, I. J., and S. A. Billings. 1985. "Input-Output Parametric Models for Non-Linear Systems Part I: Deterministic Non-Linear Systems." *International Journal of Control* 41 (2): 303–28. https://doi.org/10.1080/0020718508961129.

Leva, S., A. Dolara, F. Grimaccia, M. Mussetta, and E. Ogliari. 2017. "Analysis and Validation of 24 Hours Ahead Neural Network Forecasting of Photovoltaic Output Power." *Mathematics and Computers in Simulation* 131: 88–100. https://doi.org/10.1016/j.matcom.2015.05.010.

Levenberg, Kenneth. 1944. "A Method for the Solution of Certain Non-Linear Problems in Least Squares." *Quarterly of Applied Mathematics* 2 (2): 164–68. https://doi.org/10.1090/qam/10666.

Li, Lianbing, Yazun Zhu, Linglong Wang, Donghua Yue, and Duo Li. 2019. "Indirect Remaining Useful Life Prognostics for Lithium-Ion Batteries." *2018 24th International Conference on Automation and Computing (ICAC)*, no. September: 1–5. https://doi.org/10.23919/iconac.2018.8748973.

Li, Lyu, Yu Peng, Yuchen Song, and Datong Liu. 2019. "Lithium-Ion Battery Remaining Useful Life Prognostics Using Data-Driven Deep Learning Algorithm." *Proceedings - 2018 Prognostics and System Health Management Conference, PHM-Chongqing 2018*, 1094–1100. https://doi.org/10.1109/PHM-Chongqing.2018.00193.

Li, Yi, Kailong Liu, Aoife M. Foley, A. Zülke, Maitane Berecibar, E. Nanini-Maury, J. Van Mierlo, and Harry E. Hoster. 2019. "Data-Driven Health Estimation and Lifetime Prediction of Lithium-Ion Batteries: A Review." *Renewable and Sustainable Energy Reviews*. Elsevier Ltd. https://doi.org/10.1016/j.rser.2019.109254.

Lipu, Molla S.Hossain, Mahammad A. Hannan, Aini Hussain, Mohamad H.M. Saad, Afida Ayob, and Frede Blaabjerg. 2018. "State of Charge Estimation for Lithium-Ion Battery Using Recurrent NARX Neural Network Model

Based Lighting Search Algorithm." *IEEE Access* 6: 28150–61. https://doi.org/10.1109/ACCESS.2018.2837156.

Liu, Datong, Hong Wang, Yu Peng, Wei Xie, and Haitao Liao. 2013. "Satellite Lithium-Ion Battery Remaining Cycle Life Prediction with Novel Indirect Health Indicator Extraction." *Energies* 6 (8): 3654–68. https://doi.org/10.3390/en6083654.

Liu, Datong, Wei Xie, Haitao Liao, and Yu Peng. 2015. "An Integrated Probabilistic Approach to Lithium-Ion Battery Remaining Useful Life Estimation." *IEEE Transactions on Instrumentation and Measurement* 64 (3): 660–70. https://doi.org/10.1109/TIM.2014.2348613.

Liu, Jian, and Ziqiang Chen. 2019. "Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Health Indicator and Gaussian Process Regression Model." *IEEE Access* 7: 39474–84. https://doi.org/10.1109/ACCESS.2019.2905740.

Liu, Zhenbao, Gaoyuan Sun, Shuhui Bu, Junwei Han, Xiaojun Tang, and Michael Pecht. 2017. "Particle Learning Framework for Estimating the Remaining Useful Life of Lithium-Ion Batteries." *IEEE Transactions on Instrumentation and Measurement* 66 (2): 280–93. https://doi.org/10.1109/TIM.2016.2622838.

Lu, Languang, Xuebing Han, Jianqiu Li, Jianfeng Hua, and Minggao Ouyang. 2013. "A Review on the Key Issues for Lithium-Ion Battery Management in Electric Vehicles." *Journal of Power Sources* 226 (March): 272–88. https://doi.org/10.1016/J.JPOWSOUR.2012.10.060.

Ma, Yanying, Lifeng Wu, Yong Guan, and Zhen Peng. 2020. "The Capacity Estimation and Cycle Life Prediction of Lithium-Ion Batteries Using a New Broad Extreme Learning Machine Approach." *Journal of Power Sources* 476 (November): 228581. https://doi.org/10.1016/j.jpowsour.2020.228581.

Macdonald, J. R., and M. K. Brachman. 1955. "The Charging and Discharging of Nonlinear Capacitors." *Proceedings of the IRE* 43 (1): 71–78. https://doi.org/10.1109/JRPROC.1955.277920.

Marchildon, Jacques, Mamadou Lamine Doumbia, and Kodjo Agbossou. 2015. "SOC and SOH Characterisation of Lead Acid Batteries." In *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*, 001442–46. IEEE. https://doi.org/10.1109/IECON.2015.7392303.

Martinez-Laserna, E., I. Gandiaga, E. Sarasketa-Zabala, J. Badeda, D.-I. Stroe, M. Swierczynski, and A. Goikoetxea. 2018. "Battery Second Life: Hype,

Hope or Reality? A Critical Review of the State of the Art." *Renewable and Sustainable Energy Reviews* 93 (October): 701–18. https://doi.org/10.1016/J.RSER.2018.04.035.

Martinez-Laserna, E., E. Sarasketa-Zabala, D. I. Stroe, M. Swierczynski, A. Warnecke, J. M. Timmermans, S. Goutam, and P. Rodriguez. 2016. "Evaluation of Lithium-Ion Battery Second Life Performance and Degradation." In *ECCE 2016 - IEEE Energy Conversion Congress and Exposition, Proceedings*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ECCE.2016.7855090.

Miao, Qiang, Lei Xie, Hengjuan Cui, Wei Liang, and Michael Pecht. 2013. "Remaining Useful Life Prediction of Lithium-Ion Battery with Unscented Particle Filter Technique." *Microelectronics Reliability* 53 (6): 805–10. https://doi.org/10.1016/j.microrel.2012.12.004.

Mittal, Sparsh. n.d. "A Survey of FPGA-Based Accelerators for Convolutional Neural Networks." *Neural Computing and Applications* 32 (4): 1109–39. Accessed December 7, 2020. https://www.academia.edu/37491583/A_Survey_of_FPGA_based_Accelerators_for_Convolutional_Neural_Networks.

Mohan, Ned., Tore M. Undeland, and William P. Robbins. 2003. *Power Electronics : Converters, Applications, and Design*. John Wiley & Sons. https://www.wiley.com/en-us/Power+Electronics%3A+Converters%2C+Applications%2C+and+Design%2C+3rd+Edition-p-9780471226932.

Montaru, M., and S. Pelissier. 2010. "Frequency and Temporal Identification of a Li-Ion Polymer Battery Model Using Fractional Impedance." *Oil and Gas Science and Technology* 65 (1): 67–78. https://doi.org/10.2516/ogst/2009056.

Moseley, Patrick T., and David A. Rand. 2012. "Partial State-of-Charge Duty: A Challenge but Not a Show-Stopper for Lead-Acid Batteries!" 41 (13): 3–16. https://doi.org/10.1149/1.3691907.

NaitMalek, Y., M. Najib, M. Bakhouya, and M. Essaaidi. 2019. "Forecasting the State-of-Charge of Batteries in Micro-Grid Systems." *Proceedings of 2019 IEEE World Conference on Complex Systems, WCCS 2019*, 1–6. https://doi.org/10.1109/ICoCS.2019.8930731.

Ng, Kong Soon, Chin Sien Moo, Yi Ping Chen, and Yao Ching Hsieh. 2009. "Enhanced Coulomb Counting Method for Estimating State-of-Charge and State-of-Health of Lithium-Ion Batteries." *Applied Energy* 86 (9): 1506–11.

https://doi.org/10.1016/j.apenergy.2008.11.021.

Nital David, Sandeep. 2016. "Pulse Power Characterisation for Lithium Ion Cells in Automotive Applications Small and Large Signal Cell Impedance Analysis."

Ohsaki, Takahisa, Takashi Kishi, Takashi Kuboki, Norio Takami, Nao Shimura, Yuichi Sato, Masahiro Sekino, and Asako Satoh. 2005. "Overcharge Reaction of Lithium-Ion Batteries." In *Journal of Power Sources*, 146:97–100. Elsevier. https://doi.org/10.1016/j.jpowsour.2005.03.105.

Omar, Noshin, Peter Van den Bossche, Thierry Coosemans, and Joeri Van Mierlo. 2013. "Peukert Revisited-Critical Appraisal and Need for Modification for Lithium-Ion Batteries." *Energies* 6 (11): 5625–41. https://doi.org/10.3390/en6115625.

Ospina, Juan, Alvi Newaz, and M. Omar Faruque. 2019. "Forecasting of PV Plant Output Using Hybrid Wavelet-Based LSTM-DNN Structure Model." *IET Renewable Power Generation* 13 (7): 1087–95. https://doi.org/10.1049/iet-rpg.2018.5779.

Pajovic, Milutin, Philip Orlik, and Toshihiro Wada. 2018. "Remaining Useful Life Estimation of Batteries Using Dirichlet Process with Variational Bayes Inference." In *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, 2729–35. IEEE. https://doi.org/10.1109/IECON.2018.8592767.

Pajovic, Milutin, Zafer Sahinoglu, Yebin Wang, Philip V. Orlik, and Toshihiro Wada. 2017. "Online Data-Driven Battery Voltage Prediction." *Proceedings - 2017 IEEE 15th International Conference on Industrial Informatics, INDIN 2017*, 827–34. https://doi.org/10.1109/INDIN.2017.8104879.

Pascoe, Phillip E., and Adnan H. Anbuky. 2004. "VRLA Battery Discharge Reserve Time Estimation." *IEEE Transactions on Power Electronics* 19 (6): 1515–22. https://doi.org/10.1109/TPEL.2004.836680.

Piller, Sabine, Marion Perrin, and Andreas Jossen. 2001. "Methods for State-of-Charge Determination and Their Applications." In *Journal of Power Sources*, 96:113–20. Elsevier. https://doi.org/10.1016/S0378-7753(01)00560-2.

Plett, Gregory L. 2004. "Extended Kalman Filtering for Battery Management Systems of LiPB-Based HEV Battery Packs - Part 2. Modeling and

Identification." *Journal of Power Sources* 134 (2): 262–76. https://doi.org/10.1016/j.jpowsour.2004.02.032.

Ragone, D. V. 1968. "Review of Battery Systems for Electrically Powered Vehicles." In . https://doi.org/10.4271/680453.

Rahe, Christiane. 2016. "Lead-Acid Batteries and Lithium-Ion Batteries in Parallel Strings for an Energy Storage System for a Clinic in Africa," 1–7.

Rahmoun, Ahmad, Helmuth Biechl, and Argo Rosin. 2013. "Evaluation of Equivalent Circuit Diagrams and Transfer Functions for Modeling of Lithium-Ion Batteries." *Electrical, Control and Communication Engineering*, no. 2. https://doi.org/10.2478/ecce-2013-0005.

Ren, Lei, Li Zhao, Sheng Hong, Shiqiang Zhao, Hao Wang, and Lin Zhang. 2018. "Remaining Useful Life Prediction for Lithium-Ion Battery: A Deep Learning Approach." *IEEE Access* 6: 50587–98. https://doi.org/10.1109/ACCESS.2018.2858856.

Rodríguez, Gustavo Pérez, Iñigo Gandiaga Ondaro, Maitane Garmendia Elorza, Jean François Reynaud, and Unai Viscarret. 2013. "Modelling of Li-Ion Batteries Dynamics Using Impedance Spectroscopy and Pulse Fitting: EVs Application." *World Electric Vehicle Journal* 6 (3): 644–52. https://doi.org/10.3390/wevj6030644.

Saft Batteries. 2007. "Sunica.plus Technical Manual."

Saha, Bhaskar, Kai Goebel, Scott Poll, and Jon Christophersen. 2009. "Prognostics Methods for Battery Health Monitoring Using a Bayesian Framework." *IEEE Transactions on Instrumentation and Measurement* 58 (2): 291–96. https://doi.org/10.1109/TIM.2008.2005965.

Sahinoglu, Gozde O., Milutin Pajovic, Zafer Sahinoglu, Yebin Wang, Philip V. Orlik, and Toshihiro Wada. 2018. "Battery State-of-Charge Estimation Based on Regular/Recurrent Gaussian Process Regression." *IEEE Transactions on Industrial Electronics* 65 (5): 4311–21. https://doi.org/10.1109/TIE.2017.2764869.

Sanz-Alcaine, José Miguel, Eduardo Sebastián, Iván Sanz-Gorrachategui, Carlos Bernal-Ruiz, Antonio Bono-Nuez, Milutin Pajovic, and Philip V. Orlik. 2021. "Online Voltage Prediction Using Gaussian Process Regression for Fault-Tolerant Photovoltaic Standalone Applications." *Neural Computing and Applications*, June, 1–14. https://doi.org/10.1007/s00521-021-06254-6.

Sanz-Gorrachategui, Iván, Carlos Bernal, Estanis Oyarbide, Erik Garayalde,

Iosu Aizpuru, Jose María Canales, and Antonio Bono-Nuez. 2018. "New Battery Model Considering Thermal Transport and Partial Charge Stationary Effects in Photovoltaic Off-Grid Applications." *Journal of Power Sources* 378: 311–21. https://doi.org/10.1016/j.jpowsour.2017.12.058.

Sanz-Gorrachategui, Iván, Carlos Bernal Ruiz, Estanis Oyarbide Usabiaga, Antonio Bono Nuez, Sergio Jesús Artal Sevil, Erik Garayalde Pérez, Iosu Aizpuru Larrañaga, and Jose María Canales Segade. 2019. "Partial State-of-Charge Mitigation in Standalone Photovoltaic Hybrid Storage Systems." *Energies* 12 (22): 4393. https://doi.org/10.3390/en12224393.

Sanz-Gorrachategui, Iván, Pablo Pastor-Flores, Alejandro Guillén Asensio, J. Sergio Artal-Sevil, Antonio Bono-Nuez, Bonifacio Martin-del-Brio, and Carlos Bernal-Ruiz. 2020. "Unsupervised Clustering of Battery Waveforms in Off-Grid PV Installations." In *2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER).* https://doi.org/978-1-7281-5641-5/20.

Schiffer, Julia, Dirk Uwe Sauer, Henrik Bindner, Tom Cronin, Per Lundsager, and Rudi Kaiser. 2007. "Model Prediction for Ranking Lead-Acid Batteries According to Expected Lifetime in Renewable Energy Systems and Autonomous Power-Supply Systems." *Journal of Power Sources* 168 (1): 66–78. https://doi.org/10.1016/J.JPOWSOUR.2006.11.092.

Schmidhuber, Jürgen. 2015. "Deep Learning in Neural Networks: An Overview." *Neural Networks*. Elsevier Ltd. https://doi.org/10.1016/j.neunet.2014.09.003.

Schuster, Simon F., Tobias Bach, Elena Fleder, Jana Müller, Martin Brand, Gerhard Sextl, and Andreas Jossen. 2015. "Nonlinear Aging Characteristics of Lithium-Ion Cells under Different Operational Conditions." *Journal of Energy Storage* 1 (June): 44–53. https://doi.org/10.1016/J.EST.2015.05.003.

Schweiger, Hans Georg, Ossama Obeidi, Oliver Komesker, André Raschke, Michael Schiemann, Christian Zehner, Markus Gehnen, Michael Keller, and Peter Birke. 2010. "Comparison of Several Methods for Determining the Internal Resistance of Lithium Ion Cells." *Sensors* 10 (6): 5604–25. https://doi.org/10.3390/s100605604.

Sepasi, Saeed, Reza Ghorbani, and Bor Yann Liaw. 2015. "Inline State of Health Estimation of Lithium-Ion Batteries Using State of Charge Calculation." *Journal of Power Sources* 299: 246–54.

https://doi.org/10.1016/j.jpowsour.2015.08.091.

Severson, Kristen A., Peter M. Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H. Chen, et al. 2019. "Data-Driven Prediction of Battery Cycle Life before Capacity Degradation." *Nature Energy* 4 (5): 383–91. https://doi.org/10.1038/s41560-019-0356-8.

Shahriari, Mehrnoosh, and Mohammad Farrokhi. 2013. "Online State-of-Health Estimation of VRLA Batteries Using State of Charge." *IEEE Transactions on Industrial Electronics* 60 (1): 191–202. https://doi.org/10.1109/TIE.2012.2186771.

Shen, W.X. 2009. "Optimally Sizing of Solar Array and Battery in a Standalone Photovoltaic System in Malaysia." *Renewable Energy* 34 (1): 348–52. https://doi.org/10.1016/J.RENENE.2008.03.015.

Shrestha, G. B., and L. Goel. 1998. "A Study on Optimal Sizing of Stand-Alone Photovoltaic Stations." *IEEE Transactions on Energy Conversion* 13 (4): 373–78. https://doi.org/10.1109/60.736323.

Sokolova, Marina, Nathalie Japkowicz, and Stan Szpakowicz. 2006. "Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation." In *AAAI Workshop - Technical Report*, WS-06-06:24–29. https://doi.org/10.1007/11941439_114.

Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, and Ruslan Salakhutdinov. 2014. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." *Journal of Machine Learning Research*. Vol. 15. http://jmlr.org/papers/v15/srivastava14a.html.

Subudhi, Bidyadhar, and Raseswari Pradhan. 2013. "A Comparative Study on Maximum Power Point Tracking Techniques for Photovoltaic Power Systems." *IEEE Transactions on Sustainable Energy* 4 (1): 89–98. https://doi.org/10.1109/TSTE.2012.2202294.

Sun, Yu Hua, Hurng Liahng Jou, and Jinn Chang Wu. 2011. "Aging Estimation Method for Lead-Acid Battery." *IEEE Transactions on Energy Conversion* 26 (1): 264–71. https://doi.org/10.1109/TEC.2010.2040478.

Swingler, Andrew, and George Colgate. 2018. "12 Years of Residential ' Off-Grid ' PV Hybrid System Operation and Evolution in Nemiah Valley , Canada." In .

Tang, Xiaopeng, Boyang Liu, and Furong Gao. 2017. "State of Charge Estimation of LiFePO 4 Battery Based on a Gain-Classifier Observer." *Energy Procedia* 105: 2071–76.

https://doi.org/10.1016/j.egypro.2017.03.585.

Thounthong, Phatiphat, Stephane Raël, and Bernard Davat. 2009. "Energy Management of Fuel Cell/Battery/Supercapacitor Hybrid Power Source for Vehicle Applications." *Journal of Power Sources* 193: 376–85. https://doi.org/10.1016/j.jpowsour.2008.12.120.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Tong, Shijie, Matthew P. Klein, and Jae Wan Park. 2015. "On-Line Optimization of Battery Open Circuit Voltage for Improved State-of-Charge and State-of-Health Estimation." *Journal of Power Sources* 293: 416–28. https://doi.org/10.1016/j.jpowsour.2015.03.157.

Wang, Limei, Chaofeng Pan, Liang Liu, Yong Cheng, and Xiuliang Zhao. 2016. "On-Board State of Health Estimation of LiFePO4 Battery Pack through Differential Voltage Analysis." *Applied Energy* 168: 465–72. https://doi.org/10.1016/j.apenergy.2016.01.125.

Wei, Jingwen, Guangzhong Dong, and Zonghai Chen. 2018. "Remaining Useful Life Prediction and State of Health Diagnosis for Lithium-Ion Batteries Using Particle Filter and Support Vector Regression." *IEEE Transactions on Industrial Electronics* 65 (7): 5634–43. https://doi.org/10.1109/TIE.2017.2782224.

Wright, R.B, J.P Christophersen, C.G Motloch, J.R Belt, C.D Ho, V.S Battaglia, J.A Barnes, T.Q Duong, and R.A Sutula. 2003. "Power Fade and Capacity Fade Resulting from Cycle-Life Testing of Advanced Technology Development Program Lithium-Ion Batteries." *Journal of Power Sources* 119–121 (June): 865–69. https://doi.org/10.1016/S0378-7753(03)00190-3.

Wu, Jun, Kui Hu, Yiwei Cheng, Haiping Zhu, Xinyu Shao, and Yuanhang Wang. 2020. "Data-Driven Remaining Useful Life Prediction via Multiple Sensor Signals and Deep Long Short-Term Memory Neural Network." *ISA Transactions* 97: 241–50. https://doi.org/10.1016/j.isatra.2019.07.004.

Wyatt, John Lanier. 1978. *Foundations of Nonlinear Network Theory.*

Xiong, Rui, Linlin Li, and Jinpeng Tian. 2018. "Towards a Smarter Battery Management System: A Critical Review on Battery State of Health Monitoring Methods." *Journal of Power Sources* 405 (November): 18–29. https://doi.org/10.1016/J.JPOWSOUR.2018.10.019.

Yang, Duo, Xu Zhang, Rui Pan, Yujie Wang, and Zonghai Chen. 2018. "A Novel Gaussian Process Regression Model for State-of-Health Estimation of Lithium-Ion Battery Using Charging Curve." *Journal of Power Sources* 384. https://doi.org/10.1016/j.jpowsour.2018.03.015.

Yu, Jinsong, Weiqi Tang, Diyin Tang, and Jingjing Liu. 2018. "A Prediction Method for Discharge Voltage of Lithium-Ion Batteries under Unknown Dynamic Loads." *Microelectronics Reliability* 88–90 (May): 1206–11. https://doi.org/10.1016/j.microrel.2018.07.022.

Yuasa Battery Corporation. 1999. "NP Valve Regulated Lead Acid Battery Manual." www.yuasa-battery.co.uk.

———. 2006. "NPC Series VRLA Batteries Cyclic Application." www.yuasa-battery.co.uk.

Zelinsky, Michael A., John M. Koch, and Kwo Hsiung Young. 2017. "Performance Comparison of Rechargeable Batteries for Stationary Applications (Ni/MH vs. Ni–Cd and VRLA)." *Batteries* 4 (1): 1. https://doi.org/10.3390/batteries4010001.

Zhong, Zhifeng, Jianjun Tan, Tianjin Zhang, and Linlin Zhu. 2014. "PV Power Short-Term Forecasting Model Based on the Data Gathered from Monitoring Network." *China Communications* 11 (14): 61–69. https://doi.org/10.1109/CC.2014.7085385.

Zou, Yuan, Xiaosong Hu, Hongmin Ma, and Shengbo Eben Li. 2015. "Combined State of Charge and State of Health Estimation over Lithium-Ion Battery Cell Cycle Lifespan for Electric Vehicles." *Journal of Power Sources* 273 (January): 793–803. https://doi.org/10.1016/j.jpowsour.2014.09.146.