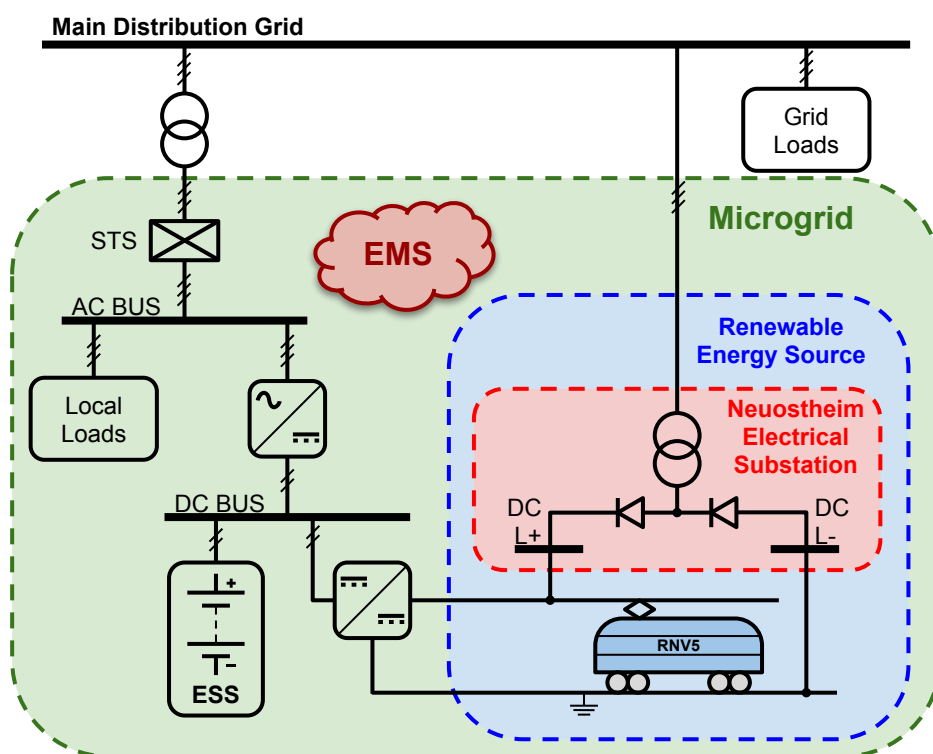


Master Thesis

A Robust Optimization-based Model Predictive Control of the energy management within Grid-connected Microgrids

cand. el. Roussineau Eduardo Esteban



Project manager: Dipl.-Ing. Philip Otto

Master Thesis

Mr. cand. el. Roussineau Eduardo Esteban

Matr.-No.: 1992036

A Robust Optimization-based Model Predictive Control of the energy management within Grid-connected Microgrids

The thesis is to be elaborated in close interaction with the institute and stays its property.
Any publication or distribution requires the institute's permission.

Topic issued on: 8th September, 2017

Thesis handed in on: 8th January, 2018

Supervisor:

Project manager:

(Prof. Dr.-Ing. Peter Gratzfeld)

(Dipl.-Ing. Philip Otto)

Author:

Address:

(cand. el. Roussineau Eduardo Esteban)

City 76139 Karlsruhe
Street Insterburger Str. 2 (Zi. 913)
Phone +49 157 8750 9845

Definition of task

for

Mr. cand. el. Roussineau Eduardo Esteban

(Matr.-No.: 1992036)

A Robust Optimization-based Model Predictive Control of the energy management within Grid-connected Microgrids

With the possibility of regenerative braking light rail trains can decelerate almost wearless. By integrating energy storage devices into a rail system a reduction of unused braking energy is obtained and a remarkable part of the braking energy can be further used. In many applications an economical standalone operation of these energy storage devices is not possible. Therefore, a higher degree of utilization is necessary. With the integration of the energy storage devices into a microgrid it offers storage capacity within the grid for further usage. The objective of this master thesis is the design of an operational strategy for a microgrid-connected wayside energy storage device within a light rail application. The operation should occur optimal to achieve high profit.

Declaration of Originality

I hereby formally declare that I have written the submitted thesis independently. I did not make use of any outside support except for the use of the quoted literature and the sources mentioned in the paper. Furthermore, I assure that all quotations and statements that have been inferred literally or in a general manner from published or unpublished writings are marked as such. I am aware that the violation of this regulation will lead to failure of the thesis.

Karlsruhe, 8th January, 2018

Acknowledgment

I would like to give thanks to my thesis tutor Philip Otto, who provided all the measurements used throughout the present work, as well as advice and guidance throughout the project.

I extend my gratitude to my master directors Cecilia Smoglie and Martin Gabi, for influencing and motivating me to join this master program and for helping along the way.

I thank my family for all the support and love provided regardless of the distance, and for having convinced me that this master's degree was the right step in my professional career.

I also thank my girlfriend for having understood and accompanied this decision.

Finally, I would like to dedicate this work to my grandfather. I am sure that we will meet again...

Abstract

With the possibility of regenerative braking light rail trains can decelerate almost wearless. By integrating energy storage devices into a rail system a reduction of unused braking energy is obtained and a remarkable part of the braking energy can be further used. In many applications an economical standalone operation of these energy storage devices is not possible. Therefore, a higher degree of utilization is necessary. The integration of the energy storage devices into a microgrid offers storage capacity within the microgrid for further usage.

This master thesis presents an operational strategy for a microgrid-connected wayside energy storage device within a light rail application. This strategy is based on a mixed-integer linear program optimization formulated over a rolling horizon framework; the optimization takes into account forecasted values for renewable energy generation and electricity usage within the microgrid, as well as the prices for energy exchange with the main distribution grid. Robust optimization theory is used to create a robust counterpart formulation of the optimization problem so that the uncertainties present in forecasted values are handled.

The operation should occur optimal to achieve high profit. Power profile shaping oriented goals provide the decision-maker with tools to meet performance sub-objectives related to the peak demand and profile smoothing. Simulations are conducted to demonstrate the effectiveness of various features of the proposed strategy in different scenarios.

The problem of how the short-term forecasts are created is also addressed. The lower upper bound estimation method is used to train artificial neural networks with the simulated annealing algorithm, so that this neural networks can be used to create prediction intervals for the different parameters that need to be forecasted. To demonstrate the effectiveness of the procedure, the prediction intervals forecasted for a electric load signal are compared with the actual realizations.

Zusammenfassung

Mit der Möglichkeit der regenerativen Bremsung können Straßenbahnen nahezu verschleißfrei abgebremst werden. Durch die Integration von Energiespeichern in Schienensysteme wird eine Reduzierung der ungenutzten Bremsenergie erzielt und ein beachtlicher Teil der Bremsenergie kann weiter genutzt werden. In vielen Anwendungen ist ein wirtschaftlicher Stand-alone-Betrieb dieser Energiespeicher nicht möglich. Daher ist ein höherer Nutzungsgrad erforderlich. Die Integration der Energiespeicher in ein Microgrid ermöglicht es, die Speicherkapazität besser auszunutzen.

In dieser Masterarbeit wird eine Betriebsstrategie für einen stationären Energiespeicher in einer Straßenbahnanwendung vorgestellt, der in ein Microgrid eingebunden ist. Diese Strategie basiert auf einer gemischt-ganzzahligen linearen Programmoptimierung, die über ein Rolling Horizon Framework formuliert wird. Die Optimierung berücksichtigt die prognostizierten Werte für die ungenutzte Bremsenergie, den Eigenbedarf an Energie im Microgrid, sowie die Preise für den Energieaustausch mit dem Hauptverteilungsnetz. Die robuste Optimierungstheorie wird verwendet, um die in den prognostizierten Werten vorhandenen Unsicherheiten zu berücksichtigen.

Um einen hohen Gewinn zu erzielen, soll der Betrieb möglichst optimal erfolgen. Anforderungen, die an die Anpassung der Leistungsprofile gestellt werden sind ein Werkzeug für den Entscheider, mit dem verschiedene Unteroptimierungskriterien, wie Spitzennachfrage und Profiglättung, erfüllt werden. Für die vorgeschlagene Strategie werden Simulationen durchgeführt, die die Wirksamkeit der verschiedenen Merkmale in unterschiedlichen Szenarien demonstrieren.

Die Problematiken bei der Erstellung von Kurzzeitprognosen werden ebenfalls diskutiert. Es werden künstliche neuronale Netze verwendet, um Prädiktionsintervalle für die verschiedenen Parameter zu erzeugen. Ein Schätzverfahren wird verwendet, um die künstlichen neuronale Netze mit dem Simulated-Annealing-Algorithmus zu trainieren. Um die Effektivität des Verfahrens zu demonstrieren, werden die für ein elektrisches Lastsignal die erstellten Vorhersageintervalle mit der tatsächlichen Realisierungen verglichen.

Resumen

La implementación de un sistema de frenado regenerativo permite que los trenes ligeros puedan desacelerar casi sin desgaste. Además, la integración de dispositivos de almacenamiento de energía en un sistema ferroviario permite reducir el porcentaje de la energía de frenado no utilizada; en muchas aplicaciones, no es posible una operación autónoma y rentable de estos dispositivos. Por lo tanto, es necesario que posean un mayor grado de utilización; la integración de los dispositivos de almacenamiento de energía en una *microgrid* ofrece capacidad de almacenamiento de energía dentro de la *microgrid* para su uso posterior.

Esta tesis de maestría presenta una estrategia operativa para un dispositivo de almacenamiento de energía dentro de una *microgrid*; el dispositivo opera dentro de un sistema de trenes ligeros y se encuentra instalado a la orilla de las vías férreas. La estrategia se basa en la optimización de un programa lineal de variables enteras y continuas formulado en un marco de horizonte móvil; la optimización toma en cuenta los pronósticos para los valores de regeneración de energía mediante el frenado de los trenes y uso de electricidad dentro de la *microgrid*, así como los precios para el intercambio de energía con la red principal de distribución. La teoría de optimización robusta es utilizada para crear una contraparte robusta del problema de optimización, de modo que las incertidumbres presentes en los valores pronosticados sean tenidas en cuenta.

La operación del sistema deberá ser óptima para obtener ganancias económicas. Objetivos orientados a la configuración de los perfiles de potencia brindan al responsable de la toma de decisiones herramientas para cumplir con sub-objetivos de rendimiento relacionados con la demanda máxima y el suavizado de los perfiles. Una serie de simulaciones demuestra la efectividad de varias características de la estrategia propuesta en diferentes escenarios.

El problema de cómo se crean los pronósticos a corto plazo es también abordado. Redes neuronales artificiales son utilizadas para crear intervalos de predicción para los diferentes parámetros que deben pronosticarse; estas redes son entrenadas mediante un método de estimación que utiliza el algoritmo de recocido simulado. Para demostrar la efectividad del procedimiento, los intervalos de predicción pronosticados para una señal de carga eléctrica se comparan con las realizaciones reales.

Contents

1	Introduction	1
1.1	Motivations	1
1.1.1	Challenges in conventional power grids	1
1.1.2	Diversification and integration of distributed energy sources	2
1.1.3	Microgrids	6
1.1.4	Dynamic braking as a renewable energy source (RES)	8
1.2	Objective	11
1.3	Outline	11
2	Control of systems integrating logic, dynamics, and constraints	13
2.1	Mathematical optimization	13
2.1.1	Applications	15
2.2	Linear programming (LP)	16
2.2.1	Mixed-integer linear programming (MILP)	17
2.2.2	MINLP to MILP conversion	17
2.3	Duality in mathematical optimization	19
2.3.1	The Lagrange dual function	20
2.3.2	The Lagrange dual problem	22
2.3.3	Weak and strong duality	22
2.4	Optimization under data uncertainty	23
2.4.1	Basic principles of robust optimization	26
2.4.2	Approaches for the implementation of robust optimization	26
2.4.3	Uncertainty set-induced robust optimization	29
2.5	Rolling horizon framework	38
2.5.1	Future horizon impact	40
2.5.2	Prediction horizon of varied duration	42
3	Optimal energy storage control strategy	43
3.1	Deterministic MILP optimization-based MPC	43
3.1.1	Objective function	43

3.1.2	Energy storage system (ESS)	46
3.1.3	Interaction between the microgrid and the main grid	50
3.1.4	Load balance	53
3.1.5	Deterministic MILP problem formulation	54
3.2	Robust optimization approach	54
3.2.1	Uncertain parameters in the model	55
3.2.2	Previous steps	55
3.2.3	Robust counterpart of the cost constraint	55
3.2.4	Robust counterpart of the load balance constraint	57
3.2.5	Robust counterpart of the load-shedding constraint	58
3.2.6	Robust MILP problem formulation	59
4	Implementation	61
4.1	Optimization model	61
4.2	Rolling horizon scheme	61
5	Results and discussion	65
5.1	Output profiles	65
5.2	Robustness of the solution	67
5.2.1	Feasibility against data uncertainty	67
5.2.2	Operation cost against data uncertainty	70
5.3	Power profile shaping goals	70
5.3.1	Grid power profile flattening	71
5.3.2	ESS power profile smoothing	72
5.4	Optimal sizing of the ESS	73
5.5	Computational load	76
5.5.1	Effect of uniform/variable time-steps and constraint relaxation	76
6	Background on rapid transit systems	79
6.1	Transferring energy to guided mass transport systems	79
6.1.1	Nominal voltage levels	80
6.1.2	Rectifying stations	81
6.1.3	Typical waveforms	83
6.2	Regenerative braking	83
6.2.1	Maximizing the usage of recovered braking energy	86
6.2.2	Techniques for estimating of the amount of energy recovered	90
6.2.3	Energy storage systems for urban rail applications	91
6.3	The Rhine-Neckar transportation system	93

6.3.1	Rolling stock operated by the RNV	94
6.3.2	Upper bound for the amount of energy recovered	94
6.3.3	Neuostheim electrical substation	97
6.4	Measurement analysis	97
6.4.1	Input data acquisition	97
6.4.2	Implications of the current measurement scheme	99
6.4.3	Regenerative braking action	100
6.4.4	Daily energy recovery profile	102
6.4.5	Over-current protection	104
6.4.6	Thevenin equivalent of the electrical substation	105
7	Short term forecasting in electric power systems using artificial neural networks	107
7.1	Introduction to forecasting	107
7.1.1	Classification	108
7.1.2	Dependencies of the signals	108
7.1.3	Models	109
7.2	Multilayer feed forward artificial neural networks (ANNs)	110
7.2.1	Artificial neurons	110
7.2.2	Network architecture	112
7.2.3	Network training	114
7.3	Prediction intervals (PIs) estimation	115
7.4	Lower upper bound estimation (LUBE) method	117
7.4.1	Measures for the quantitative assessment of PIs	118
7.4.2	Cost function: the coverage width-based criterion (CWC)	121
7.4.3	LUBE algorithm	123
7.4.4	Issues in implementing the NN LUBE method	127
7.5	Case study: electric load forecasting	132
7.5.1	Experiment methodology	132
7.5.2	Parameters	133
7.5.3	Data	134
7.5.4	Training process	134
7.5.5	Finding the best ANN architecture	137
7.5.6	Prediction interval construction	139
7.5.7	Computational load	141
7.6	Conclusions and further reading	141

8 Conclusion and prospect147

 8.1 Conclusion147

 8.2 Prospect149

A List of Figures151

B List of Tables157

C Bibliography159

Register of abbreviations and symbols

Sets

$\mathcal{B} \triangleq \mathbb{Z} \cap [0, 1]$	set of possible values for binary control variables
$\mathcal{C} \triangleq \mathbb{R} \cap [0, 1]$	set of possible values for coefficients
$\mathcal{H} \triangleq \mathbb{Z} \cap [1, N_h]$	set of N_h time-steps in the prediction horizon
\mathcal{U}	predefined uncertainty set for the random variables ξ
J_i	indices of variables with uncertain coefficients

Indices

i	problem constraint
j	optimization variable
k	rolling horizon time-step

Parameters

$H \in \mathbb{R}_+$	total amount of hours in the prediction horizon	h
$h \in \mathbb{R}_+$	duration of the time-step	h
$N_h \in \mathbb{N}$	number of time-steps in the prediction horizon	—
$\Delta p_{ss}^{\max} \in \mathbb{R}_+$	ESS allowed maximum instantaneous power-step	kW
$\varepsilon \in \mathcal{C}$	tolerance allowed for the deviation from E_{ss}^{final}	—
$\eta_c \in \mathcal{C}$	ESS charge efficiency	—
$\eta_d \in \mathcal{C}$	ESS discharge efficiency	—
$E_{ss}^{\max} \in \mathbb{R}_+$	maximum safety energy capacity	kW h
$E_{ss}^{\min} \in \mathbb{R}_+$	minimum safety energy capacity	kW h
$E_{ss}^{\text{nom}} \in \mathbb{R}_+$	nominal energy capacity	kW h
$p_{\text{grid}}^{\max} \in \mathbb{R}_+$	main grid maximum safe power exchange	kW
$p_{ss}^{\text{loss}} \in \mathbb{R}_+$	ESS self-discharge power loss	kW
$p_{ss}^{\max} \in \mathbb{R}_+$	ESS allowed maximum charge or discharge power	kW
$SOC^{\max} \in \mathcal{C}$	ESS state of charge maximum safety level	—
$SOC^{\min} \in \mathcal{C}$	ESS state of charge minimum safety level	—

$\Gamma_b \in \mathbb{R} \cap [1, N_h]$	polyhedral set size robust parameter	—
$\Gamma_{lb} \in \mathbb{R} \cap [1, 2]$	polyhedral set size robust parameter	—
$\Gamma_{sh} \in \mathbb{R} \cap [1, 3]$	polyhedral set size robust parameter	—
$\Gamma_s \in \mathbb{R} \cap [1, N_h]$	polyhedral set size robust parameter	—

Rolling horizon framework

$E_{ss} \in \mathbb{R}_+$	amount of energy present in the ESS	kWh
$P_{grid}^{base} \in \mathbb{R}_+$	main grid power usage baseline	kW
$P_g \in \mathbb{R}_+$	average power output of the RES present in the microgrid	kW
$P_l \in \mathbb{R}_+$	average power demand for all consumers inside the microgrid	kW
$c_{grid}^{flat} \in \mathbb{R}_+$	penalization cost for main grid power profile flattening	ct/kWh
$c_{grid}^{peak} \in \mathbb{R}_+$	penalization cost for main grid power usage above baseline	ct/kWh
$c_{grid}^{sm} \in \mathbb{R}_+$	penalization cost for main grid power profile smoothing	ct/kWh
$c_{sh} \in \mathbb{R}_+$	penalization cost for load shedding	ct/kWh
$c_{ss}^{sm} \in \mathbb{R}_+$	penalization cost for ESS power profile smoothing	ct/kWh
$c_{ss} \in \mathbb{R}_+$	ESS usage cost	ct/kWh
$c_b \in \mathbb{R}_+$	cost of buying electricity	ct/kWh
$c_s \in \mathbb{R}_+$	cost of selling electricity	ct/kWh
$E_{ss}^{final} \in \mathbb{R}_+$	energy level in the ESS at the end of the horizon	kWh
$E_{ss}^o \in \mathbb{R}_+$	energy level in the ESS at the beginning of the horizon	kWh
$P_b^o \in \mathbb{R}_+$	power bought from the main grid in the previous control horizon	kW
$P_c^o \in \mathbb{R}_+$	power charged to the ESS in the previous control horizon	kW
$P_d^o \in \mathbb{R}_+$	power discharged to the ESS in the previous control horizon	kW
$P_s^o \in \mathbb{R}_+$	power sold to the main grid in the previous control horizon	kW

Optimization variables

$p_b^{max} \in \mathbb{R}_+$	main grid power buying flattening	kW
$p_{grid}^{ob} \in \mathbb{R}_+$	main grid power usage above baseline	kW
$p_{sh} \in \mathbb{R}_+$	load disconnected from the electricity supply	kW
$p_s^{max} \in \mathbb{R}_+$	main grid power selling flattening	kW
$p_b \in \mathbb{R}_+$	power bought from the main grid	kW
$p_c \in \mathbb{R}_+$	power charged to the ESS	kW
$p_d \in \mathbb{R}_+$	power discharged from the ESS	kW
$p_s \in \mathbb{R}_+$	power sold to the main grid	kW
$u_{grid} \in \mathbb{R}_+$	main grid power profile smoothing	kW

$u_{ss} \in \mathbb{R}_+$	ESS power profile smoothing	kW
$\delta_b \in \mathcal{B}$	control variable for power bought from the main grid	—
$\delta_c \in \mathcal{B}$	control variable for power charged to the ESS	—
$\delta_d \in \mathcal{B}$	control variable for power discharged from the ESS	—
$\delta_s \in \mathcal{B}$	control variable for power sold to the main grid	—
$y_b \in \mathbb{R}_+$	linearization variable for the power bought from the main grid	kW
$y_c \in \mathbb{R}_+$	linearization variable for the power charged to the ESS	kW
$y_d \in \mathbb{R}_+$	linearization variable for the power discharged from the ESS	kW
$y_s \in \mathbb{R}_+$	linearization variable for the power sold to the main grid	kW
$q \in \mathbb{R}$	robust counterpart auxiliary variable for the objective function	ct/kWh
$w_b \in \mathbb{R}_+$	robust counterpart auxiliary variable for the objective function	kW
$w_g \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$w_l \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$w_s \in \mathbb{R}_+$	robust counterpart auxiliary variable for the objective function	kW
$w_{lb} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$w_{lc} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$w_{ls} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$z_b \in \mathbb{R}_+$	robust counterpart auxiliary variable for the objective function	kW
$z_s \in \mathbb{R}_+$	robust counterpart auxiliary variable for the objective function	kW
$z_{lb} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$z_{shb} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$z_{shc} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW
$z_{shs} \in \mathbb{R}_+$	robust counterpart auxiliary variable for the load balance	kW

Abbreviations

ANN	artificial neural network
API	application programming interface
CI	confidence interval
EMS	energy management system
ESS	energy storage system
GHG	greenhouse gas
HVAC	heating, ventilation and air conditioning
LP	linear program

LUBE	lower upper bound estimation
MILP	mixed-integer linear program
MINLP	mixed-integer nonlinear program
MPC	model predictive control
PI	prediction interval
RES	renewable energy source
RNV	Rhein-Neckar-Verkehr GmbH
RO	robust optimization
SO	stochastic optimization
SOC	state of charge
STS	static transfer switch
TSO	transmission system operator
WESS	wayside energy storage system

Designation principles

\hat{a}_{ij}	maximum positive perturbation amplitude for the uncertain parameter
\tilde{a}_{ij}	actual value of the parameters which are possibly subject to uncertainty
a_{ij}	nominal value for the uncertain parameter
ξ	independent random variables in the range $[-1, 1]$

1 Introduction

The research of the present Master thesis consists in the methodological development of modeling and optimization solutions for the energy management problem for grid-connected microgrids with on-site energy storage systems (ESS) under the presence of uncertainty in electricity prices, predicted load and predicted renewable power generation, and in particular the deployment of these solutions for the control of a wayside ESS in a light rail system.

The introduction to this research is presented below as follows. Section 1.1 introduces the main motivations that drive this research; the problem statement and objectives for this work are defined in section 1.2; and finally, the outline of the remainder of the thesis is presented in section 1.3.

1.1 Motivations

1.1.1 Challenges in conventional power grids

The current global *electricity* generation remains highly dependent on fossil fuels, up to a proportion of 60%^{1,2} [159], mainly due to the existence of an uninterrupted and inexpensive flow of resources. This excessive consumption of fossil fuels is aggravating the effects of climate change with 25% of the global GHG emissions³ [117]. Bearing in mind that an increase of 58% in global power demand is foreseen for the year 2040⁴ [36], the use of non-renewable resources as the main energy suppliers is not sustainable over time.

The highly hierarchical architecture present in the current energy system achieves actors with clearly defined purposes on the energy market, thus facilitating system monitoring, fault detection and correction; however, one of the major vulnerabilities of this topology

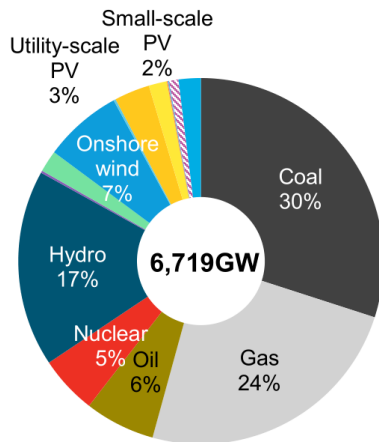
¹ Accounts for Oil (6%), Gas (24%) and Coal (30%) in 2016 (see figure 1.1).

² On the other hand, the current global *primary energy* share of fossil fuels is 85% (32% Oil, 24% Gas and 29% Coal in 2015) [40]. However, the use of primary energy as a metric (as done by BP's World Energy Outlook in [40]) is not right, because it ignores the conversion efficiency to electricity. Thus forms of energy with poor conversion efficiency, particularly the thermal sources, coal, gas and nuclear are overstated, whereas the impact of energy sources such as hydro, solar, wind, biomass and geothermal which are converted efficiently is understated.

³ For electricity and heat generation, based on global emissions from 2010.

⁴ With respect to the year 2017.

**Global cumulative installed capacity:
2016**



**Global cumulative installed capacity:
2040**

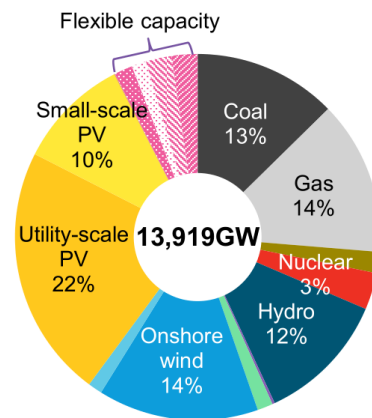


Figure 1.1: Global cumulative installed capacity, 2016 (real) and 2040 (predicted). Flexible capacity includes small-scale batteries, utility-scale batteries, demand response and other flexible capacity. Source: [159].

is the cascading failure propagation that may occur in the absence of bypass transmission [146].

1.1.2 Diversification and integration of distributed energy sources

The energetic and environmental challenges of the 21st century drive the development and adoption of new energy sources that offer sensible and sustainable alternatives to fossil fuels⁵, while helping overcome resource scarcity and meet the increasing demand; these new fostered renewable energy supplies must ideally be low carbon, environmentally sound, reliable, and affordable.

This search can be appreciated in the raw numbers of the new investments in *clean energy*⁶ and capacity installations depicted in figure 1.2. For the last seven years, the annual investment has remained constant around a third of a trillion dollars as the installed capacity has continued to grow steadily over the years⁷. This antagonism is mainly explained

⁵ Global coal-fired power generation is set to peak in 2026; growth in coal demand is centered on Asia, but is offset by sharp declines in Europe and the U.S. [36].

⁶ This broad term accounts for renewable energies, energy efficiency, power storage, smart grids and so on.

⁷ Nevertheless, it is estimated that 86% of the total world new investment in power generation capacity between 2017 and 2040 will be for zero-carbon technologies, while the remaining 14% will be directed to fossil fuel technologies [159].

by the drop in the technology costs⁸, due to the trends of the *technological learning*^{9,10} caused by innovation.

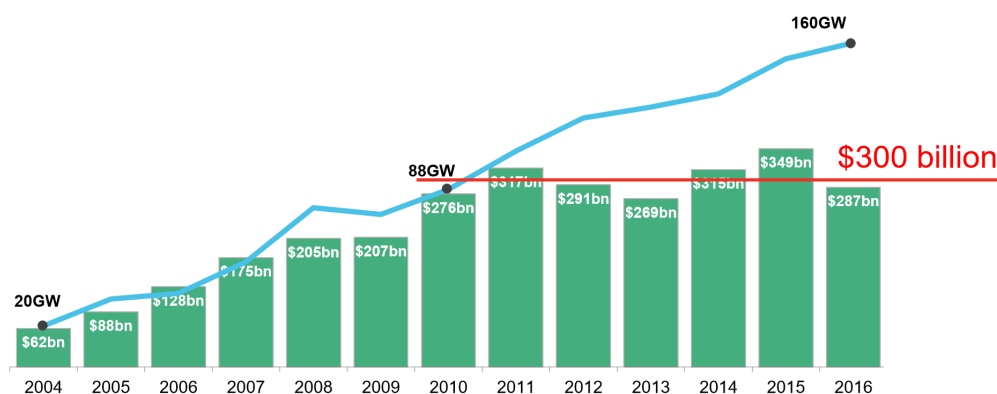


Figure 1.2: Global new clean energy investments and capacity installations. Total values include estimates for undisclosed deals. Includes corporate and government R&D, and spending for digital energy and energy storage projects (not reported in quarterly statistics). Excludes large hydro. Source: [159].

The most widely used renewable energies today can be divided into two major groups [146]: hydropower systems and other renewable energies, e.g. solar, wind, geothermal, biomass and ocean (in majority wave and tidal power) energy production systems. Figure 1.3 depicts the evolution over time of the penetration of these renewable energy sources into the electricity system, with the highest recorded peaks shown in figure 1.4.

The total renewable share of world electricity generation represented 35% in the year 2016¹¹ with a increase share prospect of 94% by the year 2040¹². As shown in figure 1.5, a prediction of 50% renewables penetration is made for Europe by 2040 [159].

Since 2010, Germany has passed legislation to support the *energy transition*¹³ to a system that heavily relies on renewable energy sources, energy efficiency and demand side management; this transition contemplates the phase-out of all nuclear reactors of the

⁸ There exist two important *tipping points* in comparing costs of installed capacity within a certain country or region; the first one is when the cost of a new capacity installation of renewable (e.g. solar, wind) is lower than the cost of a new installation of a non-renewable (e.g. coal, combined cycle gas turbine); the second one is when the cost of a new capacity installation of renewable is lower than the operational cost of a pre-existent non-renewable capacity installed. Countries like U.S. and Germany have already passed the first point; China and Germany are predicted to reach the second point sometime around 2030. For more information refer to [159].

⁹ Technological learning, learning curve, progress curve, experience curve, or learning by doing refers to the cost reductions as technology manufacturers double their accumulated experience, or installed capacity in this context, over time [168].

¹⁰ Estimated learning rates according to [159]: wind (19%), solar (24-28%), Lithium-ion batteries (19%).

¹¹ Accounts for hydropower (17%) and other renewables (18%).

¹² Accounts for a 30% decrease of the hydropower share and a 211% increase of other renewable energies.

¹³ *Energiewende* in German.

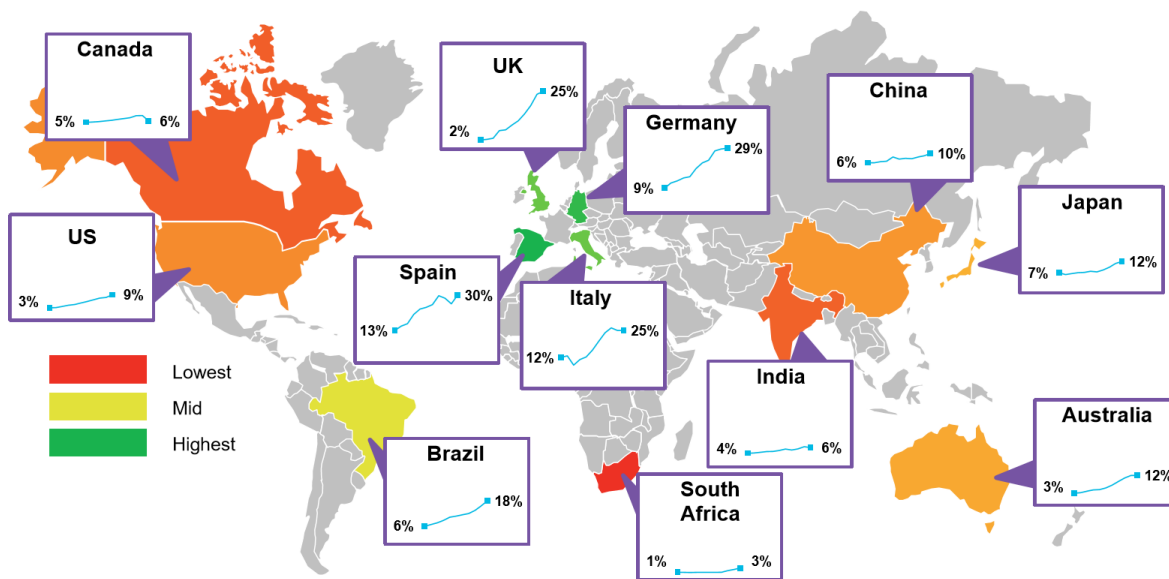


Figure 1.3: Growth of renewable energy proportion of power generation between 2006 and 2016 (excludes large hydro). Source: [159].

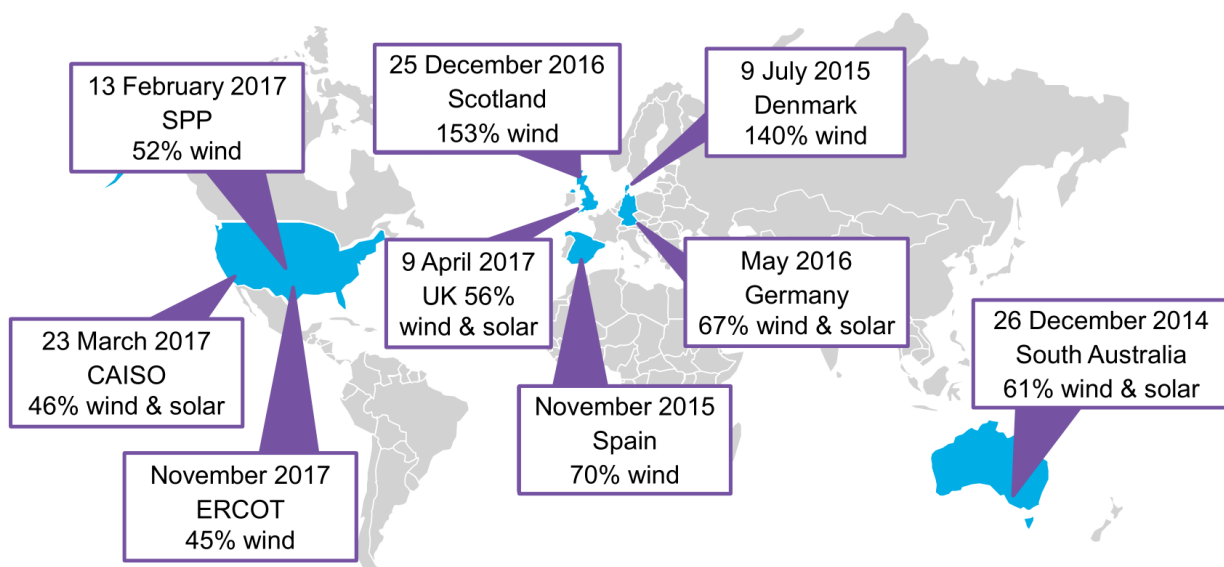


Figure 1.4: Recorded peak renewable penetration by country. Source: [159].

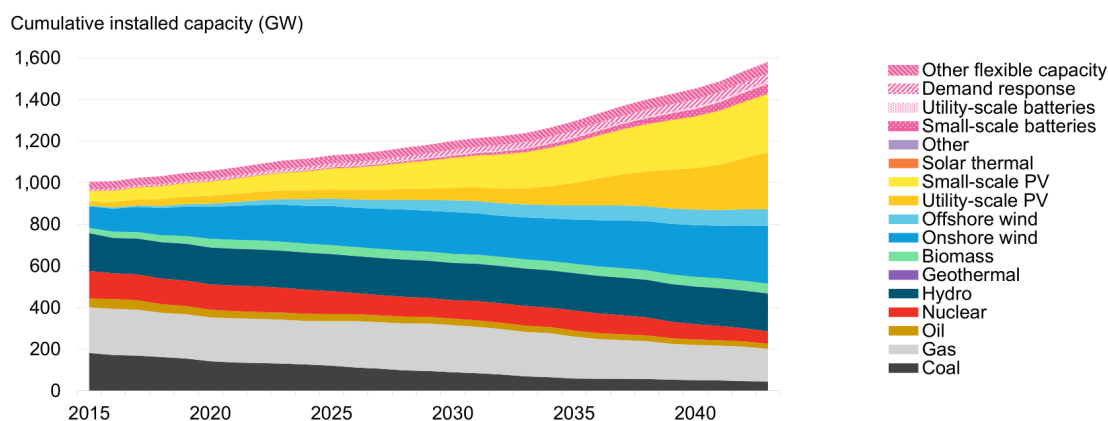


Figure 1.5: Energy Outlook for Europe. A prediction of 50% renewables penetration is made for Europe by 2040. Source: [159].

country by 2022 and the retirement of most (if not all) of the coal-fired generation facilities. Moreover, the legislation includes GHG reductions of 80-95% by 2050¹⁴ and a renewable energy target of 60% by 2050¹⁵ [217]. As exposed in [9]:

"Further market opportunities also come through innovation and cost reduction such as through big data, blockchain technologies and with the coupling of the heat and mobility sectors."

This national effort over the years has made Germany a global leading market in the area of electricity generation through renewable sources.

Nonetheless, the exploitation of renewable energy sources poses its own new technological challenges with respect to conventional transmission and distribution grids. Constraints on generation are imposed due to the geographical availability of these resources, while their intermittent nature (mainly due to local weather conditions and seasonal changes) also creates concerns regarding reliability, in contrast to the 'always-on' fossil or nuclear generation [146]. This intermittent characteristic will hinder the efficient and reliable management of the power grid in scenarios where large-scale integration of renewable energy sources is achieved.

Some of the challenges to consider for the efficient and reliable management of a grid with a high level of renewable energies penetration include, but are not limited to, the impossibility of meeting the base-load and peak demand without high-capacity energy storage systems due to the time lag between the peaks of consumption and generation present in energy sources such as solar and wind; errors present in the energy generation

¹⁴ Relative to 1990.

¹⁵ By October 2017, the country has already surpassed the 44% mark [9].

forecasts created by the current forecasting models; and decentralization of the generation and appearance of actors that are both producers and consumers¹⁶, creating a high number of nodes with bidirectional power flow, in contrast to the hierarchical system previously described. These challenges call for a large margin of backup electricity generation capacity (usually obtained with fossil fuels or nuclear reactors) in order to ensure power demand [145], which in turn increases the cost of electricity.

According to [145], the current paradigm to address these challenges is

"[...] the implementation of new emerging intelligent infrastructures, i.e., Smart grids, which are expected to improve coordination of energy generation by diverse energy sources¹⁷, transmission and distribution efficiency to meet increasing demand, protection and resilience against vulnerabilities of aging and failing components, natural disasters and human attacks."

In this regard, microgrids with energy management systems (EMS) are considered as an effective framework for the penetration of distributed energy generation into the main grid network [127].

1.1.3 Microgrids

Figure 1.6 depicts the main components of a grid-connected microgrid with EMS; Khodabakhsh provides a concise definition of a grid-connected microgrid:

"A microgrid is a small electric grid system which could include a mixture of distributed energy sources (e.g., wind turbines, solar panels, fuel cells, and microturbines), loads (including controllable loads as HVAC unit) and storage devices (e.g., batteries, ultra-capacitors and flywheels) as well as a control unit. The control unit is responsible for operating the microgrid in an efficient way with the aim of reducing the cost of electricity for the consumer(s). It essentially determines charge/discharge activities of the storage devices, controls the HVAC unit, and may also help the grid with peak reduction, load shifting etc." [127].

The EMS usually makes optimal decisions with respect to the charge/discharge profiles over time of the energy storage devices and the amount of power dispatched to every load in the system over time; these actions are taken considering factors such as predicted energy demand, predicted energy generation from renewable sources, weather forecast,

¹⁶ Usually referred to as *prosumers*.

¹⁷ Of which many were not taken into account in the original power infrastructure design and whose wide penetration could affect in unforeseen ways the operation of the power system [145].

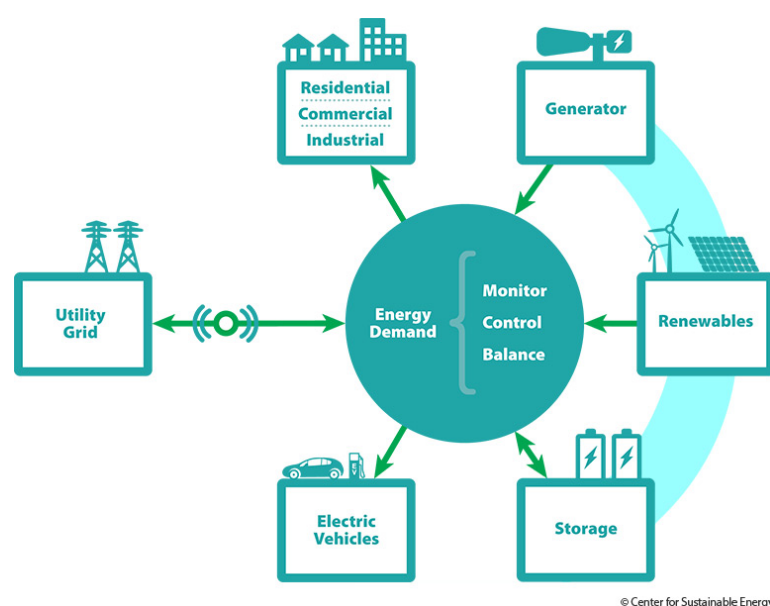


Figure 1.6: Architecture of a grid-connected microgrid showing its main components. Source: [48]

and electricity pricing. These optimal decisions can be achieved by formulating and solving relevant constrained multi-objective optimization problems, either off-line or on-line. Figure 1.7 depicts a EMS with its input parameters and output decisions.

All the predictions and forecasts used as input parameters for the optimization carried out by the EMS (i.e., values of demand, generation, and prices) can present a considerable amount of uncertainty. In consequence, the solution obtained through the optimization must be robust against all possible realizations of these prediction errors; if unaccounted, these uncertainties may produce suboptimal solutions for the optimization problem.

Thus, the goal of the present work is to use modern optimization and modeling tools to design a controller that can efficiently operate a microgrid and account for the uncertainties in the system.

The proposed controller performs as a high-level power optimizer in a hierarchical control structure by providing power profiles to the power converters of all the devices integrating the microgrid at a time scale in the order of minutes [127]. A low-level power control system is additionally needed operating at a much faster time scale to enforce these power commands, and also handle voltage and frequency regulations, as required. A block diagram of the proposed controller is presented in a model predictive control (MPC) framework in figure 1.7.

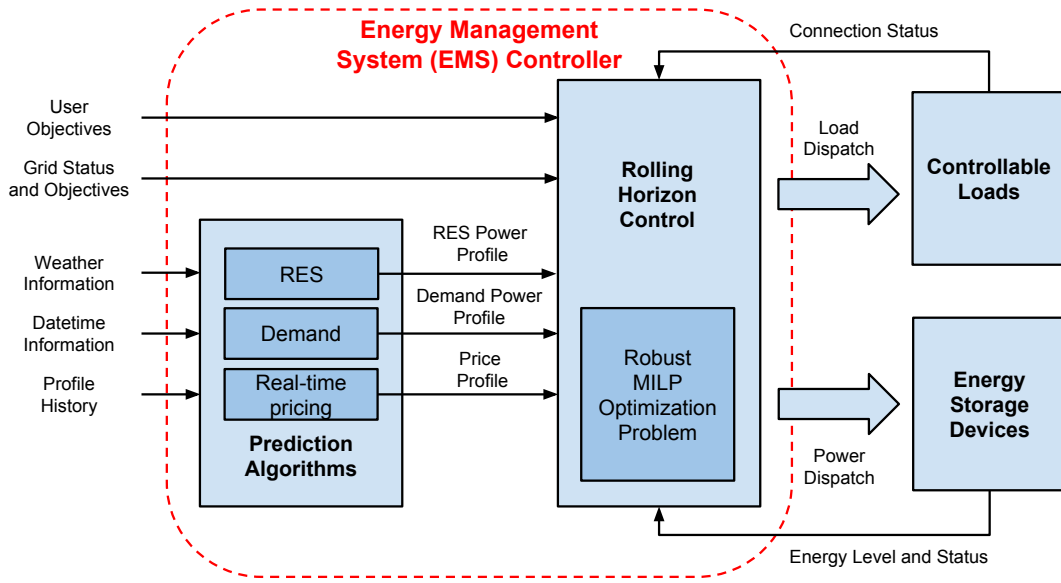


Figure 1.7: Energy Management System (EMS) Controller.

1.1.4 Dynamic braking as a renewable energy source (RES)

Pollution levels and traffic congestion related to ever-growing population levels of major cities are becoming delicate issues that could be eased by more efficient public transportation systems. Many cities have begun to incorporate electric vehicles into their transport networks to reduce their emission levels; according to the available infrastructure and the number of commuters several options can be realized (e.g., battery electric buses, trolley buses, trams, metro, light rail) [22]. Even though these mass transit vehicles allow large reductions in terms of emissions, their energy efficiency is far from optimal [22]. With electric vehicles came the implementation of dynamic braking technology¹⁸, which is nowadays widely spread in railway technology¹⁹. However, the energy recovered by the generators during the braking phase (which can reach levels up to 40% of the energy supplied to the electrical rail guided vehicles [22]) is, in most cases, dissipated as heat through rheostatic braking. The main reason for this energy waste relies on the low network receptivity when the energy is being harvested. Thus, the improvement in energy efficiency can be reached by increasing the network receptivity; one way of achieving the latter is by the implementation of wayside ESS's²⁰ which help prevent the loss of the recovered energy.

¹⁸ Dynamic braking refers to the process that converts kinetic energy into electricity as the vehicle decelerates, and it is based on the capacity the electric motors have to act as generators.

¹⁹ Mainly because it helps to limit the use of friction braking, a process that carries further investments associated with maintenance and replacement of parts subject to wear.

²⁰ Stationary storage devices installed in electrical substations or along the track.

A great deal of literature has been dedicated to the optimal design and implementation of this (and other) methodologies to store and reuse the energy recovered by the braking-phase of electric vehicles in mass transport systems²¹. However, the vast majority of the research focuses on reusing the energy within the train electric network (i.e., peak power shaving and line voltage stabilization of the overhead line).

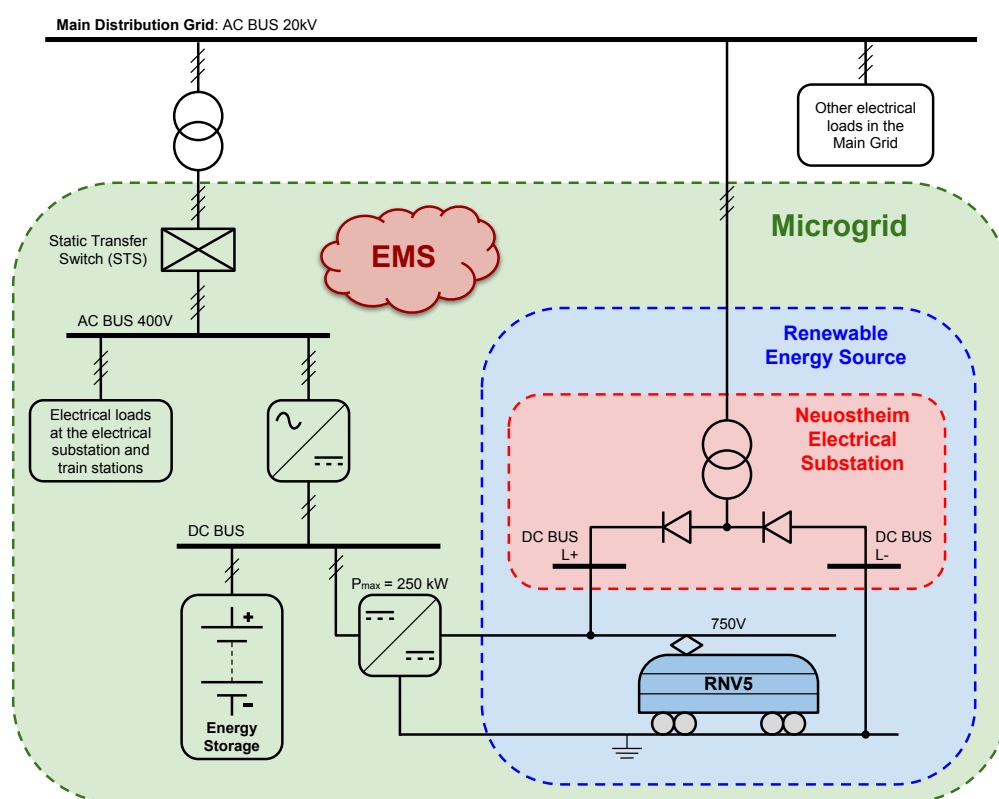


Figure 1.8: Microgrid definition for the present work.

A novel approach is to consider the recovered energy from the braking phase as the renewable energy source (RES) of a grid-connected microgrid. Figure 1.8 depicts the architecture of the proposed microgrid, where a wayside ESS is installed inside the Neuostheim electrical substation belonging to the Rhine-Neckar Transportation System in Germany²²; the energy generated by the trains arriving to the passenger stations in the vicinity of the electrical substation is recollect and stored in this ESS. The load of the microgrid is defined as the electrical load (e.g. heating and cooling, lighting, signaling, passengers information equipment, etc.) present at the electrical substation and the passenger stations in the vicinity of the substation and the ESS.

²¹ Refer to section 6.2 for a literature review on this topic.

²² Finding which is the optimal substation in the system to install the ESS (e.g. to maximize the amount of energy recovered) is a problem that remains outside the scope of the present work.

For every time-step, the EMS depicted in figure 1.7 controls the amount of energy charged or discharged from the ESS and the amount of energy exchanged with the main grid so that a series of objectives are met (e.g., minimization of operation costs, shaping of the main grid power signal, shaping of the ESS power signal, energy balance to meet the demand, amount of load-shedding in islanded mode, etc.). To find the optimal values of these decision variables, the optimization problem takes into account a range of parameters (e.g. physical limitations of the system concerning energy and power rates and capacities) and forecasts within a predefined prediction horizon (e.g. energy generation, load and electricity prices).

The *amount* of power/energy produced by this RES and *when* it is produced can not be controlled by the system; thus, this type of RES is said to be *non-dispatchable*²³. In addition to not being controllable, the energy generated at each particular time-step is not fully observable either; like any other RES, regenerative braking presents several sources of uncertainty²⁴.

Optimal solutions of optimization problems usually present a high sensitivity to data perturbation; solving the optimization problem under the assumption that all parameters are deterministic (thus ignoring possible uncertainties in the data) could cause the solution of the problem to be suboptimal or unfeasible in the worst case [157]. As a result, decision making inherently involves the consideration of such uncertainties; with this purpose, high-confidence-forecasting models for the amount of energy generated at any given point in time by the RES can be created²⁵.

Sensors installed in the Neuostheim electrical substation are used to measure the voltage and current levels delivered to the trains that travel along the four track segments connected to said substation. These data allow to calculate a percentage of the energy generated by the regenerative braking²⁶, and this history of measurements allows to generate the prediction models used by the real-time EMS controller to estimate the amount of energy generated by the RES at each time-step.

²³ On the other hand, fuel generators and Combined Heat and Power (CHP) plants are classified as *dispatchable* units, because the *amount* of power/energy and *when* it is produced can be controlled by the system.

²⁴ Main sources of uncertainty in the energy estimation include, but are not limited to, variations in train schedules, type of train used (electric motors and power electronics topology implemented), amount of passengers inside the train, variation of the driving style between different drivers, weather variables that influence the electric consumption inside the train (AC and heating systems).

²⁵ Refer to chapter 7 for more details.

²⁶ Refer to chapter 6 for more details.

1.2 Objective

The research of the present Master thesis aims at the methodological development of modeling and optimization solutions for the energy management problem for grid-connected microgrids with on-site energy storage systems under the presence of uncertainty in electricity prices²⁷, predicted load and predicted renewable power generation.

This paradigm is deployed to design a high-level control strategy, or energy management system, for an ESS in a light rail system; for this purpose, the Neuostheim electrical substation that belongs to the Rhine-Neckar Transportation system in Germany is used as a case study. The installation of an ESS together with the upgrade of the substation to a reversible substation is assumed and simulated, and a microgrid composed of the substation, the ESS, and the passenger stations powered by the substations is defined as depicted in figure 1.8.

For each time-step, the EMS as depicted in figure 1.7 optimizes the charging and discharging profile of the ESS for a predefined number of future time-steps so that the economic cost of operation of the system is minimized. This economic cost is the aggregated sum of the operational cost of several processes inside the microgrid, together with penalization terms that allow the decision-maker to *shape* the solution of the problem.

It is noted that this thesis is only concerned with the high-level power scheduling control to ensure a certain degree of robustness to uncertainties in the energy generation, demand and electricity prices; the low-level controllers that enforce these power commands and handle voltage and frequency regulations and protection system remain out of the scope of this dissertation.

1.3 Outline

The remainder of this work is organized as follows. Chapter 2 presents an overview of the theoretical background in mathematics and control systems needed to construct the optimization problem and the rolling horizon controller in which the optimization is embedded. The mixed-integer linear programming (MILP) framework is introduced as a means to model optimization problems; robust optimization (RO) is then presented to make these MILP formulations robust against sources of uncertainty in the optimization parameters, making the MILP formulation viable for practical applications; finally, the rolling-horizon framework is presented to embed the robust MILP formulation in a time horizon.

²⁷ Electricity price forecasts are mainly based on statistical analysis of historical market prices. However, actual electricity market prices are influenced by many factors including demand and supply variations. Consequently, the actual prices may differ from the predicted values.

The theory is applied in chapter 3 to formulate two mathematical models (that is, a deterministic MILP-based rolling horizon controller and its robust counterpart) that solve the problem stated in section 1.2 regarding the robust optimal control of the energy flows in a grid-connected microgrid.

The implementation of the models for the EMS controller is discussed in chapter 4; the rolling horizon controller is developed using MATLAB, which links to the MILP formulation created in AMPL and solved with the IBM ILOG CPLEX solver.

The most significant results derived from the implementation are presented in chapter 5; several graphs of interest showing the features and effectiveness of the controllers are analyzed, making a comparison of the two models realized.

Chapter 6 gives an introductory explanation of the railroad system, focusing on the tramway system being used to perform the measurements used throughout this work. Processes and equipments involved in delivering energy to the vehicles on the tracks are discussed, together with different approaches for introducing energy savings through regenerative braking. The voltage and current measurements taken at the electrical substation are analyzed and some signal processing is applied to estimate the amount of energy recovered with regenerative braking throughout the day.

Chapter 7 makes an introduction to the data forecast problem, and one of the most widely used forecasting methods is briefly explained, i.e., the use of artificial neural networks (ANNs). The implementation of the lower upper bound estimation (LUBE) method using ANNs for the generation of interval forecasts is analyzed and, as a proof of concept, an example is implemented for the generation of interval forecasts for the amount of electrical load present in a system.

Finally, some conclusions and closing remarks are presented in chapter 8, together with possible directions for future work.

2 Control of systems integrating logic, dynamics, and constraints

This chapter presents an overview of the theoretical background in mathematics and control systems needed to construct the models developed in chapter 3.

The following sections are organized as follows: section 2.1 introduces, with the work done by Boyd and Vandenberghe in [39] as a main reference, the general definition of mathematical optimization together with some examples of real applications; section 2.2 focuses on linear programs, which are a special class of mathematical optimization problems. Problem realizations with continuous, discrete and hybrid continuous/discrete variables are discussed. Some simple linearization techniques for bilinear problems are discussed as well; section 2.3 talks about the Lagrange dual problem, which will allow to obtain a lower bound on the solution of an optimization problem and, if strong duality holds, prove that the bound is tight, i.e., the bound equals the solution of the original optimization problem; in section 2.4 data uncertainty is taken into account when designing the optimization model, obtaining in consequence a non-deterministic optimization problem. To address this issue, the robust optimization methodology is introduced and implemented together with the results from section 2.4 to create a robust counterpart of the original optimization problem, which will allow to obtain solutions that remain optimal under all the possible realizations of data uncertainty predefined in an uncertainty set included in the uncertain space; finally, section 2.5 explains how the MILP problem formulation can be embedded into an integrated receding horizon framework involving problem-specific forecasting techniques (for the uncertain parameters) to deal with the control problem when the cost function has to be minimized in a given time horizon and the system is subject to constraints.

2.1 Mathematical optimization

A *mathematical optimization problem*, or just *optimization problem*, has, by convention, the following standard form¹ [39]:

¹ There are different standard forms in use in the scientific community. The one presented in this work is used by Boyd and Vandenberghe in [39]. By using different mathematical transformations it is possible to convert the problem from one standard form to another; for example, the implementation of the so-called

$$\begin{aligned}
& \underset{x}{\text{minimize}} && f_0(x) \\
& \text{subject to} && f_i(x) \leq b_i, \quad i = 1, \dots, m.
\end{aligned} \tag{2.1}$$

where:

- the vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ contains the *optimization variables* of the problem. This is a set of variables that represent actions that can be taken in the system being modeled;
- the function $f_0 : \mathbb{R}^n \mapsto \mathbb{R}$ is the *objective* or *cost* function, which maps each possible set of decisions into a single score (often representing a total economic cost incurred or revenue gained) that assesses the quality of the solution. Thus, this function is optimized by either finding its minimum or maximum value;
- the functions $f_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \dots, m$, are the (inequality) *constraint functions*. These restricting functions represent the limitations present in the system being modeled. They can be set to be equal to, not more than, or not less than, a certain numeric value;
- the constants b_1, \dots, b_m are the limits, or bounds, for the constraints.

As Boyd and Vandenberghe expose in [39], the optimization problem 2.1 is:

"[...] an abstraction of the problem of making the best possible choice of a vector in \mathbb{R}^n from a set of candidate choices. The variable x represents the choice made; the constraints $f_i(x) \leq b_i$ represent firm requirements or specifications that limit the possible choices, and the objective value $f_0(x)$ represents the cost² of choosing x . A solution of the optimization problem corresponds to a choice that has minimum cost (or maximum utility), among all choices that meet the firm requirements."

A solution that satisfies all constraints is called a *feasible solution*; of all the feasible solutions, the one that maps to the best³ objective function value is referred to as *optimal*

slack variables allows for the transformation of an inequality constraint into an equality constraint and a nonnegativity constraint [39]. Sometimes this is important, since certain resolution algorithms require as initial condition that the optimization problem is expressed in a certain standard form; one example is the Simplex method for solving linear optimization problems [174].

² $-f_0(x)$ can also be seen as representing the value, or utility, of choosing x .

³ The definition of *best* depends on whether the problem is a minimization or a maximization of the objective function.

solution [61], i.e., a vector x^* is called *optimal* in a minimization problem if for any $z \in \mathbb{R}^n$ with $f_1(z) \leq b_1, \dots, f_m(z) \leq b_m$, it holds that $f_0(z) \geq f_0(x^*)$ [39].

Sometimes no solution exists to the optimization problem, and the problem itself is defined as *infeasible*⁴. On the other hand, some feasible problems have no optimal solution, because there is an infinite number of feasible solutions that map to the *best* objective function value. These problems are defined as *unbounded*⁵.

In general, classes of optimization problems are considered; each problem class is characterized by a particular form of the objective and constraint functions [39]. As an important example, the optimization problem 2.1 is called a *linear program* if the objective and constraint functions (f_0, \dots, f_m) are linear, i.e., satisfy

$$f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y), \quad (2.2)$$

for all $x, y \in \mathbb{R}^n$ and all $\alpha, \beta \in \mathbb{R}$. If the optimization problem is not linear, then it can be classified as a *nonlinear program*.

2.1.1 Applications

Over the years, mathematical optimization has become an important tool within a large number of different disciplines that have been able to adapt their practical problems to be cast in the form of a mathematical optimization problem⁶. Usually these problems include decision making, system design, analysis or operation processes.

It is widely used in engineering, in electronic design automation [237, 152, 123, 173, 144], automatic control systems [20, 65, 231, 41], energy savings [66, 75, 176, 218, 186], and optimal design problems arising in civil [77, 198, 43, 42, 12], chemical [84, 45, 104, 86], mechanical [192, 193, 191, 37, 87], and aerospace engineering [184, 46, 222, 141]. Optimization is also used for problems arising in network design and operation [10, 235, 38, 185, 90, 17, 94], finance [169, 229, 93, 230], supply chain management [161, 122, 35, 91, 200], scheduling [56, 179, 211, 232, 172], and many other areas. The list of applications is still steadily expanding.

Mathematical optimization is a tool that can either be used as an aid to a human in charge of the decision making (or system operation), or embedded in a control system that automatically makes real-time choices and carries out the associated actions [39]. In the former case, the process is supervised by a human, who checks the optimization results

⁴ In this case, $x^* = \infty$. This means that there is no $x \in \mathbb{R}^n$ that satisfies the constraints of the problem [39].

⁵ If the problem is unbounded below, it holds that $f_0(x^*) = -\infty$. This means that there is a sequence of $x \in \mathbb{R}^n$ values, all feasible, with the cost function being $f_0(x) = -\infty$ [39].

⁶ Or some variation, such as multicriterion optimization problems.

and modifies the problem definition (or the solution approach) when necessary; he is also in charge of carrying out any action suggested by the optimization problem. On the other hand, *embedded optimization* presents no (or little) human intervention or oversight.

2.2 Linear programming (LP)

Linear programming is one of the simplest classes of optimization problems there is, in the sense that all of the functions involved (objective and constraints) are *affine*⁷. One of the many standard forms for the general definition of this class is presented as:

$$\begin{aligned} & \underset{x}{\text{minimize}} && c^T x \\ & \text{subject to} && a_i^T x \leq b_i, \quad i = 1, \dots, m. \end{aligned} \tag{2.3}$$

where the vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ contains the n optimization variables of the problem to be determined. The values in vector $c \in \mathbb{R}^n$ are referred to as *objective coefficients*, and are often linked with the costs (for minimization problems) or revenues (for maximization problems) derived from the corresponding decisions [61]. The scalars $b_1, \dots, b_m \in \mathbb{R}$ are the right-hand-side values of the constraints, and often represent amounts of available resources (especially for \leq constraints) or requirements (especially for \geq constraints) [61]. The vectors $a_1, \dots, a_m \in \mathbb{R}^n$ contain the coefficients for the m constraints of the problem; the n components of each vector $a_i \in \mathbb{R}^n$ thus typically denote how much of resource/requirement i is consumed/satisfied by each optimization variable [61].

The feasible region of a linear program is a *convex polytope*⁸, which is a set defined as the intersection of finitely many *half spaces*⁹, each of which is defined by a linear inequality [1]. The objective function $c^T x$ is a real-valued affine (linear) function defined on this polytope. Thus, a linear programming algorithm tries to find a point in the polytope where this function has the smallest (or largest) value, if such a point exists [61].

The problem class definition in 2.3 prohibits any nonlinear term [61]. However, nonlinear terms such as the multiplication of two decision variables, the maximum of several vari-

⁷ "An affine function is the composition of a linear function followed by a translation. A linear function fixes the origin, whereas an affine function need not do so. Thus, the function $f(x) = ax$ is linear and affine, while $f(x) = ax + b$ is affine but not linear." [111]. However, in colloquial language (and in most optimization literature) these terms are interchangeable.

⁸ "A polytope is a geometric object with "flat" sides. It is a generalization in any number of dimensions, of the three-dimensional polyhedron." [3]

⁹ "In geometry, a half-space is either of the two parts into which a plane divides the three-dimensional Euclidean space. More generally, a half-space is either of the two parts into which a hyperplane divides an affine space. That is, the points that are not incident to the hyperplane are partitioned into two convex sets (i.e., half-spaces), such that any subspace connecting a point in one set to a point in the other must intersect the hyperplane." [8]

ables, or the absolute value of a variable often arise when modeling a practical problem; these nonlinearities must be addressed with different techniques if the linear problem class is to be used (see section 2.2.2).

2.2.1 Mixed-integer linear programming (MILP)

Usually when a *model* of a system is derived, differential or difference equations in continuous domains are used¹⁰, which are derived from the physical laws that govern the *dynamics* of the system under consideration. However, in many practical applications the system is not isolated, and must be controlled by subsystems described by *logic*, or discrete domains, i.e., on/off switches or valves, gears or speed selectors or evolutions dependent on if-then-else rules [25].

Therefore, many practical cases require a framework to model systems described by physical laws, logical rules and operating constraints. Within this framework, any propositional logic statement is transformed into a linear inequality that involves integer and continuous variables. In consequence, the system is finally modeled by linear dynamic equations that are subject to linear mixed-integer inequalities, that is, inequalities that contain both continuous and binary variables [25].

This framework dedicated to the modeling and optimization of *hybrid systems* is referred to as *mixed-integer optimization*. When the problem is circumscribed within the class of linear problems, we refer to *mixed-integer linear programming* (MILP); otherwise, the problem is classified as *mixed-integer nonlinear programming* (MINLP).

The following section introduces a general treatment of some often used reformulation tricks to reformulate nonlinear problems into linear ones.

2.2.2 MINLP to MILP conversion

A more comprehensive treatment of MILP and the transformations presented in this section can be found in the work of Smith and Taskin in [61]. For a rigorous exposition, see the works by Bemporad and Morari in [25] and Gorissen and co-workers in [96].

As explained in the previous section, in order to work with systems that have both logic and dynamics, it is necessary to establish a link between both domains. In most cases, this results in *mixed-integer linear inequalities*, i.e., linear inequalities involving both *continuous variables* $x \in \mathbb{R}^n$ and logical (*indicator*) variables $\delta \in \{0, 1\}$. However, as it will become clear during the problem modeling in chapter 3, in some cases the problem requires a product

¹⁰ In consequence, the evolution of the system is described by smooth linear or nonlinear state transition functions.

of variables¹¹, transforming the inequality (and the problem class) to a bilinear one, i.e., a nonlinear problem.

In these cases, a re-linearization of the model needs to be carried out; this is done through the transformation of the problem by the introduction of *auxiliary variables* (logical or continuous), together with their respective constraints.

With this in mind, consider the statement $f(x) \leq 0$, where $f: \mathbb{R}^n \mapsto \mathbb{R}$ is linear. Assume that $x \in \mathbb{X}$, where \mathbb{X} is a given bounded set, and define [25]:

$$M \triangleq \max_{x \in \mathbb{X}} f(x), \quad (2.4)$$

$$m \triangleq \min_{x \in \mathbb{X}} f(x). \quad (2.5)$$

Theoretically, an over(under)-estimate of $M(m)$ suffices for the purpose of the transformation. However, computational benefits are obtained through practical estimates [25, 61].

In the remainder, some commonly used reformulation tricks to transform bilinear constraint problems into linear ones are introduced.

Product of logical variables: The product term $\delta_1 \delta_2$ where $\{\delta_i \in \{0, 1\}, i = 1, 2\}$ can be replaced by an auxiliary logical variable $\delta_3 \triangleq \delta_1 \delta_2$; then by Boolean algebra it holds that:

$$\delta_3 = 1 \leftrightarrow (\delta_1 = 1) \wedge (\delta_2 = 1), \quad (2.6)$$

and therefore the following transformation can be defined [25]:

$$\delta_3 = \delta_1 \delta_2 \text{ is equivalent to } \begin{cases} -\delta_1 + \delta_3 \leq 0, \\ -\delta_2 + \delta_3 \leq 0, \\ \delta_1 + \delta_2 - \delta_3 \leq 1. \end{cases} \quad (2.7)$$

In this way, a bilinear constraint formed by the product of two logical variables can be linearized by its replacement with the introduction of a logical auxiliary variable and three logical linear inequalities.

¹¹ Products of logical variables, and of continuous and logical variables.

Product of continuous and logical variables: The product term $\delta f(x)$, where $f: \mathbb{R}^n \mapsto \mathbb{R}$ and $\delta \in \{0, 1\}$, can be replaced by an auxiliary real variable $y \triangleq \delta f(x)$, which satisfies:

$$\begin{cases} \delta = 0 \leftrightarrow y = 0, \\ \delta = 1 \leftrightarrow y = f(x). \end{cases} \quad (2.8)$$

Therefore, by defining M and m as in equations 2.4 and 2.5 respectively, the following transformation is defined [25]:

$$y = \delta f(x) \text{ is equivalent to } \begin{cases} y \leq M\delta, \\ y \geq m\delta, \\ y \leq f(x) - m(1 - \delta), \\ y \geq f(x) - M(1 - \delta). \end{cases} \quad (2.9)$$

Thus, a bilinear constraint formed by the product of continuous and logical variables can be linearized by its replacement with the introduction of a real auxiliary variable, two logical linear inequalities and two mixed-integer linear inequalities.

2.3 Duality in mathematical optimization

Starting from *any* general optimization problem as defined in equation 2.1 (i.e., no assumptions on convexity or the problem class are made), which is usually referred to as the *primal problem*, the *Lagrange duality* produces, through the implementation of the *Lagrange dual function*, a convex problem which is called the *Lagrange dual problem*.

The dual problem stipulates which is the best lower bound (d^*) on the optimal value (p^*)¹² of the primal problem that is achievable by using the Lagrange dual function.

If the primal problem is convex, the bound will be sharp (given that some conditions are met); this property will grant alternative methods for solving the primal problem. On the other hand, if the original problem is non-convex and, possibly, a hard problem (i.e., it can not be solved efficiently), then the dual problem will provide, with little computational effort, a lower bound for the optimal value of the primal problem. Thus, duality proves to be a very crucial method for obtaining suboptimal solutions of the primal problem.

The remainder of this section is presents a brief introduction to the concepts involved in duality and is strongly influenced by the work of Boyd and Vandenberghe in [39]. For a

¹² For the optimal value p^* of the optimization problem, it holds that $p^* = f_0(x^*)$, where $f_0(x)$ is the objective function of the optimization problem and x^* is the optimal solution to the problem.

detailed theoretical analysis on duality please refer to [39]; a more practical approach on duality of Linear Programs can be found in [148].

2.3.1 The Lagrange dual function

The Lagrangian

A general optimization problem in the standard form of equation 2.1 is considered:

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && g_i(x) \leq 0, \, i = 1, \dots, m \\ & && h_i(x) = 0, \, i = 1, \dots, p \end{aligned} \tag{2.10}$$

with variable $x \in \mathbb{R}^n$ and domain $\mathcal{D} = (\bigcap_{i=1}^m \text{dom } f_i) \cap (\bigcap_{i=1}^p \text{dom } h_i)$ nonempty. The optimal value of the problem 2.10 is denoted by p^* .

Lagrangian duality takes into account the constraints in problem 2.10 by augmenting the objective function with a weighted sum of the constraint functions [39]. Thus, the *Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ associated with the problem 2.10 is defined as [39]:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x), \tag{2.11}$$

with $\text{dom } L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$. The Lagrangian is a function of x , λ and ν . The value λ_i is known as the *Lagrange multiplier* associated with the i -th inequality constraint $f_i(x) \leq 0$; similarly the value ν_i is known as Lagrange multiplier associated with the i -th equality constraint $h_i(x) = 0$. The vectors λ and ν are called the *dual variables* or *Lagrange multiplier vectors* associated with the problem 2.10.

In some practical applications, the dual variables can have a known meaning, as prices or penalty rates applied to the constraints. For example, for a solution \tilde{x} to be feasible for the problem definition 2.10, the i -th inequality constraint ($f_i(\tilde{x}) \leq 0$) needs to be met. If this constraint (or any other constraint) is not met, \tilde{x} is unfeasible.

By replacing the constraint with its corresponding term inside the Lagrangian (multiplied by the Lagrange multiplier) in 2.11, this binary behavior observed in the constraints of problem 2.10 can be shifted to a more liberal approach, where the possibility to not meet the constraints requirements exists but, in exchange, a penalty is applied. Furthermore, the flip-side gives a subsidy for coming in under-budget and both the penalties and the subsidies are added to the objective value in the Lagrangian.

The proportionality factor for the penalty (subsidy) is given by the Lagrange multiplier. Thus, a penalty (subsidy) of $\lambda_i f_i(\tilde{x})$ is imposed if the i -th constraint is greater (lower) than

zero. For every unit of measurement that $f_i(\tilde{x})$ is positive (negative), a penalty (subsidy) of λ_i units of measurement is imposed.

The Lagrange dual function

The *Lagrange dual function* (or just *dual function*) $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ is defined as the minimum value of the Lagrangian over x [39]: for $\lambda \in \mathbb{R}^m$, $\nu \in \mathbb{R}^p$,

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right). \quad (2.12)$$

The Lagrange dual function defined in equation 2.12 minimizes the Lagrangian over all possible values for x . The Lagrangian, as a function of λ and ν , is an affine function for each possible value of x ; since the infimum of a family of affine functions is concave, even if $f_0(x)$ and $f_i(x)$ are non-convex and $h_i(x)$ are non-affine in the primal problem, it is a certainty that the dual function g is concave as a function of λ and ν .

Lower bounds on optimal value

It can be easily proved that the evaluation of the dual function g yields lower bounds on the optimal value p^* of the primal problem 2.10, i.e., for any $\lambda \geq 0$ and any ν , the following holds [39]:

$$g(\lambda, \nu) \leq p^*. \quad (2.13)$$

The verification argument is pretty straight-forward. Let \tilde{x} be a feasible solution for the primal problem 2.10, then:

$$L(\tilde{x}, \lambda, \nu) = f_0(\tilde{x}) + \underbrace{\sum_{i=1}^m \lambda_i f_i(\tilde{x})}_{\leq 0} + \underbrace{\sum_{i=1}^p \nu_i h_i(\tilde{x})}_{=0} \leq f_0(\tilde{x}), \quad (2.14)$$

and therefore:

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \leq L(\tilde{x}, \lambda, \nu) \leq f_0(\tilde{x}). \quad (2.15)$$

Since $g(\lambda, \nu) \leq f_0(\tilde{x})$ holds for every feasible point, the inequality 2.13 follows.

Thus, after finding the x value that minimizes the Lagrangian, it can be plugged in back into the Lagrangian, obtaining in this way the dual function g , which is concave as a function of λ and ν , and gives the lower bound for the objective of the primal problem.

This property is immediately useful in contexts where the problem is non-convex, or where a convex problem has a huge number of variables and constraints that impose great computational efforts. With this property, iterative procedures can be created to obtain an x that approximately satisfies the problem, i.e. a suboptimal solution.

2.3.2 The Lagrange dual problem

The Lagrange dual function is a lower bound on the optimal value p^* of the optimization problem 2.10, which is parametrized by the dual variables λ and v . For the majority of possible values for these parameters, the Lagrange dual will output $g = -\infty$ as a lower bound; this is a valid lower bound, but it is totally uninformative. For a small set of cases of λ and v , lower bounds with more information are obtained, approaching suboptimal solutions to the original problem. Thus, in order to find the *best* lower bound that can be obtained from the Lagrange dual function, the following optimization problem is defined [39]:

$$\begin{aligned} & \text{maximize} && g(\lambda, v) \\ & \text{subject to} && \lambda \geq 0. \end{aligned} \tag{2.16}$$

This problem is called the *Lagrange dual problem* associated with the problem 2.10.

The Lagrange dual problem defined in 2.16 is a convex optimization problem, since the objective to be maximized is concave and the constraint is convex. This is the case whether or not the primal problem 2.10 is convex.

The pair (λ^*, v^*) is the set of parameters that are optimal for the problem 2.16, i.e. they maximize the dual function. Thus, the optimum value of the dual problem is defined as $d^* = g(\lambda^*, v^*)$, and this value corresponds to the best lower bound for the primal problem that can be obtained from the Lagrange dual function.

2.3.3 Weak and strong duality

By definition, the optimal value of the Lagrange dual problem (d^*) is the best lower bound on p^* that can be obtained from the Lagrange dual function; even if the original problem is not convex, it *always* holds that:

$$d^* \leq p^*, \tag{2.17}$$

because any dual feasible point is a lower bound on p^* , so the best one is also a lower bound [39]. This property is referred to as *weak duality*. The *duality gap* or *optimal duality gap* is a nonnegative value defined as the difference $p^* - d^*$, and allows to quantify how suboptimal is the lower bound obtained through the resolution of the dual problem.

On the other hand, *strong duality* means that the lower bound is proven to be tight, i.e. the optimal duality gap is zero:

$$d^* = p^*. \quad (2.18)$$

Thus, strong duality means that the optimal value of the Lagrange dual problem is equivalent to the optimal value of the primal problem.

If the original problem is convex, strong duality *usually* holds; the conditions that a convex problem must meet so that strong duality holds are called *constraint qualifications*. One of the simplest constraint qualifications is *Slater's constraint qualification*, which says that if there exists a *strictly feasible*¹³ point for a convex problem, then strong duality holds. For most problems that appear in engineering, Slater's condition is usually sufficient to prove strong duality [39].

For the LP problem class, the geometric interpretation of Slater's condition says that strong duality holds if the polytope created by the feasible set has a nonempty interior. Thus, for the LP problem class strong duality always holds, except for the case when the primal and the dual are both infeasible.

2.4 Optimization under data uncertainty

Until this section, it was considered that the standard modeling process in mathematical programming involves the creation of a discrete model whose input data are exactly known and equivalent to certain nominal values when the problem is solved. Although this paradigm is perfectly valid for a variety of cases, it does not take into account the influence that possible data uncertainties¹⁴ may have on the model and on the quality and feasibility of the solution obtained.

Therefore, it is possible that as the data deviate from their nominal values, some constraints of the problem are violated. Consequently, solutions assumed to be optimal could instead be heavily suboptimal and even infeasible. In this last case, the decisions cannot be implemented in practice [69]. Ben-Tal and Nemirovski refer to this problem in [29]:

"In real-world applications of Linear Programming one cannot ignore the possibility that a small uncertainty in the data (intrinsic for most real-world LP pro-

¹³ A strictly feasible solution means not only that the inequality constraints of the problem are met, but also that each one is met with a positive margin; that is, the inequality constraints hold with strict equalities.

¹⁴ "Data can be inherently stochastic/random or it can be uncertain due to errors. The reason for data errors could be measurement/estimation errors that come from the lack of knowledge of the parameters of the mathematical model (e.g., the uncertain price or demand) or could be implementation errors that come from the physical impossibility to exactly implement a computed solution in a real-life setting." [96]

grams) can make the usual optimal solution of the problem completely meaningless from a practical viewpoint."

This scenario creates the need (mainly in practical applications) of mathematical models whose solution is immune, as far as possible, to the possible uncertainties present in the data; these models are usually referred to as *robust*.

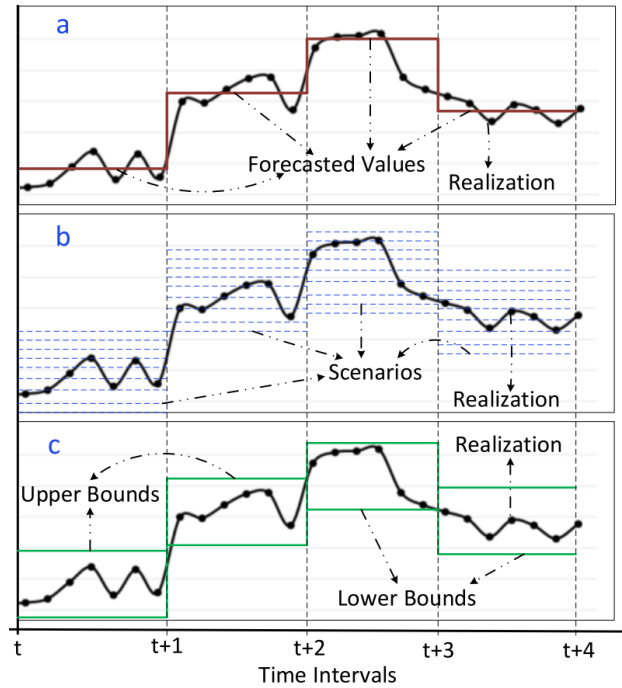


Figure 2.1: Uncertainty representation: (a) Deterministic optimization; (b) Stochastic optimization; (c) Robust optimization. Source: [112].

Approaches for scheduling problems under uncertainty can be broadly classified into reactive and proactive [205]:

- *reactive approaches*, where a nominal schedule is generated by a deterministic formulation and then it is modified when it needs to be adjusted to updated system data;
- *proactive approaches* take into account all possible cases and try to find a good solution for all of them; in consequence, the solution found is feasible for all the considered cases. However, some of these cases happen only on rare occasions, so the solution obtained could prove to be too conservative.

The most broadly used proactive approaches to address data uncertainty in optimization are robust and stochastic optimization [96, 112]:

- *Stochastic optimization* (SO) has an important assumption in its core: the true probability distribution of uncertain data has to be known or estimated (i.e., accurate information of uncertainties is required to construct reliable probability density functions, or PDFs) [223]. If this condition is met, SO only manages to provide a probabilistic guarantee for the feasibility of the solution [32]; to do so, the reformulation of the uncertain optimization problem as a SO problem requires the generation of a large number of scenarios to ensure quality of the solution (see figure 2.1), which in turn results in growth of the problem size and computational requirements. Moreover, the accuracy of the solution obtained is sensitive to the technique used for the scenario generation [124];
- *Robust optimization* (RO) does not assume that probability distributions are known, but instead addresses the parameter uncertainty based on an deterministic uncertainty set, which covers part or the whole region of the uncertain space [96]; in consequence, it only needs information about the upper and lower bounds of the uncertainty [32] (see figure 2.1). The target of RO is to select the best solution that remains feasible for any realizations of the uncertain parameters in the uncertainty set. RO has been proved to be computational tractable¹⁵ for a variety of classes of uncertainty sets and problem types [163]; the robust counterpart generally does not increase much in model size compared to the deterministic model, and the convexity of the constraints can also be preserved [234].

Due to the various limitations of stochastic optimization and the numerous advantages of robust optimization presented above, since the beginning of the century, RO has managed to find popularity as an effective method to cope with data uncertainty in optimization problems, especially due to its computational tractability, modeling power and broad applicability in a wide spectrum of domains, including, but not limited to, statistics, operations research [31], electrical engineering [202, 201], control theory [126], oil industry [206, 238, 14], finance [81], portfolio management [229, 230], logistics [233], manufacturing engineering [210], chemical engineering [30], medicine [58, 61, 239], and computer science. In the area of energy management, in addition to scheduling of microgrids [223, 124, 13, 147, 112, 166], RO has also been used for optimization of various other related objectives of power systems, which include but are not limited to, communication in microgrids [199] transmission network expansion [119], distributed generation

¹⁵ A tractable problem is one that can be solved in a reasonable amount of time; usually the distinction between tractable and intractable is drawn at the boundary between problems that can be solved in an amount of time that is polynomial; those that require exponential time are regarded as intractable.

investment [195] and placement [225], transition of electric vehicles [101], and scalability of demand response [139].

In the remainder of the current section, the principles, concepts and implementation methods involved in the robust optimization methodology are introduced; an exhaustive introduction to the theory and applications of robust optimization can be found in [26, 31].

2.4.1 Basic principles of robust optimization

The following are the basic principles on which the robust optimization methodology is based [69] (see figure 2.2):

- the *actual value* of an uncertain coefficient of the problem is unknown. Nevertheless, the modeler must have a reasonable reference for the unknown actual value, called *nominal value*. In this way, the actual value equals the sum of the nominal value and an unknown deviation;
- the modeler defines an *uncertainty set* [155] for the coefficients against which he wants to be protected. This set describes the deviations of the coefficients with respect to their nominal values;
- the modeler derives and solves a *robust counterpart* of the original optimization problem. D'Andreagiovanni and co-workers define it as:

"[...] a modified version of the original optimization problem that only considers robust feasible solutions, i.e. feasible solutions that are protected against all the deviations of the uncertainty set." [69]

Then, the *optimal robust solution* is the feasible solution that achieves the best score for the objective function under the worst data deviation;

- increasing the level of robustness of the solution is not free, i.e. the optimal solution to the problem will deteriorate as the robustness level is increased. Therefore, uncertainty sets that comprise higher levels of protection will usually lead to solutions with more deteriorated optimal values. This trade-off is usually referred to as *the price of robustness* [33].

2.4.2 Approaches for the implementation of robust optimization

To this date, several techniques have been presented to take into account the uncertainties of the data in the modeling of the problem, thus making the problem more robust:

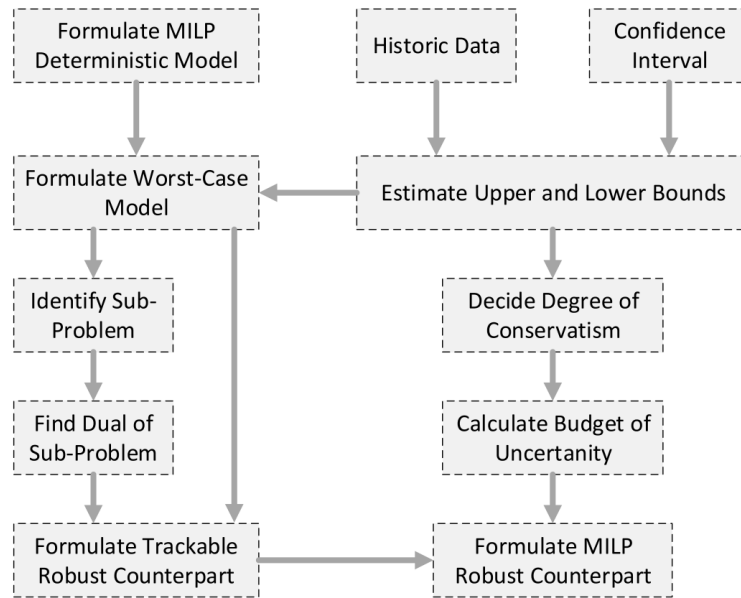


Figure 2.2: Problem formulation process for robust optimization. Source: [112].

- Soyster [207] was the one who presented the first approach with his work in the early 1970s. Bertsimas and Sim describe it as:

"[...] a linear optimization model to construct a solution that is feasible for all data that belong in a convex set. The resulting model produces solutions that are too conservative in the sense that we give up too much of optimality for the nominal problem in order to ensure robustness." [33]

Further comments can be found in the works of Ben-Tal and Nemirovski [29];

- The next big improvement was performed independently by Ben-Tal and Nemirovski [27, 28, 29] and El-Ghaoui *et al.* [78, 79]. These proposals were focused on reducing the levels of over-conservatism present in the work carried out by Soyster. The solution found relies on the use of ellipsoidal type uncertainties to model the linear problem; if properly chosen, elliptic type uncertainty intervals allow modeling of more complex uncertainties. However, this approach leads to nonlinear, although convex, models, which require a greater computing power compared to the linear models developed by Soyster [207];
- Bertsimas and Sim [33] made the next leap forward by proposing a robust optimization model that preserves the linear properties of Soyster [207] with low computing power demands, while at the same time gives absolute control over the level of conservatism of each of the constraints present in the model. The approach protects against variations of a previously specified number Γ_i of parameters in the i -th con-

straint of the model, ensuring that the solution is feasible. Moreover, in the case that the number of parameters that vary is greater than Γ_i , the approach provides a method of calculating a probabilistic feasibility guarantee, ensuring that the solution will be feasible with high probability.

- Li and co-workers in [155] extended the formulation made by Bertsimas and Sim [33] by incorporating a series of novel uncertainty sets and deriving, for each one of these new sets, the robust counterpart for both LP and MILP problem formulations. As it is pointed out by Li *et al.* [155], the robust formulation made by Bertsimas and Sim [33] is based on the combined interval and polyhedral uncertainty set.
- Jalilvand-Nejad *et al.* in [121] and Yuan *et al.* in [234] used the uncertainty sets defined by Li *et al.* in [155] and expanded the concept to present novel results on robust optimization under correlated uncertainties that appear in a single constraint; experimental results reveal that when significant correlations between the coefficients exist, the performance of this method is superior to that of the traditional uncertainty sets for independent random variables.

The advantage of working with linear optimization problems with respect to other classes of optimization problems is that they are easily generalizable to discrete optimization problems, this being of vital importance in the development of the specific optimization problem that corresponds to this work.

Therefore, the Bertsimas and Sim [33] approach gives more flexibility and is less computationally demanding than the Ben-Tal and Nemiroski [27, 28, 29] and El-Ghaoui *et al.* [78, 79] approach. Also, the Bertsimas approach presents a greater flexibility than the one of Soyster [207], since it allows the election of the desired levels of conservatism. Thus, when needed, the Bertsimas model can avoid dealing with extreme conservatism and address intermediate and more rational cases.

Due to the above advantages, the approach proposed by Bertsimas and Sim [33] (and its extensions by Li *et al.* [155]) to obtain a robust counterpart of a linear optimization problem is a well established model for real-world applications [69], and is the one that will be developed and applied in the remaining sections of this work; moreover, the findings reported by Jalilvand-Nejad *et al.* in [121] are implemented when correlation between coefficients exist.

An introduction to the concepts involved in these methodologies is made in the next section.

2.4.3 Uncertainty set-induced robust optimization

The present section aims to explain the main steps involved in the derivation of the robust counterpart of an optimization problem that is subject to data uncertainty; it is based entirely on the work carried out by Bertsimas *et al.* in [33, 31], Li *et al.* in [155, 156, 157], Gorissen *et al.* in [96], Jalilvand-Nejad *et al.* in [121] and Yuan *et al.* in [234].

In set-induced robust optimization, the uncertain data of the original problem are assumed to be varying in a given *uncertainty set*, and the aim is:

"[...] to find the best possible candidate solution that remains feasible for all realizations of the data in the uncertainty set, that is, the feasibility of the solution must be maintained no matter what value the random variables realize within a certain set that belongs to the uncertain space." [155]

The corresponding optimization problem is also called *robust counterpart optimization* problem.

Consider the following general (mixed integer) linear optimization problem with uncertainty in the left hand side constraint coefficients, right hand side and objective function coefficients:

$$\begin{aligned} & \text{maximize} && \sum_j \tilde{c}_j x_j \\ & \text{subject to} && \sum_j \tilde{a}_{ij} x_j \leq \tilde{b}_i \quad \forall i. \end{aligned} \tag{2.19}$$

where x_j can either be a continuous or an integer variable; and \tilde{a}_{ij} , \tilde{b}_i , \tilde{c}_j represent the true value of the parameters which are possibly subject to uncertainty.

Through the introduction of some auxiliary variables and some simple transformations, problem 2.19 can be rewritten as:

$$\begin{aligned} & \text{maximize} && z \\ & \text{subject to} && z - \sum_j \tilde{c}_j x_j \leq 0 \\ & && \tilde{b}_i x_0 + \sum_j \tilde{a}_{ij} x_j \leq 0 \quad \forall i \\ & && x_0 = -1. \end{aligned} \tag{2.20}$$

where all the uncertain parameter are now only on the left and side constraint coefficients.

Based on the above and without loss of generality, an analysis can be carried out on the i -th constraint of a (mixed integer) linear optimization problem with uncertain parameters *only* in the left hand side:

$$\begin{aligned} & \text{maximize} && \sum_j c_j x_j \\ & \text{subject to} && \sum_j \tilde{a}_{ij} x_j \leq b_i \quad \forall i. \end{aligned} \tag{2.21}$$

The i -th constraint is defined as follows:

$$\sum_j \tilde{a}_{ij} x_j \leq b_i, \tag{2.22}$$

where x_j can either be a continuous or an integer variable; and \tilde{a}_{ij} represent the true value of the parameters which are possibly subject to uncertainty. The uncertain coefficients \tilde{a}_{ij} can take values in the range $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$; hence \tilde{a}_{ij} can be defined as:

$$\tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \quad , \forall j \in J_i, \tag{2.23}$$

where a_{ij} represents the nominal value of the parameter; \hat{a}_{ij} represents the maximum positive perturbation amplitude; $\{\xi_{ij}\}_{j \in J_i}$ are independent random variables which are subject to uncertainty and perturbs in the range $[-1, 1]$; and J_i represents the index subset that contains the continuous and discrete variable indices whose corresponding coefficients are subject to uncertainty.

With equation 2.23, the constraint 2.22 can be rewritten by grouping the deterministic part and the uncertain part for the left hand side as follows:

$$\sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j \leq b_i, \tag{2.24}$$

where the first term of the left hand side corresponds to all the nominal values of the parameters in the constraint, while the second term has the perturbations of the parameters subject to uncertainty (i.e., all parameters for which $j \in J_i$ is met).

With a predefined uncertainty set \mathcal{U} for the random variables ξ_{ij} , the objective of the problem is to find solutions that remain feasible for any $\xi_{ij} \in \mathcal{U}$ thus avoiding possible infeasibilities [155]; in consequence, the uncertain term in the i -th constraint needs to be maximized as follows:

$$\sum_j a_{ij} x_j + \max_{\xi \in \mathcal{U}} \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j \right\} \leq b_i. \tag{2.25}$$

This equation conforms the *worst-case reformulation* of the i -th constraint, because it assures that the optimal solution found for the problem will be optimal and feasible for every possible realization of the uncertainties within the predefined uncertainty set \mathcal{U} .

To obtain the robust counterpart optimization problem, the i -th constraint of the original optimization problem 2.21 is replaced by the robust counterpart constraint 2.25:

$$\begin{aligned} & \text{maximize} && \sum_j c_j x_j \\ & \text{subject to} && \sum_j a_{ij} x_j + \max_{\xi \in \mathcal{U}} \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j \right\} \leq b_i. \quad \forall i. \end{aligned} \quad (2.26)$$

The problem formulation written above is based on a general and *undefined* uncertainty set \mathcal{U} . In consequence, the constraints and, therefore, also the problem formulation, are nonlinear. Thus, it is clear that the explicit expression of this set-induced robust counterpart formulation heavily depends on the uncertainty set \mathcal{U} defined by the decision-maker; an appropriate set definition could allow, through a series of transformations, the linearization of the constraints. It is important to mention that each constraint that presents uncertain data in the optimization problem will define its own uncertainty set \mathcal{U} where the parameters can vary.

In the remaining parts of this section, the uncertainty sets needed for the present work will be introduced together with the derivation of the robust counterpart for those particular uncertainty sets.

Uncertainty sets

In this section, the most prominent uncertainty sets proposed in literature are reviewed; as stated in section 2.4.1, one of the first steps necessary to implement a robust optimization is to define a deterministic data set within the uncertain space.

The design of the uncertainty set is related to the distribution of the uncertainty: if the uncertainty is subject to unbounded distribution, the box and polyhedral type of uncertainty set are appropriate to be used in the robust optimization framework, where the size of the uncertainty set is not restricted; instead, if the uncertainty is subject to bounded distribution, the bounds information can be involved in the uncertainty set (which will lead to more conservative solution). Hence, the *interval+ellipsoidal* uncertainty set is appropriate for bounded uncertainty solution [234].

Once the type of set is chosen, to define its size is a trade-off between the robustness of the solution (i.e. against how many physical realizations of the uncertain parameter the

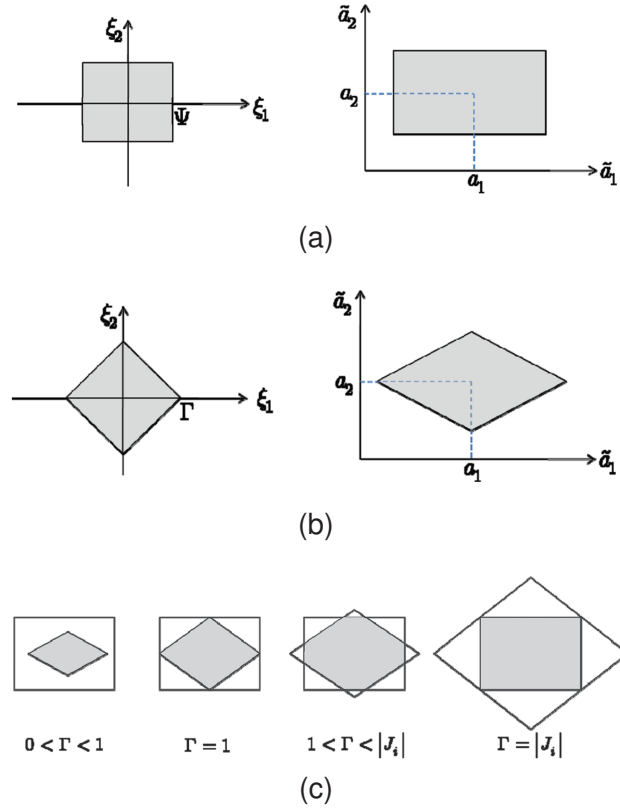


Figure 2.3: Uncertainty sets for parameter \tilde{a}_j , defined by $[\tilde{a}_j = a_j + \xi_j \hat{a}_j, j = 1, 2]$, where \tilde{a}_j denotes the true value of the parameter, a_j denotes the nominal value of the parameter, ξ_j denotes the uncertainty and \hat{a}_j represents a constant perturbation. (a): box uncertainty set; (b): polyhedral uncertainty set; (c): *box+polyhedral* uncertainty set. Source: [155].

model will be robust) and the the price of robustness (i.e., the impact this set has on the optimal value of the problem), as explained in section 2.4.1.

Box uncertainty set In an uncertain optimization problem, it is assumed that the random variables ξ_{ij} are independent and their absolute values vary between zero and a parameter $\Psi_i \leq 1$. The interaction of this type of perturbations creates a box which is called box uncertainty, which can be described as follows [155]:

$$\mathcal{U}_b \triangleq \{ \tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \mid |\xi_{ij}| \leq \Psi_i, \quad \forall j \in J_i, \forall i \}, \quad (2.27)$$

where Ψ_i is an adjustable parameter used to control the size of the uncertainty set, i.e., represents the perturbation bound for all of the uncertain coefficients related to the i -th constraint (see figure 2.3a).

As the value of Ψ_i increases, the model becomes more conservative. If $\Psi_i = 1$, then $\xi_{ij} \in [-1, 1]$ and the uncertainty set covers the whole uncertain space; in this case, the

obtained robust model is identical to the one introduced by Soyster in [207]; this special case is referred to as *interval uncertainty set* and is defined as follows [155]:

$$\mathcal{U}_i \triangleq \{ \tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \mid |\xi_{ij}| \leq 1, \quad \forall j \in J_i, \forall i \}. \quad (2.28)$$

Polyhedral uncertainty set The polyhedral uncertainty set can be formulated as follows [155]:

$$\mathcal{U}_p \triangleq \left\{ \tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \mid \sum_{j \in J_i} |\xi_{ij}| \leq \Gamma_i, \quad \forall j \in J_i, \forall i \right\}, \quad (2.29)$$

where Γ_i is the adjustable parameter controlling the size of the uncertainty set of the i -th constraint (see figure 2.3b). For the cases when the uncertainty is bounded (i.e., $\xi_{ij} \in [-1, 1]$), if $\Gamma_i \geq |J_i|$ (i.e., Γ_i is chosen to be greater or equal than the number of uncertain parameters in the i -th constraint), then the overall uncertain space is covered by the polyhedral uncertainty set.

"Box+polyhedral" uncertainty set This set is generated by intersecting the polyhedral and the interval set as follows [155]:

$$\mathcal{U}_{b \cap p} \triangleq \left\{ \tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \mid \sum_{j \in J_i} |\xi_{ij}| \leq \Gamma_i, \quad |\xi_{ij}| \leq \Psi_i, \quad \forall j \in J_i, \forall i \right\}. \quad (2.30)$$

The intersection between an adjustable box (as defined in equation 2.27) and an adjustable polyhedron (as defined in equation 2.29) does not reduce to any one of them as long as the parameters satisfy the following relationship (see figure 2.3c):

$$\Psi_i \leq \Gamma_i \leq \Psi_i |J_i|. \quad (2.31)$$

The bounded uncertain space is completely covered by the uncertainty set when the adjustable parameter (Γ_i) is set to the upper bound; increasing the parameter beyond this point might create a more conservative solution without improving the robustness of the solution.

The *interval+polyhedral* set is created when the interval set ($\Psi = 1$) is intersected with the polyhedral set [155]:

$$\mathcal{U}_{i \cap p} \triangleq \left\{ \tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \mid \sum_{j \in J_i} |\xi_{ij}| \leq \Gamma_i, \quad |\xi_{ij}| \leq 1, \quad \forall j \in J_i, \forall i \right\}. \quad (2.32)$$

On this case, if $\Gamma_i = 1$ the polyhedron is exactly inscribed by the box, and the intersection between both is the polyhedron; when $\Gamma_i = |J_i|$, the intersection between both sets is the box. This set assumes that the uncertainty interval is symmetric with respect to the nominal value of the parameter, and is particularly useful for the case of bounded uncertainty, since the uncertainty set can be restricted inside the known bounds of the uncertainty.

Correlated "interval+polyhedral" uncertainty set Independence is generally assumed among uncertainties in the parameters; however, in practice, correlations may arise in the uncertainties and if they are not addressed the optimal solution could prove to be over-conservative. Thus, this set deals with random variables that are correlated to each other (i.e., not independent). Due to the existence of correlation between coefficients, random points are gathered around the diagonal in a plot like the one in figure 2.4. The idea is to reshape the uncertainty set so that it covers the occupied spaces better than the void spaces; therefore, the borders of the polyhedral uncertainty set are bended in a way that the areas around the diagonals are covered better than the areas far from it. Moreover, the degree of bending is proportional to the level of correlation between the coefficients. This set is presented as follows [121]:

$$\mathcal{U}_{c(i \cap p)} \triangleq \left\{ \tilde{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \left| |\xi_{ij}| + \sum_{n \neq j} \left[\left(1 - \left(\frac{|J_i| - \Gamma_i}{|J_i| - 1} \right) |\rho_{ijn}| \right) |\xi_{in}| \right] \leq \Gamma_i, \quad |\xi_{ij}| \leq 1, \quad \forall j \in J_i, \forall i \right\}, \quad (2.33)$$

where ρ_{ijn} denotes the correlation between coefficients \tilde{a}_{ij} and \tilde{a}_{in} , and takes value in the range of $[-1, 1]$. If $\rho_{ijn} = 0$ (for all i, j , and $n \neq j$), the correlated polyhedral+interval uncertainty set is equivalent to the polyhedral+uncertainty set of equation 2.32; on the other hand, the correlated uncertainty set covers the diagonals of the coordinate system entirely when $|\rho_{ijn}| = 1$. For this uncertainty set, it is assumed that the estimate of the correlation matrix between uncertain coefficients is exploited from the historical data and, in consequence, is available.

Solving the robust counterpart formulation

As it was mentioned earlier during this section, the right definition for the uncertainty set could allow the linearization of the nonlinear constraints. If the uncertainty set \mathcal{U} in the formulation 2.26 is defined as *interval+polyhedral* (see equation 2.30), then the inner maximization sub-problem of the i -th constraint of problem 2.26 becomes linear [31]; nonetheless, the i -th constraint remains nonlinear.

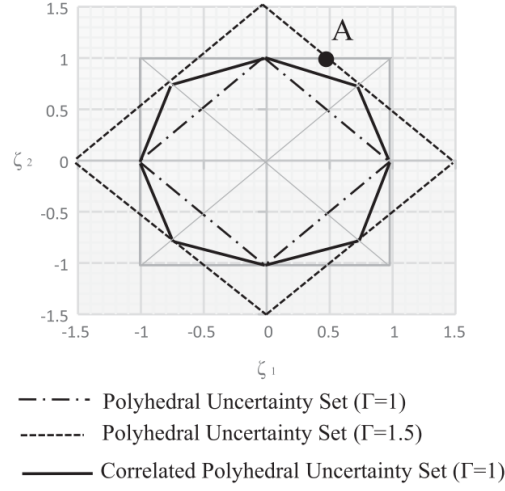


Figure 2.4: Correlated polyhedral uncertainty set vs polyhedral uncertainty set. Source: [121].

Therefore, the linearization of the i -th constraint involves the elimination of the inner maximization sub-problem. To do so, the inner maximization sub-problem is transformed into its Lagrange dual problem; due to its linear properties, the inner maximization sub-problem and its dual yield the same optimal objective value by strong duality (see section 2.3). Therefore, the incorporation of the Lagrange dual problem into the original constraint is valid, and the resulting i -th constraint is linear. For more details and a step-by-step derivation please refer to [155, 96].

Following, explicit robust counterpart constraints of equation 2.25 are investigated for the *interval+polyhedral* and the correlated interval+polyhedral uncertainty sets.

"Interval+polyhedral" uncertainty set If the uncertainty set \mathcal{U} in the formulation 2.26 is defined as *interval+polyhedral* (see equation 2.30), then the corresponding linear robust counterpart to the i -th constraint (see equation 2.25) is equivalent to the following set of linear constraints [155, 33]:

$$\sum_j a_{ij}x_j + \left[\Gamma_i z_i + \sum_{j \in J_i} w_{ij} \right] \leq b_i \quad (2.34a)$$

$$z_i + w_{ij} \geq \hat{a}_{ij} |x_j|, \forall j \in J_i \quad (2.34b)$$

$$z_i \geq 0 \quad (2.34c)$$

$$w_{ij} \geq 0. \quad (2.34d)$$

If the variables x_j are nonnegative, the nonlinearity introduced by the absolute value operator can be directly removed from the constraints, otherwise the constraints have to be redefined (see *Remark 6.4* in [155] for more details).

To derive the robust counterpart of the i -th constraint the following dual variables for the Lagrange dual problem are introduced: z_i for the uncertainty set; and w_{ij} for the x_j , $j \in J_i$ variables with uncertain parameters. Finally, the robust counterpart optimization problem is defined as follows:

$$\begin{aligned}
& \text{maximize} && \sum_j c_j x_j \\
& \text{subject to} && \sum_j a_{ij} x_j + \left[\Gamma_i z_i + \sum_{j \in J_i} w_{ij} \right] \leq b_i \\
& && z_i + w_{ij} \geq \hat{a}_{ij} x_j, \forall j \in J_i \\
& && x_j \geq 0, z_i \geq 0, w_{ij} \geq 0.
\end{aligned} \tag{2.35}$$

Correlated "interval+polyhedral" uncertainty set On the other hand, if the defined uncertainty set is the correlated *interval+polyhedral* (see equation 2.33), then the corresponding linear robust counterpart to the i -th constraint (see equation 2.25) is equivalent to the following set of linear constraints [121, 234]:

$$\sum_j a_{ij} x_j + \left[\sum_{j \in J_i} z_{ij} \Gamma_i + \sum_{j \in J_i} w_{ij} \right] \leq b_i, \forall i \tag{2.36a}$$

$$z_{ij} + \sum_{n \neq j} \left(1 - \left(\frac{|J_i| - \Gamma_i}{|J_i| - 1} \right) |\rho_{ijn}| \right) z_{in} + w_{ij} \geq \hat{a}_{ij} |x_j|, \forall i, j \tag{2.36b}$$

$$z_{ij} \geq 0 \tag{2.36c}$$

$$w_{ij} \geq 0. \tag{2.36d}$$

If the variables x_j are nonnegative, the nonlinearity introduced by the absolute value operator can be directly removed from the constraints, otherwise the constraints have to be redefined (see *Theorem 1* in [121] for more details).

To derive the robust counterpart of the i -th constraint the following dual variables for the Lagrange dual problem are introduced: z_{ij} , $j \in J_i$ for the uncertainty set; and w_{ij} for the

$x_j, j \in J_i$ variables with uncertain parameters. Finally, the robust counterpart optimization problem is defined as follows:

$$\begin{aligned}
& \text{maximize} && \sum_j c_j x_j \\
& \text{subject to} && \sum_j a_{ij} x_j + \left[\sum_{j \in J_i} z_{ij} \Gamma_i + \sum_{j \in J_i} w_{ij} \right] \leq b_i, \forall i \\
& && z_{ij} + \sum_{n \neq j} \left(1 - \left(\frac{|J_i| - \Gamma_i}{|J_i| - 1} \right) |\rho_{ijn}| \right) z_{in} + w_{ij} \geq \hat{a}_{ij} x_j \quad \forall i, j \\
& && x_j \geq 0, z_{ij} \geq 0, w_{ij} \geq 0.
\end{aligned} \tag{2.37}$$

As a recapitulation, it is important to mention that for every i -th constraint that presents uncertain parameters, a parameter Γ_i , not necessarily integer, is introduced; this parameter has a suggested range of action in the interval $[0, |J_i|]$. The main role of this parameter is to provide a way of adjusting the level of robustness of the proposed method against the level of conservatism of the solution [33]; if $\Gamma_i = 0$, the uncertain parameters only take the nominal value, and the obtained solution for the problem is the same as the deterministic optimization problem; if $0 < \Gamma_i < |J_i|$, the size of the uncertainty set taken into account grows, and the solution becomes more robust; if $\Gamma_i = |J_i|$, the uncertainty set covers the entire uncertain space, which contains all the possible realizations of uncertain parameters¹⁶; setting the parameter to $\Gamma \geq |J_i|$ does not create a more robust solution.

Both the uncertainty set and the value for the parameter Γ_i are defined by the decision maker; if $\Gamma_i = |J_i|$, then it is certain that the robust solution (if it exists) is feasible for any realizations of uncertainty (i.e., the probabilistic guarantee on constraint satisfaction is 1). However, in reality, the uncertainty set is not necessarily defined to cover the whole uncertain space (i.e., $0 < \Gamma_i < |J_i|$) because the decision maker might allow for a certain degree of constraint violation in exchange for a better objective value. In this case, the solution is guaranteed to be feasible only if less than Γ_i of the coefficients of the i -th constraint change; as it is explained at the end of the current section, a probabilistic guarantee on constraint satisfaction as a function of the Γ_i value can be computed.

Furthermore, making the solution more robust by taking into consideration a larger uncertainty set negatively affects the optimal solution of the problem. Thus, the parameter Γ_i controls the tradeoff between the probability of constraint violation (robustness of the solution) and the effect on the value of the objective function of the original problem; this tradeoff is often referred to as the *price of robustness* [33].

¹⁶ This results equals the over-conservative solution obtained by Soyster in [207].

The set-induced robust counterpart succeeds in reducing the price of robustness, since the optimal value of the objective function is marginally affected when the protection level is increased. Furthermore, the proposed robust approach is computationally tractable [33].

Probability bounds on constraint violation

The work done by Bertsimas and Sim in [33] and Li and co-workers in [156] provides a guarantee on the constraint satisfaction as a function of the value chosen for Γ_i .

Let x^* be an optimal solution of the problem formulation 2.35; the probability of constraint violation has an upper bound that is independent of the solution of the problem, and can be defined as follows [33, 156]:

$$\Pr\left(\sum_j \tilde{a}_{ij}x_j^* > b_i\right) \leq \frac{1}{2^n} \left\{ (1-\mu) \binom{n}{\lfloor v \rfloor} + \sum_{l=\lfloor v \rfloor+1}^n \binom{n}{l} \right\} \quad (2.38)$$

where $n = |J_i|$, $v = \frac{\Gamma_i + n}{2}$ and $\mu = v - \lfloor v \rfloor$.

Table 2.1: Choice of Γ_i as a function of $n = |J_i|$ so that the probability of constraint violation is less than 1%. Source: [33].

$ J_i $	Γ_i
5	5
10	8.2
100	24.3
200	33.9
2000	105

Table 2.1 illustrates, using equation 2.38, the choice of Γ_i as a function of $n = |J_i|$ so that the probability that the i -th constraint is violated is less than 1%. As the total number of uncertain parameters ($n = |J_i|$) in the i -th constraint increases, less amount of uncertain data (Γ_i) is needed to guarantee a violation probability less than 1%.

2.5 Rolling horizon framework

The rolling (or receding) horizon principle has been broadly used in control methods, such as in Model Predictive Control (MPC) as a means to deal with the control problem when a cost function has to be minimized in a given time horizon and the system is subject to constraints [143].

The rolling horizon framework has already been applied to several reactive scheduling problems under uncertainty; Kopanos and Pistikopoulos present in [143] a brief review on this subject. These approaches are based on the use of the so called *receding horizon philosophy*:

"[...] a sequence of future control actions is chosen according to a prediction of the future evolution of the system and applied to the plant until new measurements are available. Then, a new sequence is determined which replaces the previous one." [25].

An optimization procedure evaluates each possible control sequence by taking into account two objectives: minimization of the cost function, and protection of the system against possible constraint violations.

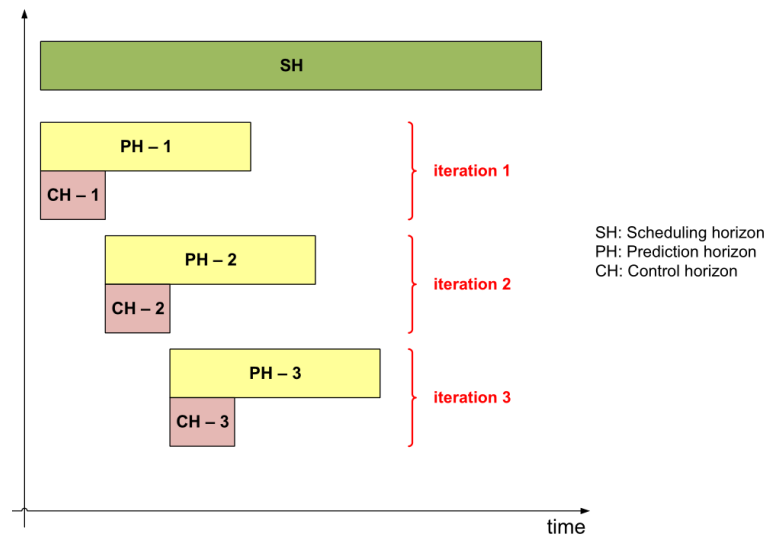


Figure 2.5: Concepts associated to rolling horizon approach: scheduling, prediction and control horizon. Source: [205].

As seen in figure 2.5, the rolling horizon implementation involves the definition of [205]:

- a *prediction horizon* (PH), where all the uncertain parameters are assumed to be known with some certainty¹⁷; and
- a *control horizon* (CH), where the decisions obtained by the optimization procedure are actually applied.

The prediction horizon length depends on the problem under consideration while the control horizon, typically equal to one time interval, is a subset of the prediction horizon [143].

¹⁷ Because the system receives feedback related to these parameters at each time-step update.

For instance, figure 2.6 shows a prediction horizon of three time intervals, while the length of the control horizon is one time interval. Thus, only the first sample of the optimal sequence is actually applied to the system at time-step k . At time-step $k + 1$, a new sequence is evaluated to replace the previous one [25]. This on-line *re-planning* provides the desired feedback control feature [25].

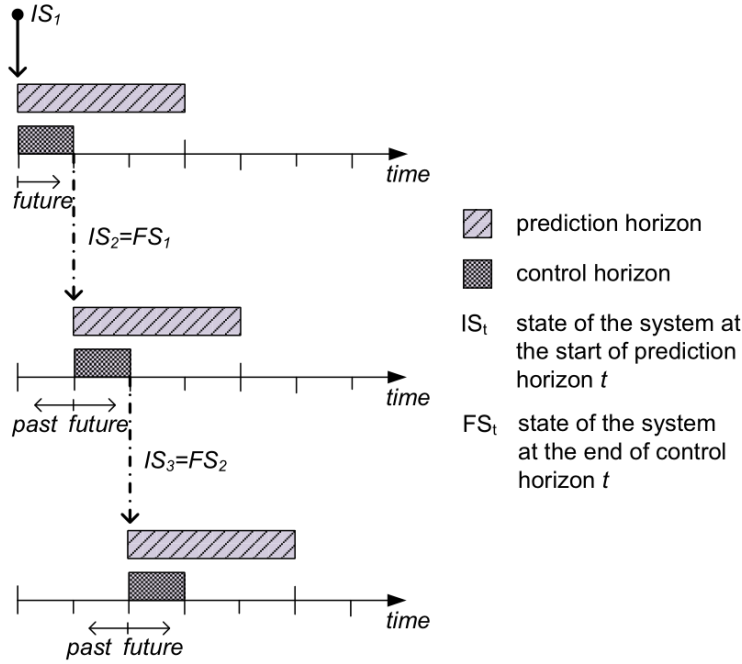


Figure 2.6: Reactive scheduling via a rolling horizon framework. Source: [143].

It is important to note that the initial state of the system in a given prediction horizon \mathcal{H} is equal to the final state of the system in the previous control horizon $\mathcal{H} - 1$ (see figure 2.6). Therefore, the outputs of the previous time-step can be used in an iterative fashion via a rolling horizon framework, such as the one shown in figure 2.5, for addressing reactive scheduling problems [143]. Moreover, the system receives feedback of all input parameters (e.g., actual demand, updated state of system) at every discrete time-step; in this manner, each optimization is done according to the current available information [143].

Figure 2.7 shows a representative algorithm for the rolling horizon by using the output of a MILP optimization problem.

2.5.1 Future horizon impact

Although the solution obtained for each of the prediction horizons is optimal for the corresponding prediction horizon, there is no theoretical guarantee for the solution of the overall scheduling problem to be either optimal or feasible [143]; the reason lies in the underlying

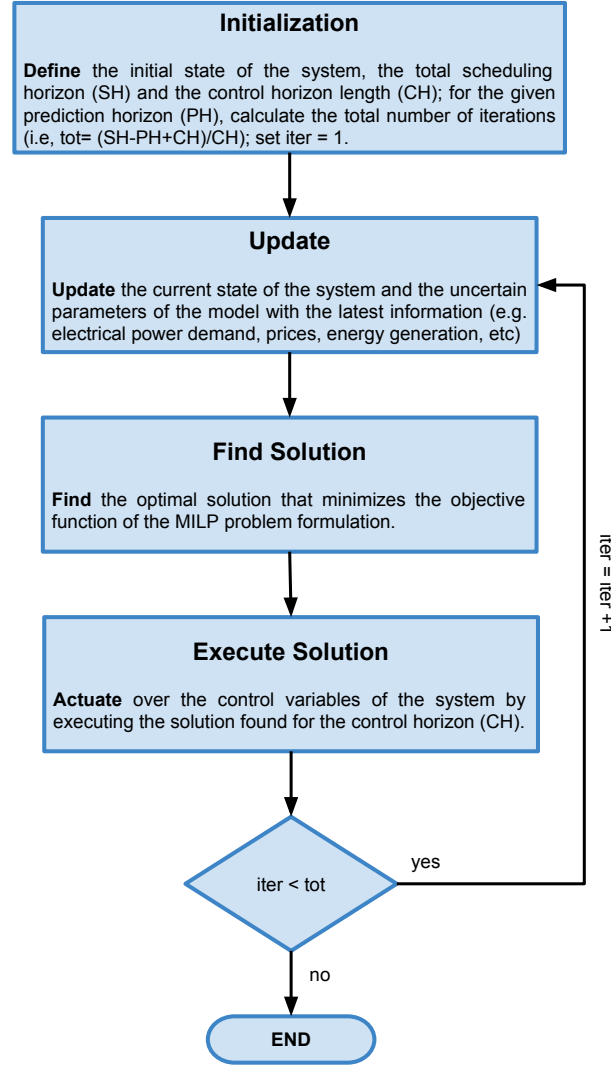


Figure 2.7: Algorithm for the rolling horizon approach. Modified from [143].

fact that no information regarding the future horizon (i.e., outside the prediction horizon) is considered when optimizing the prediction horizons [143].

While long-term optimality may not always be necessary or meaningful, feasibility is of great importance [205]. Therefore, to ensure the feasibility of the solution throughout the planning horizon, longer-term information (outside of the optimized prediction horizon) should be taken into account; keeping this in mind, problem-specific terminal constraints are added so as to ensure that the solution of the finite horizon problem is a feasible and (near-) optimal solution to the infinite horizon problem.

Once proper terminal constraints are defined, the MILP problem formulation can be embedded into an integrated MPC scheme involving problem-specific forecasting techniques (for the uncertain parameters) for improving the performance of the method [143].

2.5.2 Prediction horizon of varied duration

There is the possibility of using horizons of prediction and control of variable lengths within the receding horizon framework. In this way, the choice of these lengths is one more decision that the decision-maker must take; generally they will depend on the characteristics of the system and the forecasts of the future horizon:

- Long prediction and control horizons are recommended for a system with a low dynamic response that can estimate its uncertain parameters within a high degree of certainty for a long horizon;
- Short control horizons and larger prediction horizons are needed for highly dynamic systems that also present great variability and uncertainty in the estimation of its uncertain parameters.

Since the latter best describes the energy scheduling system considered in this study, short control horizons (i.e., one time-step) and larger (if possible) prediction horizons are implemented.

3 Optimal energy storage control strategy

In this chapter, the theory presented throughout chapter 2 is applied to formulate two mathematical models (that is, a deterministic MILP-based rolling horizon controller and its robust counterpart) that solve the problem stated in section 1.2 regarding the robust optimal control of the energy flows in a grid-connected microgrid.

This chapter is organized as follows: section 3.1 presents a multi-term and multi-period cost minimization deterministic model for the microgrid system, where every nonlinear constraint is linearized so that a MILP problem is achieved; section 3.2 takes into account the presence of data uncertainty in the MILP formulation, and each constraint that has at least one non-deterministic parameter is transformed to a min-max robust counterpart, and finally a traceable linear robust counterpart of the constraint is formulated using linear duality theory together with the uncertainty sets defined in chapter 2. The final result is a model that provides optimal solutions while remaining feasible against all realizations of the uncertainties within the predefined bounds of the uncertainty set; moreover, control parameters inherent to the robust model allow to adjust the trade-off between the robustness of the solution and the impact that said robustness has on the optimal solution.

3.1 Deterministic MILP optimization-based MPC

3.1.1 Objective function

The objective of the optimization problem is to minimize the economic costs of operation of the microgrid over a predefined time horizon \mathcal{H} . Therefore, an economic assessment function needs to be defined; consider the multi-term cost function $\phi : \mathbb{R}_+^m \mapsto \mathbb{R}$, then for $x \in \mathbb{R}_+^m$:

$$\phi(x) = \sum_{j=1}^m c_j x_j, \quad (3.1)$$

where $x_1, \dots, x_m \in \mathbb{R}_+$ represent nonnegative decision variables of the optimization problem that, for the purposes of this work, are electrical power flows in the system, measured in kW; depending on the nature of each the power flow, the objective coefficients $c_1, \dots, c_m \in \mathbb{R}$ can represent economic costs, revenues or penalties. Either way, they are all measured in

ct/kWh. Thus, the economic cost of microgrid operation ϕ for a certain set of power flows $x_1, \dots, x_m \in \mathbb{R}_+$ is measured in ct/h.

On the other hand, the rolling horizon controller scheme optimizes a certain cost function $\phi \in \mathbb{R}$ over a horizon \mathcal{H} of $N_h \in \mathbb{N}$ time-steps; in consequence, the general multi-period objective function $\chi : \mathbb{R}^{N_h} \mapsto \mathbb{R}$ is defined, then for $\phi \in \mathbb{R}^{N_h}$:

$$\chi(\phi) = \frac{1}{H} \sum_{k=1}^{N_h} h_k \phi_k, \quad (3.2)$$

where $\phi_1, \dots, \phi_{N_h} \in \mathbb{R}$ are the values of the cost function at each different time-step of the horizon; $h_k \in \mathbb{R}_+$ represents the duration in hours of the k -th time-step in the horizon; and $H \in \mathbb{R}_+$ is the total amount of hours in the horizon, defined as

$$H \triangleq \sum_{k=1}^{N_h} h_k. \quad (3.3)$$

The structure defined for the multi-period objective function in 3.2 allows the implementation of time-steps with different durations over the horizon; the analysis of the effect this has on both the optimal solution and the execution time to solve the problem is carried out in section 5.5.1.

Finally, the general structure of a multi-term multi-period objective function $f_o : \mathbb{R}_+^{N_h \times m} \mapsto \mathbb{R}$ can be obtained by unifying both concepts presented above:

$$f_o(x_1, \dots, x_{N_h}) = \frac{1}{H} \sum_{k=1}^{N_h} \left(h_k \sum_{j=1}^m c_{jk} x_{jk} \right), \quad (3.4)$$

where $x_1, \dots, x_{N_h} \in \mathbb{R}_+^m$ are the vectors containing the m decision variables for each of the N_h time-steps.

At each iteration, the controller solves the optimization problem over the entire rolling horizon problem, i.e., finds the set of values for all the vectors $x_1, \dots, x_{N_h} \in \mathbb{R}_+^m$ such that the objective function f_o is minimized; however, as stated in section 2.5, only the actions corresponding to the first time-step of the rolling horizon window are implemented in practice, i.e., $x_1 \in \mathbb{R}_+^m$; after these actions are executed, the horizon is updated and the controller carries on with the next iteration.

When applied to the system under analysis, the general structure of equation 3.4 yields the following objective function:

$$f_o = \frac{1}{H} \sum_{k=1}^{N_h} h_k \left\{ \underbrace{c_{ss} (p_{c_k} + p_{d_k}) + c_{ss}^{sm} u_{ss_k}}_{\text{energy storage system}} + \underbrace{c_{sh} p_{sh_k}}_{\text{load-shedding}} + \right. \\ \left. \underbrace{+ c_{b_k} p_{b_k} - c_{s_k} p_{s_k} + c_{grid}^{sm} u_{grid_k} + c_{grid}^{peak} p_{grid}^{ob} + c_{grid}^{flat} (p_b^{max} + p_s^{max})}_{\text{main grid}} \right\} \quad (3.5)$$

All the decision variables in this cost function are nonnegative power flows in the micro-grid measured in kW, while all the objective coefficients are measured in ct/kWh.

The *energy storage system* section of f_o represents the usage cost of the storage system and a smoothing term for its power profile; decision variables p_{c_k} and p_{d_k} represent the charging and discharging power of the storage system, respectively; the ESS usage cost c_{ss} is incorporated to the objective to avoid unnecessary activity at the ESS that would otherwise reduce its lifetime¹; the term $c_{ss}^{sm} u_{ss_k}$ is used as a penalization to smooth the power signal, that is, to minimize the differences on the power profile of the storage system over successive horizon time-steps.

The *load-shedding* section is a penalization term to avoid unnecessary interruptions in the electricity supply of certain loads present in the system; p_{sh_k} is a decision variable that represents the amount of load disconnected from the electricity supply, and c_{sh} is the economic cost associated with such decision.

The first two terms that make up the *main grid* section are the net electricity usage cost at the point of common coupling between the microgrid and the main grid; p_{b_k} is the power bought from the grid and p_{s_k} the power sold to the grid, with c_{b_k} and c_{s_k} being the buying and selling prices, respectively. The last three terms of this section comprise the grid signal penalty terms; $c_{grid}^{sm} u_{grid_k}$ smooths the grid signal by minimizing successive differences in the power profile at the point of common coupling; peak shaving is incorporated through the term $c_{grid}^{peak} p_{grid}^{ob}$, which represents an incremental peak usage cost above the baseline p_{grid}^{base} ; finally, the term $c_{grid}^{flat} (p_b^{max} + p_s^{max})$ flattens the power profile at the point of common coupling.

¹ Estimation ways for this parameter through the amortization of the ESS capital cost over a geometric series can be found in [166].

3.1.2 Energy storage system (ESS)

The dynamics and working limits of a general ESS are presented in this section; the ESS is characterized by the following parameters:

- nominal energy capacity in kWh: $E_{ss}^{\text{nom}} \in \mathbb{R}_+$;
- allowed maximum charge or discharge power in kW: $P_{ss}^{\text{max}} \in \mathbb{R}_+$;
- allowed maximum instantaneous power-step in kW: $\Delta P_{ss}^{\text{max}} \in \mathbb{R}_+$;
- charge and discharge efficiencies: $\eta_c, \eta_d \in \mathcal{C}$;
- coefficients for State Of Charge (SOC) maximum and minimum safety levels:
 $SOC^{\min}, SOC^{\max} \in \mathcal{C} : SOC^{\min} < SOC^{\max}$;
- self-discharge loss expressed in kW²: $P_{ss}^{\text{loss}} \in \mathbb{R}_+$.

These parameters will depend on the type of technology, capacity, and low-hierarchy electronic control system of the ESS installed in the microgrid (see section 6.2.3).

With these parameters, the safe operational lower and upper limits for the energy level (in kWh) of the ESS can be defined:

$$E_{ss}^{\min} \triangleq SOC^{\min} E_{ss}^{\text{nom}} \quad (3.6a)$$

$$E_{ss}^{\max} \triangleq SOC^{\max} E_{ss}^{\text{nom}} \quad (3.6b)$$

State of Charge

A discrete-time model for the ESS of the form

$$E_{ssk} = E_{ssk-1} + h_k \left(\eta_c p_{c_k} - \eta_d^{-1} p_{d_k} - P_{ss}^{\text{loss}} \right); \forall k \in \mathcal{H} \quad (3.7)$$

is employed in this work, where E_{ssk} represents the amount of energy in kWh present in the ESS at the time-step k of the horizon; h_k is the length of the time-step k in hours; and p_{c_k} and p_{d_k} are two nonnegative optimization variables which indicate, in kW, the average power charged and discharged from the ESS during every time-step k , respectively.

² The self-discharge of the ESS is expressed in amount of energy lost per hour, which is kWh/h, which in turn equals kW.

The amount of energy in the ESS present at time-step k is physically limited by its capacity, i.e., by the bounds defined in equation 3.6; thus, the following constraint ensures that, given the model of equation 3.7, and for every time-step in the horizon, the ESS remains within the safe operational bounds defined in equation 3.6:

$$E_{ss}^{\min} \leq E_{ss}^o + \sum_{i=1}^k h_i \left(\eta_c p_{c_i} - \eta_d^{-1} p_{d_i} - P_{ss}^{\text{loss}} \right) \leq E_{ss}^{\max}; \forall k \in \mathcal{H}, \quad (3.8)$$

where E_{ss}^o is the energy level of the ESS at the beginning of the horizon.

The desired energy level of the ESS at the end of the horizon ($E_{ss}^{\text{final}} \in [E_{ss}^{\min}, E_{ss}^{\max}]$) can be specified through the following constraint:

$$E_{ss}^{\text{final}} (1 - \varepsilon) \leq E_{ss}^o + \sum_{i=1}^{N_h} h_i \left(\eta_c p_{c_i} - \eta_d^{-1} p_{d_i} - P_{ss}^{\text{loss}} \right) \leq E_{ss}^{\text{final}} (1 + \varepsilon). \quad (3.9)$$

where $\varepsilon \in \mathcal{C}$ is the tolerance allowed for the deviation from E_{ss}^{final} ; this constraint can be disabled by setting $\varepsilon = 1$. Without this constraint, there is a chance that, for a certain set of problem parameters, the optimal solution involves a charging/discharging profile that drains all the stored energy in the ESS by the end of the horizon. However, this chance is negligible, since only the actions of the first time-step of the horizon are carried out before changing the horizon and performing a new optimization.

Binary control variables, maximum power and energy availability

As exposed in section 2.2.1, binary variables can be used in the model to express logic statements and if-then-else rules. The optimization problem within the rolling horizon controller makes decisions as to whether charge or discharge the ESS at any given time-step; to this end, the following control variables for the ESS are defined:

$$\delta_{c_k}, \delta_{d_k} \in \mathcal{B}; \forall k \in \mathcal{H}. \quad (3.10)$$

A common microgrid architecture comprises only one bidirectional power converter connected to the ESS; moreover, there is no logical sense in charging and discharging the ESS during the same time-step. In consequence, the following restriction is enforced on the model:

$$\delta_{c_k} + \delta_{d_k} = 1; \forall k \in \mathcal{H}, \quad (3.11)$$

which states that during the entire time-step k , the ESS is either charged or discharged³.

These binary control variables are also useful for setting the maximum allowed power for charging and discharging the ESS at each time-step, i.e.:

$$0 \leq p_{c_k} \leq \delta_{c_k} P_{ss}^{\max}; \forall k \in \mathcal{H}, \quad (3.12)$$

$$0 \leq p_{d_k} \leq \delta_{d_k} P_{ss}^{\max}; \forall k \in \mathcal{H}. \quad (3.13)$$

If either the charging or the discharging of the ESS is enabled, the maximum power that can be exchanged is limited by the parameter P_{ss}^{\max} ⁴.

Furthermore, these binary control variables are used for setting the energy availability during charging and discharging operations of the ESS at each time-step, i.e.:

$$0 \leq p_{c_k} h_k \leq [(E_{ss}^{\max} - E_{ss_{k-1}}) \eta_c^{-1}] \delta_{c_k}; \forall k \in \mathcal{H}, \quad (3.14)$$

$$0 \leq p_{d_k} h_k \leq [(E_{ss_{k-1}} - E_{ss}^{\min}) \eta_d] \delta_{d_k}; \forall k \in \mathcal{H}. \quad (3.15)$$

If $\delta_{d_k} = 1$ discharging the ESS in time-step k is enabled, and equation 3.15 states that the maximum amount of energy than can be withdrawn from the ESS equals all the energy available in the ESS until reaching the E_{ss}^{\min} energy level; if $\delta_{d_k} = 0$ then the constraint specifies $0 \leq p_{d_k} \leq 0$, meaning that the discharge of the ESS is forbidden during time-step k . The same analysis can be carried out for the charging constraint in equation 3.14.

For every time-step k , both energy availability constraints have in the upper bound a product of two decision variables (i.e., $E_{ss_{k-1}} \delta_{c_k}$ for the charge constraint and $E_{ss_{k-1}} \delta_{d_k}$ for the discharge constraint). In consequence, $2N_h$ nonlinear constraints are being introduced into the model. Below, the linearization transformations for these constraints are presented, according to the methods described in section 2.2.2.

Linearization of the ESS charging constraint The bilinear term in equation 3.14 is replaced with the introduction an auxiliary real variable $y_{c_k} \triangleq E_{ss_{k-1}} \delta_{c_k}, \forall k \in \mathcal{H}$; furthermore, upper and lower bounds are defined for the continuous decision variable involved in the nonlinearity such that $E_{ss_{k-1}} \in [E_{ss}^{\min}, E_{ss}^{\max}]$ ⁵. Then, the set of bilinear constraints in equation 3.14 is equivalent to the following set of linear constraints:

$$0 \leq p_{c_k} h_k \leq (E_{ss}^{\max} \delta_{c_k} - y_{c_k}) \eta_c^{-1}; \forall k \in \mathcal{H} \quad (3.16a)$$

³ It could happen that the optimal decision for time-step k is that the ESS is neither charged nor discharged. In this case $p_{c_k} = p_{d_k} = 0$ and it does not really matter which of the two control variables is enabled.

⁴ This value is usually imposed by the DC/DC converter of the ESS.

⁵ In this case, the bounds match the ones defined in equation 3.6.

$$y_{c_k} \leq E_{ss}^{\max} \delta_{c_k} \quad ; \forall k \in \mathcal{H} \quad (3.16b)$$

$$y_{c_k} \geq E_{ss}^{\min} \delta_{c_k} \quad ; \forall k \in \mathcal{H} \quad (3.16c)$$

$$y_{c_k} \leq E_{ss_{k-1}} - E_{ss}^{\min} (1 - \delta_{c_k}) \quad ; \forall k \in \mathcal{H} \quad (3.16d)$$

$$y_{c_k} \geq E_{ss_{k-1}} - E_{ss}^{\max} (1 - \delta_{c_k}) \quad ; \forall k \in \mathcal{H} \quad (3.16e)$$

Linearization of the ESS discharging constraint The bilinear term in equation 3.15 is replaced with the introduction of an auxiliary real variable $y_{d_k} \triangleq E_{ss_{k-1}} \delta_{d_k}$, $\forall k \in \mathcal{H}$; the continuous decision variable involved in the nonlinearity is the same one as in the charging constraint, so the upper and lower bounds already defined remain valid. Then, the set of bilinear constraints in equation 3.15 is equivalent to the following set of linear constraints:

$$0 \leq p_{d_k} h_k \leq (y_{d_k} - E_{ss}^{\min} \delta_{d_k}) \eta_d \quad ; \forall k \in \mathcal{H} \quad (3.17a)$$

$$y_{d_k} \leq E_{ss}^{\max} \delta_{d_k} \quad ; \forall k \in \mathcal{H} \quad (3.17b)$$

$$y_{d_k} \geq E_{ss}^{\min} \delta_{d_k} \quad ; \forall k \in \mathcal{H} \quad (3.17c)$$

$$y_{d_k} \leq E_{ss_{k-1}} - E_{ss}^{\min} (1 - \delta_{d_k}) \quad ; \forall k \in \mathcal{H} \quad (3.17d)$$

$$y_{d_k} \geq E_{ss_{k-1}} - E_{ss}^{\max} (1 - \delta_{d_k}) \quad ; \forall k \in \mathcal{H} \quad (3.17e)$$

Maximum power-step and power signal smoothness

The last constraints introduced for the ESS address the maximum power-step allowed between consecutive time-steps⁶, as well as a penalization variable to introduce control over the smoothness of the power signal in the ESS:

$$|(p_{c_k} - p_{d_k}) - (p_{c_{k-1}} - p_{d_{k-1}})| \leq u_{ss_k} \leq \Delta P_{ss}^{\max} ; \forall k \in \mathcal{H}. \quad (3.18)$$

The linearized form of this constraint is defined as follows:

$$-\Delta P_{ss}^{\max} \leq -u_{ss_k} \leq (p_{c_k} - p_{d_k}) - (p_{c_{k-1}} - p_{d_{k-1}}) \leq u_{ss_k} \leq \Delta P_{ss}^{\max} ; \forall k \in \mathcal{H}, \quad (3.19)$$

where u_{ss_k} is a nonnegative optimization variable for the time-step k that allows to control (through the election of the penalization cost c_{ss}^{sm} in the objective function) the level of

⁶ This is usually limited by the DC/DC converter of the ESS.

smoothness of the power signal of the ESS; smoothing the power profile reduces transients that could otherwise be harmful to the life of the ESS.

Once again, initial conditions are needed for the rolling horizon control:

$$k = 1 \rightarrow \begin{cases} p_{c_{k-1}} = P_c^o \\ p_{d_{k-1}} = P_d^o \end{cases} \quad (3.20)$$

where P_c^o and P_d^o are the charging and discharging power (in kW) for the last action taken by the control system (i.e., at the time-step before the beginning of the current horizon), respectively.

3.1.3 Interaction between the microgrid and the main grid

Physically there exists only one point of common coupling between the microgrid and the main grid; this bidirectional switch is actuated by the controller, and can be either connected or disconnected; in the former, the microgrid is allowed to exchange power with the grid, by either buying or selling energy; in the latter, the microgrid is isolated from the main grid, and is said to be in *island mode*.

By controlling the power flow at the point of common coupling, the controller is able to shape the power profile; this allows to generate a more predictable and stable load response from the main grid perspective [166]. In this way, the introduction of a microgrid topology does not disrupt the current scheme of energy distribution and control in the main grid.

Binary control variables

The logical variables

$$\delta_{b_k}, \delta_{s_k} \in \mathcal{B}; \forall k \in \mathcal{H} \quad (3.21)$$

are defined to control the power exchange with the main grid; these variables interact with each other through the following constraint:

$$\delta_{b_k} + \delta_{s_k} \leq 1; \forall k \in \mathcal{H}, \quad (3.22)$$

which means that for any time-step k the microgrid can either be buying energy from the main grid $[(\delta_{b_k} = 1) \wedge (\delta_{s_k} = 0)]$; selling energy to the main grid $[(\delta_{b_k} = 0) \wedge (\delta_{s_k} = 1)]$; or in island mode $[(\delta_{b_k} = 0) \wedge (\delta_{s_k} = 0)]$.

Maximum power exchanged with the grid and signal flattening

The power transfer between the main grid and the microgrid will present constraints related to the physical limitations of the transmission line infrastructure; these boundaries can be made variable in the optimization model, so that the power profile can be flattened or squeezed by the controller. In consequence, it is defined:

$$0 \leq p_{b_k} \leq p_b^{\max} \delta_{b_k}; \forall k \in \mathcal{H}, \quad (3.23)$$

$$0 \leq p_{s_k} \leq p_s^{\max} \delta_{s_k}; \forall k \in \mathcal{H}, \quad (3.24)$$

where p_b^{\max} and p_s^{\max} are nonnegative optimization variables corresponding to the maximum grid power rate.

For every time-step k , both grid signal flattening constraints have in the upper bound a product of two decision variables (i.e., $p_b^{\max} \delta_{b_k}$ for the buying constraint and $p_s^{\max} \delta_{s_k}$ for the selling constraint). In consequence, $2N_h$ nonlinear constraints are being introduced into the model. Below, the linearization transformations for these constraints are presented, according to the methods described in section 2.2.2.

Linearization of the buying constraint The bilinear term in equation 3.23 is replaced with the introduction an auxiliary real variable $y_{b_k} \triangleq p_b^{\max} \delta_{b_k}$, $\forall k \in \mathcal{H}$; furthermore, upper and lower bounds are defined for the continuous decision variable involved in the nonlinearity such that $p_b^{\max} \in [0, P_{\text{grid}}^{\max}]$ ⁷. Then, the set of bilinear constraints in equation 3.23 is equivalent to the following set of linear constraints:

$$0 \leq p_{b_k} \leq y_{b_k} \quad ; \forall k \in \mathcal{H} \quad (3.25a)$$

$$y_{b_k} \leq P_{\text{grid}}^{\max} \delta_{b_k} \quad ; \forall k \in \mathcal{H} \quad (3.25b)$$

$$y_{b_k} \geq 0 \quad ; \forall k \in \mathcal{H} \quad (3.25c)$$

$$y_{b_k} \leq p_b^{\max} \quad ; \forall k \in \mathcal{H} \quad (3.25d)$$

$$y_{b_k} \geq p_b^{\max} - P_{\text{grid}}^{\max} (1 - \delta_{b_k}) \quad ; \forall k \in \mathcal{H} \quad (3.25e)$$

Linearization of the selling constraint The bilinear term in equation 3.24 is replaced with the introduction an auxiliary real variable $y_{s_k} \triangleq p_s^{\max} \delta_{s_k}$, $\forall k \in \mathcal{H}$; furthermore, upper

⁷ The bounds for p_b^{\max} are given by the physical limits for the amount of power that can be exchanged at the point of common coupling and thorough the transmission lines belonging to the main grid. The minimum value is set to zero since the decision variable for buying power from the grid is nonnegative.

and lower bounds are defined for the continuous decision variable involved in the nonlinearity such that $p_s^{\max} \in [0, P_{\text{grid}}^{\max}]$ ⁸. Then, the set of bilinear constraints in equation 3.24 is equivalent to the following set of linear constraints:

$$0 \leq p_{s_k} \leq y_{s_k} \quad ; \forall k \in \mathcal{H} \quad (3.26a)$$

$$y_{s_k} \leq P_{\text{grid}}^{\max} \delta_{s_k} \quad ; \forall k \in \mathcal{H} \quad (3.26b)$$

$$y_{s_k} \geq 0 \quad ; \forall k \in \mathcal{H} \quad (3.26c)$$

$$y_{s_k} \leq p_s^{\max} \quad ; \forall k \in \mathcal{H} \quad (3.26d)$$

$$y_{s_k} \geq p_s^{\max} - P_{\text{grid}}^{\max} (1 - \delta_{s_k}) \quad ; \forall k \in \mathcal{H} \quad (3.26e)$$

Grid signal smoothing

A penalization constraint of the form

$$|(p_{b_k} - p_{s_k}) - (p_{b_{k-1}} - p_{s_{k-1}})| \leq u_{\text{grid}_k} ; \forall k \in \mathcal{H}, \quad (3.27)$$

with a linearized expression equal to

$$-u_{\text{grid}_k} \leq (p_{b_k} - p_{s_k}) - (p_{b_{k-1}} - p_{s_{k-1}}) \leq u_{\text{grid}_k} ; \forall k \in \mathcal{H} \quad (3.28)$$

is introduced as a way of controlling how smooth is the transition between time-steps of the power signal at the point of common coupling; u_{grid_k} is a nonnegative optimization variable for the time-step k that allows to control (through the election of the penalization cost $c_{\text{grid}}^{\text{sm}}$ in the objective function) how smooth the power profile is. Once again, initial conditions are needed for the rolling horizon control:

$$k = 1 \rightarrow \begin{cases} p_{b_{k-1}} = P_b^o \\ p_{s_{k-1}} = P_s^o \end{cases} \quad (3.29)$$

where P_b^o and P_s^o are the bought and sold power (in kW) for the last action taken by the control system (i.e., at the time-step before the beginning of the current horizon), respectively.

⁸ The infrastructure involved in both types of power exchange between the microgrid and the main grid (buying and selling) is the same. In consequence, the limits for the power exchange at the selling state are the same as those defined for the linearization of the buying constraint.

Peak shaving

By defining a constant power baseline $P_{\text{grid}}^{\text{base}}$ in kW, a constraint can be enforced to penalize any power exchange that produces a peak over this baseline:

$$p_{b_k} + p_{s_k} \leq P_{\text{grid}}^{\text{base}} + p_{\text{grid}}^{\text{ob}}; \forall k \in \mathcal{H}, \quad (3.30)$$

where $p_{\text{grid}}^{\text{ob}}$ is a global nonnegative decision variable (i.e., the variable is unique for the whole horizon) that specifies the amount of power exchanged over the baseline; a power spike over the baseline at any given point of the horizon produces a penalty cost of $c_{\text{grid}}^{\text{peak}} p_{\text{grid}}^{\text{ob}}$ in the objective function.

3.1.4 Load balance

At every time-step k of the horizon, the energy flow between all the actors in the system has to be balanced according to the following equation:

$$(p_{b_k} - p_{s_k}) + (p_{d_k} - p_{c_k}) + P_{g_k} + p_{sh_k} = P_{l_k}; \forall k \in \mathcal{H}, \quad (3.31)$$

where p_{sh_k} is a nonnegative optimization variable for controlling the amount of load-shedding⁹ during time-step k ; P_{l_k} is a nonnegative parameter that indicates the average demand of power in kW of all the consumers inside the microgrid during time-step k ; and P_{g_k} is a nonnegative parameter that indicates the average power output in kW during the time-step k of the RES present in the microgrid¹⁰.

Load shedding

The act of load-shedding produces a disturbance in the normal functioning of the affected equipment and the activities of the people who depend on them. Therefore, it is considered an undesired action in the normal operation of the microgrid and, according to the following constraint, it will only be allowed when the network is in island mode ($\delta_{b_k} + \delta_{s_k} = 0$):

$$0 \leq p_{sh_k} \leq P_{l_k} [1 - (\delta_{b_k} + \delta_{s_k})]; \forall k \in \mathcal{H}. \quad (3.32)$$

⁹ Load-shedding refers to the reduction of the amount of load present in the microgrid by interrupting the electricity supply to certain consumers in the system.

¹⁰ As explained in section 1.1, the RES used in this thesis is the non-dispatchable regenerative braking of trains approaching the passenger stations in the vicinity of the Neuostheim electrical substation. Nonetheless, this parameter of the problem can be generalized to any type of RES; renewable or not; dispatchable or not (if the energy source is dispatchable, the parameter becomes an optimization variable).

If the microgrid is connected to the main grid ($\delta_{b_k} + \delta_{s_k} = 1$), the constraint becomes $p_{sh_k} = 0$. Furthermore, the amount of load-shedding during time-step k can never be larger than the amount of load present at the same time-step.

If the load-shedding option is implemented at time-step k , an economic penalty of $c_{sh}p_{sh_k}$ is added to the objective function of the controller. The value of the economic cost c_{sh} is intimately related to the type of equipment connected to the microgrid and how critical is that it always has a stable electrical supply; due to the high penalty cost of load shedding, during islanded mode service reliability becomes more essential than operation cost [112].

3.1.5 Deterministic MILP problem formulation

The MILP optimization-based MPC problem developed throughout this section can be summarized as follows:

$$\begin{aligned} & \text{minimize} && 3.5 \\ & \text{subject to} && 3.8, 3.9, 3.11, 3.12, 3.13, 3.16, 3.17, 3.19, \\ & && 3.22, 3.25, 3.26, 3.28, 3.30, 3.31, 3.32. \end{aligned} \quad (3.33)$$

In order to minimize the objective, the MILP optimization problem 3.33 seeks to find the optimal values for the following nonnegative decision variables:

$$\begin{cases} p_{c_k}, p_{d_k}, p_{b_k}, p_{s_k}, p_{sh_k}, u_{ss_k}, u_{grid_k}, y_{c_k}, y_{d_k}, y_{b_k}, y_{s_k} & \in \mathbb{R}_+, \forall k \in \mathcal{H}; \\ \delta_{c_k}, \delta_{d_k}, \delta_{b_k}, \delta_{s_k} & \in \mathcal{B}, \forall k \in \mathcal{H}; \\ p_b^{max}, p_s^{max}, p_{grid}^{ob} & \in \mathbb{R}_+. \end{cases} \quad (3.34)$$

Thus, the MILP optimization problem 3.33 has $(15N_h + 3)$ decision variables and $(40N_h + 1)$ constraints.

3.2 Robust optimization approach

This section deals with the derivation of the robust counterpart of the deterministic MILP optimization problem defined in equation 3.33; to do so, once the sources of data uncertainty are identified, the theory presented in section 2.4 is used to transform each constraint with uncertain data to its robust counterpart.

3.2.1 Uncertain parameters in the model

Below are the model parameters that are assumed to have some component of uncertainty in their values; this uncertainty is represented with the structure shown in equation 2.23:

- price for buying energy from the main grid: $\tilde{c}_{b_k}, \forall k \in \mathcal{H}$;
- price for selling energy to the main grid: $\tilde{c}_{s_k}, \forall k \in \mathcal{H}$;
- energy generation from the regenerative braking of the trains: $\tilde{P}_{g_k}, \forall k \in \mathcal{H}$;
- demand of energy inside the microgrid (loads at the train stations): $\tilde{P}_{l_k}, \forall k \in \mathcal{H}$;

3.2.2 Previous steps

Following the guidelines of section 2.4.3, as a previous step to finding the robust counterpart, the problem formulation is transformed so that all the uncertain parameters are in the left hand side of the constraints; because some of the uncertain parameters defined in section 3.2.1 are in the objective function f_o , the optimization variable $q \in \mathbb{R}$ is defined so that problem 3.33 is cast as:

$$\begin{aligned}
 & \text{minimize} && q \\
 & \text{subject to} && f_o - q \leq 0, \\
 & && 3.8, 3.9, 3.11, 3.16, 3.17, 3.19, 3.22 \\
 & && 3.25, 3.26, 3.28, 3.30, 3.31, 3.32.
 \end{aligned} \tag{3.35}$$

3.2.3 Robust counterpart of the cost constraint

The new cost constraint introduced in section 3.2.2 (i.e., $f_o - q \leq 0$) presents $2N_h$ uncertain parameters, i.e., $\tilde{c}_{b_k}, \tilde{c}_{s_k} \forall k \in \mathcal{H}$; as explained in section 2.4.3, a constraint that contains uncertain parameters as defined in equation 2.23 can be split into a deterministic part and an uncertain part (see equation 2.24); thus, for non-deterministic electricity buying and selling prices the cost function defined in equation 3.5 can be rewritten as:

$$\tilde{f}_o = f_o + \frac{1}{H} \sum_{k=1}^{N_h} h_k (\xi_{b_k} \hat{c}_{b_k} p_{b_k} - \xi_{s_k} \hat{c}_{s_k} p_{s_k}) \tag{3.36}$$

where the deterministic term f_o remains as defined in equation 3.5. Then, the robust counterpart of the cost constraint is defined as:

$$f_o - q + \max_{\xi \in \mathcal{U}} \left\{ \frac{1}{H} \sum_{k=1}^{N_h} h_k (\xi_{b_k} \hat{c}_{b_k} p_{b_k} - \xi_{s_k} \hat{c}_{s_k} p_{s_k}) \right\} \leq 0. \quad (3.37)$$

For a *interval+polyhedral* uncertainty set, the transformation 2.34 holds; thus, the robust counterpart 3.37 of the cost constraint is equivalent to:

$$f_o + \sum_{k=1}^{N_h} (w_{b_k} + w_{s_k}) + \Gamma_{bs} z_{bs} - q \leq 0; \quad (3.38a)$$

$$z_{bs} + w_{b_k} \geq \frac{1}{H} h_k \hat{c}_{b_k} p_{b_k}, \quad \forall k \in \mathcal{H}; \quad (3.38b)$$

$$z_{bs} + w_{s_k} \geq \frac{1}{H} h_k \hat{c}_{s_k} p_{s_k}, \quad \forall k \in \mathcal{H}; \quad (3.38c)$$

$$w_{b_k} \geq 0, \quad \forall k \in \mathcal{H}; \quad (3.38d)$$

$$w_{s_k} \geq 0, \quad \forall k \in \mathcal{H}; \quad (3.38e)$$

$$z_{bs} \geq 0; \quad (3.38f)$$

where $\Gamma_{bs} \in \mathbb{R} \cap [1, 2N_h]$ is a parameter controlled by the decision-maker to adjust the level of conservatism of the optimal solution, that is, how robust the solution is against uncertainties in the price signals. This robust counterpart adds to the problem $2N_h + 1$ nonnegative optimization variables and $2N_h$ linear constraints.

As pointed out by Leiras *et al.* in [151], another possible approach is to consider two different uncertainty sets for the cost constraint, one for each of the price signals:

$$f_o + \sum_{k=1}^{N_h} (w_{b_k} + w_{s_k}) + \Gamma_b z_b + \Gamma_s z_s - q \leq 0; \quad (3.39a)$$

$$z_b + w_{b_k} \geq \frac{1}{H} h_k \hat{c}_{b_k} p_{b_k}, \quad \forall k \in \mathcal{H}; \quad (3.39b)$$

$$z_s + w_{s_k} \geq \frac{1}{H} h_k \hat{c}_{s_k} p_{s_k}, \quad \forall k \in \mathcal{H}; \quad (3.39c)$$

$$w_{b_k} \geq 0, \quad \forall k \in \mathcal{H}; \quad (3.39d)$$

$$w_{s_k} \geq 0, \quad \forall k \in \mathcal{H}; \quad (3.39e)$$

$$z_b \geq 0; \quad (3.39f)$$

$$z_s \geq 0. \quad (3.39g)$$

In this case, the decision-maker is provided with two parameters $\Gamma_s, \Gamma_b \in \mathbb{R} \cap [1, N_h]$, that allow independent control of the conservatism level for both price signals.

This approach to the robust counterpart adds to the problem $2N_h + 2$ nonnegative optimization variables and $2N_h$ linear constraints.

3.2.4 Robust counterpart of the load balance constraint

Constraint 3.31 defines the load balance of the microgrid at each time-step of the horizon \mathcal{H} ; thus, N_h load balance constraints exist in the problem, each one with two parameters prone to uncertainty ($\tilde{P}_{g_k}, \tilde{P}_{l_k} \in \mathbb{R}_+$).

Following the transformation introduced in equation 2.20, the auxiliary variables $x_l = -1$ and $x_g = 1$ are defined so that the constraint 3.31 can be rewritten as:

$$(p_{b_k} - p_{s_k}) + (p_{d_k} - p_{c_k}) + p_{sh_k} + \tilde{P}_{g_k}x_g + \tilde{P}_{l_k}x_l = 0 \quad ; \forall k \in \mathcal{H}, \quad (3.40)$$

and the robust counterpart is derived as follows:

$$(p_{b_k} - p_{s_k}) + (p_{d_k} - p_{c_k}) + p_{sh_k} + P_{g_k}x_g + P_{l_k}x_l + \max_{\xi \in \mathcal{U}} \{ \xi_{g_k} \hat{P}_{g_k}x_g + \xi_{l_k} \hat{P}_{l_k}x_l \} = 0 \quad ; \forall k \in \mathcal{H}. \quad (3.41)$$

For a *interval+polyhedral* uncertainty set, the transformation 2.34 holds; thus, the robust counterpart 3.41 of the load balance constraint is equivalent to:

$$(p_{b_k} - p_{s_k}) + (p_{d_k} - p_{c_k}) + p_{sh_k} + w_{l_k} + w_{g_k} + \Gamma_{lb_k} z_{lb_k} = P_{l_k} - P_{g_k} \quad , \forall k \in \mathcal{H}; \quad (3.42a)$$

$$z_{lb_k} + w_{l_k} \geq \hat{P}_{l_k} \quad , \forall k \in \mathcal{H}; \quad (3.42b)$$

$$z_{lb_k} + w_{g_k} \geq \hat{P}_{g_k} \quad , \forall k \in \mathcal{H}; \quad (3.42c)$$

$$z_{lb_k} \geq 0 \quad , \forall k \in \mathcal{H}; \quad (3.42d)$$

$$w_{g_k} \geq 0 \quad , \forall k \in \mathcal{H}; \quad (3.42e)$$

$$w_{l_k} \geq 0 \quad , \forall k \in \mathcal{H}; \quad (3.42f)$$

where $\Gamma_{lb_k} \in \mathbb{R} \cap [1, 2]$ is a parameter controlled by the decision-maker to adjust the level of conservatism of the optimal solution, that is, how robust the solution is against uncertainties in the price signals.

This robust counterpart adds to the problem $3N_h$ nonnegative optimization variables and $2N_h$ linear constraints. Note for this set of constraints exist N_h optimization parameters (i.e., Γ_{lb_k}); the decision-maker can either set the same value for all of the parameters or treat them separately; the latter could prove to be useful for increasing the level of robustness for time-steps far on the horizon, as it is more likely that the uncertainty of the predictions is greater there.

3.2.5 Robust counterpart of the load-shedding constraint

Constraint 3.32 defines the load-shedding limitations at each time-step of the horizon \mathcal{H} ; thus, N_h load-shedding constraints exist in the problem, each one with one uncertain parameter (i.e. $\tilde{P}_{l_k} \in \mathbb{R}_+$) that multiplies two logical variables and a constant.

Following the transformation introduced in equation 2.20, the auxiliary variable $x_l = -1$ is defined so that the right term of constraint 3.32 can be rewritten as:

$$p_{sh_k} + \tilde{P}_{l_k} x_l + \tilde{P}_{l_k} \delta_{b_k} + \tilde{P}_{l_k} \delta_{s_k} \leq 0; \forall k \in \mathcal{H}, \quad (3.43)$$

and the robust counterpart is derived is as follows:

$$p_{sh_k} + P_{l_k} x_l + P_{l_k} \delta_{b_k} + P_{l_k} \delta_{s_k} + \max_{\xi \in \mathcal{U}} \{ \xi_{l_k} \hat{P}_{l_k} x_l + \xi_{l_k} \hat{P}_{l_k} \delta_{b_k} + \xi_{l_k} \hat{P}_{l_k} \delta_{s_k} \} \leq 0; \forall k \in \mathcal{H}. \quad (3.44)$$

These uncertainties can be treated like three uncertain coefficients (one for each decision variable) with a known correlation equivalent to one; thus, the correlated interval+polyhedral uncertainty set is used. For this set the transformation defined in equation 2.36 holds; the robust counterpart 3.44 of the load-shedding constraint is equivalent to the following set of linear constraints:

$$p_{sh_k} + \Gamma_{sh_k} (z_{shc_k} + z_{shb_k} + z_{shs_k}) + (w_{lc_k} + w_{lb_k} + w_{ls_k}) \leq P_{l_k} [1 - (\delta_{b_k} + \delta_{s_k})], \forall k \in \mathcal{H}; \quad (3.45a)$$

$$z_{shc_k} + \left(\frac{\Gamma_{sh_k} - 1}{2} \right) (z_{shb_k} + z_{shs_k}) + w_{lc_k} \geq \hat{P}_{l_k}, \forall k \in \mathcal{H}; \quad (3.45b)$$

$$z_{shb_k} + \left(\frac{\Gamma_{sh_k} - 1}{2} \right) (z_{shc_k} + z_{shs_k}) + w_{lb_k} \geq \hat{P}_{l_k} \delta_{b_k}, \forall k \in \mathcal{H}; \quad (3.45c)$$

$$z_{shs_k} + \left(\frac{\Gamma_{sh_k} - 1}{2} \right) (z_{shb_k} + z_{shc_k}) + w_{ls_k} \geq \hat{P}_{l_k} \delta_{s_k}, \forall k \in \mathcal{H}; \quad (3.45d)$$

$$z_{shc_k} \geq 0, \forall k \in \mathcal{H}; \quad (3.45e)$$

$$z_{shb_k} \geq 0, \forall k \in \mathcal{H}; \quad (3.45f)$$

$$z_{shs_k} \geq 0, \forall k \in \mathcal{H}; \quad (3.45g)$$

$$w_{lc_k} \geq 0, \forall k \in \mathcal{H}; \quad (3.45h)$$

$$w_{lb_k} \geq 0, \forall k \in \mathcal{H}; \quad (3.45i)$$

$$w_{ls_k} \geq 0, \forall k \in \mathcal{H}; \quad (3.45j)$$

where $\Gamma_{sh_k} \in \mathbb{R} \cap [1, 3]$ is a parameter controlled by the decision-maker to adjust the level of conservatism of the optimal solution, that is, how robust the solution is against uncertainties the load demand signal.

This robust counterpart adds to the problem $6N_h$ nonnegative optimization variables and $3N_h$ linear constraints. Note for this set of constraints exist N_h optimization parameters (i.e., Γ_{sh_k}); the decision-maker can either set the same value for all of the parameters or treat them separately; the latter could prove to be useful for increasing the level of robustness for time-steps far on the horizon, as it is more likely that the uncertainty of the predictions is greater there.

3.2.6 Robust MILP problem formulation

The robust counterpart of the MILP optimization-based MPC problem of equation 3.35 is obtained by replacing each constraint that has non-deterministic parameters with its robust counterpart; thus, the robust counterpart is defined as:

$$\begin{aligned} & \text{minimize} && q \\ & \text{subject to} && 3.8, 3.9, 3.11, 3.12, 3.13, 3.16, 3.17, 3.19, \\ & && 3.22, 3.25, 3.26, 3.28, 3.30, 3.39, 3.42, 3.45. \end{aligned} \quad (3.46)$$

In order to minimize the objective, the robust MILP optimization problem 3.46 seeks to find the optimal values for the following decision variables:

$$\left\{ \begin{array}{ll} p_{c_k}, p_{d_k}, p_{b_k}, p_{s_k}, p_{sh_k}, u_{ss_k}, u_{grid_k}, y_{c_k}, y_{d_k}, y_{b_k}, y_{s_k} & \in \mathbb{R}_+, \forall k \in \mathcal{H}; \\ \delta_{c_k}, \delta_{d_k}, \delta_{b_k}, \delta_{s_k} & \in \mathcal{B}, \forall k \in \mathcal{H}; \\ p_b^{max}, p_s^{max}, p_{grid}^{ob} & \in \mathbb{R}_+; \\ q & \in \mathbb{R}; \\ w_{b_k}, w_{s_k}, w_{l_k}, w_{g_k}, w_{lc_k}, w_{lb_k}, w_{ls_k} & \in \mathbb{R}_+, \forall k \in \mathcal{H}; \\ z_{lb_k}, z_{shc_k}, z_{shb_k}, z_{shs_k} & \in \mathbb{R}_+, \forall k \in \mathcal{H}; \\ z_b, z_s & \in \mathbb{R}_+. \end{array} \right. \quad (3.47)$$

Thus, the MILP optimization problem 3.33 has $(15N_h + 3)$ decision variables and $(40N_h + 1)$ constraints.

Furthermore, the decision-maker is provided with the following $(2N_h + 2)$ parameters to control the level of robustness of the optimal solution:

$$\left\{ \begin{array}{ll} \Gamma_b, \Gamma_s & \in \mathbb{R} \cap [1, N_h]; \\ \Gamma_{lb_k} & \in \mathbb{R} \cap [1, 2], \forall k \in \mathcal{H}; \\ \Gamma_{sh_k} & \in \mathbb{R} \cap [1, 3], \forall k \in \mathcal{H}. \end{array} \right. \quad (3.48)$$

4 Implementation

This chapter briefly describes in section 4.1 the techniques and software used to implement the EMS models developed in the previous chapter, and in section 4.2 how these models are embed and solved within a rolling horizon scheme.

4.1 Optimization model

Practical mathematical programming can be fairly iterative as described as in figure 4.1 [85]. A *Mathematical Programming Language*, or AMPL, is a modeling language designed specially designed to help people through this process [85].

AMPL is classified as an *algebraic modeling language*, because it allows the use of traditional mathematical notation to describe objective and constraint functions [85]. Therefore, the language syntax proves to be intuitive and easy to read and understand for humans. Furthermore, there is no need for additional conversions steps (made by the modeler) to input the model into the software; this saves time, money and reduces the probability of errors in the model. Figures 4.2a and 4.2b show two fragments of code that illustrate the high level syntax of AMPL.

AMPL interfaces with several well-known commercially available solvers, each one of them specializing in different classes of mathematical problems. The solver used for LP and MILP formulations (among other classes of problems) is the *IBM ILOG CPLEX Optimizer*.

Even though there exist other modeling languages for solving mathematical optimization problems (*General Algebraic Modeling System*, or GAMS, is another popular option), it was found that AMPL provides to the user the most complete documentation, as well as a very functional interface with external programs.

4.2 Rolling horizon scheme

AMPL has an Application Programming Interface (API) that provides an object-oriented callable library that allows access to AMPL models and run AMPL commands from external programs [18]. Through this API data and results can be easily exchanged between AMPL and external languages, allowing the embedding of the AMPL models in external

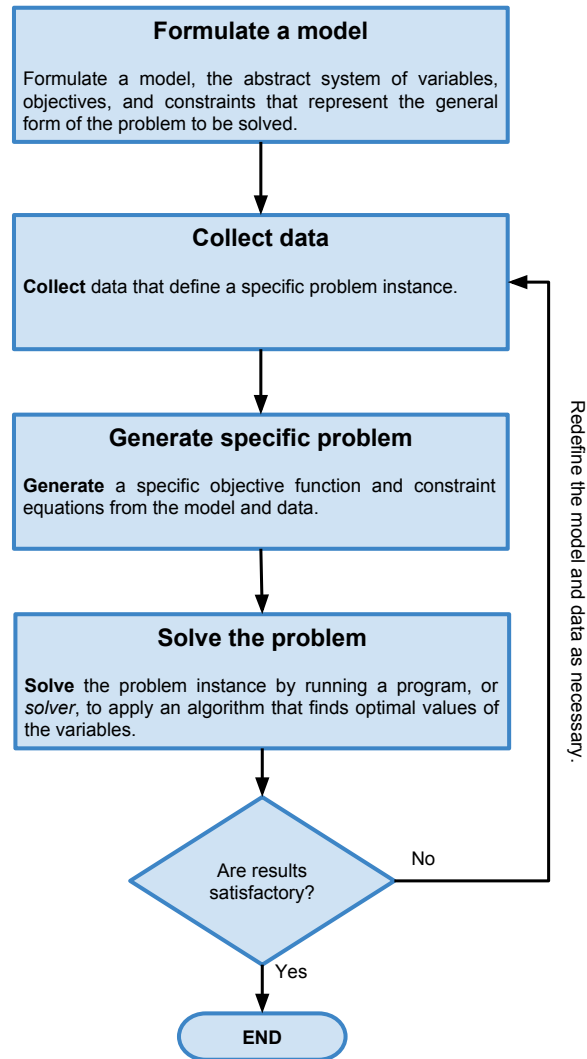


Figure 4.1: Algorithm for practical mathematical programming. Modified from [85].

```

## HORIZON PARAMETERS
param Nh > 0 integer;
param h {1..Nh} > 0;
param H = (sum{k in 1..Nh}(h[k]));

```

(a)

```

## MAIN GRID BINARY CONTROL VARIABLES
var Dbuy {1..Nh} binary;
var Dsell {1..Nh} binary;
subject to GridStatus {k in 1..Nh}:
    Dbuy[k] + Dsell[k] <= 1;

```

(b)

Figure 4.2: Fragments of AMPL code for the models created in chapter 3. (a) Definition of horizon related parameters (see equation 3.3); (a) Definition of the grid binary control variables (see equation 3.21) and its related constraint (see equation 3.22).

applications and the programming of complex algorithm schemes. Therefore, the AMPL API is used to embed the AMPL model into a rolling horizon scheme programmed in a MATLAB script; during each iteration of the rolling horizon algorithm, the MATLAB script, using the AMPL API, updates the parameters of the model and calls the AMPL solver that solves the problem using the IBM ILOG CPLEX solver.

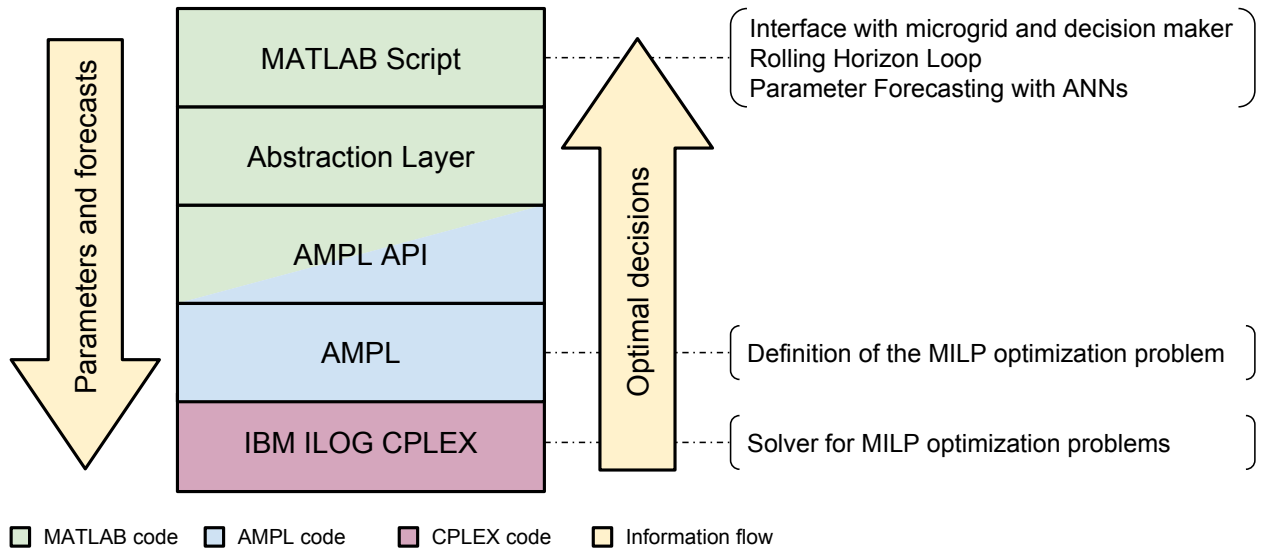


Figure 4.3: Information flow between the different layers of the EMS controller.

Moreover, an abstraction layer is programmed in MATLAB to make the code of the EMS controller independent from the API provided by AMPL. The information flow between the different layers of the EMS architecture can be depicted in figure 4.3; the general flow of the EMS code can be seen in the flowchart of figure 4.4.

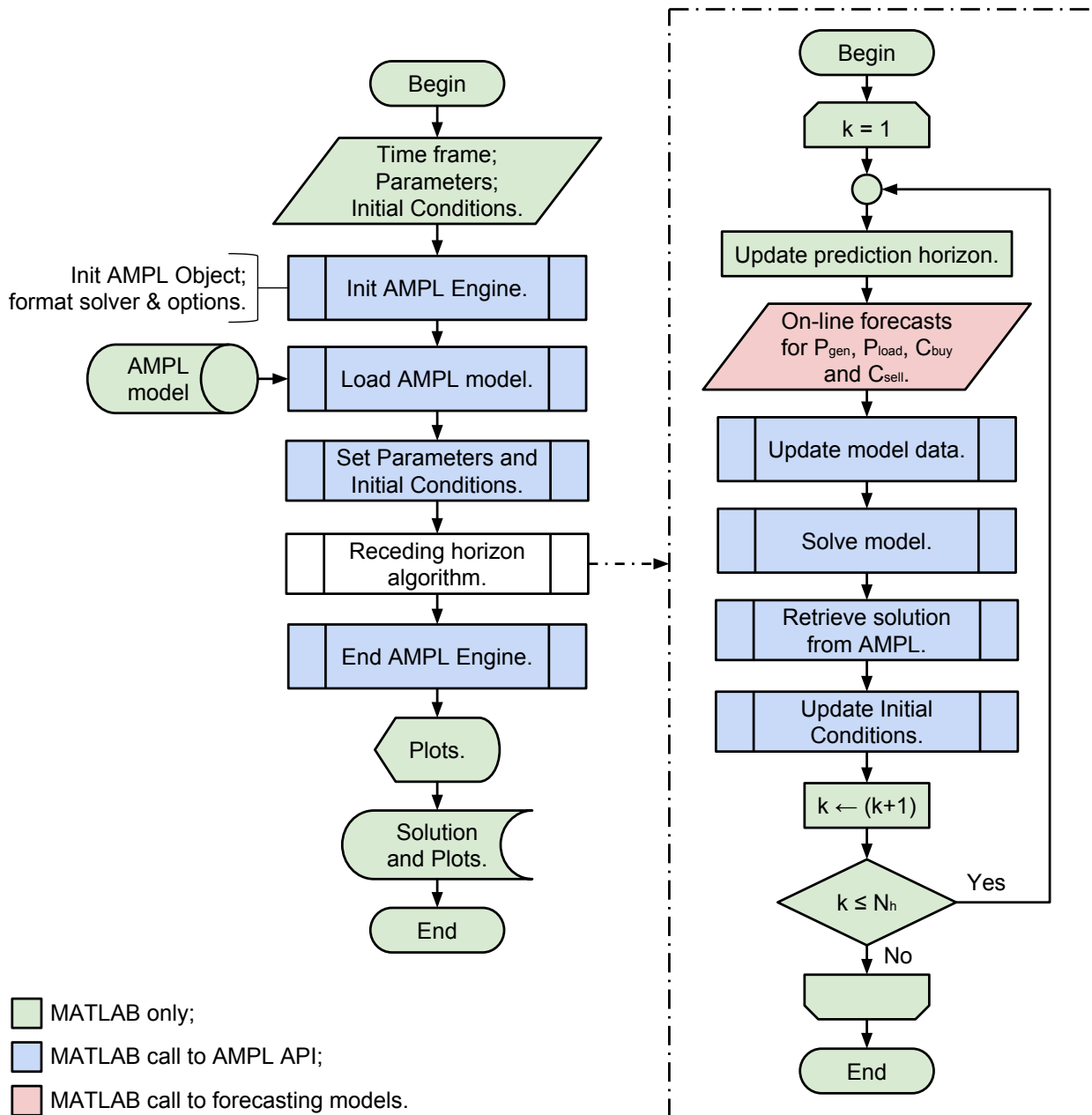


Figure 4.4: Flowchart for the EMS controller.

5 Results and discussion

In order to demonstrate the different features of the mathematical models derived for the EMS throughout chapter 3, simulations are performed under a variety of conditions.

This chapter is organized as follows: section 5.1 shows the output signals from the EMS and how they relate with the different stages of the rolling horizon control; section 5.2 investigates the robustness of the EMS by checking the feasibility of its decisions when they are exposed to the actual realizations of the forecasted values; section 5.3 presents the different penalization terms the decision-maker has available, and shows how their usage changes the power profiles given by the EMS; section 5.4 explains how the EMS can be used as a design tool to find the optimal capacity and power rating for the ESS; finally, section 5.5 studies the computational load of the EMS to determine if it is suitable for use in real-time applications.

5.1 Output profiles

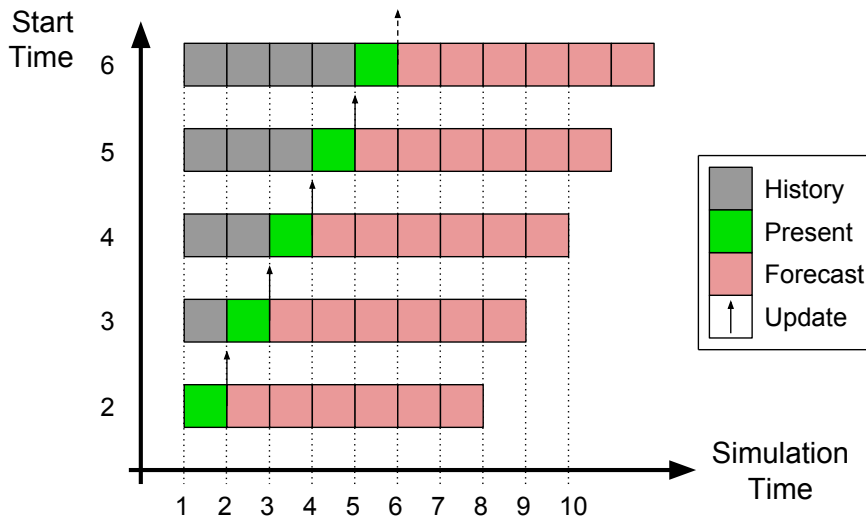


Figure 5.1: Rolling horizon framework.

Figure 5.1 re-visits the theory of the rolling horizon framework exposed in section 2.5. The control horizon is represented by the *present* time-step, and for the simulations performed throughout this chapter it will vary between 30 and 60 minutes. On the other hand,

the prediction horizon, or forecasting horizon, will range from 12 up to 96 hours. Finally, the scheduling horizon will be as long as it is necessary.

An implementation example is shown below with the control horizon set to 30 minutes, a prediction horizon of 24 hours and a scheduling horizon of 1 week; all the data use in the simulation is recollected from a week belonging to the winter of 2016.

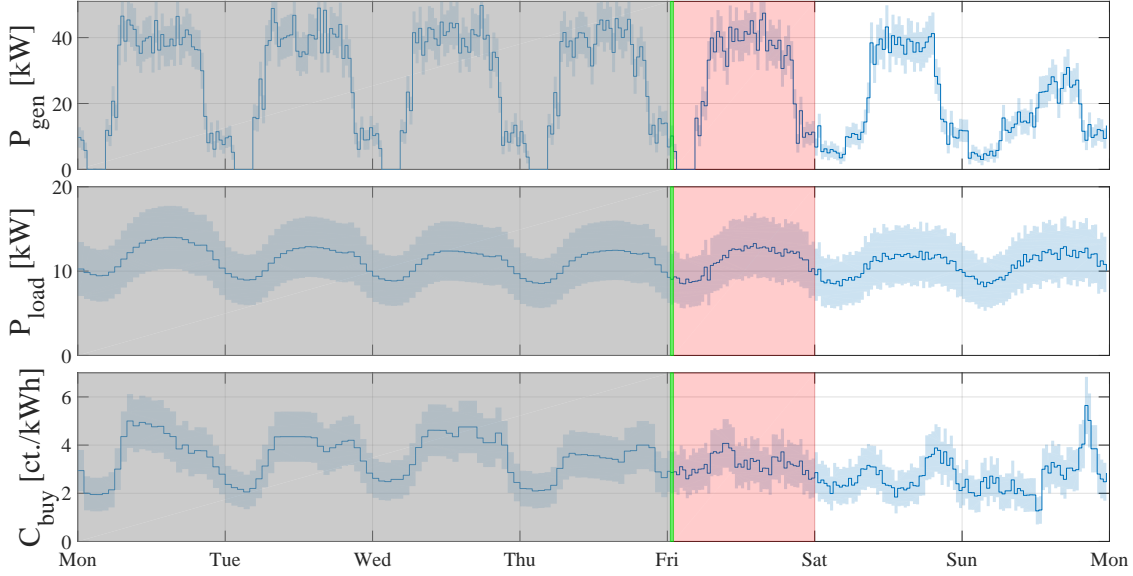


Figure 5.2: Some of the EMS input parameters that need to be forecasted by the prediction algorithms. This example shows a control horizon of 30 minutes, a forecasting horizon of 24 hours and a scheduling horizon of 1 week; this frame of the EMS shows the control horizon on the first 30 minutes of the Friday.

Figure 5.2 depicts some of the EMS input parameters that need to be forecasted by prediction algorithms, the color-coding matches the one in figure 5.1, showing the current time-step, as well as the prediction horizon and all the past time-steps.

These data is used as input to the EMS, which outputs, for the entire prediction horizon, the optimal power profiles so that the operational cost of the microgrid is minimized (see figure 5.3). The first plot of figure 5.3 shows the profile for the power exchanged with the ESS at each time step. This profile is the result of subtracting at each time-step the decision variable p_{s_k} from the decision variable p_{c_k} ; since both decision variables are nonnegative, a positive power in the power profile means that the ESS is charging, while a negative power means that it is discharging.

On the other hand, the second plot of figure 5.3 shows the profile for the power exchanged with the main distribution grid at each time-step. This profile is the result of subtracting at each time-step the decision variable p_{b_k} from the decision variable p_{s_k} ; since both decision variables are nonnegative, a positive power in the power profile means that

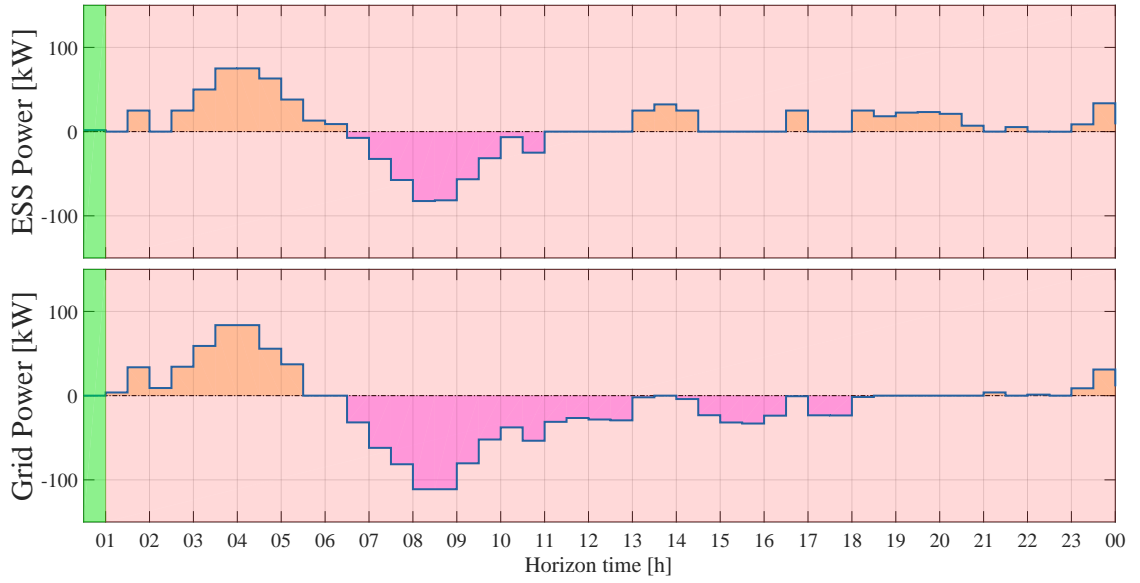


Figure 5.3: Optimal set of decisions found by the EMS controller for the entire forecasting horizon. Only the decisions corresponding to the control horizon will be implemented, while the decisions for the rest of the horizon will be discarded before updating the time-step.

the microgrid is buying power from the main grid, while a negative power means that it is selling power to the main grid.

As described in section 2.5 only the actions for the control horizon are saved and implemented, while the rest of the the optimal solution found by the EMS is discarded before updating the time-step and starting a new iteration. Figure 5.4 shows the history of decisions implemented by the EMS controller at each one of the different control horizons up to the *present* time-step, together with the evolution over time of the state of charge of the ESS.

5.2 Robustness of the solution

In this section the optimal solution given by the robust optimization approach is compared to the deterministic case; the feasibility and quality of the solution are studied for different scenarios of uncertainty.

5.2.1 Feasibility against data uncertainty

To check how robust the decisions made by the EMS truly are, a set of decisions is derived from the EMS for a predefined uncertainty set; this set of decisions is then exposed to the actual realizations of the forecasted parameters, and the feasibility of the solution is

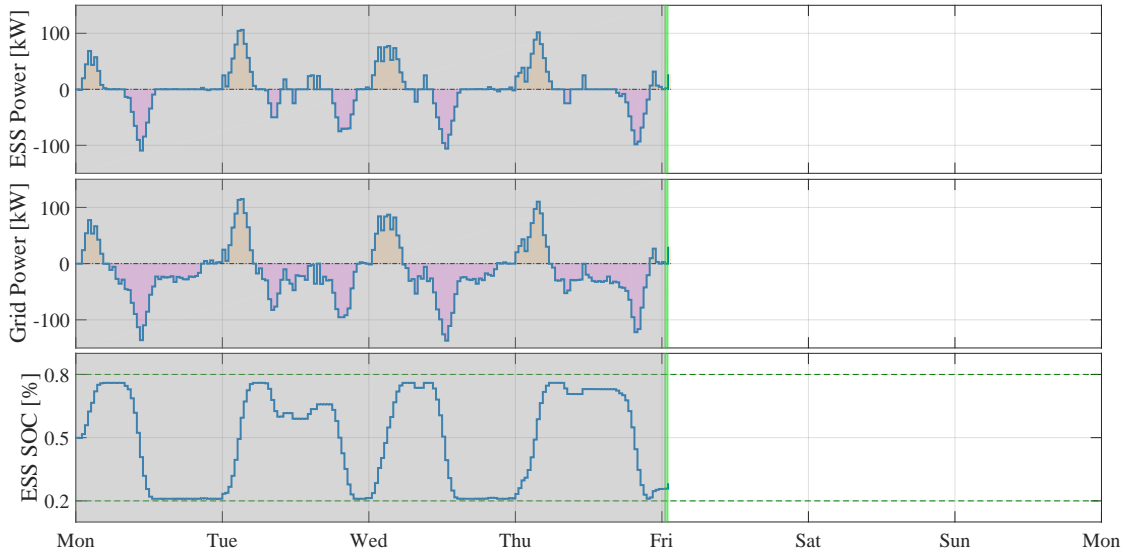


Figure 5.4: History of decision implemented by the EMS controller at each one of the different control horizons up to the *present* time-step. The evolution over time of the state of charge of the ESS (affected by these decisions) is also shown in the figure; for the ESS operation the safe operational bounds SOC^{\min} and SOC^{\max} are set to 0,2 and 0,8, respectively.

checked. Monte-Carlo simulations are carried out with the objective of verifying the feasibility of the solution for different realizations of the parameters, both inside and outside the predefined uncertainty set; this is achieved by adding to the realization a Gaussian noise of zero mean and a three-sigma value that ranges from zero to four times the length of the uncertainty set. For clarity, the uncertainty sets are normalized in the following figures.

Figure 5.5a shows the results for three different predefined uncertainty sets; the first uncertainty set is normalized to one, while the others have one half of the length and one-tenth of the length of the first uncertainty set. The EMS solution for each of these three uncertainty sets is exposed to the actual realizations of the forecasted parameters and the probability that the solution remains feasible is calculated. The probability at each point of the graph is the result of Monte-Carlo simulations with a different range of variation for the actual realization of the parameter. Figures 5.5b to 5.5d show three examples of these ranges of variation; as the realization of the parameters starts to fall outside the predefined uncertainty set, the probability that the solution of the EMS remains feasible starts to drop. Most importantly, it is a certainty that the solution given by the EMS will remain feasible as long as the realizations stay within the predefined uncertainty set. This results clearly shows the importance that appropriate uncertainty sets have on the performance of the solution, as it is explained in chapter 7.

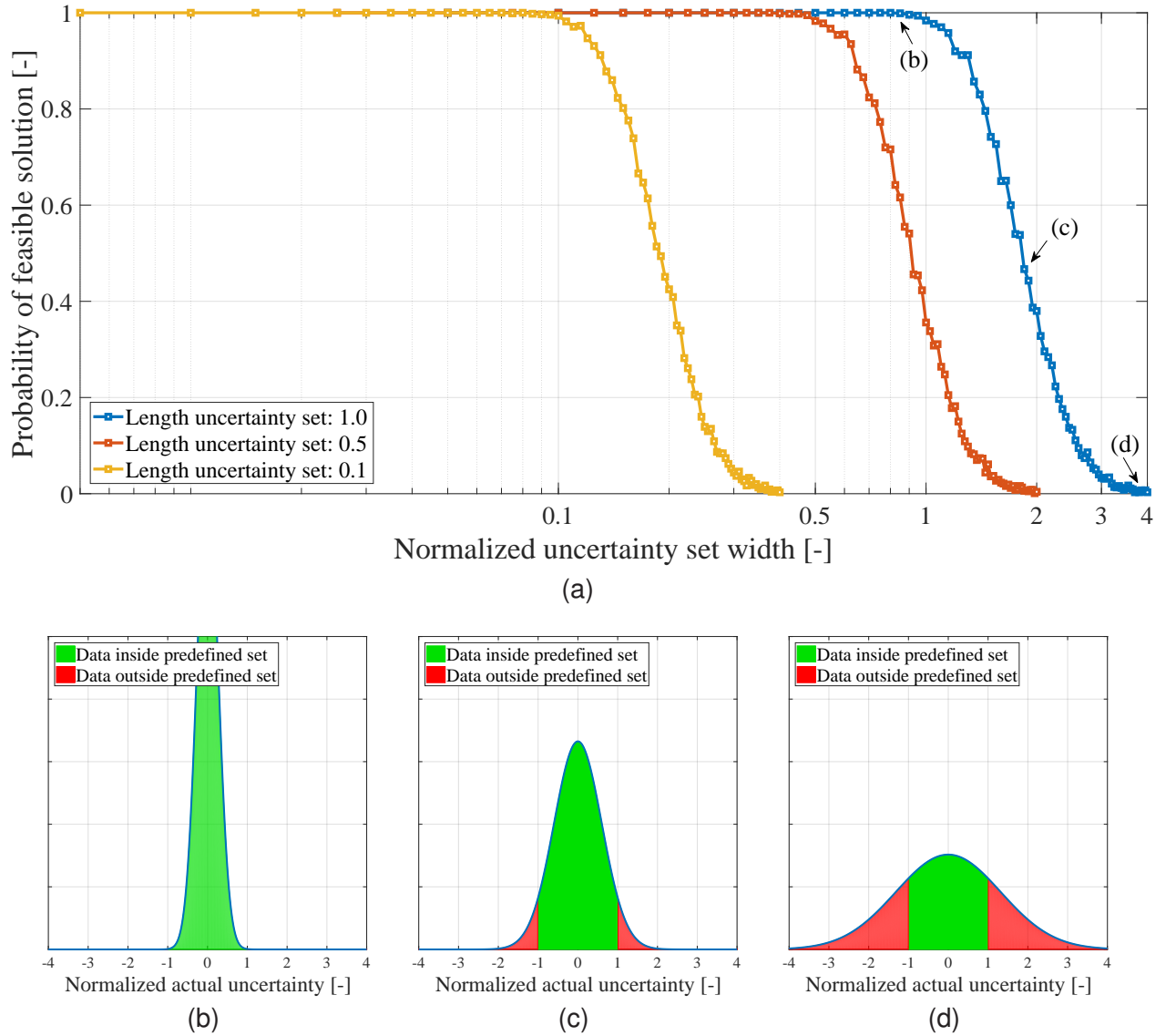


Figure 5.5: (a) Probability that the solution given by the EMS remains feasible for different realizations of the forecasted parameter, shown for three different uncertainty sets. When the uncertainty set is not defined properly, the realization of the parameter starts to fall outside of this set and the probability that the solution remains feasible starts to fall. Different Monte-Carlo simulations are shown for the actual realization of the forecasted parameter: (b) good uncertainty set definition, where the realization of the parameter always falls within the predefined interval; (c) the realization of the parameter starts to fall outside the predefined uncertainty set, causing infeasible solutions; (d) the choice of the uncertainty interval is quite poor in this example, since the realization of the parameter falls quite often outside of the interval and the probability the the solution is feasible is close to zero.

5.2.2 Operation cost against data uncertainty

As it was explained in section 2.4, it is not advisable to make an excessive increase in the size of the uncertainty interval, since this increment is accompanied by a deterioration of the optimal solution found by the EMS. Figure 5.6 shows the relationship between the length of the uncertainty set and the deterioration of the optimal solution; if there is no uncertainty in the forecasts made for the parameters (an ideal state), the profit given by the deterministic approach is greater than the profit given by the robust approach. However, as the system evolves to a more realistic implementation, the forecasts become uncertain and the robust approach gives more profit than the deterministic approach. In consequence, in a system where forecasts are not deterministic and are likely to contain uncertainties, the robust approach to the optimization gives more profit than the deterministic case.

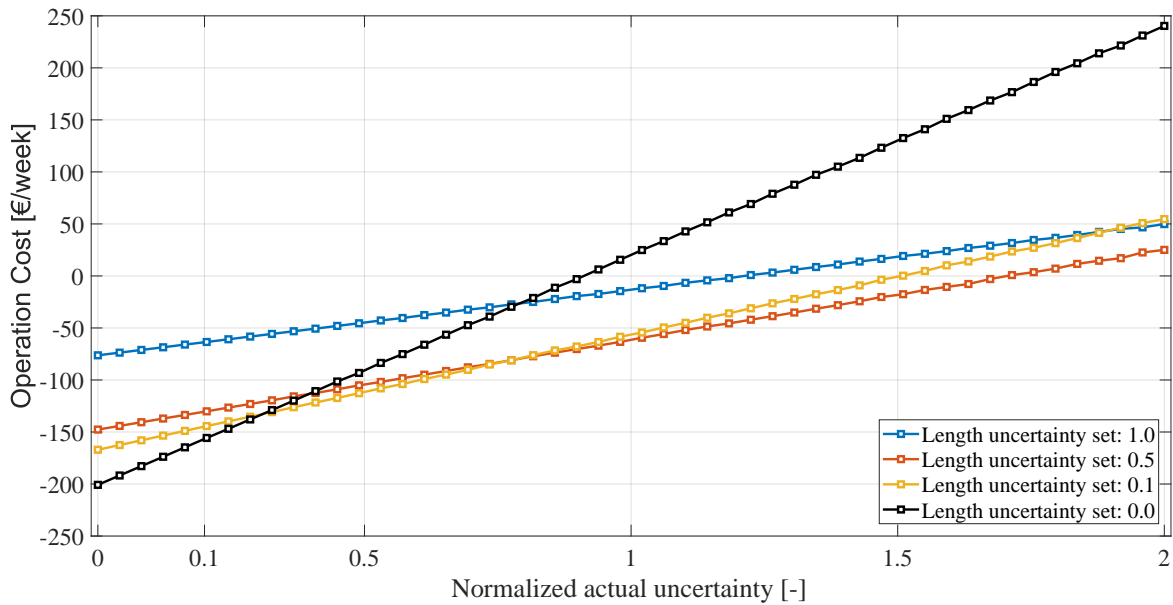


Figure 5.6: Deterioration of the optimal solution as the uncertainty in the forecasts increases. When the forecasts present no uncertainty (ideal case) the deterministic approach presents a better solution than the robust cases; as the uncertainty of the forecasts increases (real case), the robust optimization presents a solution of better quality.

5.3 Power profile shaping goals

The penalization terms added in the objective function of equation 3.5 together with their corresponding constraints allow the decision-maker to shape the power profiles given as output by the EMS; he does this by choosing the values of the penalization costs. This

grants the decision-maker the ability to meet sub-objectives of the system related to peak demand and profile smoothing, as the following sections describe.

5.3.1 Grid power profile flattening

The penalization cost for flattening the profile of the power exchanged with the main grid ($c_{\text{grid}}^{\text{flat}}$) allows to control the maximum peak-to-peak power exchanged; this is achieved by multiplying in equation 3.5 this penalization cost with the maximum peak-to-peak value for the power exchanged throughout the whole prediction horizon.

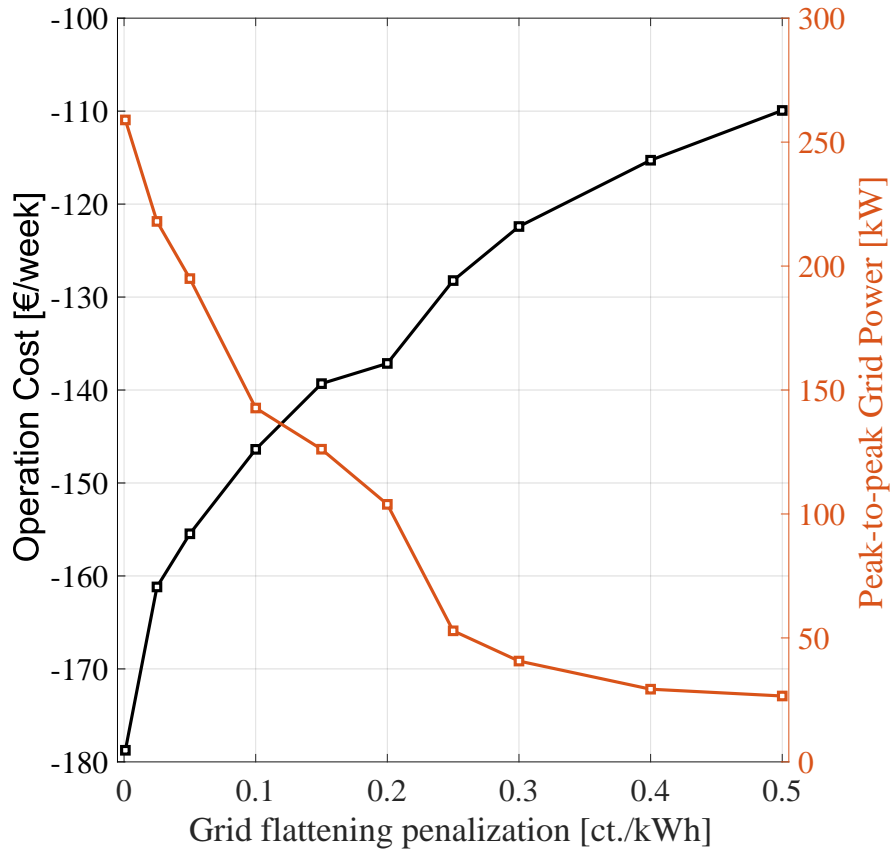


Figure 5.7: Trade-off for different values of $c_{\text{grid}}^{\text{flat}}$ between controlling the peak-to-peak power exchange with the main distribution grid and the savings from the operation of the micro-grid.

Figure 5.7 shows the trade-off between controlling the value of the peak-to-peak power exchange and the savings from the operation of the microgrid; increasing the penalization cost will give the decision-maker more control over the peak-to-peak power exchanged with the main grid, but this will decrease the savings obtained by the EMS.

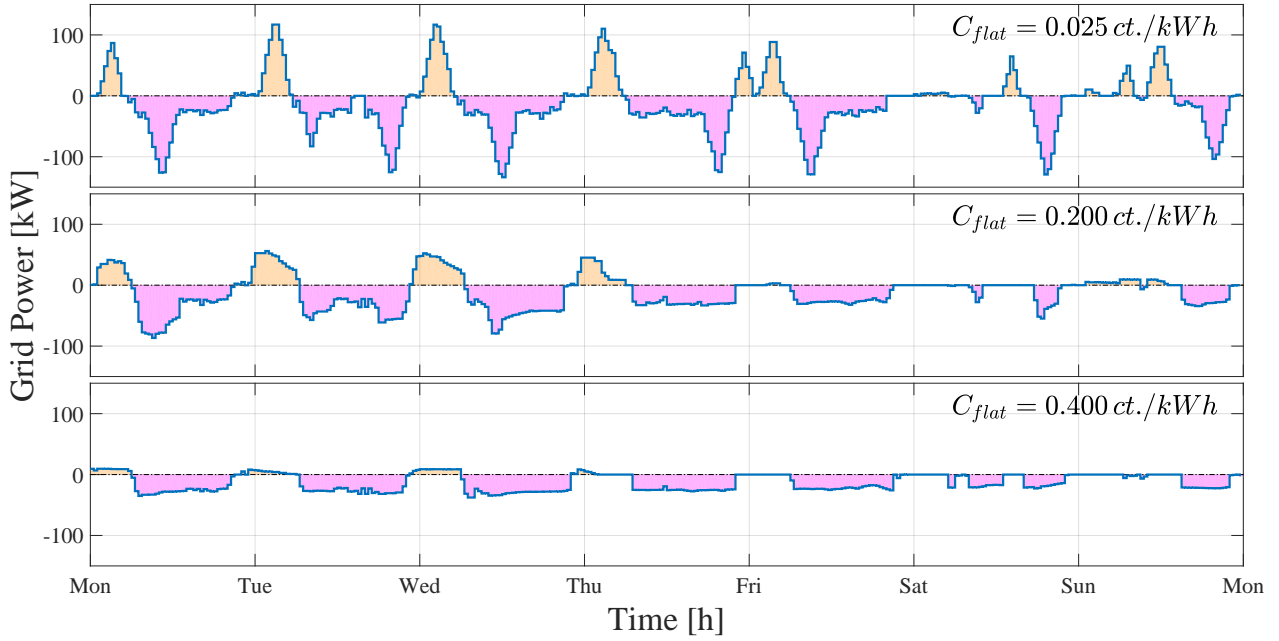


Figure 5.8: Evolution of the power profile for the exchange with the main distribution grid for different values of the penalization cost $c_{\text{grid}}^{\text{flat}}$; as the penalization cost is increased, the peak-to-peak power exchanged with the main grid is decreased.

Figure 5.8 shows how the power profile for the main grid changes for different values of $c_{\text{grid}}^{\text{flat}}$ while all other parameters and input data are left unchanged between simulations.

5.3.2 ESS power profile smoothing

Through the incorporation of the smoothing penalization cost ($c_{\text{ss}}^{\text{sm}}$), the decision-maker has some influence over the final shape of the power profile for the exchange with the ESS; this is achieved by adding to the cost function of equation 3.5 a term where the penalization cost multiplies the decision variable $u_{\text{ss}k}$, which is linked to the constraint of equation 3.18.

This penalization allows to control the maximum number of charging and discharging cycles for the ESS over a period of time, allowing in consequence to extend the lifetime of the ESS.

Figure 5.9 shows the trade-off between controlling the number of ESS charging and discharging cycles and the savings from the operation of the microgrid; increasing the penalization cost will give the decision-maker more control over the number of cycles over a period of time so that the lifetime of the ESS is extended, but this will decrease the savings obtained by the EMS.

Figure 5.10 shows how the power profile and the state of charge (SOC) of the ESS change for different values of $c_{\text{ss}}^{\text{sm}}$ while all other parameters and input data are left un-

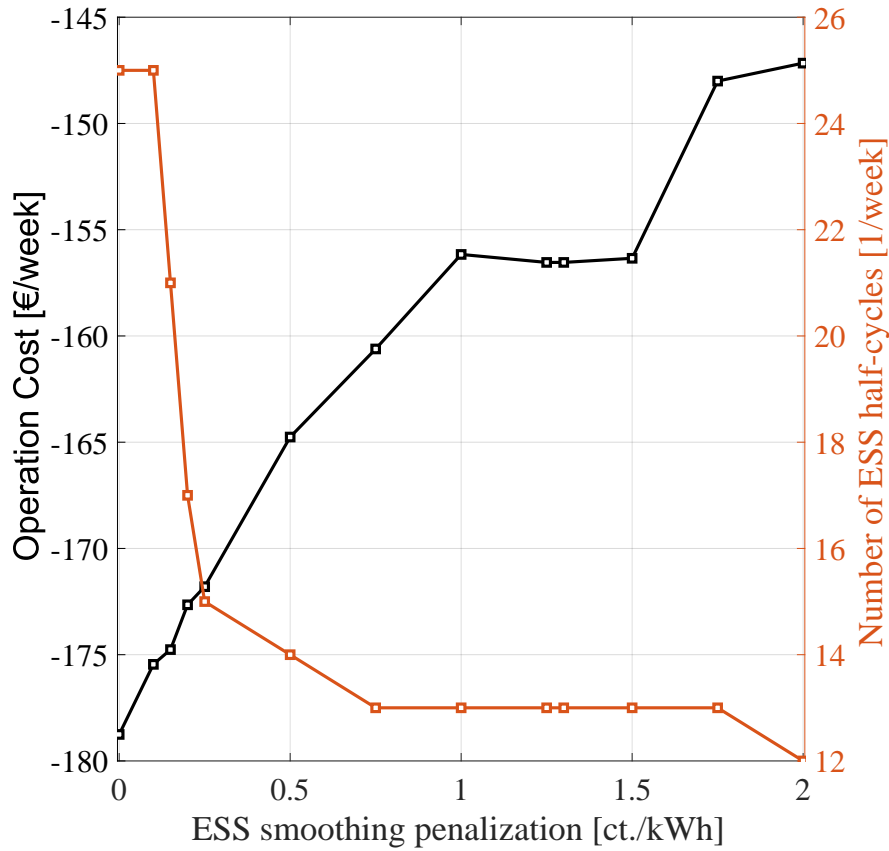


Figure 5.9: Trade-off between controlling the number of ESS charging and discharging cycles and the savings from the operation of the microgrid; increasing the penalization cost will give the decision-maker more control over the number of cycles over a period of time so that the lifetime of the ESS is extended, but this will decrease the savings obtained by the EMS.

changed between simulations. As the penalization increases, several small charging and discharging cycles done to optimize the energy transfer and operational cost are no longer implemented by the EMS.

5.4 Optimal sizing of the ESS

The proposed EMS can also be used as a designing tool to derive the capacity and power rating needed for the ESS that is going to be installed in the microgrid. With this purpose, the impact that the capacity and power rating of the ESS have on the operational cost of the microgrid is investigated in this section by performing a sensitivity analysis; to accomplish

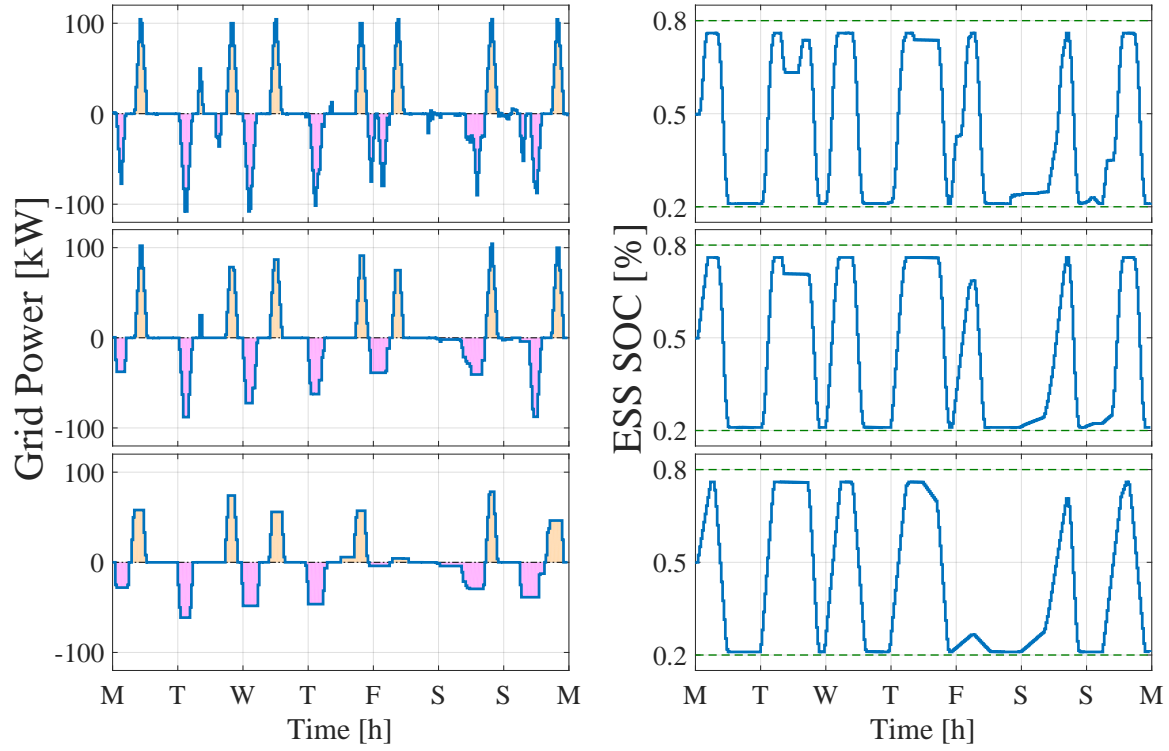


Figure 5.10: Evolution of the power profile and state of charge of the ESS for different values of the penalization cost c_{ss}^{sm} ; as the penalization increases, several small charging and discharging cycles done to optimize the energy transfer and operational cost are no longer implemented by the EMS. Top: $c_{ss}^{sm} = 0,10 \text{ ct./kWh}$; middle: $c_{ss}^{sm} = 0,25 \text{ ct./kWh}$; bottom: $c_{ss}^{sm} = 1,00 \text{ ct./kWh}$.

this task, the same optimization problem is solved for a variety of ESS capacities and power ratings.

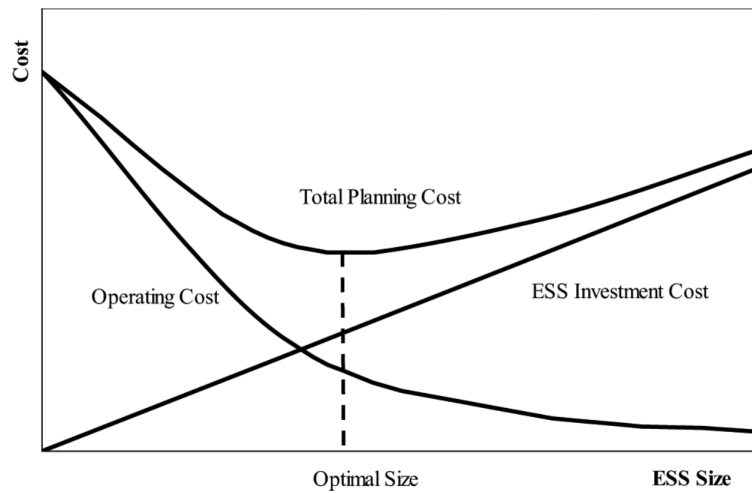


Figure 5.11: Optimal ESS Sizing. Source: [21].

Figure 5.11 depicts the *total microgrid cost* as a function of the ESS capacity; this total cost is obtained by performing the summation of the microgrid *operating cost* and *ESS capital investment*. The investment cost is directly proportional (in a linear fashion) to the ESS capacity, while the microgrid operating cost is inversely proportional to the ESS capacity. Hence, the minimum value of the total microgrid cost gives the optimal ESS size for the microgrid, as shown in figure 5.11.

Apart from being proportional to the capacity and power rating, the capital investment cost for the ESS is also a function of the technology used to storage the energy; since this work considers a generic storage system of no particular technology, the capital investment cost is unknown and the optimal ESS capacity and power rating is not determined. However, the analysis can be performed when a specific storage system technology is applied to the problem.

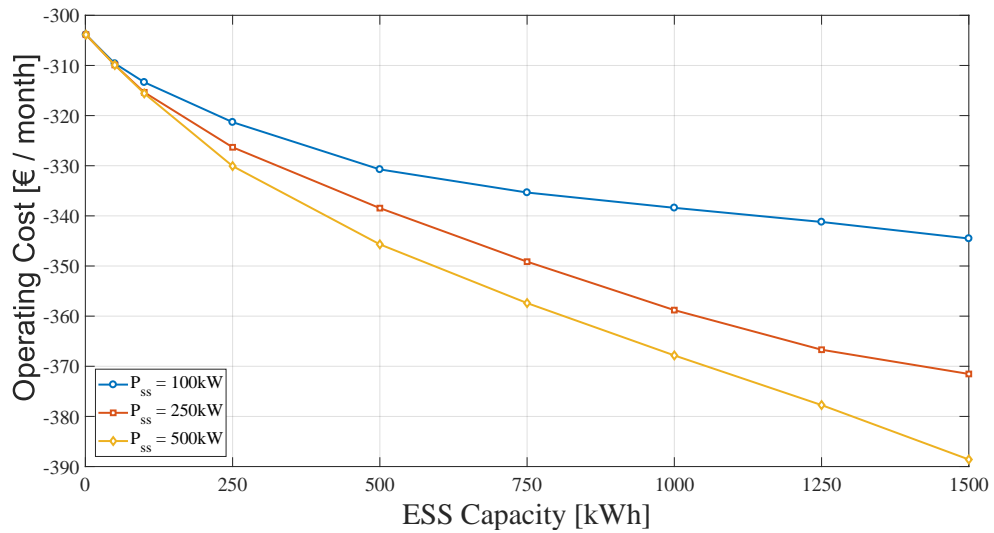


Figure 5.12: Optimal ESS sizing: sensitivity of the operational cost to the ESS capacity and power rating.

Figure 5.11 shows that the revenues from the microgrid increase proportionally to the capacity and power rating of the ESS. This is verified through simulations in figure 5.12. As the capacity of the ESS increase, the EMS can optimize the moment of buying and selling of a greater number of energy quanta; as Bahramirad *et al.* explains it in [21]:

"A larger ESS requires higher power import (as well as local generation) in low prices hours, thus increasing the cost of power import. On the other hand, a larger ESS increases the power export to the grid at times of high electricity prices and also reduces the units generation cost. Therefore, it would result in reduced operating costs."

Furthermore, the increase of the power rating gives the storage (and the control system) a higher dynamic response, because it can respond in a shorter period of time to a transfer of a certain number of energy quanta.

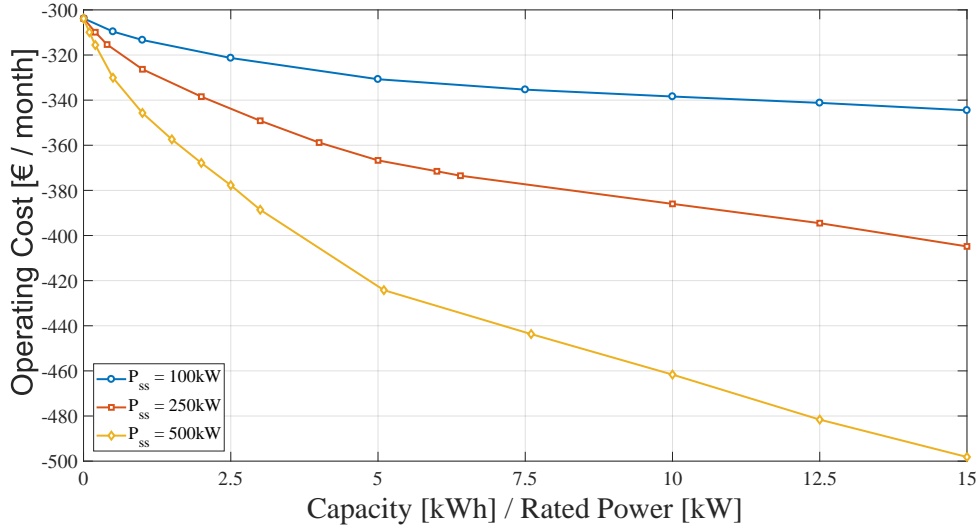


Figure 5.13: Optimal ESS sizing: sensitivity of the operational cost to the *minimum* number of hours needed to reach the *maximum* capacity of the ESS.

The same information of figure 5.12 can be shown in a different manner, as in figure 5.13. Here, the horizontal axis represents the *minimum* number of hours needed for the ESS to reach its *maximum* capacity; by choosing the desired power rate and minimum charging time, the needed capacity for the ESS can be derived from the graph.

5.5 Computational load

5.5.1 Effect of uniform/variable time-steps and constraint relaxation

This section explores different techniques to improve the computational time of the EMS controller; with this purpose, the correlation between the computational load and several properties of the system is studied, mainly:

- length of the prediction horizon;
- uniformity of the prediction horizon;
- relaxation of binary constraints.

The length of the prediction horizon window is studied between 12 and 96 hours; uniformity of the time-steps is studied by comparing a *uniform* window where all time-steps are

equal to 1h with a *variable* time-step window, where the horizon vector is created from all or a truncated portion of the following vector:

$$h = [1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 6, 6, 6, 6, 12, 12, 12, 12].$$

Finally, the effect *constraint relaxation*¹ has on the computational load is explored. To do so, the binary constraints on all but the first time-step in the rolling horizon window are relaxed.

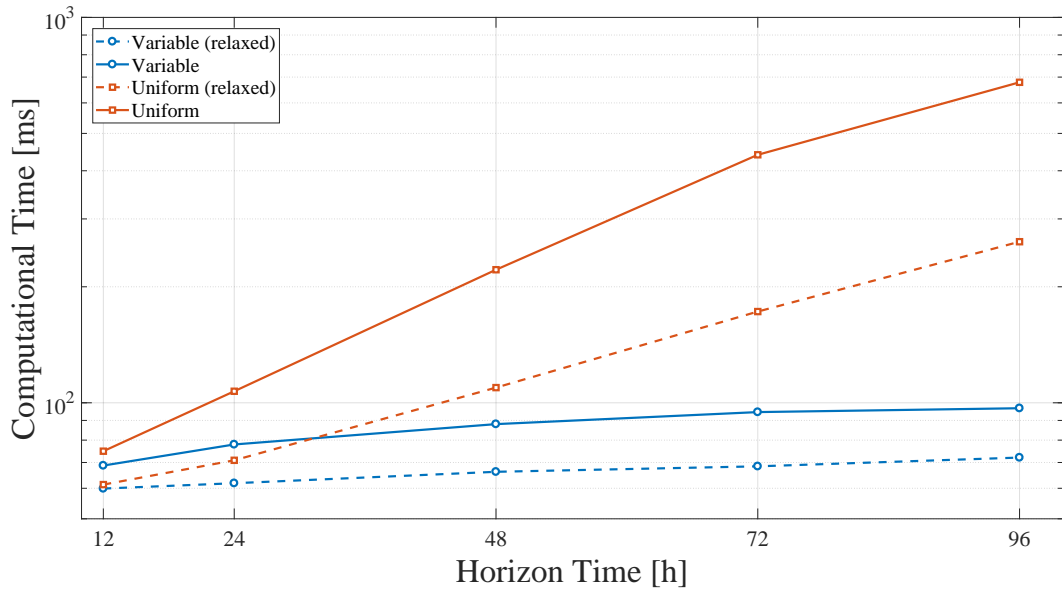


Figure 5.14: .

Figure 5.14 shows the results of the simulations; each point in the plot represents the computational time for the optimization of one time-step, averaged over 720 iterations (one month).

It is clear that the best solution in terms of computational speed involves a variable horizon with relaxed binary constraints. However, even though there is a difference of one order of magnitude between the average computational times of the different strategies presented in figure 5.14, all of them are at least four orders of magnitude smaller than the length of the control horizon. In consequence, real-time application would not be an issue for this controller.

¹ The relaxation of a MILP program is the problem that arises by "[...]replacing the constraint that each binary variable must be either 0 or 1 by a weaker constraint, where each variable now belongs to the interval $[0, 1]$. Thus, the relaxation technique transforms an NP-hard optimization problem (MILP) into a related problem that is solvable in polynomial time (LP)." [2]

6 Background on rapid transit systems

This chapter gives an introductory explanation of the railroad system, focusing on the tramway system being used to perform the measurements used hitherto in this work.

This chapter is organized as follows: the processes and equipments involved in delivering energy to the vehicles on the tracks are discussed in section 6.1; different approaches for introducing energy savings through *regenerative braking* are presented in section 6.2, together with different strategies for estimating the amount of energy recovered; finally, section 6.3 introduces the case-study of the RNV5 tramway line in the state of Baden-Württemberg in Germany, where several measurement probes were installed; an analysis of these measurements is carried out in section 6.4 to estimate the amount of energy harvested through regenerative braking action.

6.1 Transferring energy to guided mass transport systems

There are mainly two widely implemented approaches used to transfer energy to railroad vehicles [76, 216]:

- Use of an *overhead catenary* manufactured in copper as a phase conductor together with the track as a return conductor. This is the oldest and most frequently implemented method [76], used in both tram and train electrification all over the world. The reason for this is that the high voltage is on the catenary, at a safe distance from the users, and the voltage present in the tracks is low enough to be considered safe.
- Use of a *third rail* running parallel to the track with a relatively high voltage level. This method is only used when the tracks are completely separated from users and other means of transport. The third rail is usually manufactured in steel, with a cross-sectional area much greater than that of the catenary, allowing the conduction of higher currents with lower resistive losses and voltage drops.

Regardless of which of the two methods is implemented, the driving power supply needed for a direct-current (DC) commuter train has the basic structure according to figure 6.1¹.

¹ If the vehicles are fed with alternating-current (AC), the rectifier at the output of the power supply is omitted.

As shown in the figure, the systems that make up the structure for an overhead catenary system consist in particular of: i) supply lines, ii) return lines, iii) switching points and iv) catenary system² [220]. In addition to the electrical power needed to operate the trains, the power supply also feeds power to a variety of electrical equipment and installations along the train route, specially in the passenger stops, providing energy for luminaries, signaling and user information systems.

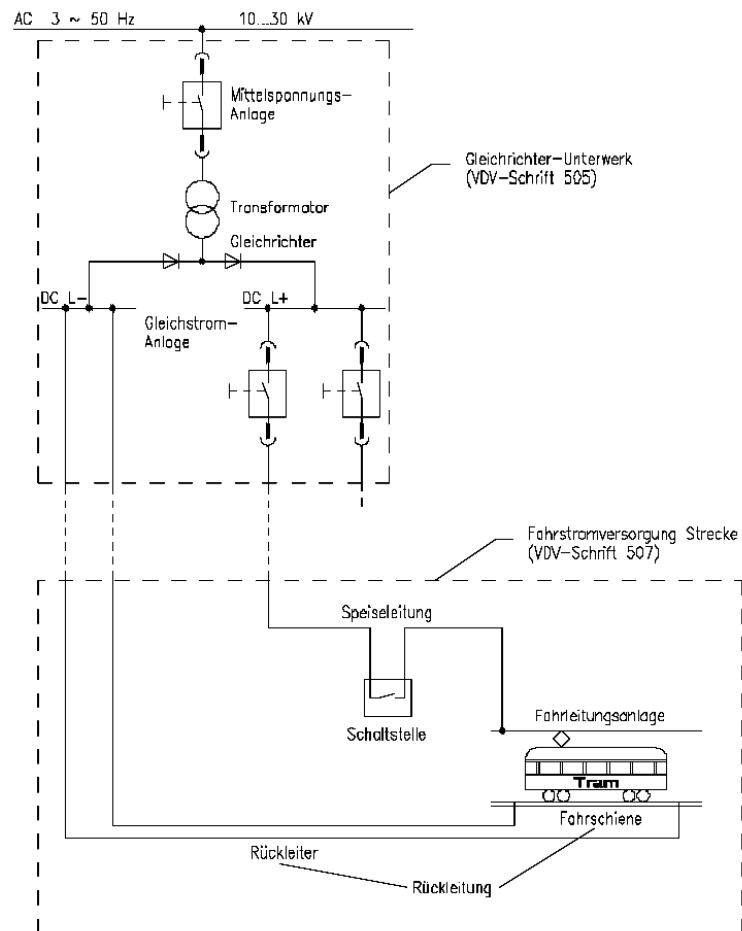


Figure 6.1: One-line diagram of the power supply of DC trains with overhead catenary.
Source: [220].

6.1.1 Nominal voltage levels

The nominal voltage levels recommended for railroad systems are defined by the *International Electrotechnical Commission* (IEC) in its standard *IEC60850* (see table 6.1) and these are usually the ones implemented in railway, tramway and subway systems.

² This also includes electrical systems which are fed directly from the catenary, like heating systems for the railroad switches, so that they remain operational in cold weathers.

Table 6.1: Standarized nominal voltage levels for railroad systems. Source: [116].

DC				AC	
600V	750V	1,5kV	3,0kV	15kV (16,7Hz)	25kV (50Hz)

The urban electrified public transport is usually supplied by a Low Voltage Direct Current (LVDC) system, generally at 750V or 1,5kV [76]. LVDC systems are preferred over others due to their simplicity in design, safety, and because there is no need to have large and heavy transformers on-board the vehicles, as the *pantograph* connects the catenary voltage directly to the power inverters of the train by means of filter capacitors [59]. Thus, LVDC systems are usually the most economic option.

The main drawback of a LVDC line with respect to a line with higher voltage is the increment in conduction losses; in order to transmit the same amount of power with lower voltage, more current is needed, according to the relation $P = VI$. Furthermore, conduction losses in copper (P_{loss}) increase with the square of the current:

$$P_{\text{loss}} = I^2 R_{\text{catenary}}, \quad (6.1)$$

and the voltage drop over the copper overhead catenary (U_{drop}) is directly proportional to the increment in current:

$$U_{\text{drop}} = IR_{\text{catenary}} = I \frac{\rho_{\text{copper}} l_{\text{catenary}}}{A_{\text{catenary}}}, \quad (6.2)$$

where ρ_{copper} is the electrical resistivity of copper, l_{catenary} is the length of the catenary between the feeding point and the vehicle, and A_{catenary} is the cross-sectional area of the catenary.

6.1.2 Rectifying stations

To reduce both copper losses and voltage drops, the electrical resistance of the catenary (R_{catenary}) needs to be addressed. Following the definition presented in equation 6.2, a drop in the resistance value can be achieved with two different approaches:

1. Increment of the cross-sectional area of the catenary (A_{catenary}).
2. Reduction of the conductor length (l_{catenary}) between the power supply and the vehicle through the deployment of load sharing rectifying stations along the tracks, assuring that all segments of the line are fed by one or two substations (see figure 6.2).

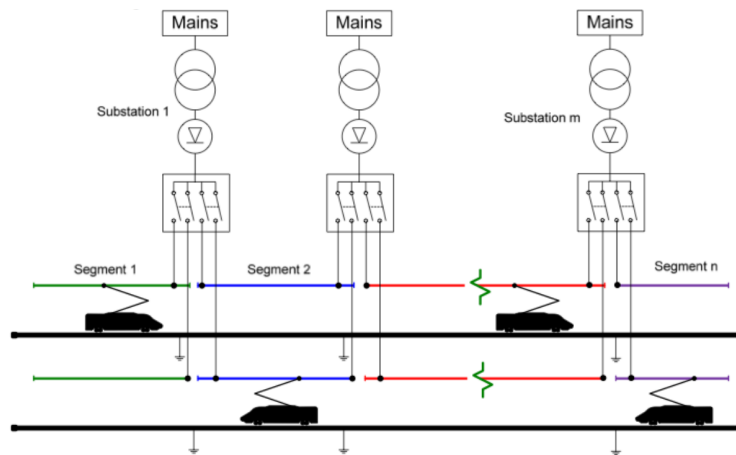


Figure 6.2: Schematic of a tramway line fed by several unidirectional substations. Source: [216].

If the substation deployment strategy is to be implemented, the contact line and the rails are divided in segments (see figure 6.2), mainly for operating, maintenance and security reasons. These segments are to be supplied and separable with circuit breakers and disconnectors, that can be actuated with a manual or remote command [59].

Furthermore, it is easy to see that the distance between electrical substations depends on the nominal voltage of the catenary, as well as the intensity of the traffic. For LVDC systems, the average distance between substations can be very limited. As Ciccarelli exemplifies in [59], a 750 V tramway system with heavy traffic would need to deploy substations every 1300 to 1500 meters. In contrast, a 3kV system would need substations every 20 to 45 km, depending on the amount of traffic.

As depicted in figure 6.1, each one of this rectifying substations is connected to the AC distribution network used within cities. These networks usually work with voltage values ranging from 10 to 30 kV³ [59]. The three-phase AC voltage delivered by this network is converted inside the substation to the nominal LVDC by means of a three-phase transformer and a 6-pulse diode rectifier. To further reduce the DC voltage ripple, some substations incorporate a capacitor connected in parallel to the output of the rectifier.

Because the AC/DC converter has a passive topology (i.e., only passive components like diodes are implemented), the current can only flow in one direction. This means that these substations work in only one power quadrant and are not capable of feeding back to the distribution network the energy harvested by the regenerative braking of modern vehicles. Furthermore, if the configuration of the railway does not include energy storage systems (ESSs) for electrical energy, the energy harvested by a braking train needs to be consumed by another load on the line at the same instant is being generated (i.e.,

³ This voltage values reach a good compromise between conduction losses and physical size of the equipment needed to transform the voltage down to the customer level.

another tram accelerating in the vicinity). Otherwise, the line voltage will rise over safety values⁴, and the braking vehicle will automatically redirect all the excess energy to its internal braking resistors, where it will be wasted as heat. Thus, in order to be able to feed back energy to the distribution network, an active converter that works in two power quadrants needs to be installed in the substations. A more detailed description of these processes is presented in section 6.2.1.

6.1.3 Typical waveforms

Figure 6.3 shows the typical waveforms measured at a rectifying substation that feeds one section of the tracks. The shapes described by the voltage and current are characteristic for a vehicle in the vicinity of the substation accelerating from rest to a constant travel speed. The cycle can be analyzed through the following stages:

1. initially, the vehicle is at rest at a passenger station, so it only consumes electricity to power the on-board equipment;
2. when the vehicle accelerates at the moment of departure, the current consumed by the motors increases proportionally with the speed. This stage always has the highest current consumption and the line voltage drops due to the load on the system;
3. when the desired speed is reached, the acceleration phase ends and the current decreases to a much lower value with the vehicle traveling at constant speed;
4. when the train brakes, the on-board motors act as generators and they deliver current to the overhead line. Thus, during this stage the current measured is negative. If there is no load in the system that can use this energy, the voltage of the line will increase until values near 900V, point where the on-board resistors will burn the recovered energy into heat.

Even though the stages remain the same, the general shape of these waveforms may change for vehicles with different power electronics, electronic control systems and electric motors; the driving technique implemented by different drivers also has a significant correlation with the shapes of these stages.

6.2 Regenerative braking

Dynamic braking refers to the process that converts kinetic energy into electricity, and it is based on the capacity the electric motors have to act as generators. The implementation

⁴ Typically 900V.

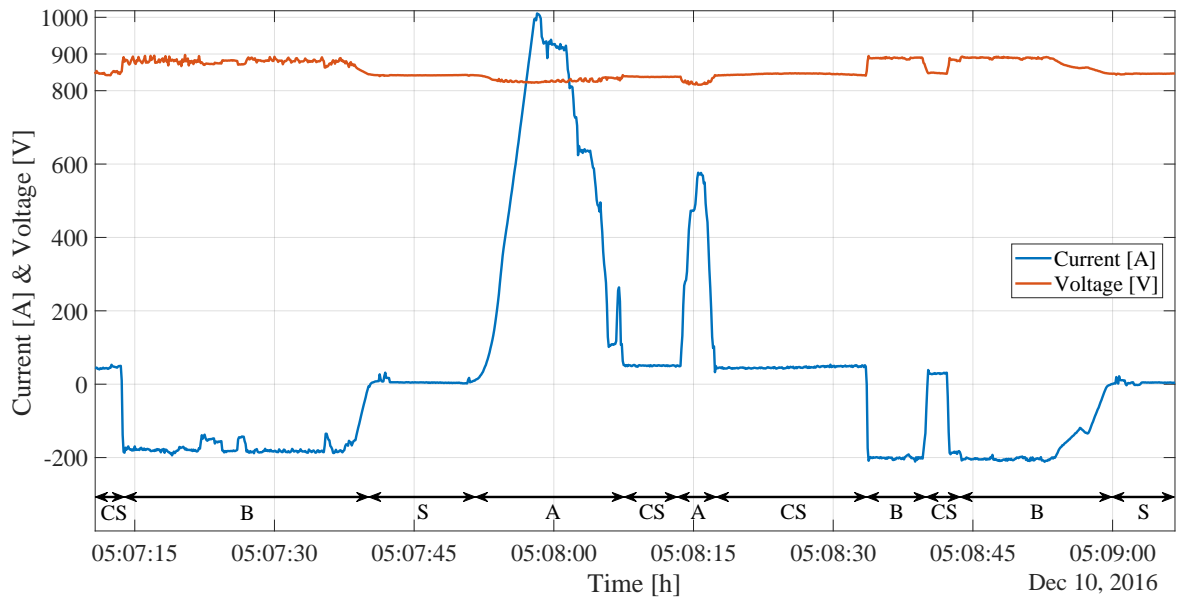


Figure 6.3: Voltage and current measurements as a train travels through one of the track segments of the Neuostheim electrical substation. CS: constant speed; B: braking; S: stop at passenger station; A: accelerating.

of this braking strategy is widely spread in railway technology nowadays, as it helps to limit the use of *friction braking*, a process that carries further investments associated with maintenance and replacement of parts subject to wear. The lack of consumable parts added to the fact that it does not generate wear and tear, dust, smell, heat or noise [95], makes dynamic braking a good system for urban and suburban lines, which are characterized by the shortness of their individual path sections and high deceleration rates [59].

The energy recovered by the generators during the dynamic break may either be dissipated in banks of variable resistors commonly installed on the rooftops of trains (*rheostatic braking*) or it may be reused within the transport network itself (*regenerative braking*), the latter becoming a viable option only after the power electronics development of the last decades.

A diagram of the energy flow in a vehicle during the braking of an electric traction unit is presented in figure 6.4. During the braking stage, the traction motor (TM) switches into a generator mode and the kinetic energy is converted to electrical energy as the vehicle reduces its speed. In order to be able to return this electrical energy to the overhead line and the supply system, a voltage higher than the one present in the line needs to be generated. For this purpose, during braking, the traction inverter (TC) increases the voltage generated by the traction motor.

The recovered energy is primarily used to supply the auxiliary and comfort functions of the vehicle itself. If there is a energy surplus (which is usually the case), the train will attempt to feed it back through its braking converter (BC) and pantograph back to the power supply line.

However, with electrical substations working in only one power quadrant (see section 6.1.2), the regenerated energy will only be accepted by the supply network when a simultaneous consumption is taking place, for instance when another train is accelerating in the same line segment.

In the absence of reception of the braking energy, voltage in the overhead lines (U_{CN}) on the vehicle's current collector exceeds the maximum recuperation voltage U_{\max} [24]. This triggers the braking converter BC, which directs the energy generated in the drive unit to the braking resistor. Hence, this energy is dissipated as heat through rheostatic braking.

The *network receptivity* is a figure of merit defined as:

"[...] the ratio of the total energy returned back to the line over the potential energy that could be regenerated in the braking process" [95].

As stated in section 6.1.1, the most commonly used urban rail systems are the DC electrified rail networks, and these type of networks are said to be *unreceptive*, meaning that most of the time they are reluctant to accept an injection of energy.

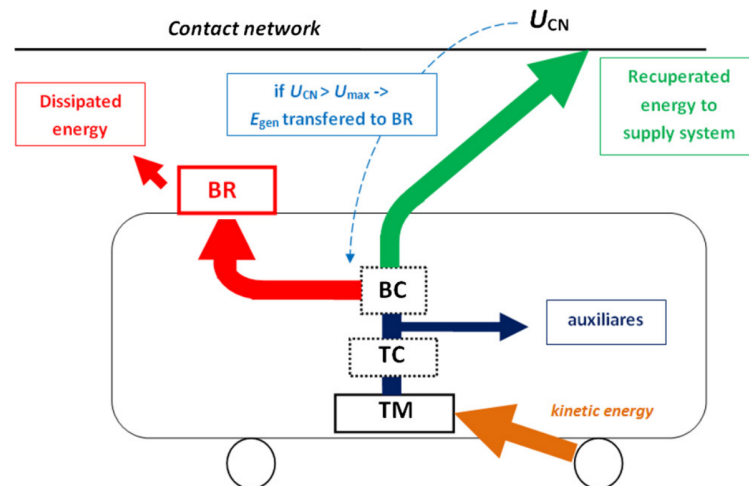


Figure 6.4: Braking energy flow in vehicle, U_{CN} - voltage of contact supply network, E_K - kinetic energy of vehicle, TM - traction motor, TC - traction converter, BC - braking converter, BR - braking resistors. Source: [24].

6.2.1 Maximizing the usage of recovered braking energy

Even though regenerative braking is a proven technology, its application in urban rail systems remains relatively unexploited. According to several studies found in the literature (see [95, 59] and its references), the application of regenerative braking in urban railways has the potential to reduce net energy consumption from 10% to 45%, depending on the characteristics of the system. Furthermore, regenerative braking could mitigate some problems typically associated with electrified transport systems such as voltage drops at the catenary distribution network or high power peak consumptions. In contrast, nowadays recovered braking energy is mainly used in rheostatic braking and only a small portion of it is used to supply the auxiliary systems of vehicles or returned to the distribution network.

Two main strategies have been developed over the years to maximize the use of the recovered energy while minimizing at the same time the need for on-board braking resistors [170, 95, 59] (see figure 6.5):

- Reduction of the energy losses by temporarily accumulating the excess of regenerated energy inside ESSs installed on-board the vehicles. This energy can then be released in the next acceleration phase, reducing the peak power consumption from the supply network during the acceleration of the motors [22, 113, 208, 23, 105, 50, 197].
- Improvement of the network receptivity. This can be achieved in several ways:
 - Introduction of additional loads in the system demanding energy at the same time that is being harvested by the braking process. The easiest and cost-efficient way of performing this task is to optimize the scheduled timetables so as to synchronize acceleration and deceleration of trains as far as possible [15, 55, 153, 82, 160, 183, 165].
 - Installation of storage devices in electrical substations (*stationary ESS*) or along the track (*wayside ESS*). They could absorb the surplus of energy harvested and deliver it when it is required for the acceleration of other vehicles [196, 60, 114, 209, 23, 149, 50].
 - Equipment of electrical substations with DC/AC inverters, transforming their operation from one to two power quadrants (*reversible or active substations*). With this technique, the regenerated energy can be fed back to the medium voltage distribution network, which by definition is naturally receptive [88, 89, 115, 181, 171, 226, 63].

The remainder of this section is devoted to briefly introducing each of these four strategies.

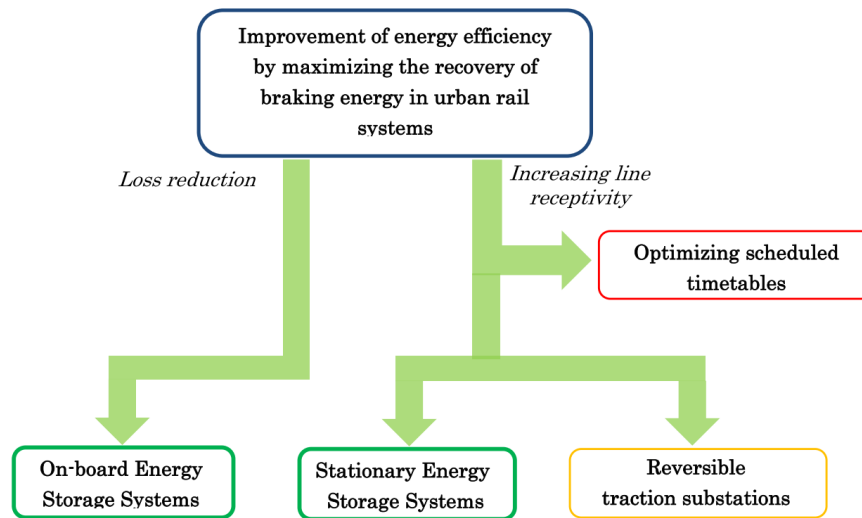


Figure 6.5: Strategies to maximize the use of the recovered energy. Source: [59].

Optimizing scheduled timetables

Increasing the number of trains accelerating and braking simultaneously is the approach that involves the least amount of capital cost for improving line receptivity. This is based on the principle of instantaneous power generation and consumption (see figure 6.6). In order to be successful in generating energy savings, this strategy demands a careful design of the operation schedule of trains. Several studies dealing with this approach can be found in the literature, where both frequency of the service and stop durations are considered as the main optimization parameters [15, 55, 153, 82, 160, 183, 165].

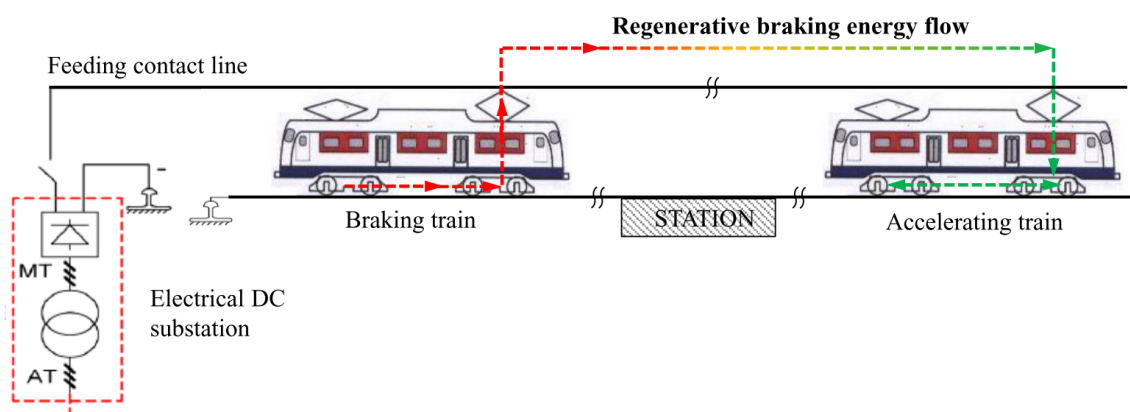


Figure 6.6: Energy exchange between vehicles in urban rail. Source: [59].

On-board ESSs

This approach does not rely on line receptivity, but instead contributes to energy savings by storing the recovered braking energy in ESSs on-board the vehicle (see figure 6.7). This energy can then be utilized to power the train itself during the next acceleration stages. Furthermore, this methodology contributes with other benefits like shaving the power peaks demanded to the supply network during acceleration (leading to a reduction in resistive losses in the supply line), limiting the voltage drops in the network and giving some power autonomy to the train in case it is needed (i.e., for free-catenary extensions of the track).

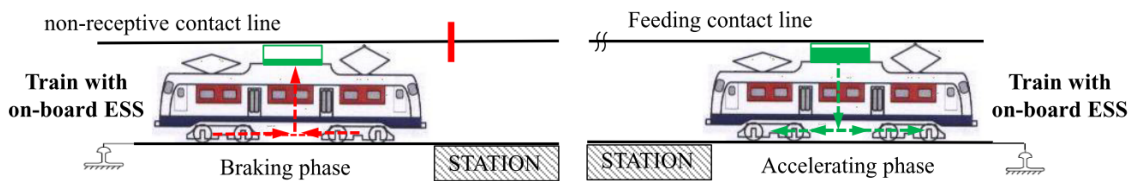


Figure 6.7: Energy exchange between vehicle and on-board ESS. Source: [59].

In order to achieve an optimal design for the technology and size of the on-board ESSs, a detailed analysis of the entire railroad network is needed; the design requirements will be different when optimizing each one of the benefits listed above [113, 208, 209, 105]. Nonetheless, the on-board ESSs are always installed together with a rheostatic braking option in case the recovered energy exceeds the storage capacity [95]. To minimize the times when rheostatic braking is needed, the control systems need to ensure that the ESS has enough energy to power the vehicle during accelerations and that it remains at a sufficiently low SoC level when the vehicle is traveling at high speeds so as to accept the highest amount of energy during braking.

The main advantage of this method with respect to the stationary ESSs relies on the higher efficiency during operation; since the ESS is on-board the vehicle the resistive losses of the supply line are avoided. The main disadvantage is the additional weight and space required by the on-board ESS. According to the literature, the energy consumption of rolling stock can be increased by 1 – 2% due to on-board ESSs [22].

It should be noted that generally the on-board ESS strategy is only considered feasible when designing or buying new rolling stock units, and is rarely implemented to modify existing vehicles that do not have storage systems [95].

Stationary ESSs

Stationary or wayside ESSs increase line receptivity by absorbing the regenerated braking energy when there is no load on the line willing to accept it. Thus, avoiding its waste in

rheostatic braking. Later on, it delivers the stored energy when it is required for the acceleration of any vehicle in its same electric section or on a neighboring one (see figure 6.8).

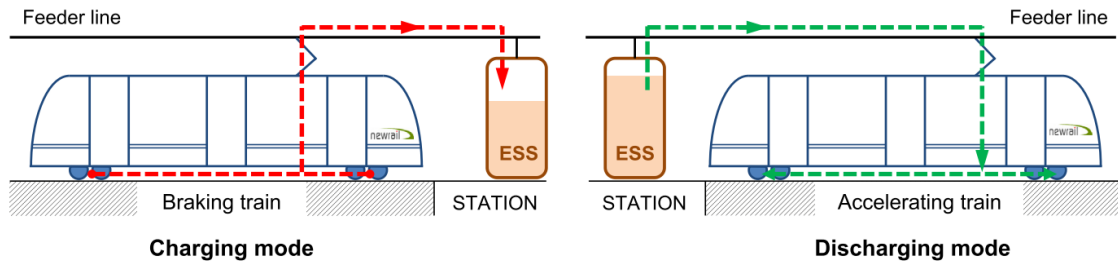


Figure 6.8: Energy exchange between vehicle and wayside ESS. Source: [95].

The advantages of a wayside system with respect to on-board ESSs rely on the fewer restrictions in terms of weight and space. Furthermore, one single stationary system can recover energy from several vehicles braking in its vicinity at the same time, and its implementation and maintenance does not affect normal line operations and time-tables. In contrast, stationary systems tend to be less efficient than on-board ESSs due to the transmission losses added when transferring energy back and forth between vehicles and the stationary ESS. Hence, in order to maximize the benefits from this approach, existing substations or places where the catenary voltage presents the strongest variation (e.g., in the vicinity of passenger stations) are the places usually chosen to install stationary ESSs. Thus, it is indispensable to perform a careful and detailed analysis of the whole network in order to determine the optimal number and positions of the stationary ESSs along the tracks [95, 60, 221, 228, 162, 98].

Reversible substations

As explained in section 6.1.2 conventional substations used in DC rail networks are based on diode rectifiers, which only permit a unidirectional power flow (i.e., they work in the first power quadrant). In contrast, *reversible substations* include a DC/AC inverter⁵, enabling a two power quadrant operation of the substation (see figure 6.9). This topology allows for any energy excess to be fed back to the railway operator's network (feeding energy to passenger stations, luminaries, escalators, offices, etc.) or to be sold back to the energy provider if the local legislation enables it.

The main advantage of this approach is that the medium voltage AC distribution network is naturally receptive and, in consequence, all the regenerated energy may be potentially

⁵ Inverters used in railway substations are typically made with thyristor-controlled rectifiers (RTCR), which enable a bidirectional current flow. Additionally, RTCR provide better voltage regulation and fault current limiting with respect to common diode rectifiers [88].

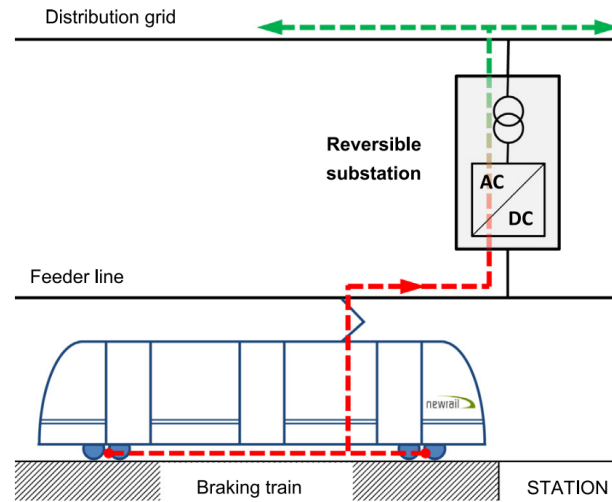


Figure 6.9: Energy exchange between vehicle and reversible substation. Source: [95].

recovered [95]. Conversely, the main drawback of this approach is related to the large cost of new equipment and the strong impact on the layout of existing substations [59].

As with the wayside ESSs, several studies can be found in the literature analyzing the feasibility of this methodology and determining the optimal number and positions of these substations in a variety of existing railroads around the globe, considering that the objective is to reduce the resistive losses due to the transmission of energy through the catenary [88, 89, 63, 171, 226]. It should be borne in mind that although reversible substations are installed to maximize the use of the recovered energy (through its injection into the upstream network), the priority should still be the exchange of recovered energy between the rolling stock.

6.2.2 Techniques for estimating of the amount of energy recovered

Knowing the amount of electric energy that will be recovered through regenerative braking in a certain future time interval will prove throughout the following chapters to be of vital importance in order to implement a reliable control system that achieves a good performance. This will help the integration of the ESS in a *microgrid* by improving its interaction with the other agents present in the network. In consequence, estimation methods must be implemented to be able to generate forecasts of future values of these quantities.

A first rough estimate of the recovered braking energy of each train can be made by calculating the change of kinetic and potential energy of the vehicle during its transition to the state of rest. Several assumptions are made in this calculation, but the result obtained will be the maximum amount of energy that can be recovered in the best case. More details are presented in section 6.3.2.

A more sophisticated procedure was proposed by Lakovskis and Grigans in [150], where the untapped regenerative braking energy in urban electric transportation networks is estimated. This method is based on probabilistic principles, transport schedules and measured power diagrams. Further improvement of the technique was introduced in [97] with the implementation of a stochastic modeling approach. Measurements of the power consumption of an electric vehicle are used together with the transportation schedule and speed profiles to estimate the unused braking energy in a particular section of the network, such as the area covered by an electrical substation. This unused braking energy can then be stored in energy storage devices integrated in the electric vehicle or stationary in the electrical substation.

Although the results obtained with this calculation method seem promising, various voltage and current sensors are required on-board the electric vehicle, as well as a detailed speed profile as the vehicle travels along the track. Since the system under test in this work does not present this configuration, other types of forecasting methods must be implemented, with input signals corresponding to data available in the system.

Therefore, in chapter 7 the process of creating the so called *prediction intervals* is analyzed. With this technique, artificial neural networks (ANNs) are trained through meta-heuristic algorithms to estimate the upper and lower bound of a prediction interval where the actual value will lie with a predefined confidence level. This method uses as input variables the history of voltage and current values measured at the output of the electrical substation together with meteorological information and train schedules to make a real-time estimate of the value of the recovered energy. The estimation is done in a fast and simple way, allowing the use of these forecasts in the decision making process performed by the different electronic control systems of the microgrid.

6.2.3 Energy storage systems for urban rail applications

ESSs typically consist of three main functions as depicted in figure 6.10, regardless of whether they are used for mobile or stationary applications:

- The energy storage device itself. Some of the technologies being considered as feasible options by the literature include, but are not limited to, electromechanical double layer capacitors⁶ (EDLC) [22, 196, 60, 114, 113, 208, 209, 154], flywheels [197, 89], superconducting magnetic energy storage (SMES) [118, 213, 212, 16] and batteries⁷.

⁶ Also known as ultracapacitors or supercapacitors.

⁷ Combinations of these technologies have also shown good results, like EDLCs with batteries; EDLCs provide a fast dynamic response while batteries provide a larger storage capacity increasing the autonomy of the system [105].

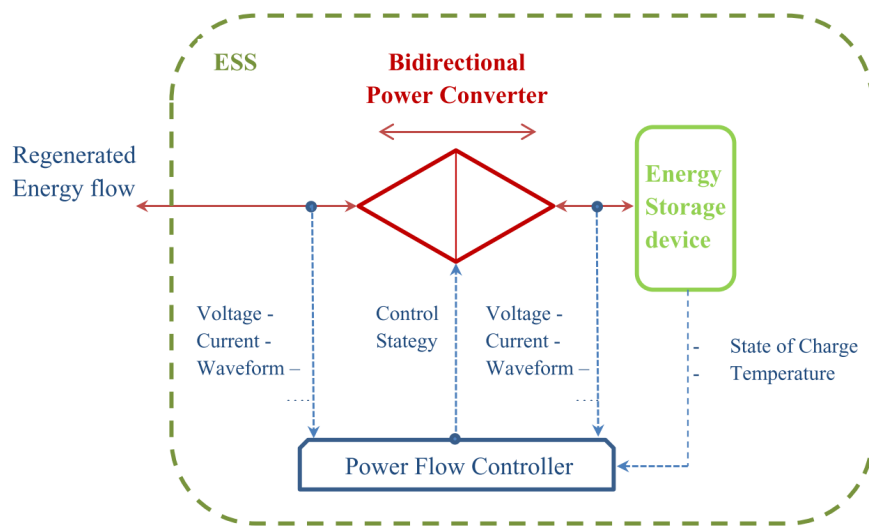


Figure 6.10: Energy storage systems for urban rail applications. Source: [59].

The latter contains sub-classifications like Lead-acid batteries [89, 100], Nickel-based batteries⁸ [89, 180], Lithium-based batteries⁹ [51, 49, 50, 149], Sodium-based batteries¹⁰ [44, 80] and Redox Flow Storage systems¹¹ [57, 106, 68, 99]. The selection of the technology depends on the specific characteristics of each application. Variables taken into account include, but are not limited to, technical maturity, energy and power density, efficiency of charge-discharge cycles, self-discharge rates, durability and capital costs. Detailed reviews on each of these technologies and decision variables can be found in [54, 164, 194, 95, 142] and its references.

- Power conversion systems. They are needed to guarantee a proper operation of the energy storage, as they normally work with different input and output conditions than those required by the railway distribution network¹².
- Power flow controllers. A controller is needed regardless of the storage technology implemented. Its main objective is to optimize the ESS performance by managing the charge and discharge cycles, which depend mainly on the line voltage and the state of charge (SoC) of the ESS [60, 114, 113, 105, 23, 154]; the line voltage rises if a braking vehicle attempts to feed back energy to the line when there is no load available, and a voltage drop on the network may be produced by a sudden power

⁸ Nickel-cadmium (NiCd) and nickel-metal hydride (NiMH).

⁹ Lithium-ion (Li-ion) and lithium-polymer (Li-poly).

¹⁰ Sodium-sulphur (NaS) and sodium-nickel-chloride (also known as ZEBRA).

¹¹ Vanadium Redox batteries (VRB).

¹² Some special arrangements of wayside ESSs made with batteries have been tested connected to directly to the DC catenary voltage without a DC/DC converter [89, 180].

peak demand from an accelerating vehicle. In consequence, the electronic controller is programmed with the nominal voltage thresholds for the charging and discharging phases.

6.3 The Rhine-Neckar transportation system

The Rhine-Neckar Transportation system¹³ is a company operating public transport in the Rhine-Neckar region, which is located in the north-east part of the Baden-Württemberg state, in south-west Germany. The company connects mainly the cities of Heidelberg, Mannheim and Ludwigshafen on the Rhine through a variety of tramways and bus routes. One of its tramway lines, the RNV5, operates connecting the cities of Mannheim, Weinheim and Heidelberg. The trace of this route can be seen in figure 6.11.



Figure 6.11: Tram line number 5 of the RNV company, connecting the cities of Heidelberg, Weinheim and Mannheim (red line). Source: [6].

The two segmented tracks are powered by a series of electrical substations; the tramways operate with an overhead catenary system that delivers a nominal voltage of 750 V. All the electrical substations in this line have passive energy converters, meaning that the electricity flow can only be unidirectional and no energy can be fed back into the medium voltage AC distribution network. Additionally, no electrical *Energy Storage Systems* (ESSs) are currently installed on this railroad.

¹³ *Rhein-Neckar-Verkehr GmbH* in German, or RNV.

6.3.1 Rolling stock operated by the RNV

Table 6.3.3 shows a list of tram models being operated by the RNV company¹⁴. It is clear in the table that as the years advance and cities progress, demand for public transport increases, so more transportation capacity is needed. In order to meet this demand, over the last decades trams became heavier and got more powerful motors, which in turn has led to an increase of the power needed to operate them. Moreover, this development has allowed to incorporate new technologies on-board (e.g., HVAC units, lights, ticket machines and electrical doors) which increase comfort as much as power consumption. Nevertheless, as seen in figure 6.3 the electric motor still are the main power consumers, mainly during the acceleration stage.

6.3.2 Upper bound for the amount of energy recovered

Knowing the average speed profile of a vehicle in the railway, the maximum speed for the trams and the approximate weights for each vehicle (table 6.3.3), it is possible to estimate an upper bound to the amount of energy one vehicle is capable of harvest when it is using a regenerative braking system to reduce its speed. This is achieved through a kinematic analysis.

The absolute maximum value of harvested energy with one single brake is obtained by calculating the change in kinetic energy (see equation 6.3). Here it is assumed a RNV8 train at full capacity (76,5t) braking from its maximum speed (80km/h) to rest condition. Furthermore, a loss-less system is assumed, meaning that there is no energy loss as heat or noise when converting the kinetic energy to electrical energy.

$$\Delta E_{k_{MAX}} = \frac{1}{2}m(v_f^2 - v_i^2) = 5,2 \text{ kWh} \quad (6.3)$$

In figure 6.12 the surface delimiting the upper bound for the recovered energy for different speeds and masses is drawn. Additionally, a log-log plot parametric with the initial speed of the vehicle is presented in figure 6.13.

This analysis assumes a fairly plane route, which seems to be the case for the majority of urban and suburban tramways [51, 22]. However, if the track profile presents important altitude gradients, the potential energy must be also considered in the calculations, since it will affect (either positively or negatively) the amount of electrical energy generated. As explained in section 6.2, once recovered, this energy can be transferred to different locations and used for different applications.

¹⁴ Up to and including the year 2017.

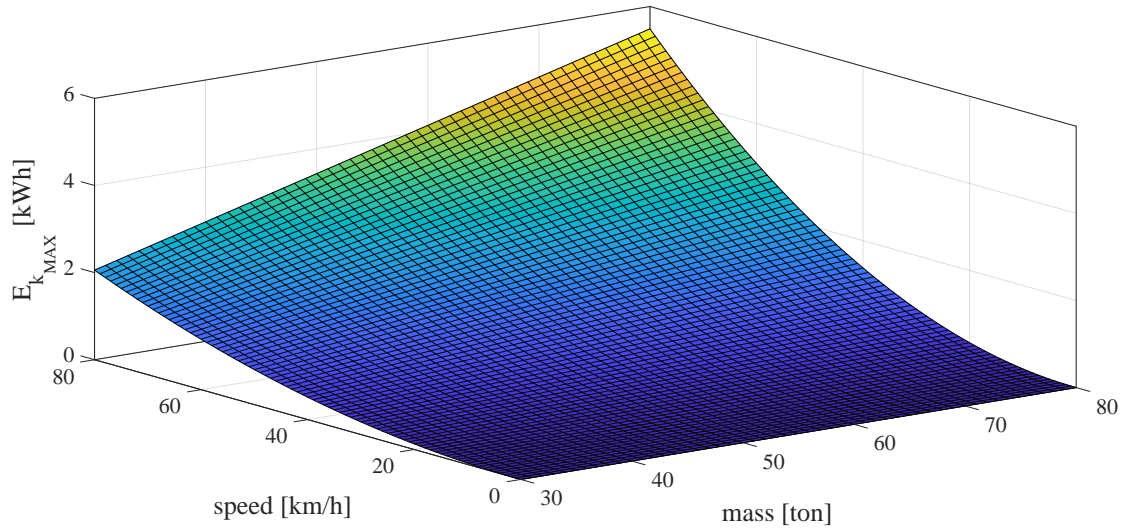


Figure 6.12: Upper bound for the electrical energy harvested during regenerative braking for a train with mass m braking from speed v to rest condition.

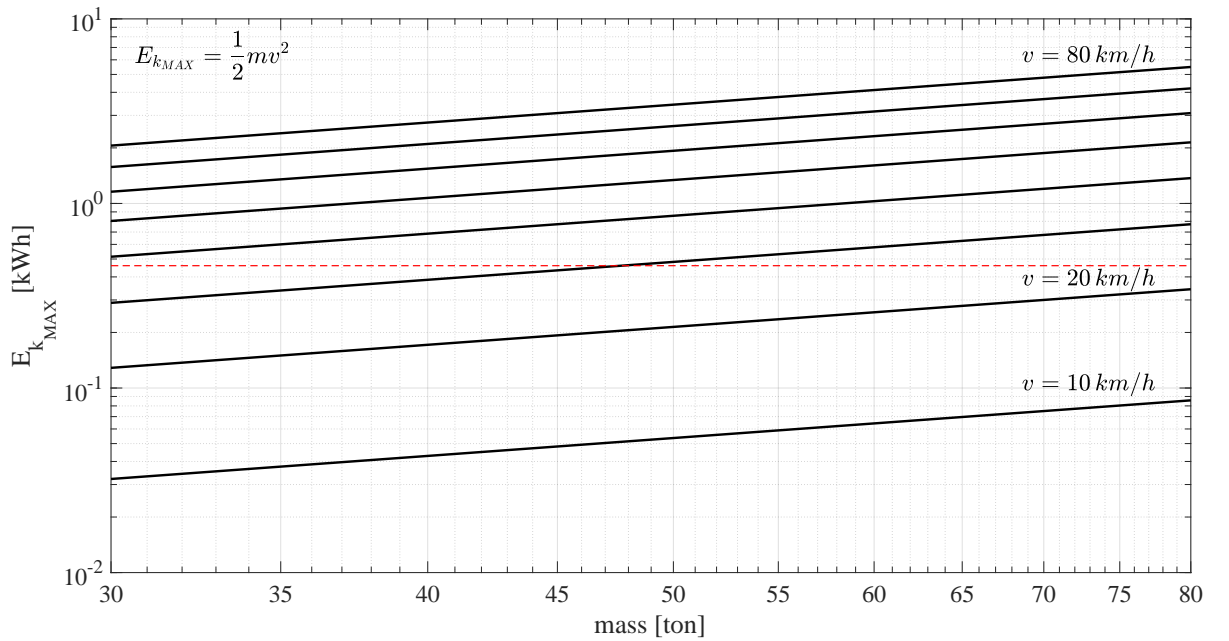


Figure 6.13: Upper bound for the electrical energy harvested during regenerative braking. Curves parametric with respect to the speed v the vehicle has before the braking begins. The red dashed line indicates the average energy recovered by a braking phase according to the measurements at Neuostheim electrical substation (see section 6.4).

						
Model	M8C-NC	MGT6D	6MGT	ET8N	RNV6	RNV8
Manufacturer	DUEWAG	DUEWAG	DUEWAG	DUEWAG	Bombardier	Bombardier
Year Fabrication	1985-86	1994-95	1994-95	1994-95	1996-16	2002-10
Quantity	8	12	14	5	57	32
Length	26,58 m	28,6 m	29,2 m	40,5 m	30,5 m	40,0 m
Weight (empty)	36 t	32 t	32 t	42 t	38 t	50 t
Sitting capacity	54	72	85	97	88	129
Standing capacity ¹⁵	86	100	119	141	90	133
Weight (full)	49,3 t	49 t	54 t	74,1 t	55,6 t	76,5 t
Motors	2×150 kW	4×95 kW	4×95 kW	4×95 kW	4×95 kW	6×95 kW
Maximum speed	70 km/h	70 km/h	70 km/h	70 km/h	80 km/h	70 km/h
Traction converter	Chopper	13SG31	13SG31	13SG31	13SG31/44/76	13SG44/76

Table 6.2: Current rolling stock operated by the Rhine-Neckar Transportation system. Sources: [5, 7, 11].

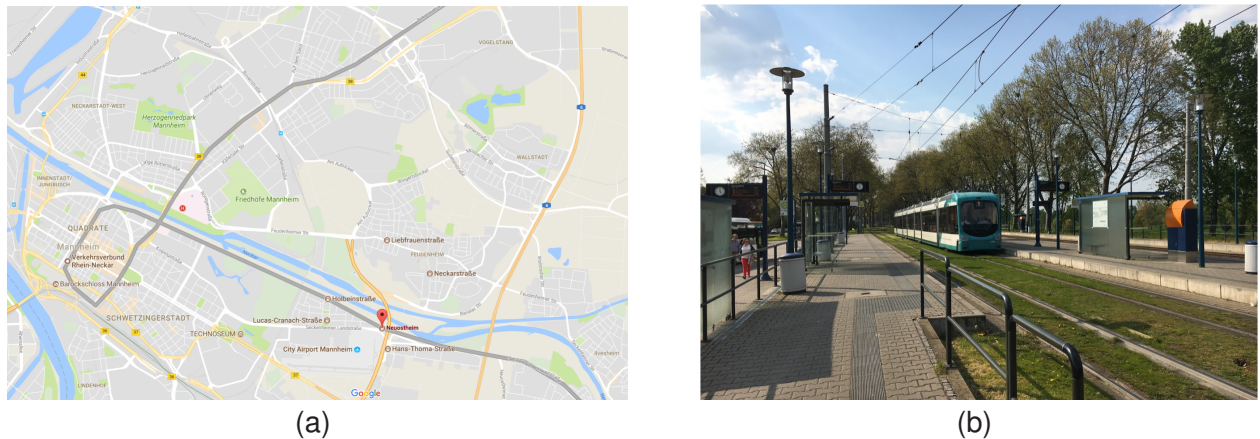


Figure 6.14: (a) Neuostheim electrical substation; (b) Neuostheim passenger station. Source: Google Maps.

6.3.3 Neuostheim electrical substation

The map in figure 6.14a shows the geographical location of the Neuostheim electrical substation and its position in the RNV5 tramway. All the voltage and current measurements for this work were made at this electrical substation. A photo of the Neuostheim passenger station with a RNV8 tram from the RNV5 line passing by can be seen in figure 6.14b.

A schematic of the railroad segment under study can be seen in figure 6.15, where the *Neuostheim* electrical substation (see figure 6.16) feeds four different track segments numbered from 211 to 214. Track segments 211 and 212 connect *Neuostheim* and *Pforzheimer Str.* electrical substations, with three passenger stations in between¹⁶. On the other hand, track segments 213 and 214 connect *Neuostheim* and *O41* electrical substations, with five passenger stations in between¹⁷.

6.4 Measurement analysis

6.4.1 Input data acquisition

All the data gathered and used in the present work can be sub-classified in the following data-sets:

- Neuostheim substation: Four voltage and current probes are installed at the *Neuostheim* electrical substation (see section 6.3.3), each one sampling the catenary DC voltage and current of every track segment with a 100 ms sample rate;

¹⁶ Neuostheim, Duale Hochschule and Pforzheimer Str.

¹⁷ Neuostheim, Holbinstr., Fernmeldeturm, Lessingstr. and Collini-Center.

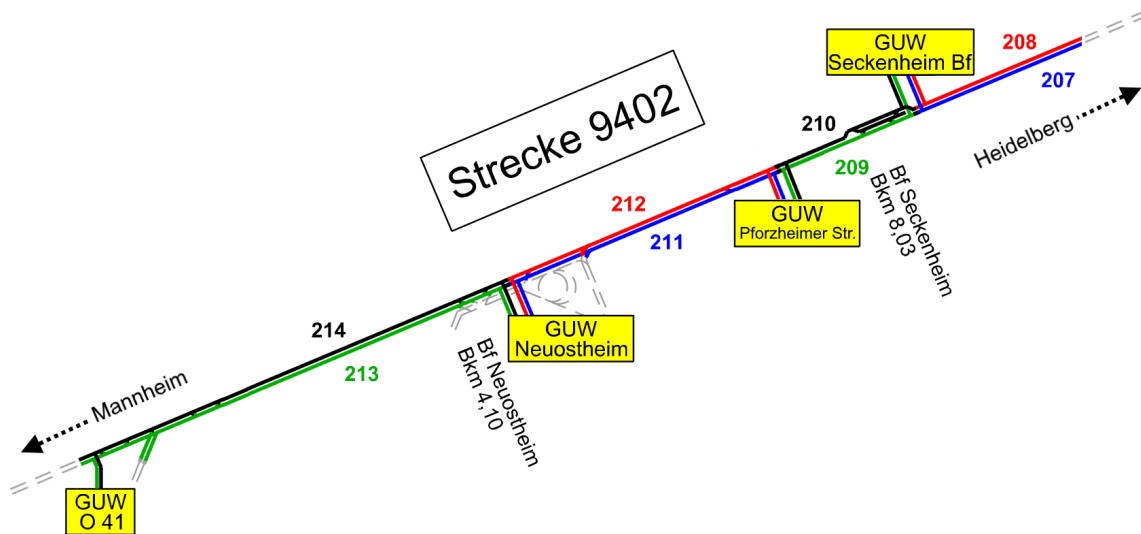


Figure 6.15: Railroad segment under study of the RNV5 line. The measurements are taken at the Neuostheim electrical substation (*G UW Neuostheim*), which is directly connected to the track segments 211, 212, 213 and 214. Source: RNV (Modified).

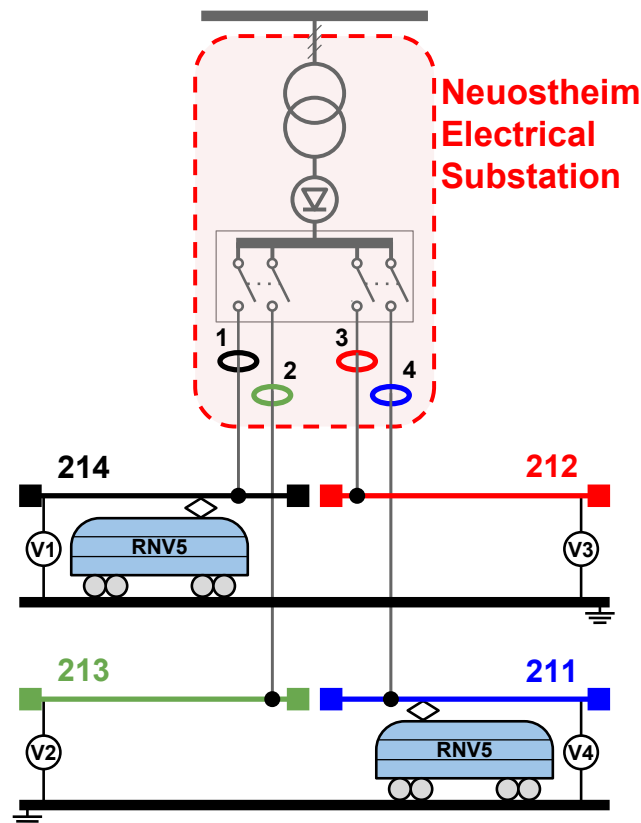


Figure 6.16: Current and voltage sensors installed at the Neuostheim electrical substation, which is directly connected to the track segments 211, 212, 213 and 214.

- **Scheduled time-tables:** Real time information regarding the time-tables of the vehicles and departure times is recorded from the RNV API. This information is vital to analyze the voltage and current signals recorded at the substation and to be able to decipher the power consumption and energy recovery of the trains as they travel through the tracks. More details are presented throughout this section.
- **Weather:** Meteorological variables are stored and used as inputs to a predictive model that correlates them with other information such as the amount of energy recovered during the braking. This model allows to predict (with a predefined confidence interval) the amount of energy that the system will generate in a certain interval of time in the future. More detailed information on this procedure is presented in chapter 7.

6.4.2 Implications of the current measurement scheme

Equation 6.4 describes the possible destinations of the total amount of energy harvested through dynamic braking ($E_{\text{gen}_{\text{tot}}}$) with the current system configuration. They include the energy burned through rheostatic braking¹⁸ ($E_{\text{gen}_{\text{rheostatic}}}$), the energy used for on-board equipments ($E_{\text{gen}_{\text{on-board}}}$), the energy transferred to another vehicle in the same track segment ($E_{\text{gen}_{\text{segment}}}$), the energy transferred to vehicles traveling on the other three track segments of the Neuostheim substation ($E_{\text{gen}_{\text{sameSS}}}$), the energy transferred to vehicles traveling on neighboring track segments connected to other substations ($E_{\text{gen}_{\text{otherSS}}}$) and a term accounting for the transmission losses between source and load ($E_{\text{gen}_{\text{loss}}}$).

$$E_{\text{gen}_{\text{tot}}} = E_{\text{gen}_{\text{rheostatic}}} + E_{\text{gen}_{\text{on-board}}} + E_{\text{gen}_{\text{segment}}} + E_{\text{gen}_{\text{sameSS}}} + E_{\text{gen}_{\text{otherSS}}} - E_{\text{gen}_{\text{loss}}} \quad (6.4)$$

However, the current system configuration does not allow to record all of these variables. Different scenarios can be found in the data recorded at the substation:

- if the power is delivered from the Neuostheim substation to a train in one of its four segments, the recorded current value is positive;
- an accelerating train can be fed energy from the electrical substations of the same segment, as well as neighboring substations. Energy delivered from the *O41* substation to a train in segments 211 or 212 (see figure 6.15) is recorded as negative currents with equal value in segments 213 and 214 and a positive current in the line segment where the train lies (an analog situation exists with the *Pforzheimer Str.* substation);

¹⁸ Rheostatic braking is used only when the line receptivity is not good enough to absorb all the recovered energy.

- if the Neuostheim substation feeds energy to a train in a segment not directly connected to it¹⁹, it is recorded as a positive current value in both line segments of the same side (211 and 212 or 213 and 214);
- the energy recovered by a braking vehicle in one segment and exchanged with other vehicle in one of the remaining three line segments ($E_{\text{gen_sameSS}}$) is recorded as a negative current at the sensor from the segment where the braking vehicle is, and as a positive current in the line segment with the vehicle that consumes the energy (see figure 6.17);
- the energy recovered by a braking vehicle and burned through rheostatic braking ($E_{\text{gen_rheostatic}}$) is not recorded;
- the energy recovered by a braking vehicle and used on the on-board systems ($E_{\text{gen_on-board}}$) of the same vehicle is not recorded;
- the energy recovered by a braking vehicle and transferred to other vehicle in the same line segment ($E_{\text{gen_segment}}$) is not recorded;
- the energy recovered by a braking vehicle and transferred to other vehicle in a line segment of other substation ($E_{\text{gen_otherSS}}$) is not recorded.

These observations mean that only a fraction of the energy harvested through dynamic braking ($E_{\text{gen_tot}}$) is recorded by the installed sensors, and this fraction corresponds to the energy transferred between at least two trains which are traveling in two of the four line segments directly connected to the Neuostheim substation ($E_{\text{gen_sameSS}}$). Nonetheless, these measurements are valid (they represent a lower bound for the total amount of recovered energy) and are used as input parameters for the energy management system developed throughout this work. Increasing the receptivity of the line through the implementation of a ESS controlled by an EMS lowers the amount of energy wasted in rheostatic braking ($E_{\text{gen_rheostatic}}$) and, in consequence, increases energy savings and system efficiency.

6.4.3 Regenerative braking action

Figures 6.18 and 6.18 show (for different time periods) the amount of energy that is consumed and recovered in the system defined by the electrical substation, its four track segments and all the vehicles that travel through them. The average amount of energy recovered for each vehicle is 0,46kWh, with an average recovery of 0,68MWh each day for the

¹⁹ For example, segment 209.

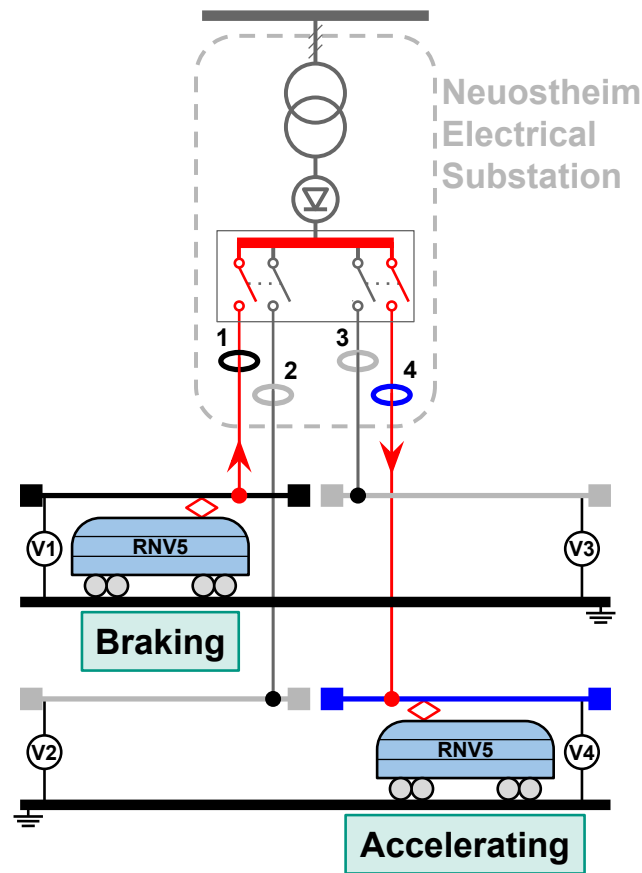


Figure 6.17: The energy recovered by a braking vehicle in one segment and exchanged with other vehicle in one of the remaining three line segments ($E_{\text{gen_sameSS}}$) is recorded as a negative current at the sensor from the segment where the braking vehicle is ($i_1 < 0$), and as a positive current in the line segment with the vehicle that consumes the energy ($i_4 > 0$). The current measurements for braking and accelerating phases in one track segment can be seen in figure 6.3.

defined system. This amount of energy represents approximately the 28,3% of the total consumed by this section of the railroad²⁰.

The reason for the existence of the great discrepancy between these measured values and the value obtained by the kinematic analysis of equation 6.3 lies in the fact that, as explained in section 6.4.2, the amounts of recovered energy that are recorded by the sensors represent only a fraction of the total value. In addition, unlike the model proposed in the kinematic analysis, the real system presents non-ideal energy conversion efficiencies, with energy losses in the different stages of recovery and transmission.

²⁰ The total amount of energy consumed by the system is obtained by adding the energy contribution of the three substations (Neuostheim, O41 and Pforzheimer Str.) together with all the energy that was recovered and re-utilized by the vehicles.

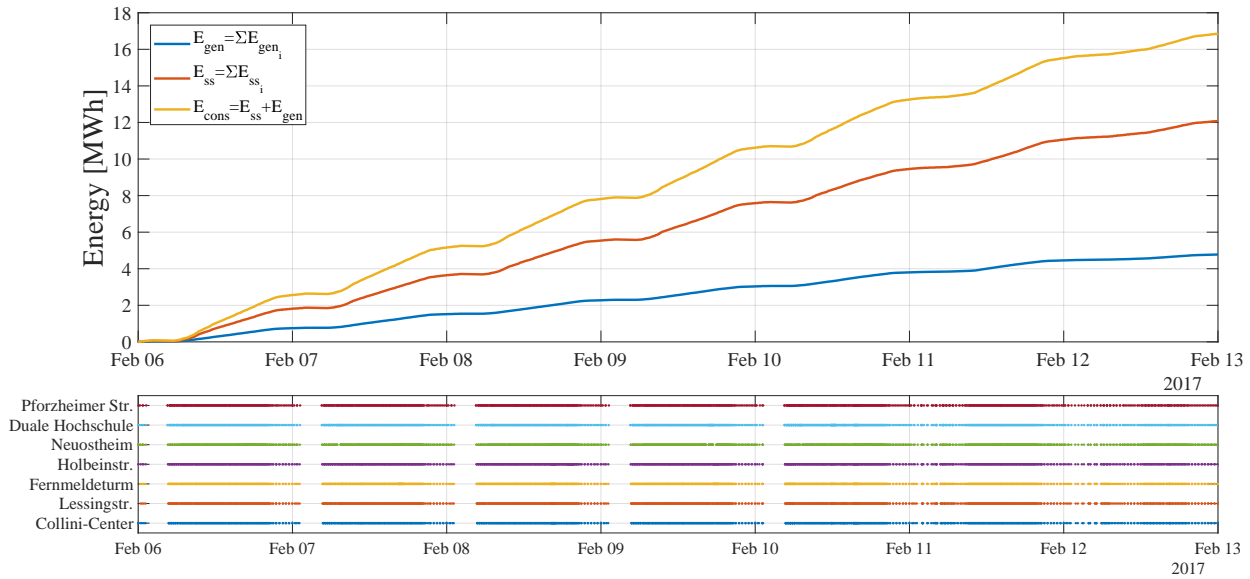


Figure 6.18: Detail of the energy consumption and generation during one week, corresponding to the four track segments directly connected to the Neuostheim substation. Top: energy recovered by braking trains and transferred between line segments (E_{gen}), energy delivered to trains in the four line segments by the Neuostheim and neighboring substations (E_{ss}), and total amount of energy consumed by the trains whilst traveling through the Neuostheim segments (E_{cons}). E_{gen} represents 28,3% of the total energy consumed at the four segments; the day average for E_{gen} is 0,68 MWh; the average for each train is 0,46 kWh. Bottom: time-table for the seven passenger stations inside the Neuostheim line segments. Each point corresponds to a train arriving at the station. The service interruption at nights and the extended service during weekends is clearly appreciable.

Nevertheless, even a partial measurement of the total energy recovered yields very significant results in energy savings²¹, and they could be considerably improved if the line receptivity is improved (see section 6.2.1). The latter would allow to take advantage of the energy that is currently being wasted by rheostatic braking ($E_{gen_{rheosatic}}$).

6.4.4 Daily energy recovery profile

Figure 6.20 shows how the amount of energy recovered varies during the day. These plots are made averaging the values recorded for 103 days between the months of December 2016 and February 2017. As the figure shows, a clear distinction is found in the generation patterns for the working days and the weekend days, which related to the traffic intensity and the amount of passengers. As it is shown in chapter 7, this distinction between working

²¹ As section 6.2.1 states, application of regenerative braking in urban railways has the potential to reduce net energy consumption from 10% to 45%, depending on the characteristics of the system.

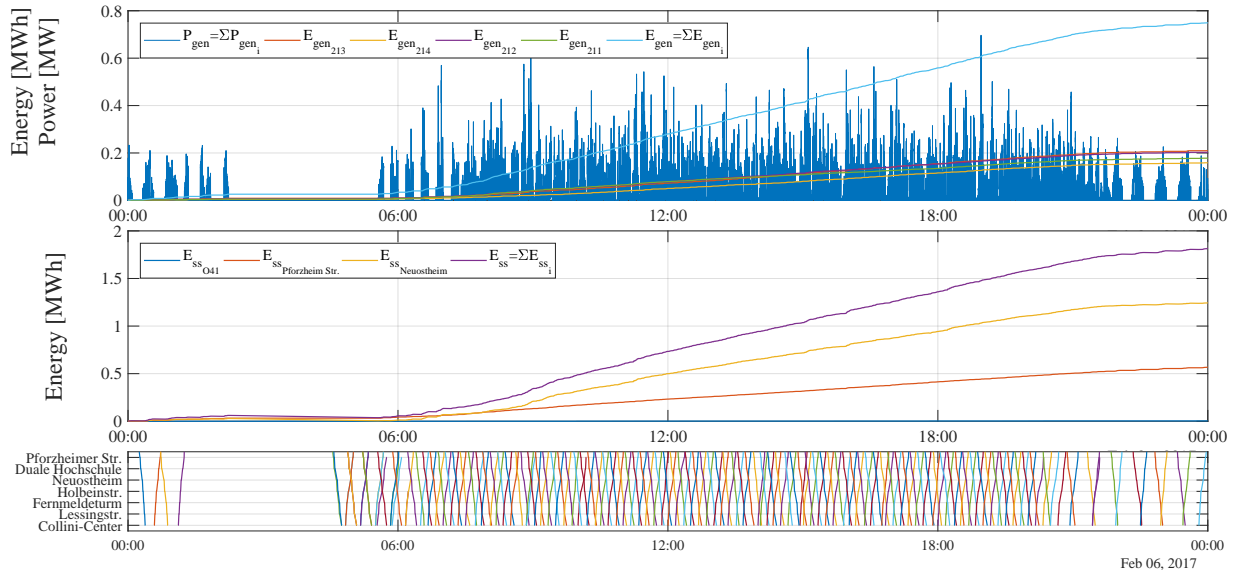


Figure 6.19: Detail of the energy consumption and generation during one day, corresponding to the four track segments directly connected to the Neuostheim substation. Top: recovered energy at the $i - th$ line segment and transferred to one of the three remaining segments ($E_{gen_i}, i = 211, 212, 213, 214$), total amount of recorded recovered energy (E_{gen}), and total instantaneous recovered power (P_{gen}). Middle: energy delivered by the different substations to the trains in the segments 211 – 214 ($E_{ss_i}, i = O41, Neuostheim, Pforzheimer$), and total amount of energy delivered by all the substations (E_{ss}). Bottom: time-table for the seven passenger stations inside the Neuostheim line segments. Each line corresponds to the route followed by a train through the stations in the segments. The service interruption at nights is clearly appreciable.

days and weekend profiles is important to achieve good performances in the creation of prediction models for recovered energy.

As it is exposed in section 1.2, the information presented in figure 6.20 of the recovered energy can be used as an input to an electronic system that controls the energy flow between a microgrid (defined by renewable energy generation systems, energy storage devices and electric loads) and the main distribution grid in order to meet a predefined set of objectives and constraints.

Moreover, these profiles will be vital for determining the optimal size for the ESS to be implemented. The whole untapped energy can be saved (excluding losses in ESS and wires) if both the rated power and the rated energy of the ESS are larger than the total braking power and braking energy of the maximum number of vehicles running within the feeding zone of the substation. However, the situation when all vehicles simultaneously or sequentially brake and no one accelerates is very seldom, so there is no sense in such a

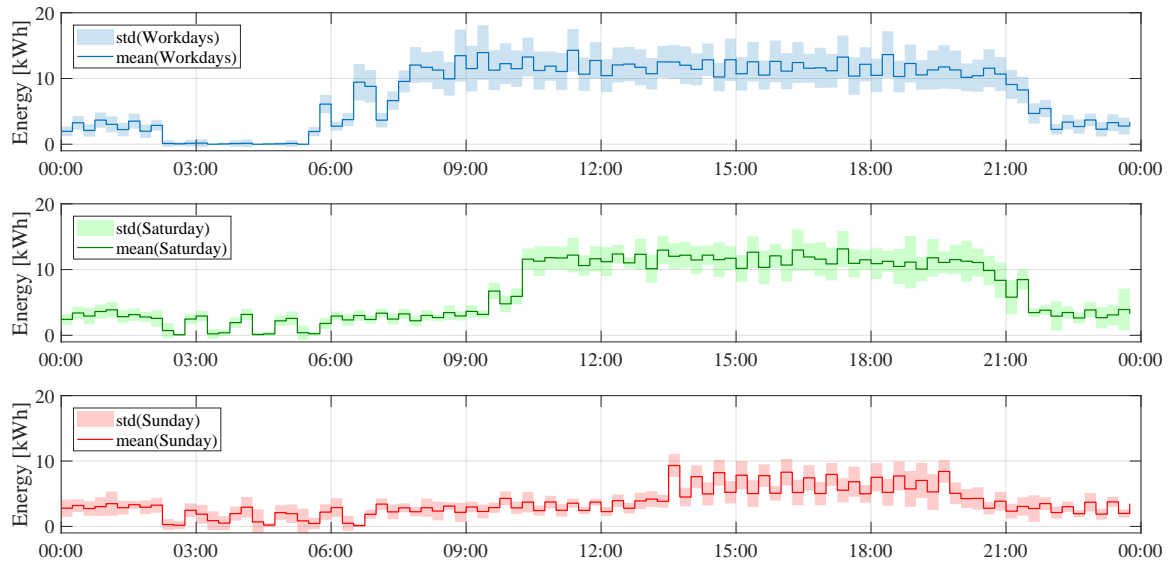


Figure 6.20: Mean and standard deviation values for the amounts of energy recovered in fifteen minute intervals throughout the day; 103 days of recorded data (from 10-12-2016 to 22-03-2016) are used in these plots. There is a clear distinction in the patterns of energy generation for working days, Saturday and Sunday. This distinction is directly correlated with the social patterns with which the means of transport are used, which in turn impacts on service time-tables and the number of people aboard the vehicle (influencing the total mass and kinetic energy).

sizing of the ESS to capture all the untapped energy. As explained in section 5.4, in order to find the optimum power and energy requirements for an ESS it is necessary to investigate the influence of its limited power rating and energy capacity on the energy savings [97]. Thus, the sizing procedure for the ESS is an optimization problem that searches for the option that presents the highest economic profitability while meeting at the same time a set of objectives and constraints like perspectives of energy savings, installation and operation costs [105, 221, 228, 175, 83].

6.4.5 Over-current protection

Figure 6.21 shows how the over-current protection of the Neuostheim electrical substation is activated due to a high load demand when a train is accelerating; after the protection is enabled, the substation needs about 15s to normalize its output and then the train accelerates with a less demanding profile.

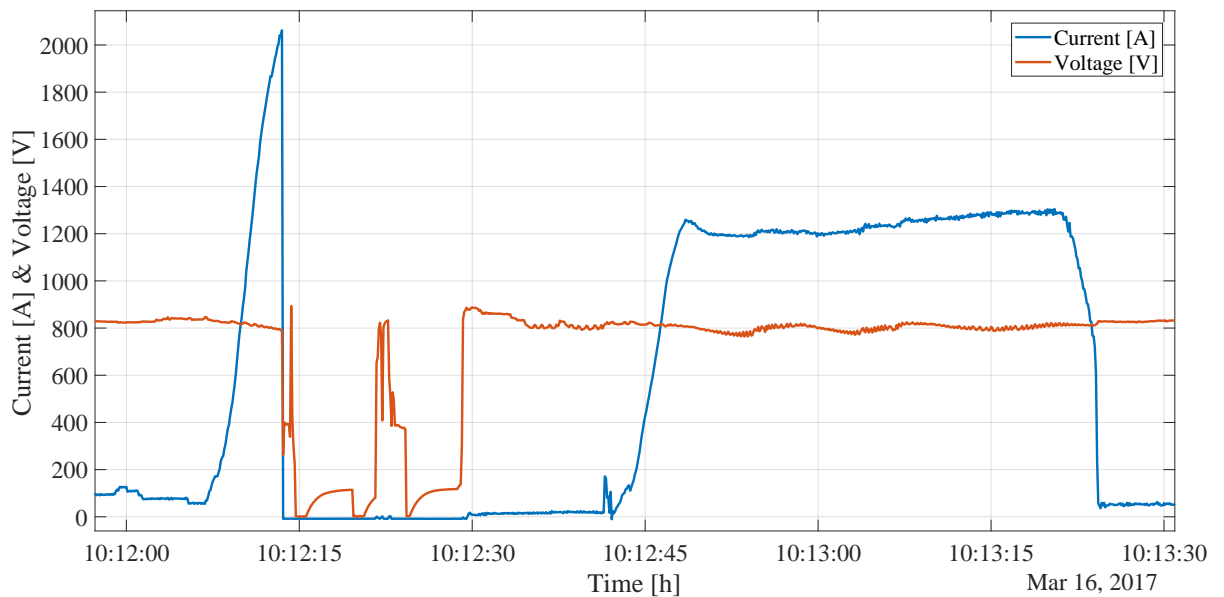


Figure 6.21: Over-current protection of the Neuostheim electrical substation; after the protection is enabled, the substation needs about 15s to normalize its output voltage.

6.4.6 Thevenin equivalent of the electrical substation

As described in section 6.1.2, the DC voltage value at the output of the electrical substation is given by a six-pulse rectifier, and since it is out of the scope of this work to evaluate effects of harmonics, only the DC component of this source is of interest. Therefore, the electrical substation can be reduced to the well known Thevenin equivalent circuit with a DC source of value V_{th} and a series resistor R_{th} . An ideal diode is connected in series with the resistor to match the non-reversible behavior of the passive converter (see figure 6.22).

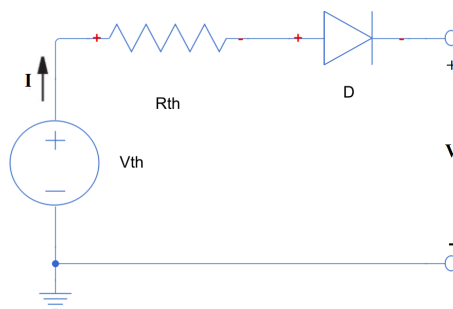


Figure 6.22: Equivalent circuit of the electrical substation.

In order to approximate the Thevenin equivalent values, a random day of data is chosen and the plot from figure 6.23 showing the current against the voltage is generated. This model requires to know the no-load voltage and the resistance values. A linear regression

of the data is performed for currents bigger than 150 A, obtaining an equivalent resistance of $R_{th} = 18\text{m}\Omega$ and a no-load voltage of $V_{th} = 817\text{V}$. For current values lower than 150 A, there is a big voltage fluctuation. The reason relies in the regenerative braking effect. As explained in section 6.2, when trains brake and recover energy the traction converter raises the voltage of the line to feed back the energy. The actual voltage level will depend mainly on the line receptivity at that particular moment.

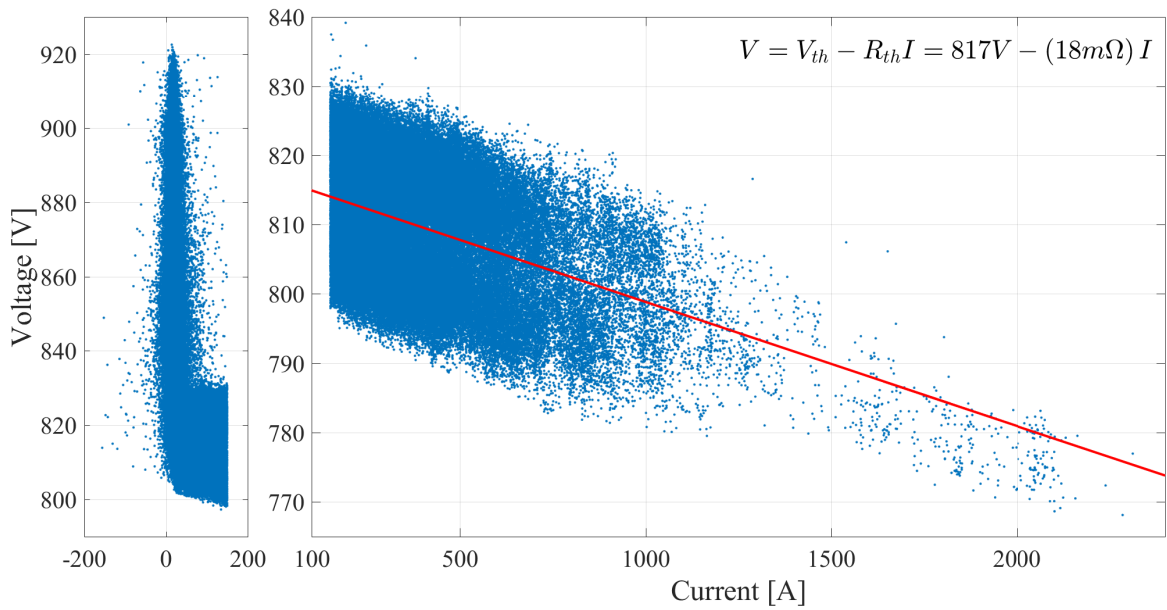


Figure 6.23: Voltage vs. current at the output of the Neuostheim electrical substation for a random day.

7 Short term forecasting in electric power systems using artificial neural networks

As described in section 1.2, the EMS controller depicted in figure 1.7 needs as input several uncertain parameters that must be predicted within a certain error margin. In consequence, this chapter presents an introduction to the general data forecast problem, and one of the most widely used forecasting methods is briefly explained, i.e., the use of *artificial neural networks* (ANNs).

The chapter is organized as follows: section 7.1 enumerates the different classifications for the forecasting problem concerning time frames and mathematical models implemented; section 7.2 presents a theoretical synthesis of the concepts needed to apply one of these mathematical models (i.e., ANNs); section 7.3 describes the need to perform an *interval forecast* instead of a *point forecast*; the general implementation of the *lower upper bound estimation* (LUBE) method using ANNs for the generation of interval forecasts is explained in detail in section 7.4; the LUBE method is applied for the forecast generation in a case study throughout section 7.5; finally some conclusions and future prospects are presented in section 7.6.

7.1 Introduction to forecasting

The level of uncertainty present in the electrical power system has been increasing since the advent of competitive energy markets¹, in which the system evolves from a centralized to a decentralized scheme with a greater penetration of renewable energies. As a result, forecasting systems become both more complex and more necessary to perform a good system administration. Thus, the forecasting of electricity demand, generation and price has become one of the major research fields in electrical engineering for smart grid applications [107, 188].

¹ Many countries have recently privatized and deregulated their power systems, and electricity has been turned into a commodity to be sold and bought at market prices.

7.1.1 Classification

The hierarchy of control systems used throughout the entire generation and distribution chain require forecasts with different lead times depending on their position on the chain. Based on time-scale, forecasts can be broadly classified into three main categories [227]:

- Short-Term Forecast² (STF): The STF time period lays between a few minutes or hours, up to a day-ahead or a week. STF is a fundamental factor in day-to-day operations (economic dispatch), optimal generator unit commitment (UC), estimation of available transfer capability, and scheduling functions (i.e, spinning reserve and load shedding decisions), real-time control and security assessment (i.e., stability margins) [107, 188, 219, 138].
- Mid-Term Forecast (MTF): The MTF time period lays between a month to a year or two. MTF is used for maintenance scheduling, fuel scheduling, hydro reservoir management and price settlement to balance demand and generation [188].
- Long-Term Forecast (LTF): The LTF time period lays between a few years to 10-20 years ahead. The purpose of these forecasts is to be used for system expansion planning (i.e., generation, transmission and distribution) and maintenance scheduling [138].

According to this definition, the resolution of the problem defined in section 1.2 implies the implementation of STF systems. A review of the state of the art of medium- and long-term forecast systems can be found in [138].

7.1.2 Dependencies of the signals

Regardless of the time-scale used, the forecast of load, generation, or price signals is not a simple task; this is due to a number of factors:

- the statistical analysis shows that these signals are complex time series with different levels of seasonality³;

² Due to the advance of the deregulation of the electric markets, the time slots traded in the intra-day market are becoming smaller. This has led certain authors to further differentiate the STF in *very short-term forecast* (VSTF) for some minutes, and *short-term forecast* for a few hours, up to a week [219, 190].

³ For example, the value of the signal for a given time-step can depend on: the value in the previous time-step, the value in the same time-step in the previous day, and in the value in the same time-step in the previous week in the day with the same denomination. For medium- and long-term predictions, other seasonalities such as changes in the signal due to calendar seasons also become important.

- exogenous variables also influence the values taken by the signals⁴.

According to Hippert *et al.* [107], it is relatively easy to achieve forecasts with about 10% mean absolute percent error (MAPE); however, the economic costs associated with each percentage point are so high that any research focused on diminishing the extent of the error is duly justified.

7.1.3 Models

The problem of forecasting load, energy generation and price signals has already been addressed with most of the forecasting models and methods in existence, obtaining different degrees of success. The most popular models can be classified into three categories [188]:

- *statistical* models: such as the auto-regressive (AR), AR integrated moving average (ARIMA), ARIMA with Exogenous Variables (ARIMAX), and exponential smoothing (ES) models, to name a few [107, 215, 62, 71, 167];
- *artificial intelligence* models: such as Neural Networks (NNs) [136, 134, 187], fuzzy logic systems (FLSs) [130, 137, 120], and expert systems [92];
- *hybrid* models: like neuro-fuzzy systems [70, 53, 19].

Since the first reports on their application to the load forecasting problem were published in the late 1980's and early 1990's [67], artificial neural networks (ANNs) are the models that have received the largest share of attention on the application of artificial intelligence techniques to the forecasting problem [107].

As a preview of the next section, it can be said that some of the reasons for this preference for ANNs over other forecasting methods are [135]:

- being an analytical technique inspired in biology, ANNs have the capability to learn and model complex nonlinear relationships⁵ [135];
- ANNs do not require *a priori* model to be assumed or *a priori* assumptions to be made on the properties of data [135, 34];
- ANNs have been successfully implemented to solve modeling, prediction, classification, optimization, and control problems across a wide range of disciplines [135].

⁴ Examples of these variables are social and weather-related variables, like temperature and humidity [177]: a particularly cold or rainy day during the summer could obtain the incorrect predictions if only a statistical analysis of the historical data is made without taking into account the meteorological variables in the analysis.

⁵ According to Kolmogorov's *Universal approximation theorem* [108, 103], *feedforward* NNs with one *hidden layer* that comprises an adequate number of *neurons* are considered to be universal approximators, and as such, have an excellent ability of approximating any nonlinear mapping to any degree of accuracy.

7.2 Multilayer feed forward artificial neural networks (ANNs)

The amount of knowledge linked to the research, development and application of neural networks is really vast and its ramifications cover various fields of study; a complete treatment of the subject may be found in [34, 103]. This section briefly introduces only those concepts that are required for the application of ANNs to the forecasting problem in smart grid applications.

7.2.1 Artificial neurons

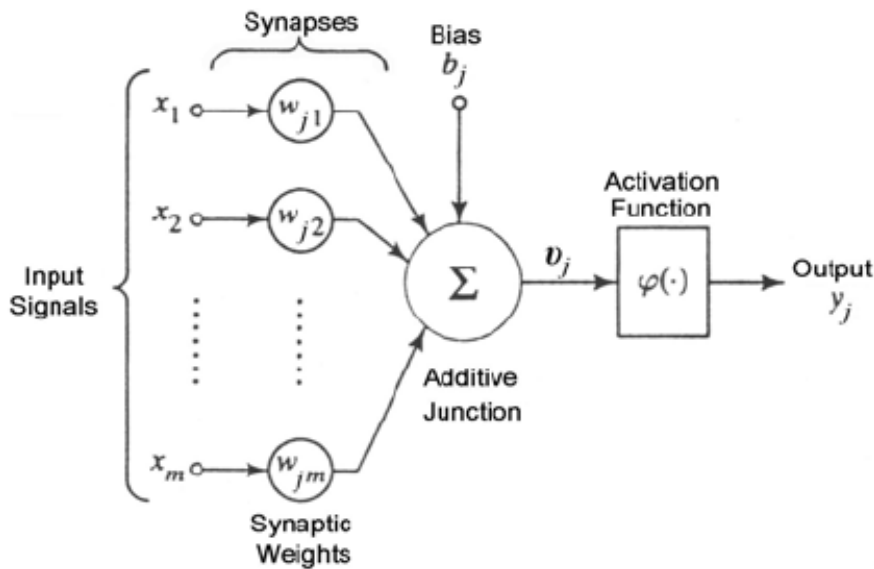


Figure 7.1: Schematic of an artificial neuron. Source: [103].

ANNs are mathematical tools originally inspired by the way the human brain processes information, and their elementary building block is an information processing unit called *artificial neuron*; this mathematical function is schematically represented in figure 7.1. It receives (numerical) information through a number of input nodes; this information is then processed internally and an output response (or activation) is generated. The processing is usually done in three stages:

1. the signal x_i at the input of the i -th synapse of the j -th neuron is multiplied by the *synaptic weight* w_{ji} ⁶;

⁶ The subscripts of the weight w_{ji} mean that this particular weight corresponds to the i -th input of the j -th neuron.

2. the m weighted *input signals* of the j -th artificial neuron are linearly combined, together with a constant *bias* term⁷ b_j , represented in figure 7.1 by the weight of a connection with a fixed input equal to +1;
3. the result of the linear combination is used as the argument of a nonlinear function known as an *activation function* or *transfer function*. The output value of this activation function constitutes the output of the artificial neuron.

Thus, according to the diagram shown in figure 7.1, we may describe an artificial neuron j in mathematical terms with the following equations:

$$\begin{cases} v_j = \sum_{i=1}^m (w_{ji}x_i) + b_j \\ y_j = \varphi(v_j) \end{cases} \quad (7.1)$$

where x_1, x_2, \dots, x_m are the input signals; $w_{j1}, w_{j2}, \dots, w_{jm}$ are the synaptic weights of the j -th neuron; b_j is the bias; $\varphi(\cdot)$ is the activation function; v_j is the output of the linear combiner, called *activation potential* or *induced local field* of the j -th neuron; and y_j is the output of the neuron. Hence, the output of the j -th neuron is:

$$y_j = \varphi \left(\sum_{i=1}^m (w_{ji}x_i) + b_j \right). \quad (7.2)$$

Types of activation function

The activation function must comply with certain requirements, such as being monotonically increasing, continuous, differentiable and bounded [103]; the most common choices are either the identity function ($y = x$), or bounded sigmoid (s-shaped) functions. However, the activation function may also take the form of other nonlinear function as shown in figure 7.2.

Nonlinear activation functions are preferred in nonlinear problems. However, the use of sigmoid activation functions may cause saturation problems, since this function is highly nonlinear outside the region $[-1, 1]$. Therefore, in order to avoid saturation problems, the input and output signals of the ANN are normalized using equation 7.3 so that they lay inside the $[-1, 1]$ region [203, 219]:

$$\hat{x} = a + \frac{b-a}{x_{max} - x_{min}} (x - x_{min}) \quad (7.3)$$

⁷ "The bias has the effect of increasing or lowering the net input of the activation function, depending on whether it is positive or negative, respectively." [103]

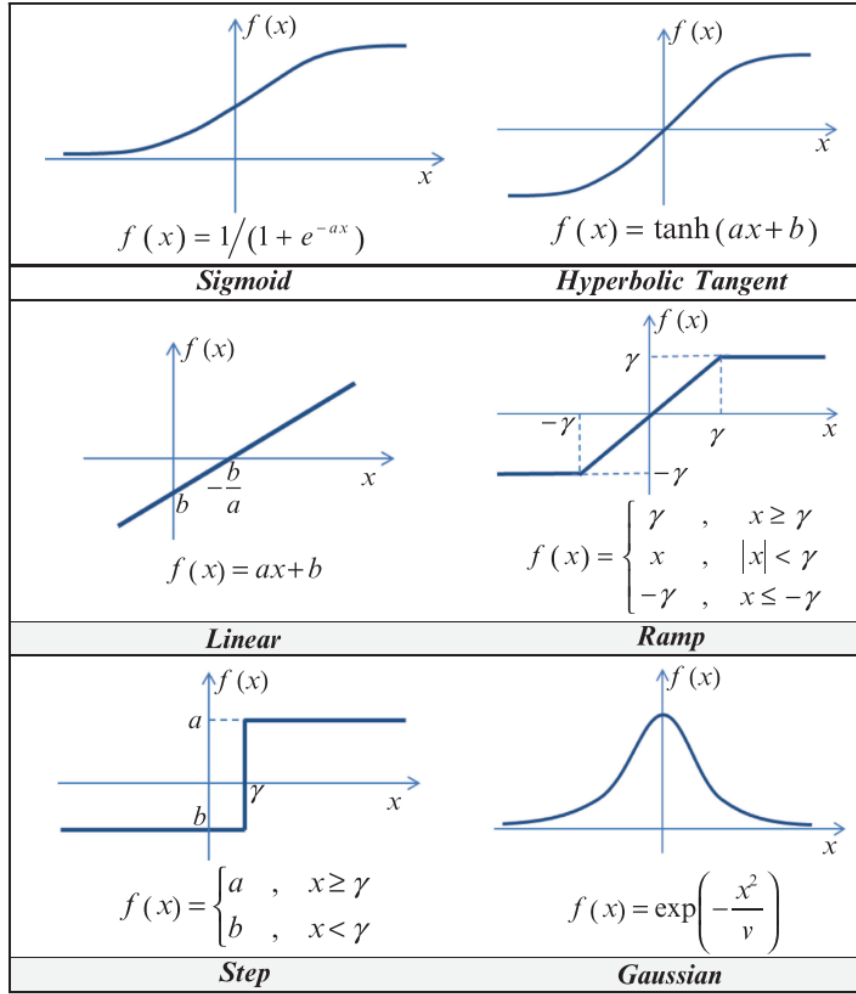


Figure 7.2: Commonly used activation functions. Source: [219].

where \hat{x} is the normalized value of variable x ; x_{min} and x_{max} are the lower and the upper bounds of x ; and a and b are the respective values of the normalized variable. Once the output of the network is obtained, it is denormalized by solving for x in the same equation.

7.2.2 Network architecture

The neurons are organized in a way that defines the network *architecture*, and the type of architecture used depends on the problem to be solved. Several authors in the specialized literature have come to the conclusion that the implementation of a *feedforward multilayer perceptron* (MLP) architecture is the best option for the preparation of forecasts [107, 132].

The general diagram of this architecture can be depicted in figure 7.3, where the neurons are organized in one *input layer* of input nodes, one or more *hidden layers* of artificial neurons, and one *output layer* of artificial neurons. The MLP is characterized as a *feedfor-*

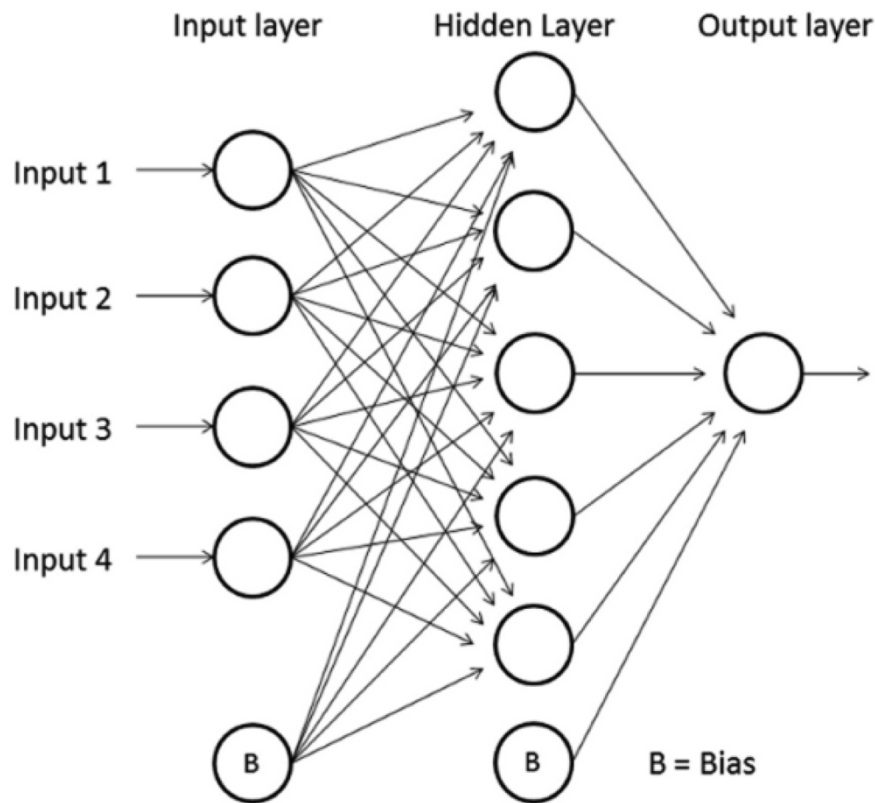


Figure 7.3: Fully connected Multilayer Perceptron (MLP) network. Source: [64].

ward network, since the information flow is only defined on the direction from the input to the output⁸. Furthermore, the neurons in each layer have as their inputs only the output signals of the preceding layer.

The network is classified as *fully connected* when every neuron in each layer is connected to every neuron in the adjacent forward layer; if one or more links are missing, the network is defined as *partially connected*.

The presence of hidden layers allows the network to learn complex tasks by performing high-order statistics and, therefore, extract more significant features from the input signals [103]. In this architecture it is possible to have more than one hidden layers, but according to Kolmogorov's *Universal approximation theorem* [108, 103], feedforward NNs with one hidden layer (that comprises an adequate number of neurons) are considered to be universal approximators, that is, they can approximate nonlinear mappings to any desired degree of accuracy. Therefore, in most practical cases only one hidden layer is sufficient. Nonetheless, the number of neurons belonging to that layer must be chosen appropriately based on the requirements of the problem to be solved.

⁸ That is, this architecture is not *recurrent* (i.e., no feedback is present in the network).

Figure 7.3 shows an example of a fully connected MLP network with four input nodes and two layers (that is, a hidden layer with five neurons and an output layer with only one neuron).

Consider a $N_i - N_h - N_o$ MLP network. That is, a network with N_i input signals, one hidden layer with N_h neurons, and N_o neurons in the output layer. Then, the parameters for this network are the weight matrix $W_{N_h \times N_i}$ (containing the weights w_{ji} that connect the j -th hidden neuron to the i -th input signal), the weight matrix $U_{N_o \times N_h}$ (containing the weights u_{kj} that connect the output of the j -th hidden neuron to the k -th neuron of the output layer), and the bias vectors $B_{N_h \times 1}$ and $B_{N_o \times 1}$ for the hidden and output layers.

The general mathematical model represented by this network is the following:

$$y_k = \varphi_o \left(\sum_{j=1}^{N_h} u_{kj} \cdot \varphi_h \left(\sum_{i=1}^{N_i} w_{ji} x_i + b_j \right) + b_k \right) ; k = 1, \dots, N_o \quad (7.4)$$

where y_k represents the output of the k -th neuron in the output layer, $\varphi_h(x)$ and $\varphi_o(x)$ are the activation functions of the neurons in the hidden and output layers, respectively. Equation 7.4 shows how complex and flexible even a small network can be.

A common implementation in the forecasting literature uses logistic functions for the activation of the hidden layer, and linear functions for the output layer [107, 135, 136]. The resulting model is as follows⁹:

$$y_k = \sum_{j=1}^{N_h} \left(u_{kj} \cdot \frac{1}{1 + \exp \left(- \sum_{i=1}^{N_i} w_{ji} x_i + b_j \right)} \right) + b_k ; k = 1, \dots, N_o. \quad (7.5)$$

7.2.3 Network training

The estimation of the network parameters is called the *training* of the network, and is done by the minimization of a predefined cost function. Therefore, through the implementation of a *learning algorithm*, the network is able to change its internal parameters in function of external stimuli, *learning* from its environment and *improving* its performance.

This learning procedure suggests that neural networks are data-driven methods. This means that when creating and training the network, the designer makes no assumptions about the model or properties of the data. By using only a sample of input and output signals, the network is able to learn the relationship between these two sets of signals, and map it into its synaptic weights [34]. Therefore, these networks are particularly useful when

⁹ The slope a and intercept b of the linear activation function, as well as the scale factor a in the sigmoid activation function shown in figure 7.2 are set to one in this example.

one has a lot of data, but little information about the system that generated this data, and the laws that govern its behavior.

There is a diverse variety of learning algorithms, *back-propagation* being the first one to be formulated and one of the most popular. The main characteristic that differentiates these algorithms is the way in which the synaptic weights are adjusted in each iteration [103].

7.3 Prediction intervals (PIs) estimation

As mentioned in previous sections, the field of forecasting is a popular application of neural networks. However, most applications reported in the literature are focused on point forecasts [107], i.e., at time t the task is to predict the signal value for time $t + h$, where h is the forecasting horizon [190]. In practical applications, this approach presents two limitations when the uncertainty in the signal increases¹⁰ [135, 187]:

- The reliability of the point forecast decreases significantly, obtaining a low prediction performance. The reason is purely theoretical, and in the words of Khosravi *et al.*:

"Statistically, the NN output approximates the average of the underlying target conditioned on the NN input vector. If the target is multivalued, the NN conditional averaged output can be far from the actual target, and is therefore unreliable." [136]

- Point predictions yield no information about the accuracy of the prediction performed [136]. It would be very useful for decision-makers and control systems if the predictor system could establish a measure of the uncertainty of the deterministic approximation performed by the neural network, since the prediction would increase its reliability and credibility.

To solve these drawbacks, various methods have been proposed for constructing the so-called *prediction intervals* (PIs) [135, 136].

As can be seen in figure 7.4, most existing methods for creating prediction intervals use past observations of the signal to create a point forecast value (the *model output*). This point forecast is then used to create a *prediction interval*, composed of an *upper bound* and a *lower bound*, which contains a future unknown value, or *target*, with a certain prescribed probability denominated *confidence level* $[(1 - \alpha) \%]$ or *coverage probability* [52]. Thus, PIs

¹⁰ Events that cannot be properly predicted in advanced like machine breakdowns and unexpected repairs, abrupt changes in demand or supply and sudden weather changes may be possible sources of uncertain patterns in the signal. Case studies of various industries dealing with this problem can be found in [135, 136] and its references.

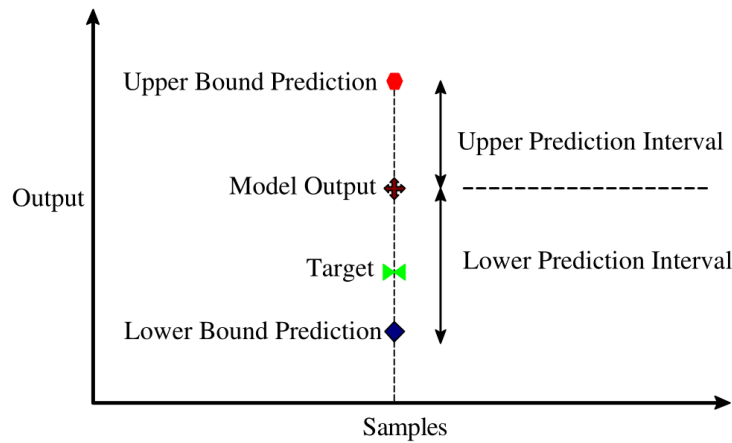


Figure 7.4: Prediction Interval (PI) concept. Source: [131].

tell not only the range of values that the target will fall in, but also have an indication of their accuracy, indicating on which percentage targets will be covered.

Therefore, it can be asserted that the PIs present two fundamental characteristics that define them and differentiate them from the point forecasts: width and coverage probability. As Chatfield points out in [52], this additional information allows decision-makers to consider a wide range of solutions/scenarios for the best and worst conditions and plan appropriate strategies with more confidence, maximizing their benefits. For example, as Khosravi *et al.* describe in [135], the width of the PI is an indicator of the uncertainty present in the prediction, and its value may influence decision makers to avoid or face risky actions in the future. These properties make them especially useful for, among other things, applications that require supply and demand balancing, such as in electricity markets [190].

Since it can become a source of confusion, before proceeding it is necessary to explain the difference between a *prediction interval* (PI) and a *confidence interval* (CI). Khosravi *et al.* clearly explain these differences:

"Confidence intervals (CIs) and prediction intervals (PIs) are two well-known tools for quantifying and representing the uncertainty of predictions. While a CI describes the uncertainty in the prediction of an unknown but fixed value, a PI deals with the uncertainty in the prediction of a future realization of a random variable. By definition, a PI accounts for more sources of uncertainty (model misspecification and noise variance) and is wider than the corresponding CI." [136].

The most widespread and traditional methods found in the literature for the construction of PIs include the delta, Bayesian, mean-variance estimation, and bootstrap techniques. A thorough explanation of them can be found in [135].

However, several difficulties are encountered when trying to apply these methods to the resolution of practical problems. These include, but are not limited to, high complexity of the resolution algorithms, need to hypothesize about the nature and distribution of data, and the need for high computational power. As exemplified by Quan *et al.*:

"For example, Jacobian matrix and Hessian matrix need to be calculated in delta and Bayesian methods separately. Delta method assumes that the noises are normally distributed and t-distribution is applied. Bootstrap method assumes that an ensemble of NN models will produce a less biased estimate of the true regression of the targets." [187]

Thus, the complexity of traditional methods hinders widespread applications of PIs [188]. Furthermore, regardless of their implementation differences, all of these traditional methods construct NN-based PIs using as input data the point forecast calculated by a data regression to a specified model or function [187]. Moreover, these NNs perform a minimization of an error-based cost function, such as the sum of squared errors or weight decay cost function [136]. According to Khosravi *et al.* in [136], this methodology of PI construction is questionable, since the objective of the described algorithms is to minimize the error, instead of improving the quality of the PIs. Therefore, the obtained PIs will not be optimal in terms of their two key features: width and coverage probability. This is why a new method of constructing PIs has recently been proposed and is gaining considerable traction in the community and already has examples of applications in several fields [204, 189, 214, 224, 128, 110, 109, 236, 190]. This method, named *lower upper bound estimation*, or LUBE, was initially developed by Khosravi *et al.* in [136] and aims to maximize the quality of the PIs.

7.4 Lower upper bound estimation (LUBE) method

LUBE [136] is a *nonparametric method* for construction of PIs. Thus, unlike traditional methods for PI construction, it does not make any assumptions regarding the probability distributions of the variables being assessed, and there is no need to calculate computationally demanding matrices¹¹, like Jacobian or Hessian matrices.

To accomplish this task, the method implements a fully connected feedforward NN model with two outputs, which correspond to the values of the upper and lower bounds of the PI (see figure 7.5). In this way, the PIs are directly constructed in one step¹², gaining speed

¹¹ Thus avoiding singularity problems.

¹² As mentioned in section 7.3, traditional methods first make a point forecast, and the prediction interval is estimated through the evaluation of the mean and variance values of the deterministic forecasts.

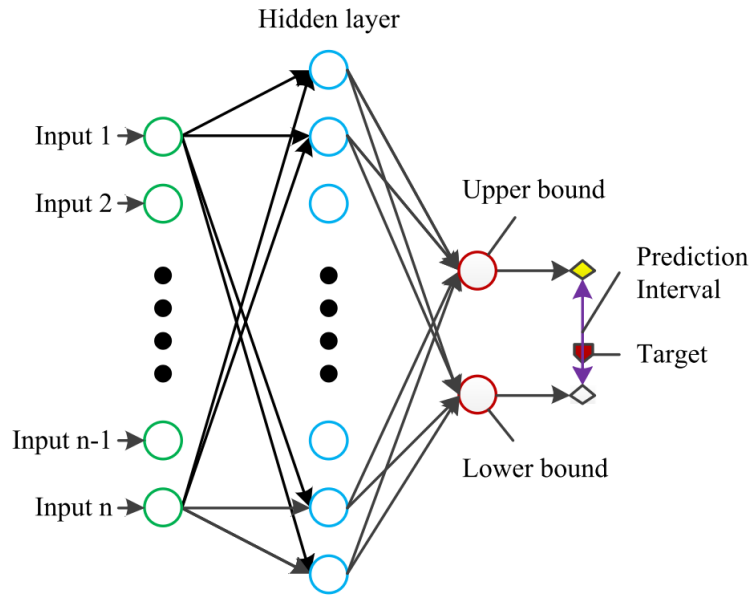


Figure 7.5: NN architecture of the LUBE method. Source: [188].

and simplicity. Furthermore, since the input signals are the only information used to try to approximate the lower and upper bounds of the PIs, LUBE does not suffer from practical concerns about data distribution [128].

In order to generate high quality PIs at the output of the NN, this technique performs the training process using a specially designed cost function that focuses on enhancing the important features of PIs¹³. This is not an error-based cost function like the traditional methods, and since it is highly nonlinear, complex, and discontinuous [136], the process of training the MLP through Back-propagation has to be avoided. Instead, meta-heuristic optimization methods such as Simulated Annealing (SA) [190], Genetic Algorithms (GA) [189], Bat Algorithm (BA) [125] or Particle Swarm Optimization (PSO) [187] are applied for the minimization of the cost function.

The creators of this PI construction technique reported through several synthetic and real-world applications that the quality of the PIs achieved with this technique is superior compared to the traditional methods [136].

7.4.1 Measures for the quantitative assessment of PIs

The paradigm shift of the cost function as compared to the classical methods is the key piece of this technique. In order to derive it, it is necessary to define previously some measures that help to evaluate quantitatively the quality of the constructed PIs.

¹³ Width and coverage probability.

In this section, evaluation indices¹⁴ for both the coverage probability and width of PIs are presented. Finally, a symmetry index¹⁵ that studies the geometric structure of the PIs by assessing the symmetry between the target values and the PIs is introduced.

PI Coverage Probability (PICP): The coverage probability of a PI is a measure indicating the probability that the target values (see figure 7.4) will be within the region established by the lower and upper bounds. On the other hand, according to the definition of prediction interval presented in section 7.3, the targets values have to fall within the PI with a probability equal to the confidence level $[(1 - \alpha) \%)$. Hence, it is expected that, during the training process, the PICP value evolves in an asymptotic manner towards the pre-established nominal level of confidence. Due to this evolutionary behavior, the PICP can be used to assess the quality of the PIs constructed: a PICP with a larger value relates with with more targets lying within the constructed PI and vice versa [136].

The PICP is calculated empirically by counting, for a given dataset, the number of target values that fall within the bounds. Mathematically, is defined as follows [134]:

$$PICP = \frac{1}{n} \sum_{i=1}^n c_i \quad (7.6)$$

where n is the number of target values inside the dataset, and c_i is a boolean variable that indicates if the i -th target value y_i is within the lower bound L_i and upper bound U_i of the i -th PI:

$$c_i = \begin{cases} 1, & \text{if } y_i \in [L_i, U_i]; \\ 0, & \text{otherwise.} \end{cases} \quad (7.7)$$

In practical application, the confidence level is a pre-established value, and through the training process, the model must be able to achieve values of $PICP \geq ((1 - \alpha) \%)$. Otherwise, the constructed PIs are not reliable and should be discarded [128]. The ideal case for PICP is to obtain full coverage for the target values ($PICP = 1$).

Prediction Interval Normalized Averaged Width (PINAW): A simple method to achieve the required minimum PICP consists in widening the interval. However, in practice, wider intervals are not useful for decision makers, since they do not have any information regarding target variation [136]. The width of PIs determines their informativeness [188], and a quantitative measure of the width is necessary to correctly assess the quality of the con-

¹⁴ Initially introduced by Khosravi *et al.* in [134].

¹⁵ Developed by Zhang *et al.* in [236].

structed PIs. Taking into account the above, the Prediction Interval Normalized Averaged Width (PINAW) is defined as follows [134]:

$$PINAW = \frac{1}{nR} \sum_{i=1}^n (U_i - L_i) \quad (7.8)$$

where R is the normalization factor and equals the range of values of the underlying targets ($\max(target) - \min(target)$) [110]. By normalizing the range, the comparison between different data-sets is possible.

It can be easily observed that PICP and PINAW have a direct relationship, since both increase their value when the constructed PI grows in width. Under equal conditions, a larger PINAW will usually result in a higher PICP [136].

Prediction Interval Normalized Root-Mean-Square Width (PINRW): Quan *et al.* [187, 188] draw an analogy with the Mean Square Errors (MSE) used for training traditional error-based cost functions, and proposed a quantitative measure of the width of PIs that gives more weight in the normalized sum to PIs that have wider intervals. Experimental results show that this measure achieves better training performances and higher quality PIs than PINAW. The Prediction Interval Normalized Root-Mean-Square Width (PINRW) is defined as follows [187]:

$$PINRW = \frac{1}{R} \sqrt{\frac{1}{n} \sum_{i=1}^n (U_i - L_i)^2}. \quad (7.9)$$

In view of these results, they propose to use PINRW during the training process, and PINAW to validate the quality of the PIs made after training.

Prediction Interval Symmetry (PIS): This measure assesses the symmetry between the geometric structure of the band formed by the bounds of the PI and the target values. This is an important property of the PI, since PIs symmetric around the target value are known to be more useful for decision-makers [236]. The PIS index is defined as the ratio of the difference between the target value and the mean value of upper and lower value of PI to the actual bandwidth. The mathematical expression is as follows [236]:

$$PIS = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \frac{1}{2}(U_i + L_i)|}{U_i - L_i} \quad (7.10)$$

where U_i , L_i , y_i and n are the same as in PICP. Different scenarios can be distinguished in function of the target values with respect to the limits of the interval:

$$\begin{cases} PIS \in [0, 0.5], & y_i \in [L_i, U_i] \\ PIS \in (0.5, \infty), & y_i \notin [L_i, U_i]. \end{cases} \quad (7.11)$$

If the target value lies within the PI bounds, the index is nonnegative and less or equal than 0.5, with the ideal case being that the PI is completely symmetrical around the target y_i ($PIS = 0$).

7.4.2 Cost function: the coverage width-based criterion (CWC)

As explained in the previous section, the coverage probability (PICP) and the width (PINRW) of the intervals have a direct relation of growth. In consequence, the PI construction process has a conflicting objective, since an ideal PI has a coverage probability of not less than the confidence level $((1 - \alpha) \%)$, and a width small enough to convey some degree of information. But reducing the width of the interval often results in a decrease in PICP, due to some observations becoming outliers by dropping out of the intervals [136].

The coverage width-based criterion (CWC) introduced in [136] is a PI-based cost function used to comprehensively evaluate both characteristics in a balanced manner¹⁶, resulting in an index that evaluates the overall quality of PIs. Its definition is as follows:

$$CWC_K = PINAW \left(1 + \gamma(PICP) e^{-\eta(PICP - \mu)} \right) \quad (7.12)$$

where $\gamma(PICP)$ is a variable that, *enables* or *disables* the contribution of the exponential term to the CWC index. $\gamma(PICP)$ is given by the following Boolean step function:

$$\gamma(PICP) = \begin{cases} 0, & PICP \geq \mu \\ 1, & PICP < \mu. \end{cases} \quad (7.13)$$

η and μ in equation 7.12 are two *hyper-parameters* that determine the penalty term in case of unsatisfying the required conditions:

- μ stands for the preassigned PICP that must be satisfied, and in practice is usually set to the confidence level, $((1 - \alpha) \%)$ [188];
- η exponentially magnifies the difference between the PICP and μ , and is set to a value between 50 to 100 to highly penalize invalid PIs when the preassigned PICP is not satisfied [187].

The design of this cost function is based on two principles [128]:

¹⁶ Focusing on only one of these characteristics may lead to misleading results.

1. if $PICP$ is less than the nominal confidence level, ($PICP \leq \mu = (1 - \alpha)\%$), CWC should be large regardless of the width¹⁷ of PIs, since the constructed intervals are not reliable. This is achieved by setting ($\gamma(PICP) = 1$) (thus, enabling the exponential term in the function) and the high values reached by the exponential due to the penalization parameter η . The influence of the PIs width in the optimization (through $PINAW$) is negligible as the exponential term becomes higher than $PINAW$.
2. if $PICP$ is greater than or equal to its corresponding confidence level, ($PICP \geq \mu = (1 - \alpha)\%$), then the pre-required coverage probability has been achieved, and the PIs width should now be the influential factor in the optimization process. This is achieved by disabling the exponential term in the CWC index, ($\gamma(PICP) = 0$), resulting in $CWC = PINAW$.

It is because of the above mentioned that the CWC index achieves a balance between the usefulness or informativeness of the constructed PIs (narrow width) with their correctness (acceptable coverage probability) [190]. A thorough analysis of the CWC evaluation function can be found in [136].

To obtain better training results and PIs performance, during the NN training procedure the step function γ is set to one in all cases. It is only implemented as described during the test cases. This conservative approach avoids an excessive reduction of PI's width during training, which in turn would result in a $PICP$ below the nominal value for the test samples [136].

Equation 7.12 is the CWC index initially proposed by Khosravi *et al.* [136], and because of this the sub-index K is implemented. Quan *et al.* proposed in [187] a new CWC index (here referred to as CWC_Q) that implements the $PINRW$ (see equation 7.9) instead of the $PINAW$ during the NN training:

$$CWC_Q = PINRW \left(1 + \gamma(PICP) e^{-\eta(PICP - \mu)} \right) \quad (7.14)$$

where the same definitions and explanations as equation 7.12 hold.

Furthermore, Zhang *et al.* in [236] improved this cost function by including a symmetry constraint in the optimization procedure, and they did this by adding the PIS index (see equation 7.10). This new cost function, referred to as the coverage width symmetry-based criterion ($CWSC_Z$) is defined as follows:

$$CWSC_Z = \gamma(PIS) e^{-\eta_3(PIS - \mu_2)} + \eta_2 PINRW + \gamma(PICP) e^{-\eta_1(PICP - \mu_1)}. \quad (7.15)$$

¹⁷ Measured by $PINAW$.

The three hyper-parameters η_1 , η_2 , μ_1 and the function $\gamma(PICP)$ are the same as in previous CWCs. The last two terms of CWC_Z are identical to the cost function CWC_Q , except that multiplication is replaced with addition. It is proven in [129] and its references that replacing the multiplication with the addition addresses the problem of zero-width intervals¹⁸ whilst the index remains to be a strictly proper score, meaning that it is a valid PI-based cost function.

On the other hand, hyper-parameters η_3 , μ_2 and the Boolean step function $\gamma(PIS)$ are added herein to control the PIS index. Because of the PIS index definition (see equation 7.10), is preferred to the optimization to have a PIS value lower than μ_2 . Thus, the function is defined as follows [236]:

$$\gamma(PIS) = \begin{cases} 0, & PIS < \mu_2 \\ 1, & PIS \geq \mu_2. \end{cases} \quad (7.16)$$

Following the same reasoning as for the creation of the CWC cost function, when the constructed PI meets the requirements of coverage probability and symmetry, their respective terms will be disabled through the gamma functions, and the only remaining parameter to optimize will be the width of the interval. On the other hand, if these requirements are not met, the exponential functions and the η hyper-parameters will ensure that the penalties are large enough so that the weight of the width term is negligible. In addition, because the coverage probability is the main characteristic that defines the quality of constructed PIs [136], its η hyper-parameter should have a value higher than the penalty for not meeting the requirements of symmetry [236].

At the beginning of the training procedure, the PICP index will have a smaller value than the hyper-parameter μ_1 . Thus, a large penalization is set for the PICP term. As the training process evolves, the PICP gradually increases to its nominal confidence level (lowering the PICP penalization) while, at the same time, the PIS index decreases to its own nominal confidence level. Later on, the three criteria would transfer into a conflicting status, resulting in a comprehensive PI that eventually takes into account all indices. During this training process, $\gamma(PIS) = 1$, just like $\gamma(PICP)$ [236].

7.4.3 LUBE algorithm

The main idea is to use the PI-based cost function developed in the previous section to train a NN for PI construction. The network learns to compute PIs on a training set and is then used to predict PIs on a new set of examples. The proposed architecture for the NN

¹⁸ In this situation, a minimum CWC value of zero can occur by finding a zero PI width.

is a feedforward MLP with two output neurons (one for each bound of the PIs) as depicted in figure 7.5, and it can be modified to fit any number of hidden layers, neurons per hidden layer and input signals, depending on the problem to be solved.

Traditional gradient descent-based NN training algorithms like Back-propagation cannot be implemented for this cost function minimization for two main reasons:

- the cost functions defined in equations 7.12, 7.14 and 7.15 are nonlinear, complex, and non-differentiable;
- gradient descent-based algorithms are notorious for being trapped in local minima, resulting in a suboptimal set of NN parameters [136].

Therefore, stochastic gradient-free-based methods are the best candidates for global optimization of the PI-based cost function [136]. Examples of this methods are Simulated Annealing (SA), Particle swarm Optimization (PSO), Bat Algorithm (BA) and Genetic Algorithms (GA).

Before starting the training algorithm, the available data sets are split into three sub-sets:

- *training set*: used to adjust the connection weights of NNs;
- *validation set*: applied to determine the optimal NN structure and other undetermined parameters;
- *test set*: this set will evaluate the final performance of the algorithm once the training is completed.

After splitting the original set, the three sub-sets are normalized to $[-1, 1]$ using equation 7.3.

Figure 7.6 shows the procedure to train the NN parameters by optimization of the cost function through the Simulated Annealing algorithm.

The first step is to initialize the parameters involved in the algorithm. This includes the starting solution (NN weights) w_0 , the starting cost function value CWC_0 , and the starting temperature T_0 for the simulated annealing. After this, the training iterations begin. In any of these iterations, PIs are constructed for all the samples in the training set, and then the cost function CWC_{new} is calculated.

Depending on which of the above mentioned optimization algorithms is used, the NN parameters are adjusted and the training procedure continues. The SA algorithm combines *hill-climbing* and *random walk* strategies, where the latter is used to escape local minima [190]. If there is an improvement in the cost function calculated, i.e. $CWC_{new} < CWC$,

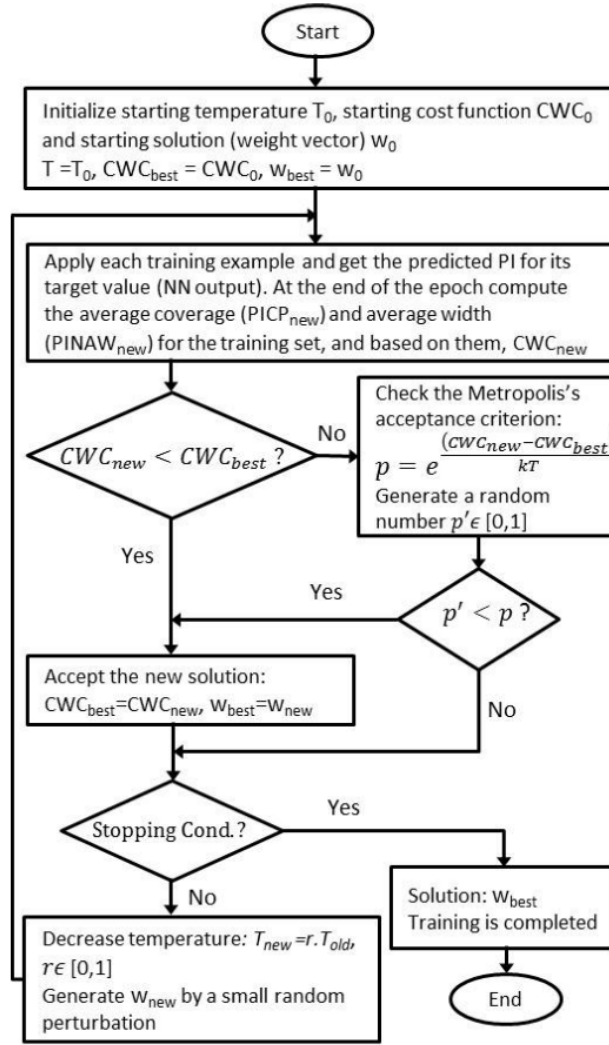


Figure 7.6: LUBE training method using the Simulated Annealing algorithm. Source: [190].

then CWC_{new} and w_{new} are accepted (hill-climbing step). Otherwise, the *Metropolis criterion* is used to test for acceptance (random walk step). The acceptance depends on the temperature and the badness of the new solution, and since the temperature is reduced as the algorithm progresses, the probability that a bad solution is accepted is higher at the beginning of the training process [190]. The SA algorithm is explained in more detailed in the following paragraphs of this section.

If the stopping criterion is not satisfied at the end of the iteration, a new set of weights is obtained by creating a small perturbation in their values, and the training continues. Stopping conditions include reaching a predefined maximum number of iterations or reaching a *plateau* in the cost improvement (i.e., no improvement is obtained in the CWC value for a predefined number of iterations). Once the stopping condition is reached, the optimization

process (training) is complete, and the trained NN can be used for estimating the lower and upper bounds and construction of PIs with the test set.

A detailed discussion about the LUBE method and its training procedure can be found in [133, 136, 190]. More details about the parameter setting and the implementation of the training algorithm with SA are presented in section 7.4.4.

Simulated Annealing (SA)

Simulated Annealing [140] is a gradient-free-based¹⁹ meta-heuristic optimization method to approximate the global optimum of a given function. It has been shown to have an excellent efficiency in finding the optimal solution for complex optimization problems [132].

It is inspired in the annealing process²⁰; in the algorithm the slow cooling of metals is represented by a slow decrease in the probability of accepting worse solutions as the solution space is explored [4]. At each step, the exploring of the solution space is done by a stochastic sampling method, where the current state is conservatively altered in order to find a neighboring state. In the LUBE NN training, this involves the introduction of a random perturbation in a randomly chosen NN weight.

By only using a simple heuristic like hill-climbing, the optimum solution that is found might be a local minimum, since the hill-climbing procedure only accepts new solutions from neighbors that are better than the current solution. Instead, SA combines hill-climbing with random walk; this meta-heuristic allows the algorithm to accept, with some degree of probability, a neighbor that represents a worse solution than the current one, creating a better search of the solution space to find the global optimum. The temperature drops as the algorithm progresses, meaning that the probability of accepting a solution that is worse than the best stored solution decreases. This ultimately leads the system to move to states of lower energy.

The SA algorithm is initialized at a temperature T_0 with a configuration²¹ (x_{old}) whose energy²² is evaluated to be E_{old} . A new configuration (x_{new} with energy level E_{new}) is constructed by applying a random change. Decision about acceptance or rejection of the new configuration is made based on the difference in the energy level ($\Delta E = E_{new} - E_{old} \leq 0$):

- If the energy of the system is lowered ($\Delta E \leq 0$), E_{new} is accepted unconditionally.

¹⁹ Meaning that its implementation does not require the calculation of derivatives.

²⁰ Annealing is a process in metallurgy involving heating and controlled cooling of a material to change its crystallographic properties.

²¹ In the LUBE NN training, this configuration is equivalent to the NN weights w .

²² In the LUBE NN training, the energy of the configuration is represented by the CWC value.

- If the energy of the system is increased ($\Delta E > 0$), the random probability $P = e^{-(E_{\text{new}} - E_{\text{old}})/(\kappa T)}$ is calculated, where κ is the Boltzmann factor. E_{new} is accepted if $P \geq r$, where r is a random number chosen from a uniform distribution between 0 and 1.

The temperature remains unchanged for a predefined amount of iterations while sampling the search space, and then it is decreased based on a *cooling schedule*. This procedure continues until one of the stopping criterion is met. Examples of cooling schedules are geometric and exponential. Generally, the higher the temperature, the more likely the acceptance of an uphill transition. This means that in early stages of optimization, SA behaves like a random walk. Mathematically, it should be chosen so that $\forall (x_{\text{old}}, x_{\text{new}}), e^{-(E_{\text{new}} - E_{\text{old}})/(\kappa T)} \simeq 1$. As T decreases, SA becomes a greedy optimization search looking for global optimum. When $T = 0$, SA becomes totally greedy and only accepts good changes.

Many authors in the literature fail to explain the method in which the temperature of the algorithm is updated. They talk about the implementation of geometric or exponential cooling schedules, but few mention that the temperature update process is not performed with each iteration of the training algorithm. In contrast, several iterations are performed at the same temperature in order to obtain better samples of the search space. If a geometric update of temperature is assumed, the number of iterations between updates (*subiter*) can be determined by the equation 7.17, where *maxiter* is the maximum number of iterations in the algorithm, T_0 is the initial temperature, T_{min} is the final temperature that the algorithm is set to reach at *maxiter*, and β is the geometric cooling factor.

$$\text{subiter} = \left\lceil \frac{\text{maxiter}}{\log_{\beta} \left(\frac{T_{\text{min}}}{T_0} \right)} \right\rceil \quad (7.17)$$

The initial temperature should be chosen to be high enough so that the first part of the algorithm takes a random walk strategy, while the final temperature should be low enough so that at the last iterations the algorithm becomes completely greedy by switching to a hill-climbing strategy. Further information about SA and its fundamental theories can be found in [140].

7.4.4 Issues in implementing the NN LUBE method

This section exposes a series of issues (and the corresponding proposed solutions) encountered when implementing the LUBE algorithm with ANNs for the resolution of practical problems.

Data pre-processing and feature extraction

A potentially critical issue that has not yet been mentioned is what are the input signals of the neural network, and how they are chosen. In order to make the forecasting problem more manageable, the input data is pre-processed before entering the NN [107]. Usual pre-processing stages include the following tasks [34]:

- *clean* the data, by removing outliers, missing values or any irregularities, since NNs are sensitive to such defective data [107, 103];
- *feature selection* to reduce the dimension of the input vector, so as to avoid the *curse of dimensionality*²³. Reducing the dimensionality of the input space also allows for faster training of the prediction models and faster online use of these models [190].
- divide the input data set into three groups of *training* set, *validation* set, and *test* set and normalize each set so that they fall in the interval $[-1,1]$. Normalization ensures a uniform distribution of values for each net input and output [102]. Note that if a normalization procedure takes place in the pre-processing, the output signals of the ANN will also be normalized, and a de-normalization has to be done as a post-processing procedure to transform the solution space to its original state.

The data used to train NNs applied to forecasting in the electricity market can be classified into two groups:

- the time series created with the history of measurements made;
- the exogenous variables that modify, with a certain correlation, the time series.

The exogenous variables may include meteorological variables such as temperature and humidity, and social variables such as holidays; the functions relating the exogenous variables with the time series are clearly not linear, and this is one of the main motivations to use NNs in this context, since NNs can easily deal with nonlinear relationships [107].

The objective of the feature selection process is to perform an analysis on the available data to find which components are the ones that affect the most (i.e., have the highest correlation) the forecast of the next value. In this way, only the most important components are used as the input of the NN, reducing the dimension of the input space and the complexity of the problem to be optimized.

²³ The curse of dimensionality in Machine Learning refers to the exponential growth in the complexity of the problem that results from an increase in the number of dimensions. For example, by increasing the number of inputs of the NN by one, the number of weights to be optimized increases by the number of neurons in the first hidden layer, if the NN is fully connected.

For example, several authors in the literature [203, 72, 73, 102] have performed these analyzes and have determined, for the problem of load forecasting, that the components of the time series and the exogenous variables that have the highest correlation with the determination of the next forecast are the eight variables listed in the table 7.1 and depicted in figure 7.7.

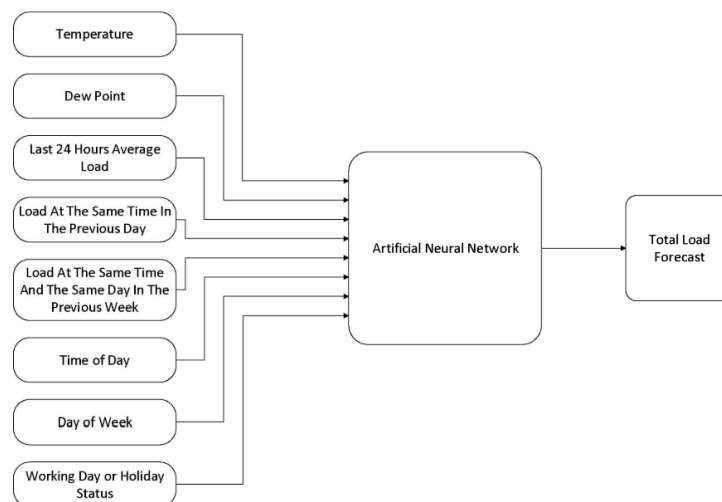


Figure 7.7: LUBE NN input features for total aggregated load forecasting. Source: [203].

Table 7.1: Input features for the LUBE NN.

Symbol	Description
x_1	temperature
x_2	dew Point
x_3	average load in the previous 24hs period
x_4	load at the same hour in the previous day
x_5	load at the same hour in the same day of week in the previous week
x_6	time of day
x_7	day of week (1-7)
x_8	1 for working days, 0 for holidays and weekends

Overfitting and overparametization

The term *overfitting* in machine learning usually refers to models that fit the data-set so well that they end up including in their parameters some of the random errors present in

the data. As a consequence, the model will fit well to the data of the training set, but will have a low performance forecasting other data-sets. This can happen in MLPs for two reasons: because the model was over-trained, or because it was too complex [107].

One way to avoid over-training is by implementing a *cross-validation* scheme. As explained before, the input samples are usually split into three different sets: training set, validation set and test set. When using cross-validation, the NN parameters are estimated on the training set, and the performance of the model is tested, every few iterations, on the validation set. When this performance starts to deteriorate (which means the NN is overfitting the training data), the iterations are stopped, and the last set of parameters to be computed is used to produce the forecasts of the test set [107].

On the other hand, overfitting can be a consequence of *overparametrization*, i.e., an excessive complexity in the model. This usually happens when a large number of hidden neurons and input features are implemented, leading to a number of parameters to be optimized that is comparable in size to the number of available training samples. As a solution, methods can be implemented to experimentally estimate the optimum number of neurons in the hidden layer.

Selecting the optimal number of hidden neurons

The data-sets available can be used to find the optimal number of neurons in the hidden layer. To perform this step is important, since with a low number of hidden neurons the NN will not be flexible enough to model the data well, while with a large number of neurons the NN will overfit the data [107]. To obtain the best possible ANN structure, Rana *et al.* [190] implement an ensemble of NNs instead of a single NN. This method helps to reduce the sensitivity the LUBE NN presents to the network architecture, the initialization of the weights at the beginning, and the random perturbations of the weights during the training process. Thus, the constructed PIs are more reliable and stable.

This technique implements feedforward MLPs with one hidden layer²⁴ to define 30 different NN architectures ($A_i, i = 1, \dots, 30$), where the i -th architecture has i neurons in the hidden layer. Then, the method builds 30 ensembles ($E_i, i = 1, \dots, 30$) of m LUBE NNs each (in this case, $m = 100$). The whole process for one ensemble is depicted in figure 7.8. Every NN that is a member of the ensemble E_i is built using the architecture A_i . All NNs of all ensembles behave independently during training. That is, the processes of weight initialization and update of one NN does not have effects on other NNs. Each LUBE NN is trained in the training set. Once the training process is completed, the ensemble behaves

²⁴ As described in section 7.2.2, Kolmogorov's theorem assures that one-layer-MLPs (with a sufficient number of neurons) are universal approximations for any nonlinear mapping.

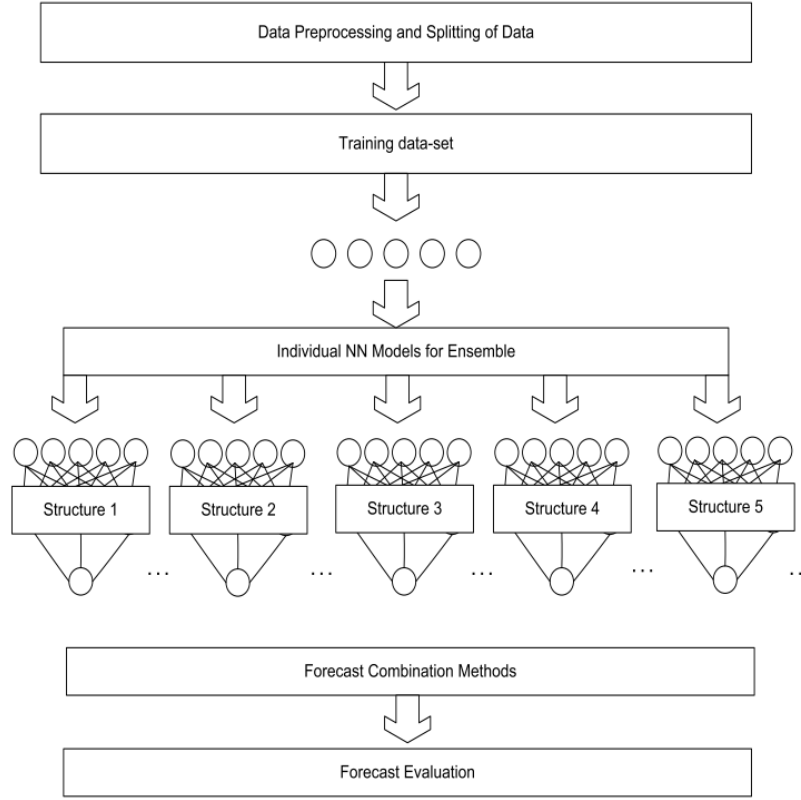


Figure 7.8: LUBE NN Ensemble. Source: [102].

collectively. That is, to construct a PI for the new input data, the ensemble E_i uses an aggregation algorithm that combines the forecasts of its members by taking the median of their lower and upper bounds:

$$PI = [\text{median}(L_1, \dots, L_m), \text{median}(U_1, \dots, U_m)]. \quad (7.18)$$

To choose the best architecture, the 30 ensembles are evaluated using the validation set, and the one that achieves the smallest CWC is chosen to predict the test data. For a thorough explanation of the method refer to [190]. Hassan *et al.* give in [102] a review on different aggregation algorithms and their performances when applied to forecasting in the electricity market.

Activation functions

Choosing the activation functions for the different layers is not a complicated task. This functions must be differentiable and non-decreasing. Hippert *et al.* showed in [107] that one of the most popular configurations due to its good results in the field of forecasting is to use either the logistic (sigmoid in figure 7.2) or hyperbolic tangent functions for the

neurons in the hidden layers, and linear functions in the neurons of the output layer. With the *logistic+linear* configuration, the NN architecture represents the mathematical model of equation 7.5.

Weight Initialization

If the initial NN, weights are set in a way that the value of the cost function is in the vicinity of the global minimum so that the time needed for training the NN is greatly reduced. Thus, the initialization of the NN weights, that is, the method implemented to set random initial weights for the NNs, becomes crucial for reducing the number of training cycles (or *epochs*) needed [182]. A variety of initialization algorithms were investigated and compared in [182]. From the results of these experiments, the authors concluded that the Nguyen-Widrow (NW) layer initialization function obtains the best and stablest results, and already has been tested in forecasting environments with good performances [188]. This algorithm produces values for the individual weights so that the active region of each one of the neurons is evenly distributed over the input space [178]. Thus, the NW method is chosen for NN weight initialization in this algorithm. A more detailed description of the operation of this algorithm can be found in [182, 178].

7.5 Case study: electric load forecasting

This section presents one case studies used to exemplify the implementation of the forecasting method developed throughout this chapter. A well known data-set of a Transmission System Operator (TSO) is implemented to verify the correct functioning of the algorithm by predicting the demand for electric consumption.

7.5.1 Experiment methodology

At this point in the chapter, it is clear that the performance of the constructed PIs will depend mainly on the architecture chosen for the ANNs and the training process carried out. Therefore, it is necessary to find the best architecture and simulation parameters for the problem. ANNs with a single hidden layer (see section 7.2.2) are considered for the following experiments. The optimization of the architecture is carried out by the process described in section 7.4.4. The NN training is done through the minimization of the cost function CWC_Q defined in equation 7.14 (with $\gamma = 1$ during the training) using a simulated annealing algorithm and following the steps described in section 7.4.3. After completing the training stage and obtaining the ensemble E_i with the best performance, said ensemble

is evaluated by means of a test data-set, calculating the measures that evaluate the quality of constructed PIs (those being: $PICP$ as defined in equation 7.6, $PINAW$ as defined in equation 7.8 and CWC_K as defined in equation 7.12, with γ being now the step function).

7.5.2 Parameters

Table 7.2: Parameters used with the case studies.

Parameter	Numerical value
α	0,1
η	50
μ	0,90
CWC_Q^{\min}	1×10^{-3}
κ	1
T_0	5
T_{min}	1×10^{-10}
β	0,9
<i>maxiter</i>	5000
<i>subiter</i>	22

Table 7.2 summarizes the parameters and its numerical values used in the simulations. The confidence level associated with all constructed PIs is $[(1 - \alpha) \%) = 0,9$. The first hyper-parameter of the cost function, μ , is set to 0,9 to match the prescribed level of confidence of PIs. The second hyper-parameter of the cost function, η , is selected to be 50 in order to highly penalize PIs with a coverage probability lower than 90% [136]. The SA training process has a geometric cooling schedule where $T_{k+1} = \beta T_k$, with a initial temperature of $T_0 = 5$ and a cooling factor β of 0,9.

According to the parameters defined in table 7.2 and equation 7.17, the number of iterations *subiter* between temperature updates is fixed to 22. The training is terminated when at least one of the following conditions is met: the maximum number of epochs (*maxiter*) is reached, the cost function reaches a value lower than CWC_Q^{\min} , or no further improvement is observed in the cost function value after a certain number of epochs.

The NNs weights and biases are initialized using the Nguyen-Widrow algorithm described in section 7.4.4 and the NN parameter perturbation function called in each iteration

of the LUBE method generates random numbers whose elements are normally distributed with mean zero and unit variance as proposed in [136].

In both case studies, the data set is partitioned in a percentage proportion 70-15-15, i.e., a randomly chosen 75% of the samples of the data-set is used as the training set and the remaining samples are divided into two groups of equal size, forming the validation and test sets.

7.5.3 Data

This case study involves a data-set taken from the Transmission System Operator (TSO) *ISO New England* in the United States of America. The importance of this data-set is twofold, since it provides all the needed input attributes to the LUBE NN, and has already been tested in previous research [73, 72].

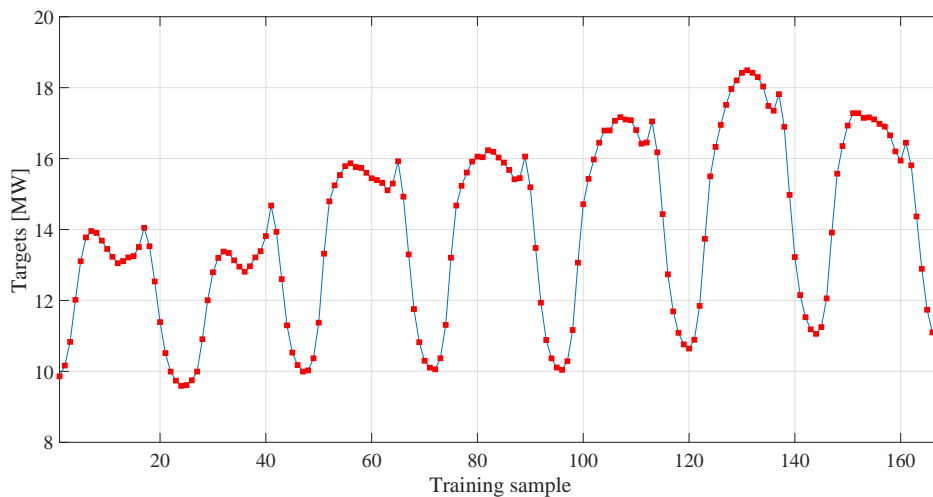


Figure 7.9: Fragment of the *ISO New England* data-set of electric load. The fragment corresponds to one week of samples at a sample rate of one sample per hour.

This set includes hourly measurements for a period of six years, from January 1, 2004 through December 31, 2009 inclusive. A one day extract can be seen in table 7.3, and the general waveform of the load that this particular TSO has during an entire week is shown in figure 7.9.

7.5.4 Training process

Figure 7.10 describes through a log-log graph the convergence behavior of the cost function and the profile of the cooling temperature. This is done by showing how the perfor-

Table 7.3: Extract of the data-set for the input features and target for the LUBE NN. This particular fragment corresponds to the 24 hours of a Friday (day 6). The attributes are defined in table 7.1. Source: [73, 72].

Input Attributes								Target [kW]
$x_1 [^{\circ}C]$	$x_2 [^{\circ}C]$	$x_6 [h]$	$x_7 [-]$	$x_8 [-]$	$x_5 [kW]$	$x_4 [kW]$	$x_3 [kW]$	
8	-3	1	6	1	11728	14186	17681.2	14592
6	-5	2	6	1	11275	13757	17699.6	14199
6	-6	3	6	1	11083	13593	17719.2	14064
5	-8	4	6	1	11071	13612	17739.0	14088
3	-8	5	6	1	11304	13963	17760.9	14488
2	-9	6	6	1	12053	15158	17781.9	15661
1	-10	7	6	1	13372	17498	17807.1	18103
0	-11	8	6	1	14656	18924	17835.9	19616
0	-13	9	6	1	15658	19042	17870.5	19871
0	-15	10	6	1	16472	19072	17907.2	19953
1	-16	11	6	1	16889	19063	17946.8	20015
2	-17	12	6	1	17015	18934	17988.0	19922
4	-16	13	6	1	16869	18694	18029	19678
5	-15	14	6	1	16595	18543	18067.5	19467
6	-15	15	6	1	16387	18381	18103.2	19239
7	-14	16	6	1	16414	18447	18138.7	19297
6	-14	17	6	1	17327	19587	18165.8	20238
4	-14	18	6	1	18164	20991	18185.8	21471
3	-14	19	6	1	17766	20978	18200.6	21334
2	-14	20	6	1	17144	20458	18216.8	20846
1	-14	21	6	1	16461	19800	18234.8	20232
1	-15	22	6	1	15562	18602	18264.5	19316
0	-15	23	6	1	14357	17094	18302.5	18006
-1	-15	24	6	1	13086	15565	18351.3	16735

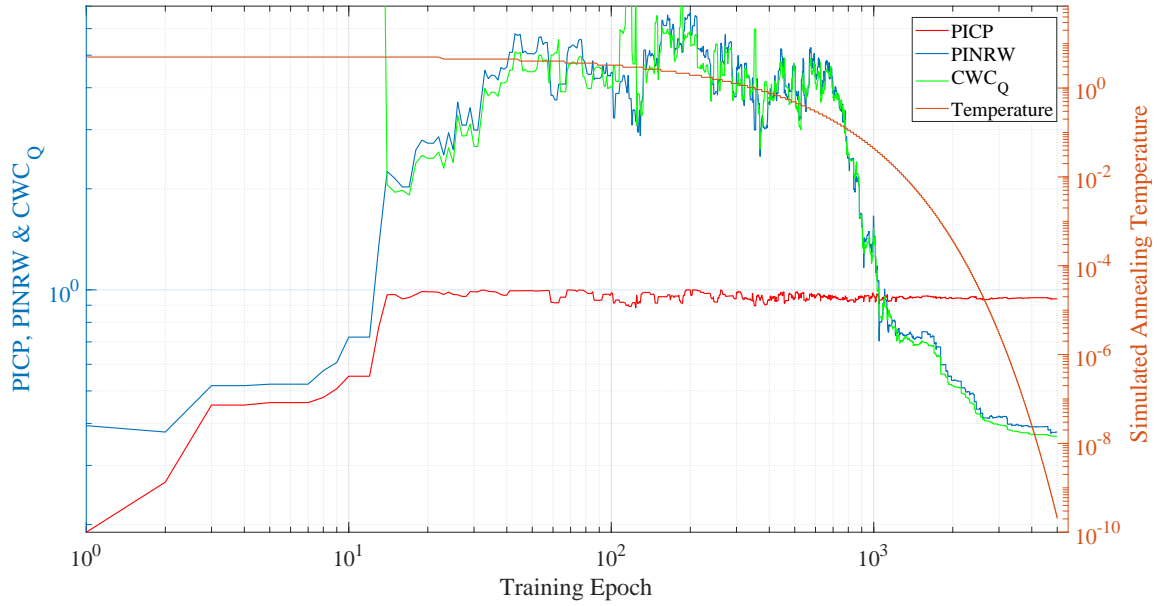


Figure 7.10: Evolution of the performance indexes of the best solution (i.e., the weights and biases of the LUBE NNs) found on each epoch of the training process. ($T_0 = 5, \mu = 0.9$).

mance indexes of the constructed PIs change as the training proceeds and the temperature of the SA algorithm decreases.

As explained in section 7.4.4, the initial condition of the training process is determined by the NW initialization method. Although the performance of this initialization process is better than a random one, we can see that the cost of the initial solution is still outside the scale of the graph, and reaches values close to infinity. The cause of this high value is attributed to the penalty imposed by the exponential term in the cost function, since the PICP index has a value much lower than the hyper-parameter μ .

The presence of this penalty forces the algorithm to first worry about obtaining a coverage probability not lower than the value imposed by the μ hyper-parameter. This is why it is observed that in the first 15 iterations the PICP approaches its nominal value, and little importance is given to the value of the PINRW index, since its weight in the cost function is negligible compared to the penalty. The penalty decreases exponentially as PICP approaches μ . When the value of the penalty is comparable to the PINRW term, the value of the cost function enters the scale of the graph. Once the PICP outgrows μ in iteration 12 approximately, the exponential term becomes a nonpositive value close to zero, so the cost function has a slightly lower value than PINRW. From this point in the training, the hard constraint concerning construction of the prediction intervals is fulfilled. Therefore, as long

as this restriction continues to be fulfilled, the following iterations will aim to find the set of synaptic weights that minimize the width of the intervals (i.e., minimizes PINRW).

As calculated in section 7.5.2, this training process updates the temperature every 22 iterations. From the point where the width of the intervals becomes the main target to be minimized, it is observed how the value of PINRW evolves while PICP stays above its nominal value. Due to the high temperatures, the probability of accepting worse solutions in the Metropolis acceptance criteria is high, and the early stages of optimization behave like a random walk, allowing a better search within the solution space. This results in the random jumps observed in the PINRW and CWC values. Some of these jumps accepting worse solutions produce a decrease in the PICP below its nominal value, causing an abrupt increase in the value of CWC with respect to PINRW. In this condition, the algorithm refocuses its attention on improving the PICP value and reducing the penalty associated with it.

As the temperature drops, the probability of accepting worse solutions decreases, and the algorithm shifts from a random walk to a hill-climbing strategy, accepting only better solutions. The beginning of this change is easily noticeable around the iteration 900, when the temperature reaches values close to 0.1 and the *CWC* starts to decrease gradually, but non-monotonically. Once the iteration 1500 is reached, the temperature is sufficiently low to transform the algorithm into totally greedy, accepting only better solutions.

7.5.5 Finding the best ANN architecture

The process described in section 7.4.4 is implemented to search for the optimum ANN architecture. Initially, the equation 7.14 is implemented as the cost function for the training via simulated annealing, while equation 7.12 is used to compute the cost with the test data-set. After completing the training process of each ensemble E_i ($i = 1, \dots, 30$), a selection process is carried out. The performance of each ensemble is analyzed using the test data-set, and out of the 100 NNs that make up each ensemble, only those 20 NNs that have the lowest cost value survive.

Figures 7.11a to 7.11c show the performance results of the ensembles after the selection process. The hyper-parameter μ in equations 7.12 and 7.14 is set to the confidence level of 0.90, and all the trained NNs have a coverage probability greater than this value.

The box and whiskers plot (or *Box plot*) is used here to show the distribution of the results by presenting a five-number summary of the results on a plot. The edges of the box are the 25th and 75th percentiles²⁵, the mark in the middle is the median (the 50th

²⁵ The 25th percentile is the value at which 25% of the data values are below this value. Thus, the middle 50% of the data values fall between the 25th percentile and the 75th percentile. This distance between percentiles

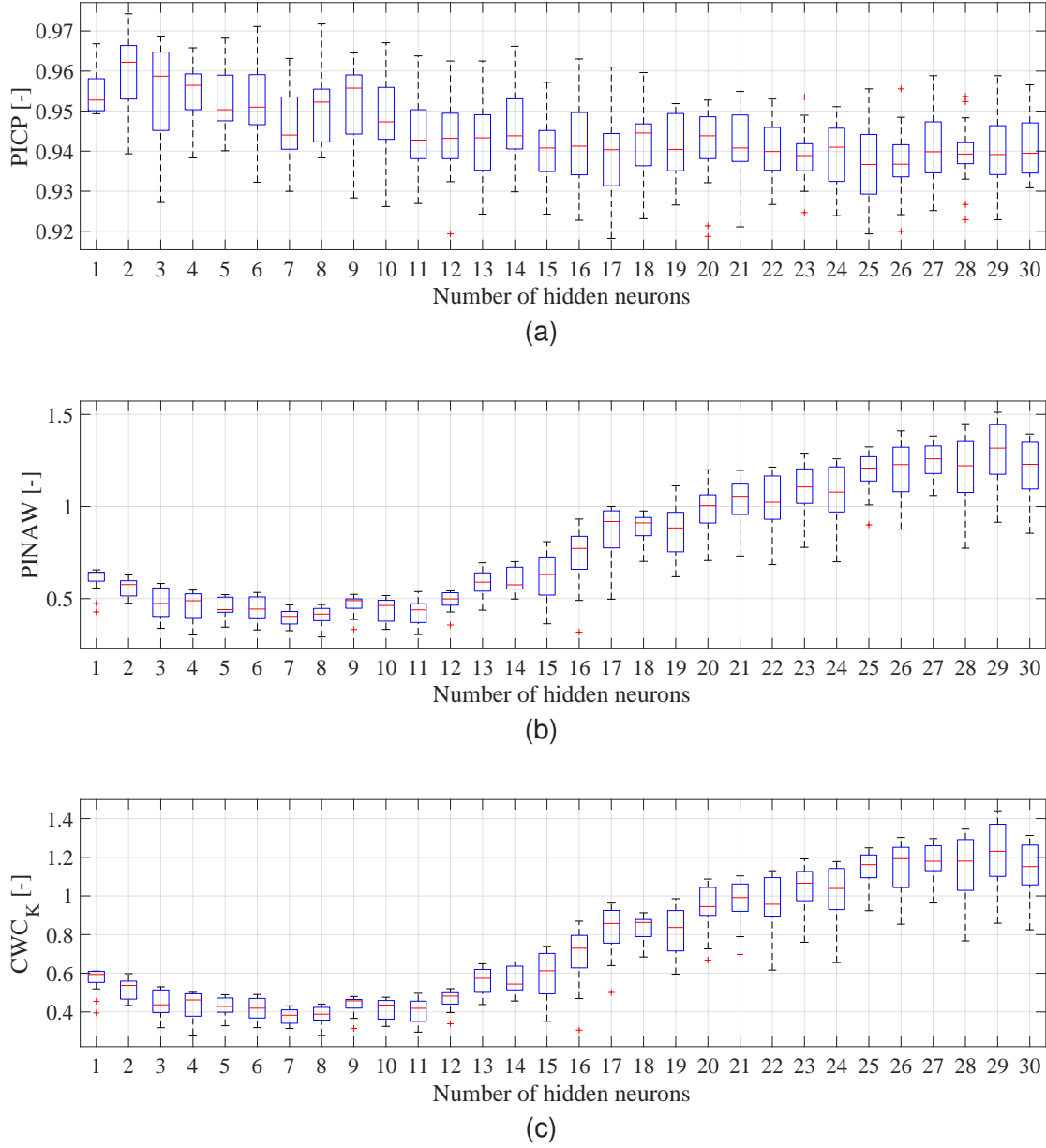


Figure 7.11: Box and whiskers plots showing the statistical quality assessment of the constructed PIs. These PIs were constructed with the test data-set using 30 different ensembles E_i , each containing 20 NNs of architecture A_i . (a) $PICP$; (b) $PINAW$; (c) CWC_K .

percentile), and the ends of the whiskers are the minimum and maximum of the five values, that are not considered outliers; outliers are shown with crosses. The figure shows that all of the trained NNs constructed PIs with a coverage probability higher than the specified confidence level. This means that the penalization imposed by the exponential term in the cost function is negligible, and the numerical value of the function approaches the value of the *PINAW* index.

In table 7.4 a numerical summary of the graphs presented in figures 7.11a to 7.11c is detailed. The table shows that the coverage probability *PICP* is higher than the nominal confidence level of 90% for all the NNs in the ensembles, meaning that the constructed PIs can be accepted as valid. Furthermore, the *PICP* values of the constructed PIs considerably outperformed the prescribed confidence level, with values ranging from 91,8% to 97,4%.

In terms of the width of the constructed intervals, we can observe that the lowest *PINAW* value obtained with the test data-set is 29,2% for E_8 , while the maximum values obtained exceed 150%, probably due to the large over-parameterization of the ensembles that possess a high number of hidden neurons. A comparison between these two cases is analyzed in figure 7.13.

A quick examination of the columns corresponding to the total value of the CWC_K cost function reveals that the cost values are the same as the *PINAW* values for all ensembles. This is as expected and follows from the definition of the CWC_K cost function in equation 7.12. When the NN is used with test cases, the γ function is defined as the step function from equation 7.13, so when the *PICP* is higher than the hyper-parameter μ , the function $\gamma(PICP) = 0$, and $CWC_K = PINAW$.

Therefore, it can be concluded for this example that the LUBE NN with better performance belongs to the ensemble E_8 , and in turn the ensemble E_7 has the lowest mean of the cost value. For the construction of PIs, either of these two options will be a good alternative, since their quality indicators are very similar.

7.5.6 Prediction interval construction

The ensemble E_8 consisting of 20 NNs of architecture A_8 (i.e., eight neurons in the hidden layer) is used to construct PIs from a data-set corresponding to one week of samples. The constructed PIs together with the actual targets are shown in figure 7.12a. A fragment of the samples is showed in detail in figure 7.12b. A comparison between the constructed PIs of this LUBE NN and the ones constructed by one NN from the ensemble E_{29} is shown in

marked by the height of the box is a popular measure of spread and it is referred to as the *inter-quartile range* (IQR).

Table 7.4: Statistics for the quality assessment measurements of PIs constructed from the test data-set for the best 20 NNs in each trained ensemble. Minimum values of each column are in bold.

E_i	$PICP$ [%]			$PINAW$ [%]			CWC_K [%]		
	min	mean	max	min	mean	max	min	mean	max
E_1	94.9	95.5	96.7	42.7	60.8	65.6	42.7	60.8	65.6
E_2	93.9	95.9	97.4	47.5	56.0	62.8	47.5	56.0	62.8
E_3	92.7	95.5	96.9	33.8	47.8	58.3	33.8	47.8	58.3
E_4	93.8	95.4	96.6	30.3	46.2	54.7	30.3	46.2	54.7
E_5	94.0	95.2	96.8	34.5	45.7	52.1	34.5	45.7	52.1
E_6	93.2	95.3	97.1	33.0	44.7	53.3	33.0	44.7	53.3
E_7	93.0	94.6	96.3	32.6	39.8	46.6	32.6	39.8	46.6
E_8	93.8	95.1	97.2	29.2	40.7	46.8	29.2	40.7	46.8
E_9	92.8	95.2	96.5	33.2	46.9	52.3	33.2	46.9	52.3
E_{10}	92.6	94.8	96.7	33.3	43.9	51.7	33.3	43.9	51.7
E_{11}	92.7	94.4	96.4	30.5	42.8	53.8	30.5	42.8	53.8
E_{12}	91.9	94.4	96.3	35.6	49.1	54.2	35.6	49.1	54.2
E_{13}	92.4	94.3	96.3	43.7	58.6	69.4	43.7	58.6	69.4
E_{14}	93.0	94.7	96.6	49.8	60.3	70.0	49.8	60.3	70.0
E_{15}	92.4	94.0	95.7	36.3	61.6	80.9	36.3	61.6	80.9
E_{16}	92.3	94.1	96.3	31.8	72.8	93.2	31.8	72.8	93.2
E_{17}	91.8	93.9	96.1	49.7	86.5	100.0	49.7	86.5	100.0
E_{18}	92.3	94.2	96.0	70.1	87.9	97.6	70.1	87.9	97.6
E_{19}	92.7	94.1	95.2	61.9	85.7	111.2	61.9	85.7	111.2
E_{20}	91.9	94.2	95.3	70.6	98.7	119.9	70.6	98.7	119.9
E_{21}	92.1	94.1	95.5	73.1	103.1	119.6	73.1	103.1	119.6
E_{22}	92.7	94.1	95.3	68.4	101.5	121.4	68.4	101.5	121.4
E_{23}	92.5	93.9	95.4	77.8	109.5	129.0	77.8	109.5	129.0
E_{24}	92.4	94.0	95.1	69.9	107.0	125.9	69.9	107.0	125.9
E_{25}	91.9	93.7	95.6	90.1	119.0	132.4	90.1	119.0	132.4
E_{26}	92.0	93.7	95.6	87.7	119.9	141.1	87.7	119.9	141.1
E_{27}	92.5	94.1	95.9	106.0	124.7	138.3	106.0	124.7	138.3
E_{28}	92.3	93.9	95.4	77.4	121.4	144.9	77.4	121.4	144.9
E_{29}	92.3	94.0	95.9	91.5	129.8	151.2	91.5	129.8	151.2
E_{30}	93.1	94.1	95.7	85.5	120.9	139.3	85.5	120.9	139.3

figure 7.13. All the constructed PIs by both NNs are valid PIs with respect to their PICP values. However, the over-parametrization of the E_{29} NN led to a poorly optimized PINAW value during the training process.

7.5.7 Computational load

One ensemble of random architecture A_i with 100 NNs is trained and the time duration of each iteration of the algorithm is recorded. In order to estimate the *probability density function* (pdf) of the duration of each iteration, the finite data sample recorded is passed through a *Kernel density estimator* for one-dimensional data. To apply this estimator, a Gaussian kernel is assumed, and the bandwidth is chosen automatically by the estimator. Figure 7.14 illustrates the probability density function (pdf), as well as the *cumulative density function* (cdf) of the random variable. The most frequently occurring value (or *mode*) for one training iteration is 23,4 ms, the mean value is calculated in 24,3 ms, and the standard deviation is 1,7 ms. All iterations take less than 35 ms to be completed. This means that, with $maxiter = 5000$, the training of each LUBE NN in the ensemble is completed in less than 150 s in the worst case.

The same analysis is carried out with a LUBE NN that is already trained, with the objective of estimating the mean time of PI construction. The 100 NNs of the ensemble previously trained in this section now estimate 1×10^3 PIs each, and the resulting 1×10^5 time values are passed through the same Kernel density estimator as before. Results are shown in figure 7.15. The mode for one PI estimation is 6,3 ms, the median 6,6 ms, and the standard deviation 1,0 ms. On the other hand, if the ensemble is implemented collectively using as the PI estimation the mean value of all the constructed PIs in the ensemble, the total time for one PI estimation is, in the worst case, 180 ms for an ensemble with 20 NNs and 900 ms for an ensemble with 100 NNs (assuming no parallel computing is implemented).

Due to the simplicity of the calculations involved, the training and implementation times of the LUBE method are three orders of magnitude lower than the times of the traditional PI construction methods described in section 7.3 [136]. Therefore, the LUBE method is easily implementable in real-time applications, specially in the energy industry, where the rate of the decision-making keeps increasing over time.

7.6 Conclusions and further reading

The main thesis of this chapter is that "in the energy control system of an electric grid, variables that present uncertainties (such as energy purchase and sale prices, electric demand and power generation) can be estimated in a fast and simple way, without making any

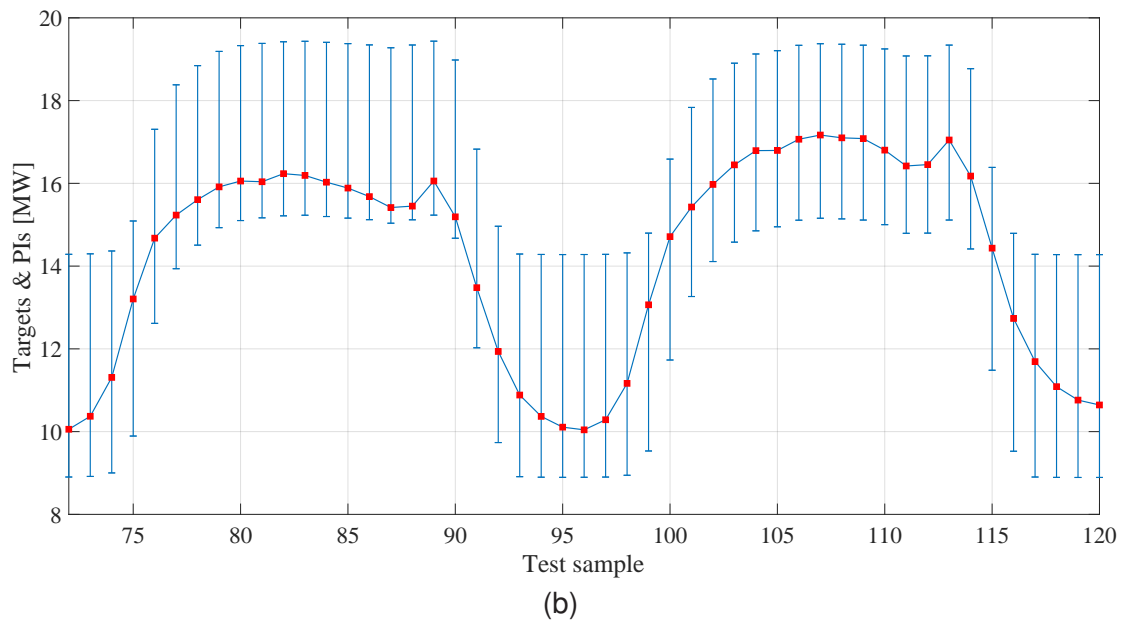
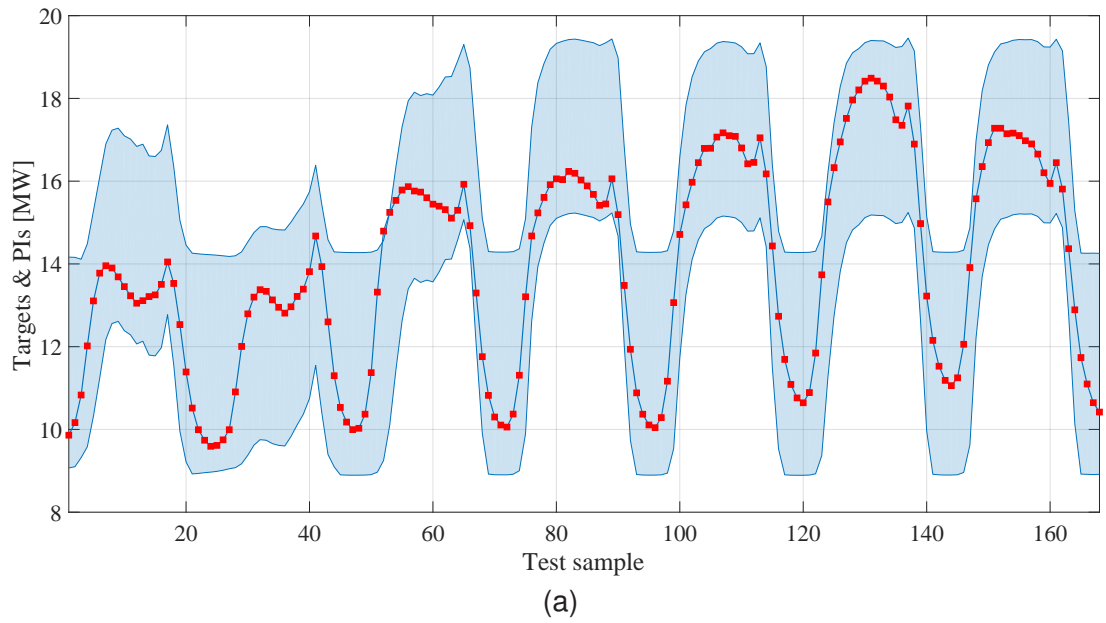


Figure 7.12: Target values of the electric load signal and PIs constructed with the ensemble E_8 consisting of 20 NNs of architecture A_8 (i.e., eight neurons in the hidden layer). (a) one week fragment; (b) two days fragment.

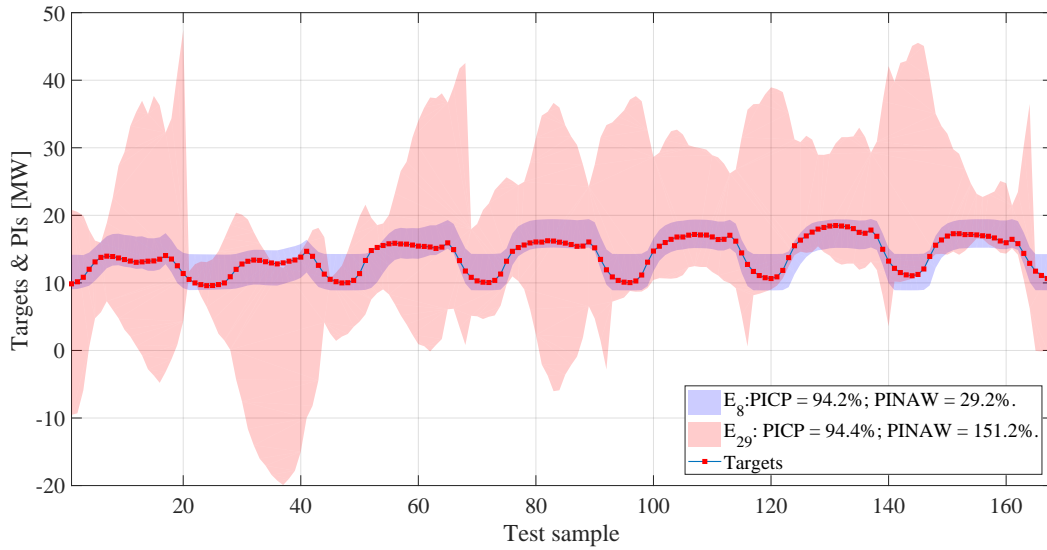


Figure 7.13: Comparison between PIs constructed by a LUBE NN with a successful optimization of the cost function (E_8) and one with a poorly optimized cost function due to over-parametrization of the NN (E_{29}). All constructed PIs have valid PICP values.

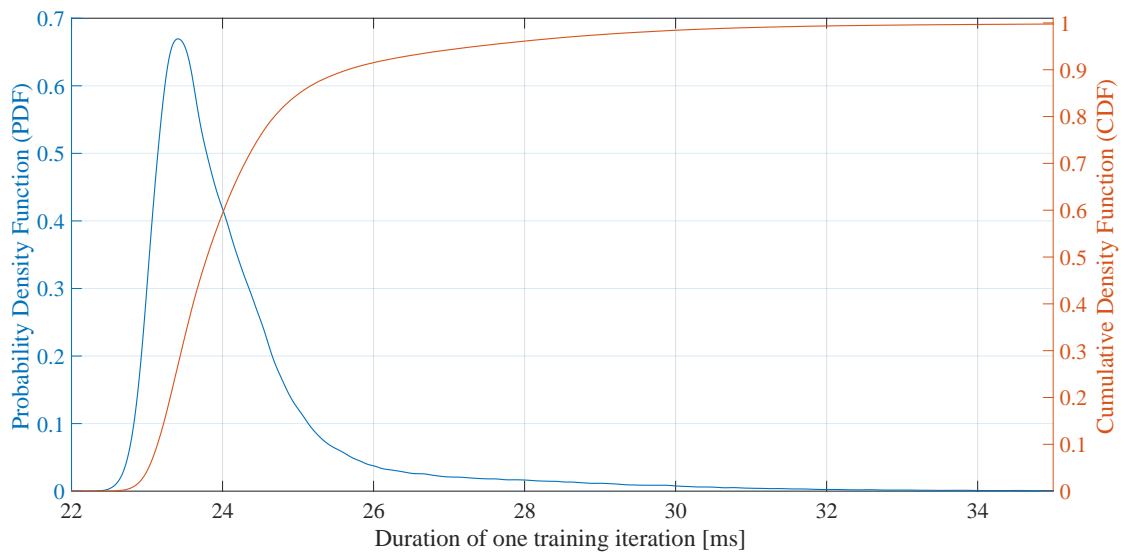


Figure 7.14: PDF and CDF of the time taken for a training iteration to be completed.

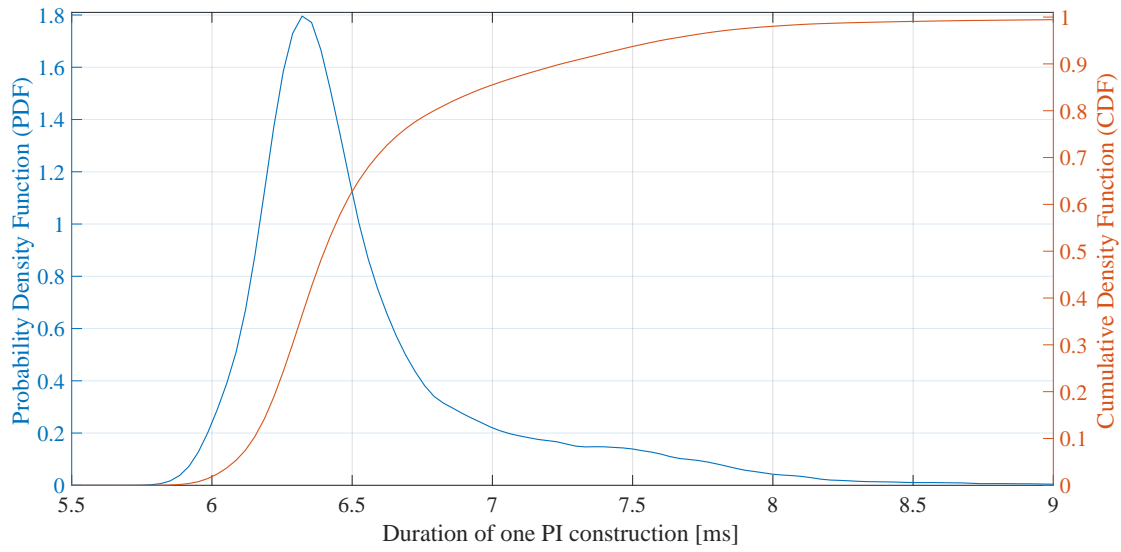


Figure 7.15: PDF and CDF of the time taken to construct one Prediction Interval.

assumptions on the input data, and allowing real-time forecasts to be used to improve decision making". To demonstrate this, the concept of prediction interval (PI) is introduced and the advantages that present its implementation in the decision processes in comparison to a point forecast are described. Then the *lower upper bound estimation* (LUBE) method is introduced, which uses an ensemble of neural networks to construct the prediction intervals. In order to generate high quality PIs at the output of the NN, this technique performs the training process using a specially designed cost function that focuses on enhancing the important features of PIs (width and coverage probability). The implementation of this method together with a meta-heuristic process of neural network training such as simulated annealing (SA) a real-world application showed results that prove the thesis right. The constructed PIs have a coverage probability of no less than the specified confidence level, and the required calculation time is small enough to allow this method to be implemented in most real-time applications.

However, similar to any other optimization problem, the quality of developed models and constructed intervals directly depends on the performance of the applied optimization method. Since this chapter has as an objective only to demonstrate that it is possible to construct these types of intervals under the described conditions, one of the simplest optimization methods is implemented for tuning the neural network parameters, i.e., simulated annealing. Within this implementation, the set of hyper-parameters used was hand-picked for the specific problems. However, hyper-parameters that create LUBE NNs with better performance at constructing PIs can be obtained by implementing hyper-parameter optimization techniques, like brute force strategies (e.g. grid search, random search), or

more intelligent strategies like Bayesian Optimization, which looks most promising [47, 74]. Bayesian Optimization picks a prior belief about how the hyper-parameters will behave and then searches the hyper-parameters space by enforcing and updating that prior belief based on ongoing measurements. In this way, Bayesian Optimization lets prior beliefs influence its predictions.

Furthermore, the introduction of more advanced evolutionary algorithms such as simulated annealing mixed with the mutation operator present in genetic algorithms, particle swarm optimization, genetic algorithms or firefly algorithm can significantly enhance the quality of constructed prediction intervals. To the authors best knowledge, one of the most promising options is a hybrid evolutionary algorithm called PSOGSA, a combination of particle swarm optimization (PSO) and gravitational search algorithm (GSA). This algorithm is implemented by Lian *et al.* in [158] to construct PIs together with an artificial neural network with *random hidden weights*. This particular area of research formulated a theorem which states that for *any* random set of synaptic weights and biases of the hidden layer, there is a combination of synaptic weights in the output layer that guarantees that the constructed PIs will contain the target value with probability equal to one. Therefore, it is possible to successfully train the NN by optimizing only the synaptic weights of the output layer.

Other options to implement the LUBE method have also gained some attention, like hybrid systems mixing ANN with Fuzzy Logic as demonstrated by Kavousi-Fard *et al.* in [125]. This option allows the online modification of some parameters, providing more versatility to the decision-maker to analyze different scenarios.

8 Conclusion and prospect

8.1 Conclusion

This work presents modeling and optimization solutions for the energy management problem for grid-connected microgrids with on-site energy storage systems (ESS) under the presence of uncertainty in electricity prices, predicted load and predicted renewable power generation; this paradigm is deployed to design a high-level control strategy, or energy management system (EMS), for an energy storage system (ESS) in a light rail system.

The correlation between the current and voltage measurements performed at the Neuostheim electrical substation of the RNV5 light rail line (see figure 6.16) and the time schedules for the trains is analyzed; the results show (figure 6.20) that the recovered energy through the braking phases represents no less than 28% (figure 6.18) of the total energy consumed at the four track segments of the substation.

According to [22], the recovered energy can reach levels up to 40% of the energy supplied to the electrical rail guided vehicles. While the vast majority of the literature focuses on reusing this energy within the train electric network (i.e., peak power shaving and line voltage stabilization of the overhead line), a novel approach is to consider the recovered energy from the braking phase as the renewable energy source (RES) of a grid-connected microgrid. With this in mind, the microgrid architecture shown in figure 1.8 is proposed to further increase the energy savings and system efficiency by increasing the receptivity of the overhead catenary line of the light rail system; this is achieved through the installation of a wayside ESS that helps to prevent the loss of the recovered energy (mainly the $E_{\text{gen}_{\text{rheostatic}}}$ term of equation 6.4) during the braking phase of the trains in the vicinity of the Neuostheim electrical substation.

All the power flows inside the proposed microgrid architecture are controlled by the EMS, whose structure is presented in figure 1.7; the controller performs as a high-level power optimizer in a hierarchical control structure by providing power flow commands to the power converters of the different sub-systems inside the microgrid at a time scale in the order of minutes.

The optimization is carried out by the mixed-integer linear program (MILP) defined in equation 3.33, whose main decision variables are all the power flows inside the microgrid

for a predefined number of future time-steps; the optimal solution for this problem is the one that presents the highest economic profitability while meeting at the same time a set of constraints and sub-objectives (e.g., minimization of operation costs, shaping of the main grid power signal, shaping of the ESS power signal, energy balance to meet the demand, amount of load-shedding in islanded mode, etc.). To find the optimal values of these decision variables, the optimization problem takes into account a range of parameters (e.g. physical limitations of the system concerning energy and power rates and capacities) and forecasts within a predefined prediction horizon (e.g. energy generation, load and electricity prices).

Because the cost function has to be minimized in a given time horizon and the system is subject to constraints, the MILP problem formulation is embedded into an integrated receding horizon framework involving problem-specific forecasting techniques for the uncertain parameters. Thus, the system receives feedback of all input parameters (e.g., actual demand, updated state of system) at every discrete time-step and each optimization is done according to the current available information.

Optimal solutions of optimization problems usually present a high sensitivity to data perturbation; solving the optimization problem under the assumption that all parameters are deterministic (thus ignoring possible uncertainties in forecasted parameters) could cause the solution of the problem to be suboptimal or unfeasible in the worst case. To address this issue, the robust optimization methodology is introduced and implemented to create a robust counterpart of the original deterministic MILP formulation, as shown in equation 3.46; this new optimization problem allows to obtain solutions that remain feasible under all the possible realizations of data uncertainty predefined in an uncertainty set included in the uncertain space.

The optimization problem is modeled in the AMPL programming language and solved by the IBM ILOG CPLEX solver; then this model is embedded in the rolling horizon framework created with MATLAB; the different features of the controller are demonstrated by performing several simulations under a variety of conditions. Figures 5.5 and 5.6 compare the robust optimization approach to the deterministic case in terms of feasibility and quality of the solution under different scenarios of uncertainty; the first figure shows that it is a certainty that the solution granted by the robust approach of the EMS remains feasible as long as the realizations of the parameters remain within the predefined bounds of the uncertainty set; the second figure shows that in a system where forecasts are not deterministic and are likely to contain uncertainties, the robust approach to the optimization gives more profit than the deterministic case. The ability the decision-maker has to shape the output of the EMS controller is studied in figures 5.7 and 5.9; by changing the economic costs of the penalization terms inside the objective function, the decision-maker is able to meet

sub-objectives of the system related to extending the lifetime of the ESS and generating a more predictable and stable load response from the main grid perspective at the point of common coupling. In this way, the introduction of a microgrid topology does not disrupt the current scheme of energy distribution and control in the main grid, and it can be considered as an effective framework for the penetration of distributed energy generation into the main grid network.

Parametric plots for optimal ESS sizing in terms of capacity and power rating are also derived in figures 5.12 and 5.13, allowing for a quick determination of the optimal parameters for the ESS once its technology is decided.

Simulations are also conducted to study the computational time needed by the EMS controller; the best solution in terms of computational speed involves the resolution of the optimization problem with a variable prediction horizon with relaxed binary constraints (figure 5.14). Nonetheless, the average computational time of all strategies investigated is at least four orders of magnitude smaller than the length of the control horizon. In consequence, regardless of the strategy implemented real-time application does not present an issue for this particular controller.

As decision making in the EMS inherently involves the consideration of forecasting uncertainties, for completeness sake the problem of creating real time high-confidence-forecasting models is also addressed in this work. An ensemble of artificial neural networks is used to create prediction intervals for the different parameters that need to be forecasted; the lower upper bound estimation method is used to train this ensemble with the simulated annealing algorithm. Figure 7.12 shows for the electrical load of a transmission system operator the results of this implementation for the creation of prediction intervals compared with the actual realizations of the signal; the constructed PIs have a coverage probability of no less than the specified confidence level. Furthermore, as shown in figure 7.15 the required calculation time of 6,3ms is small enough to allow this method to be implemented in most real-time applications.

8.2 Prospect

As suggestion for further research a list of topics that remain to be covered is presented below.

- Perform a deeper and more detailed study of the behavior of the controller under different scenarios and parameter configurations.
- Replicate the microgrid architecture in several electrical substations along the rail track and integrate these architectures into a higher-level controller involving multi-

microgrid optimization models that also allows the exchange of energy between the different microgrids with the ultimate goal of creating a *virtual power plant* to exchange power with the main distribution grid.

- Further cost reduction and energy savings could be achieved by the coupling of the heat and mobility sectors to the microgrid (e.g., charging stations for electric vehicles with vehicle-to-grid capabilities).
- Measurements of the power consumption of a single electric vehicle can be used together with the transportation schedule and speed profiles to better estimate the unused braking energy in a particular section of the network, such as the area covered by an electrical substation. More information about this procedure proposed by Lakovskis and Grigans in [150] is found in section 6.2.2.
- The quality of the constructed prediction intervals directly depends on the performance of the applied optimization method. The introduction of more advanced evolutionary algorithms such as simulated annealing mixed with the mutation operator present in genetic algorithms, particle swarm optimization, genetic algorithms or firefly algorithm can significantly enhance the quality of constructed prediction intervals (more information in section 7.6).
- Hyper-parameter optimization techniques can be implemented to increase the performance of the LUBE NNs forecasting model; several options include brute force strategies (e.g. grid search, random search), or more intelligent approaches like Bayesian Optimization (more information in section 7.6).

A List of Figures

1.1	Global cumulative installed capacity, 2016 (real) and 2040 (predicted). Flexible capacity includes small-scale batteries, utility-scale batteries, demand response and other flexible capacity. Source: [159].	2
1.2	Global new clean energy investments and capacity installations. Total values include estimates for undisclosed deals. Includes corporate and government R&D, and spending for digital energy and energy storage projects (not reported in quarterly statistics). Excludes large hydro. Source: [159].	3
1.3	Growth of renewable energy proportion of power generation between 2006 and 2016 (excludes large hydro). Source: [159].	4
1.4	Recorded peak renewable penetration by country. Source: [159].	4
1.5	Energy Outlook for Europe. A prediction of 50% renewables penetration is made for Europe by 2040. Source: [159].	5
1.6	Architecture of a grid-connected microgrid showing its main components. Source: [48]	7
1.7	Energy Management System (EMS) Controller.	8
1.8	Microgrid definition for the present work.	9
2.1	Uncertainty representation: (a) Deterministic optimization; (b) Stochastic optimization; (c) Robust optimization. Source: [112].	24
2.2	Problem formulation process for robust optimization. Source: [112].	27
2.3	Uncertainty sets for parameter \tilde{a}_j , defined by $[\tilde{a}_j = a_j + \xi_j \hat{a}_j, j = 1, 2]$, where \tilde{a}_j denotes the true value of the parameter, a_j denotes the nominal value of the parameter, ξ_j denotes the uncertainty and \hat{a}_j represents a constant perturbation. (a): box uncertainty set; (b): polyhedral uncertainty set; (c): <i>box+polyhedral</i> uncertainty set. Source: [155].	32
2.4	Correlated polyhedral uncertainty set vs polyhedral uncertainty set. Source: [121].	35
2.5	Concepts associated to rolling horizon approach: scheduling, prediction and control horizon. Source: [205].	39
2.6	Reactive scheduling via a rolling horizon framework. Source: [143].	40
2.7	Algorithm for the rolling horizon approach. Modified from [143].	41
4.1	Algorithm for practical mathematical programming. Modified from [85].	62

4.2	Fragments of AMPL code for the models created in chapter 3. (a) Definition of horizon related parameters (see equation 3.3); (a) Definition of the grid binary control variables (see equation 3.21) and its related constraint (see equation 3.22).	62
4.3	Information flow between the different layers of the EMS controller.	63
4.4	Flowchart for the EMS controller.	64
5.1	Rolling horizon framework.	65
5.2	Some of the EMS input parameters that need to be forecasted by the prediction algorithms. This example shows a control horizon of 30 minutes, a forecasting horizon of 24 hours and a scheduling horizon of 1 week; this frame of the EMS shows the control horizon on the first 30 minutes of the Friday.	66
5.3	Optimal set of decisions found by the EMS controller for the entire forecasting horizon. Only the decisions corresponding to the control horizon will be implemented, while the decisions for the rest of the horizon will be discarded before updating the time-step.	67
5.4	History of decision implemented by the EMS controller at each one of the different control horizons up to the <i>present</i> time-step. The evolution over time of the state of charge of the ESS (affected by these decisions) is also shown in the figure; for the ESS operation the safe operational bounds SOC^{\min} and SOC^{\max} are set to 0,2 and 0,8, respectively.	68
5.5	(a) Probability that the solution given by the EMS remains feasible for different realizations of the forecasted parameter, shown for three different uncertainty sets. When the uncertainty set is not defined properly, the realization of the parameter starts to fall outside of this set and the probability that the solution remains feasible starts to fall. Different Monte-Carlo simulations are shown for the actual realization of the forecasted parameter: (b) good uncertainty set definition, where the realization of the parameter always falls within the predefined interval; (c) the realization of the parameter starts to fall outside the predefined uncertainty set, causing infeasible solutions; (d) the choice of the uncertainty interval is quite poor in this example, since the realization of the parameter falls quite often outside of the interval and the probability the the solution is feasible is close to zero.	69

5.6	Deterioration of the optimal solution as the uncertainty in the forecasts increases. When the forecasts present no uncertainty (ideal case) the deterministic approach presents a better solution than the robust cases; as the uncertainty of the forecasts increases (real case), the robust optimization presents a solution of better quality.	70
5.7	Trade-off for different values of $c_{\text{grid}}^{\text{flat}}$ between controlling the peak-to-peak power exchange with the main distribution grid and the savings from the operation of the microgrid.	71
5.8	Evolution of the power profile for the exchange with the main distribution grid for different values of the penalization cost $c_{\text{grid}}^{\text{flat}}$; as the penalization cost is increased, the peak-to-peak power exchanged with the main grid is decreased. . .	72
5.9	Trade-off between controlling the number of ESS charging and discharging cycles and the savings from the operation of the microgrid; increasing the penalization cost will give the decision-maker more control over the number of cycles over a period of time so that the lifetime of the ESS is extended, but this will decrease the savings obtained by the EMS.	73
5.10	Evolution of the power profile and state of charge of the ESS for different values of the penalization cost $c_{\text{ss}}^{\text{sm}}$; as the penalization increases, several small charging and discharging cycles done to optimize the energy transfer and operational cost are no longer implemented by the EMS. Top: $c_{\text{ss}}^{\text{sm}} = 0,10\text{ct./kWh}$; middle: $c_{\text{ss}}^{\text{sm}} = 0,25\text{ct./kWh}$; bottom: $c_{\text{ss}}^{\text{sm}} = 1,00\text{ct./kWh}$	74
5.11	Optimal ESS Sizing. Source: [21].	74
5.12	Optimal ESS sizing: sensitivity of the operational cost to the ESS capacity and power rating.	75
5.13	Optimal ESS sizing: sensitivity of the operational cost to the <i>minimum</i> number of hours needed to reach the <i>maximum</i> capacity of the ESS.	76
5.14	77
6.1	One-line diagram of the power supply of DC trains with overhead catenary. Source: [220].	80
6.2	Schematic of a tramway line fed by several unidirectional substations. Source: [216].	82
6.3	Voltage and current measurements as a train travels through one of the track segments of the Neuostheim electrical substation. CS: constant speed; B: braking; S: stop at passenger station; A: accelerating.	84
6.4	Braking energy flow in vehicle, U_{CN} - voltage of contact supply network, E_K - kinetic energy of vehicle, TM - traction motor, TC - traction converter, BC - braking converter, BR - braking resistors. Source: [24].	85

6.5	Strategies to maximize the use of the recovered energy. Source: [59].	87
6.6	Energy exchange between vehicles in urban rail. Source: [59].	87
6.7	Energy exchange between vehicle and on-board ESS. Source: [59].	88
6.8	Energy exchange between vehicle and wayside ESS. Source: [95].	89
6.9	Energy exchange between vehicle and reversible substation. Source: [95]. . . .	90
6.10	Energy storage systems for urban rail applications. Source: [59].	92
6.11	Tram line number 5 of the RNV company, connecting the cities of Heidelberg, Weinheim and Mannheim (red line). Source: [6].	93
6.12	Upper bound for the electrical energy harvested during regenerative braking for a train with mass m braking from speed v to rest condition.	95
6.13	Upper bound for the electrical energy harvested during regenerative braking. Curves parametric with respect to the speed v the vehicle has before the braking begins. The red dashed line indicates the average energy recovered by a braking phase according to the measurements at Neuostheim electrical substation (see section 6.4).	95
6.14	(a) Neuostheim electrical substation; (b) Neuostheim passenger station. Source: Google Maps.	97
6.15	Railroad segment under study of the RNV5 line. The measurements are taken at the Neuostheim electrical substation (<i>GUW Neuostheim</i>), which is directly connected to the track segments 211, 212, 213 and 214. Source: RNV (Modified).	98
6.16	Current and voltage sensors installed at the Neuostheim electrical substation, which is directly connected to the track segments 211, 212, 213 and 214.	98
6.17	The energy recovered by a braking vehicle in one segment and exchanged with other vehicle in one of the remaining three line segments ($E_{\text{gen}_{\text{sameSS}}}$) is recorded as a negative current at the sensor from the segment where the braking vehicle is ($i_1 < 0$), and as a positive current in the line segment with the vehicle that consumes the energy ($i_4 > 0$). The current measurements for braking and accelerating phases in one track segment can be seen in figure 6.3.	101

6.18	Detail of the energy consumption and generation during one week, corresponding to the four track segments directly connected to the Neuostheim substation. Top: energy recovered by braking trains and transferred between line segments (E_{gen}), energy delivered to trains in the four line segments by the Neuostheim and neighboring substations (E_{ss}), and total amount of energy consumed by the trains whilst traveling through the Neuostheim segments (E_{cons}). E_{gen} represents 28,3% of the total energy consumed at the four segments; the day average for E_{gen} is 0,68MWh; the average for each train is 0,46kWh. Bottom: time-table for the seven passenger stations inside the Neuostheim line segments. Each point corresponds to a train arriving at the station. The service interruption at nights and the extended service during weekends is clearly appreciable.	102
6.19	Detail of the energy consumption and generation during one day, corresponding to the four track segments directly connected to the Neuostheim substation. Top: recovered energy at the $i - th$ line segment and transferred to one of the three remaining segments ($E_{gen_i}, i = 211, 212, 213, 214$), total amount of recorded recovered energy (E_{gen}), and total instantaneous recovered power (P_{gen}). Middle: energy delivered by the different substations to the trains in the segments 211 – 214 ($E_{ss_i}, i = O41, Neuostheim, Pforzheimer$), and total amount of energy delivered by all the substations (E_{ss}). Bottom: time-table for the seven passenger stations inside the Neuostheim line segments. Each line corresponds to the route followed by a train through the stations in the segments. The service interruption at nights is clearly appreciable.	103
6.20	Mean and standard deviation values for the amounts of energy recovered in fifteen minute intervals throughout the day; 103 days of recorded data (from 10-12-2016 to 22-03-2016) are used in these plots. There is a clear distinction in the patterns of energy generation for working days, Saturday and Sunday. This distinction is directly correlated with the social patterns with which the means of transport are used, which in turn impacts on service time-tables and the number of people aboard the vehicle (influencing the total mass and kinetic energy). . . .	104
6.21	Over-current protection of the Neuostheim electrical substation; after the protection is enabled, the substation needs about 15s to normalize its output voltage.	105
6.22	Equivalent circuit of the electrical substation.	105
6.23	Voltage vs. current at the output of the Neuostheim electrical substation for a random day.	106
7.1	Schematic of an artificial neuron. Source: [103].	110
7.2	Commonly used activation functions. Source: [219].	112

7.3	Fully connected Multilayer Perceptron (MLP) network. Source: [64].	113
7.4	Prediction Interval (PI) concept. Source: [131].	116
7.5	NN architecture of the LUBE method. Source: [188].	118
7.6	LUBE training method using the Simulated Annealing algorithm. Source: [190]. .	125
7.7	LUBE NN input features for total aggregated load forecasting. Source: [203]. .	129
7.8	LUBE NN Ensemble. Source: [102].	131
7.9	Fragment of the <i>ISO New England</i> data-set of electric load. The fragment corresponds to one week of samples at a sample rate of one sample per hour. . .	134
7.10	Evolution of the performance indexes of the best solution (i.e., the weights and biases of the LUBE NNs) found on each epoch of the training process. ($T_0 = 5, \mu = 0.9$).	136
7.11	Box and whiskers plots showing the statistical quality assessment of the constructed PIs. These PIs were constructed with the test data-set using 30 different ensembles E_i , each containing 20 NNs of architecture A_i . (a) <i>PICP</i> ; (b) <i>PINAW</i> ; (c) <i>CWC_K</i>	138
7.12	Target values of the electric load signal and PIs constructed with the ensemble E_8 consisting of 20 NNs of architecture A_8 (i.e., eight neurons in the hidden layer).(a) one week fragment; (b) two days fragment.	142
7.13	Comparison between PIs constructed by a LUBE NN with a successful optimization of the cost function (E_8) and one with a poorly optimized cost function due to over-parametrization of the NN (E_{29}). All constructed PIs have valid <i>PICP</i> values.	143
7.14	PDF and CDF of the time taken for a training iteration to be completed.	143
7.15	PDF and CDF of the time taken to construct one Prediction Interval.	144

B List of Tables

2.1	Choice of Γ_i as a function of $n = J_i $ so that the probability of constraint violation is less than 1%. Source: [33].	38
6.1	Standardized nominal voltage levels for railroad systems. Source: [116].	81
6.2	Current rolling stock operated by the Rhine-Neckar Transportation system. Sources: [5, 7, 11].	96
7.1	Input features for the LUBE NN.	129
7.2	Parameters used with the case studies.	133
7.3	Extract of the data-set for the input features and target for the LUBE NN. This particular fragment corresponds to the 24 hours of a Friday (day 6). The attributes are defined in table 7.1. Source: [73, 72].	135
7.4	Statistics for the quality assessment measurements of PIs constructed from the test data-set for the best 20 NNs in each trained ensemble. Minimum values of each column are in bold.	140

C Bibliography

- [1] WIKIPEDIA (Hrsg.): *Linear programming* - *Wikipedia*. <https://en.wikipedia.org/w/index.php?oldid=818243879>
- [2] WIKIPEDIA (Hrsg.): *Linear programming relaxation* - *Wikipedia*. <https://en.wikipedia.org/w/index.php?oldid=780713556>
- [3] WIKIPEDIA (Hrsg.): *Polytope* - *Wikipedia*. <https://en.wikipedia.org/w/index.php?oldid=815630532>
- [4] WIKIPEDIA (Hrsg.): *Simulated annealing* - *Wikipedia*. <https://en.wikipedia.org/w/index.php?oldid=820036255>
- [5] *Strassenbahn-Online*. <http://www.strassenbahn-online.de/index.html>. Version: 13/07/2017
- [6] WIKIPEDIA (Hrsg.): *Bahnstrecke Mannheim-Käfertal–Heddesheim*. <https://de.wikipedia.org/w/index.php?oldid=166877863>. Version: 18/07/2017
- [7] WIKIPEDIA (Hrsg.): *Variobahn*. <https://de.wikipedia.org/w/index.php?oldid=167165353>. Version: 29/07/2017
- [8] WIKIPEDIA (Hrsg.): *Half-space (geometry)* - *Wikipedia*. <https://en.wikipedia.org/w/index.php?oldid=698570860>. Version: 31/12/2017
- [9] *Germany's Eco-power hits a new high: 44.1 per cent of all Germany's electricity production for October was produced from renewable sources*. https://www.gtai.de/GTAI/Navigation/EN/Meta/Press/press-releases,t=germanys-ecopower-hits-a-new-high,did=1815460.html?channel=red_gtai_pressenews_en. Version: Nov 13, 2017
- [10] ABBAS, Ghulam ; GU, Jason ; FAROOQ, Umar ; ASAD, M. U. ; EL-HAWARY, M.: Solution of an Economic Dispatch Problem through Particle Swarm Optimization: A Detailed Survey – Part I. In: *IEEE Access* (2017), S. 1. <http://dx.doi.org/10.1109/ACCESS.2017.2723862>. – DOI 10.1109/ACCESS.2017.2723862. – ISSN 2169–3536

- [11] AG TRAM-INFO: *Wagenparkliste Rhein-Neckar-Verkehr GmbH (rnv)*. http://www.tram-info.de/wagenp/rnv_a.htm. Version: 29/07/2017
- [12] AKBARI, Jalal ; AHMADI, Mohammad T. ; MOHARRAMI, Hamid: Advances in concrete arch dams shape optimization. In: *Applied Mathematical Modelling* 35 (2011), Nr. 7, S. 3316–3333. <http://dx.doi.org/10.1016/j.apm.2011.01.020>. – DOI 10.1016/j.apm.2011.01.020
- [13] AKBARI, Kaveh ; NASIRI, Mohammad M. ; JOLAI, Fariborz ; GHADERI, Seyed F.: Optimal investment and unit sizing of distributed energy systems under uncertainty: A robust optimization approach. In: *Energy and Buildings* 85 (2014), S. 275–286. <http://dx.doi.org/10.1016/j.enbuild.2014.09.009>. – DOI 10.1016/j.enbuild.2014.09.009. – ISSN 03787788
- [14] AL-JAWAD, Jafar Y. ; TANYIMBOH, Tiku T.: Reservoir operation using a robust evolutionary optimization algorithm. In: *Journal of environmental management* 197 (2017), S. 275–286. <http://dx.doi.org/10.1016/j.jenvman.2017.03.081>. – DOI 10.1016/j.jenvman.2017.03.081. – ISSN 1095–8630
- [15] ALBRECHT, A.: Reducing Power Peaks And Energy Consumption In Rail Transit Systems By Simultaneous Train Running Time Control. In: *WIT Transactions on The Built Environment* 74 (2004), 885–894. <https://www.witpress.com/elibrary/wit-transactions-on-the-built-environment/74/12117>
- [16] ALI, Mohd H. ; WU, Bin ; DOUGAL, Roger A.: An Overview of SMES Applications in Power and Energy Systems. In: *IEEE Transactions on Sustainable Energy* 1 (2010), Nr. 1, S. 38–47. <http://dx.doi.org/10.1109/TSTE.2010.2044901>. – DOI 10.1109/TSTE.2010.2044901. – ISSN 1949–3029
- [17] ALVES, Felipe da S. ; SOUZA, Jame Neiva Miranda d. ; COSTA, André Luiz H.: Multi-objective design optimization of natural gas transmission networks. In: *Computers & Chemical Engineering* 93 (2016), S. 212–220. <http://dx.doi.org/10.1016/j.compchemeng.2016.06.006>. – DOI 10.1016/j.compchemeng.2016.06.006. – ISSN 00981354
- [18] AMPL: *AMPL API*. <http://ampl.com/products/api/>
- [19] ANDRADE, Luciano Carli M. ; DA SILVA, Ivan N.: Very Short-Term Load Forecasting Using a Hybrid Neuro-fuzzy Approach. In: *2010 Eleventh Brazilian Symposium on Neural Networks*, IEEE, 2010. – ISBN 978–1–4244–8391–4, S. 115–120
- [20] ANSARI, Y. ; MANTI, M. ; FALOTICO, E. ; CIANCHETTI, M. ; LASCHI, C.: Multiobjective Optimization for Stiffness and Position Control in a Soft Robot Arm Module. In: *IEEE*

- Robotics and Automation Letters* 3 (2018), Nr. 1, S. 108–115. <http://dx.doi.org/10.1109/LRA.2017.2734247>. – DOI 10.1109/LRA.2017.2734247. – ISSN 2377–3766
- [21] BAHRAMIRAD, Shaghayegh ; REDER, Wanda ; KHODAEI, Amin: Reliability-Constrained Optimal Sizing of Energy Storage System in a Microgrid. In: *IEEE Transactions on Smart Grid* 3 (2012), Nr. 4, S. 2056–2062. <http://dx.doi.org/10.1109/TSG.2012.2217991>. – DOI 10.1109/TSG.2012.2217991. – ISSN 1949–3053
- [22] BARRERO, R. ; MIERLO, J. ; TACKOEN, X.: Energy savings in public transport. In: *IEEE Vehicular Technology Magazine* 3 (2008), Nr. 3, S. 26–36. <http://dx.doi.org/10.1109/MVT.2008.927485>. – DOI 10.1109/MVT.2008.927485. – ISSN 1556–6072
- [23] BARRERO, R. ; TACKOEN, X. ; VAN MIERLO, J.: Stationary or onboard energy storage systems for energy consumption reduction in a metro network. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 224 (2010), Nr. 3, S. 207–225. <http://dx.doi.org/10.1243/09544097JRRT322>. – DOI 10.1243/09544097JRRT322. – ISSN 0954–4097
- [24] BARTŁOMIEJCZYK, Mikołaj ; POŁOM, Marcin: Multiaspect measurement analysis of breaking energy recovery. In: *Energy Conversion and Management* 127 (2016), S. 35–42. <http://dx.doi.org/10.1016/j.enconman.2016.08.089>. – DOI 10.1016/j.enconman.2016.08.089. – ISSN 01968904
- [25] BEMPORAD, Alberto ; MORARI, Manfred: Control of systems integrating logic, dynamics, and constraints. In: *Automatica* 35 (1999), Nr. 3, S. 407–427. [http://dx.doi.org/10.1016/S0005-1098\(98\)00178-2](http://dx.doi.org/10.1016/S0005-1098(98)00178-2). – DOI 10.1016/S0005-1098(98)00178-2. – ISSN 00051098
- [26] BEN-TAL, A. ; EL GHAOU, Laurent ; NEMIROVSKIĀ■, ArkadiĀ■ S.: *Robust optimization*. Princeton, N.J. and Woodstock : Princeton University Press, 2009 (Princeton series in applied mathematics). – ISBN 978–0–691–14368–2
- [27] BEN-TAL, A. ; NEMIROVSKI, A.: Robust Convex Optimization. In: *Mathematics of Operations Research* 23 (1998), Nr. 4, S. 769–805. <http://dx.doi.org/10.1287/moor.23.4.769>. – DOI 10.1287/moor.23.4.769
- [28] BEN-TAL, A. ; NEMIROVSKI, A.: Robust solutions of uncertain linear programs. In: *Operations Research Letters* 25 (1999), Nr. 1, S. 1–13. [http://dx.doi.org/10.1016/S0167-6377\(99\)00016-4](http://dx.doi.org/10.1016/S0167-6377(99)00016-4). – DOI 10.1016/S0167-6377(99)00016-4. – ISSN 01676377

- [29] BEN-TAL, Aharon ; NEMIROVSKI, Arkadi: Robust solutions of Linear Programming problems contaminated with uncertain data. In: *Mathematical Programming* 88 (2000), Nr. 3, S. 411–424. <http://dx.doi.org/10.1007/PL00011380>. – DOI 10.1007/PL00011380. – ISSN 0025–5610
- [30] BERNARDO, Fernando P. ; SARAIVA, Pedro M.: Robust optimization framework for process parameter and tolerance design. In: *AIChE Journal* 44 (1998), Nr. 9, S. 2007–2017. <http://dx.doi.org/10.1002/aic.690440908>. – DOI 10.1002/aic.690440908. – ISSN 00011541
- [31] BERTSIMAS, Dimitris ; BROWN, David B. ; CARAMANIS, Constantine: Theory and Applications of Robust Optimization. In: *SIAM Review* 53 (2011), Nr. 3, S. 464–501. <http://dx.doi.org/10.1137/080734510>. – DOI 10.1137/080734510. – ISSN 0036–1445
- [32] BERTSIMAS, Dimitris ; LITVINOV, Eugene ; SUN, Xu A. ; ZHAO, JinYE ; ZHENG, Tongxin: Adaptive Robust Optimization for the Security Constrained Unit Commitment Problem. In: *IEEE Transactions on Power Systems* 28 (2013), Nr. 1, S. 52–63. <http://dx.doi.org/10.1109/TPWRS.2012.2205021>. – DOI 10.1109/TPWRS.2012.2205021. – ISSN 08858950
- [33] BERTSIMAS, Dimitris ; SIM, Melvyn: The Price of Robustness. In: *Operations Research* 52 (2004), Nr. 1, S. 35–53
- [34] BISHOP, Christopher M.: *Neural networks for pattern recognition*. Oxford and New York : Clarendon Press and Oxford University Press, 1995. – ISBN 0198538642
- [35] BITTANTE, Alice ; JOKINEN, Raine ; PETTERSSON, Frank ; SAXÉN, Henrik: Optimization of LNG Supply Chain. Version:2015. <http://dx.doi.org/10.1016/B978-0-444-63578-5.50125-0>. In: GERNAEY, Krist V. (Hrsg.) ; HUUSOM, Jakob K. (Hrsg.) ; GANI, R. (Hrsg.): *12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering* Bd. 37. Amsterdam : Elsevier, 2015. – DOI 10.1016/B978-0-444-63578-5.50125-0. – ISBN 9780444634290, S. 779–784
- [36] BLOOMBERG NEW ENERGY FINANCE: *New Energy Outlook 2017: Bloomberg New Energy Finance's annual long-term economic forecast of the world's power sector: Executive Summary*
- [37] BODLA, Karthik K. ; MURTHY, Jayathi Y. ; GARIMELLA, Suresh V.: Optimization under uncertainty for electronics cooling design applications. In: *13th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*

- (*ITherm*), 2012. Piscataway, NJ and Piscataway, NJ : IEEE, 2012. – ISBN 978–1–4244–9532–0, S. 1191–1201
- [38] BORDIN, Chiara ; GORDINI, Angelo ; VIGO, Daniele: An optimization approach for district heating strategic network design. In: *European Journal of Operational Research* 252 (2016), Nr. 1, S. 296–307. <http://dx.doi.org/10.1016/j.ejor.2015.12.049>. – DOI 10.1016/j.ejor.2015.12.049. – ISSN 03772217
- [39] BOYD, Stephen P. ; VANDENBERGHE, Lieven: *Convex optimization*. Cambridge : Cambridge University Press, 2004. – ISBN 978–0–521–83378–3
- [40] BP P.L.C.: *BP Energy Outlook: 2017 Edition*. <http://www.bp.com/energyoutlook>
- [41] BRAUN, Philipp ; GRUNE, Lars ; KELLETT, Christopher M. ; WELLER, Steven R. ; WORTHMANN, Karl: A Distributed Optimization Algorithm for the Predictive Control of Smart Grids. In: *IEEE Transactions on Automatic Control* 61 (2016), Nr. 12, S. 3898–3911. <http://dx.doi.org/10.1109/TAC.2016.2525808>. – DOI 10.1109/TAC.2016.2525808. – ISSN 00189286
- [42] BRUNO, D. ; LONETTI, P. ; PASCUZZO, A.: An optimization model for the design of network arch bridges. In: *Computers & Structures* 170 (2016), S. 13–25. <http://dx.doi.org/10.1016/j.compstruc.2016.03.011>. – DOI 10.1016/j.compstruc.2016.03.011. – ISSN 00457949
- [43] CAO, Hongyou ; QIAN, Xudong ; CHEN, Zhijun ; ZHU, Hongping: Layout and size optimization of suspension bridges based on coupled modelling approach and enhanced particle swarm optimization. In: *Engineering Structures* 146 (2017), S. 170–183. <http://dx.doi.org/10.1016/j.engstruct.2017.05.048>. – DOI 10.1016/j.engstruct.2017.05.048. – ISSN 01410296
- [44] CAPASSO, C. ; SEPE, V. ; VENERI, O. ; MONTANARI, M. ; POLETTI, L.: Experimentation with a ZEBRA plus EDLC based hybrid storage system for urban means of transport. In: *2015 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles (ESARS)*, IEEE, 2015. – ISBN 978–1–4799–7400–9, S. 1–6
- [45] CARDELLA, U. ; DECKER, L. ; SUNDBERG, J. ; KLEIN, H.: Process optimization for large-scale hydrogen liquefaction. In: *International Journal of Hydrogen Energy* 42 (2017), Nr. 17, S. 12339–12354. <http://dx.doi.org/10.1016/j.ijhydene.2017.03.167>. – DOI 10.1016/j.ijhydene.2017.03.167. – ISSN 03603199

- [46] CASTELLINI, Luca ; LUCIDI, Stefano ; VILLANI, Marco: Design optimization of switched reluctance motor for aerospace application. In: *IEEE International Electric Machines and Drives Conference (IEMDC)*. Piscataway, NJ and Piscataway, NJ : IEEE, 2015. – ISBN 978–1–4799–7941–7, S. 1678–1682
- [47] CASTRO, P.A.D. ; ZUBEN, F. J.: Bayesian Learning of Neural Networks by Means of Artificial Immune Systems. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, IEEE, 2006. – ISBN 0–7803–9490–9, S. 4831–4838
- [48] CENTER FOR SUSTAINABLE ENERGY: *What is a Microgrid?* <https://energycenter.org/self-generation-incentive-program/business/technologies/microgrid>
- [49] CERAOLO, M. ; GIGLIOLI, R. ; LUTZEMBERGER, G. ; BECHINI, A.: Cost effective storage for energy saving in feeding systems of tramways. In: *2014 IEEE International Electric Vehicle Conference (IEVC)*, IEEE, 2014. – ISBN 978–1–4799–6075–0, S. 1–6
- [50] CERAOLO, M. ; LUTZEMBERGER, G.: Stationary and on-board storage systems to enhance energy and cost efficiency of tramways. In: *Journal of Power Sources* 264 (2014), S. 128–139. <http://dx.doi.org/10.1016/j.jpowsour.2014.04.070>. – DOI 10.1016/j.jpowsour.2014.04.070. – ISSN 03787753
- [51] CERAOLO, Massimo ; GIGLIOLI, Romano ; LUTZEMBERGER, Giovanni ; CONTE, Mario ; PASQUALI, Manlio: Use of electrochemical storage in substations to enhance energy and cost efficiency of tramways. In: *AEIT Annual Conference 2013*, IEEE, 2013. – ISBN 978–8–8872–3732–0, S. 1–6
- [52] CHATFIELD, Chris: Prediction Intervals for Time-Series Forecasting. Version: 2001. http://dx.doi.org/10.1007/978-0-306-47630-3_{_}21. In: HILLIER, Frederick S. (Hrsg.) ; ARMSTRONG, J. S. (Hrsg.): *Principles of Forecasting* Bd. 30. Boston, MA : Springer US, 2001. – DOI 10.1007/978-0-306-47630-3_21. – ISBN 978–0–7923–7401–5, S. 475–494
- [53] CHATURVEDI, D. K. ; PREMDAYAL, Sinha A.: Modified neural and neuro-fuzzy approach for short term load forecasting. In: *2012 2nd International Conference on Power, Control and Embedded Systems*, IEEE, 2012. – ISBN 978–1–4673–1049–9, S. 1–5
- [54] CHEMALI, Ephrem ; PREINDL, Matthias ; MALYSZ, Pawel ; EMADI, Ali: Electrochemical and Electrostatic Energy Storage and Management Systems for Electric Drive Vehicles: State-of-the-Art Review and Future Trends. In: *IEEE Journal of Emerging and Selected Topics in Power Electronics* 4 (2016), Nr. 3, S. 1117–1134. <http://dx>

- .doi.org/10.1109/JESTPE.2016.2566583. – DOI 10.1109/JESTPE.2016.2566583. – ISSN 2168–6777
- [55] CHEN, J.-F. ; LIN, R.-L. ; LIU, Y.-C.: Optimization of an MRT Train Schedule: Reducing Maximum Traction Power by Using Genetic Algorithms. In: *IEEE Transactions on Power Systems* 20 (2005), Nr. 3, S. 1366–1372. <http://dx.doi.org/10.1109/TPWRS.2005.851939>. – DOI 10.1109/TPWRS.2005.851939. – ISSN 08858950
- [56] CHOI, Jin W. ; LEE, Hajeong ; LEE, Jung C. ; LEE, Saram ; KIM, Yon S. ; YOON, Hyung-Jin ; KIM, Hee C.: Application of genetic algorithm for hemodialysis schedule optimization. In: *Computer methods and programs in biomedicine* 145 (2017), S. 35–43. <http://dx.doi.org/10.1016/j.cmpb.2017.04.003>. – DOI 10.1016/j.cmpb.2017.04.003. – ISSN 1872–7565
- [57] CHOI, Jongwoo ; PARK, Wan-Ki ; LEE, Il-Woo: Application of vanadium redox flow battery to grid connected microgrid energy management. In: *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, IEEE, 2016. – ISBN 978–1–5090–3388–1, S. 903–906
- [58] CHU, Millie ; ZINCHENKO, Yuriy ; HENDERSON, Shane G. ; SHARPE, Michael B.: Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. In: *Physics in medicine and biology* 50 (2005), Nr. 23, S. 5463–5477. <http://dx.doi.org/10.1088/0031-9155/50/23/003>. – DOI 10.1088/0031–9155/50/23/003. – ISSN 0031–9155
- [59] CICCARELLI, Flavio: *Energy management and control strategies for the use of supercapacitors storage technologies in urban railway traction systems*. Naples, University of Naples "Federico II", PhD Thesis, 2014
- [60] CICCARELLI, Flavio ; IANNUZZI, Diego ; KONDO, Keiichiro ; FRATELLI, Luigi: Line-Voltage Control Based on Wayside Energy Storage Systems for Tramway Networks. In: *IEEE Transactions on Power Electronics* 31 (2016), Nr. 1, S. 884–899. <http://dx.doi.org/10.1109/TPEL.2015.2411996>. – DOI 10.1109/TPEL.2015.2411996. – ISSN 0885–8993
- [61] COLE SMITH, J. ; CANER TASKIN, Z.: A Tutorial Guide to Mixed-Integer Programming Models and Solution Techniques. In: LIM, Gino J. (Hrsg.) ; LEE, Eva K. (Hrsg.): *Optimization in medicine and biology*. Boca Raton, Fla. : Auerbach and London : Taylor & Francis [distributor], 2008 (Engineering and management innovation series). – ISBN 0849305632, S. 521–548

- [62] CONEJO, A. J. ; PLAZAS, M. A. ; ESPINOLA, R. ; MOLINA, A. B.: Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. In: *IEEE Transactions on Power Systems* 20 (2005), Nr. 2, 1035–1042. <http://dx.doi.org/10.1109/TPWRS.2005.846054>. – DOI 10.1109/TPWRS.2005.846054. – ISSN 08858950
- [63] CORNIC, Daniel: Efficient recovery of braking energy through a reversible dc substation. In: *Electrical Systems for Aircraft, Railway and Ship Propulsion*, IEEE, 2010. – ISBN 978–1–4244–9092–9, S. 1–9
- [64] COS JUEZ, Francisco J. ; SÁNCHEZ LASHERAS, Fernando ; ROQUEÑÍ, Nieves ; OSBORN, James: An ANN-based smart tomographic reconstructor in a dynamic environment. In: *Sensors (Basel, Switzerland)* 12 (2012), Nr. 7, S. 8895–8911. <http://dx.doi.org/10.3390/s120708895>. – DOI 10.3390/s120708895. – ISSN 1424–8220
- [65] COSSUTTA, Pablo ; ENGELHARDT, Mathias A. ; AGUIRRE, Miguel ; PONCE, Juan ; VALLA, Maria I.: Model predictive control of a multilevel current source inverter together with its current source. In: *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, IEEE, 2017. – ISBN 978–1–5090–1412–5, S. 762–767
- [66] CUI, Yunfei ; GENG, Zhiqiang ; ZHU, Qunxiong ; HAN, Yongming: Review: Multi-objective optimization methods and application in energy saving. In: *Energy* 125 (2017), S. 681–704. <http://dx.doi.org/10.1016/j.energy.2017.02.174>. – DOI 10.1016/j.energy.2017.02.174. – ISSN 03605442
- [67] CZERNICHOW, T. ; PIRAS, A. ; IMHOF, K. ; CAIRE, P. ; JACCARD, Y. ; DORIZZI, B. ; GERMOND, A.: Short term electrical load forecasting with artificial neural networks. In: *Engineering Intelligent Syst.* vol. 2 (1996), S. 85–99
- [68] DAGOSTINO, Riccardo ; BAUMANN, Lars ; DAMIANO, Alfonso ; BOGGASCH, Ekkehard: A Vanadium-Redox-Flow-Battery Model for Evaluation of Distributed Storage Implementation in Residential Energy Systems. In: *IEEE Transactions on Energy Conversion* 30 (2015), Nr. 2, S. 421–430. <http://dx.doi.org/10.1109/TEC.2014.2369437>. – DOI 10.1109/TEC.2014.2369437. – ISSN 0885–8969
- [69] D'ANDREAGIOVANNI, Fabio ; FELICI, Giovanni ; LACALANDRA, Fabrizio: *Revisiting the use of Robust Optimization for optimal energy offering under price uncertainty*. <http://arxiv.org/pdf/1601.01728v2>
- [70] DAVLEA, Laura ; TEODORESCU, Bogdan: A neuro-fuzzy algorithm for middle-term load forecasting. In: *2016 International Conference and Exposition on Electrical and Power Engineering (EPE)*, IEEE, 2016. – ISBN 978–1–5090–6129–7, S. 005–009

- [71] DENG, Jianguang ; JIRUTITIJAROEN, Panida: Short-term load forecasting using time series analysis: A case study for Singapore. In: *2010 IEEE Conference on Cybernetics and Intelligent Systems*, 2010. – ISBN –2005, 231–236
- [72] DEORAS, Ameya: *Electricity Load and Price Forecasting Webinar Case Study*. <https://de.mathworks.com/matlabcentral/fileexchange/28684-electricity-load-and-price-forecasting-webinar-case-study>. Version:2010
- [73] DEORAS, Ameya: *Electricity Load and Price Forecasting with MATLAB*. <https://de.mathworks.com/videos/electricity-load-and-price-forecasting-with-matlab-81765.html>. Version:2010
- [74] DERNONCOURT, Franck ; LEE, Ji Y.: Optimizing neural network hyperparameters with Gaussian processes for dialog act classification. In: *2016 IEEE Spoken Language Technology Workshop (SLT)*, IEEE, 2016. – ISBN 978–1–5090–4903–5, S. 406–413
- [75] DI SOMMA, M. ; YAN, B. ; BIANCO, N. ; GRADITI, G. ; LUH, P. B. ; MONGIBELLO, L. ; NASO, V.: Multi-objective design optimization of distributed energy systems through cost and exergy assessments. In: *Applied Energy* (2017). <http://dx.doi.org/10.1016/j.apenergy.2017.03.105>. – DOI 10.1016/j.apenergy.2017.03.105. – ISSN 03062619
- [76] EDSTRAND, Jakob: *Calculation method for powering a tramway network*. Göteborg, Sweden, Chalmers University of Technology, Master of Science Thesis in Electrical Engineering, 2012
- [77] EHTERAM, Mohammed ; KARAMI, Hojat ; MOUSAVI, Sayed-Farhad ; EL-SHAFIE, Ahmed ; AMINI, Zahra: Optimizing dam and reservoirs operation based model utilizing shark algorithm approach. In: *Knowledge-Based Systems* 122 (2017), S. 26–38. <http://dx.doi.org/10.1016/j.knosys.2017.01.026>. – DOI 10.1016/j.knosys.2017.01.026. – ISSN 09507051
- [78] EL GHAOU, Laurent ; LEBRET, Herv: Robust Solutions to Least-Squares Problems with Uncertain Data. In: *SIAM Journal on Matrix Analysis and Applications* 18 (1997), Nr. 4, S. 1035–1064. <http://dx.doi.org/10.1137/S0895479896298130>. – DOI 10.1137/S0895479896298130. – ISSN 0895–4798
- [79] EL GHAOU, Laurent ; OUSTRY, Francois ; LEBRET, Hervé: Robust Solutions to Uncertain Semidefinite Programs. In: *SIAM Journal on Optimization* 9 (1998), Nr. 1, S. 33–52. <http://dx.doi.org/10.1137/S1052623496305717>. – DOI 10.1137/S1052623496305717. – ISSN 1052–6234

- [80] ELLIS, Brian L. ; NAZAR, Linda F.: Sodium and sodium-ion energy storage batteries. In: *Current Opinion in Solid State and Materials Science* 16 (2012), Nr. 4, S. 168–177. <http://dx.doi.org/10.1016/j.cossms.2012.04.002>. – DOI 10.1016/j.cossms.2012.04.002. – ISSN 13590286
- [81] FABOZZI, Frank J.: *Robust portfolio optimization and management*. Hoboken, N.J. : John Wiley, 2007 (Frank J. Fabozzi series). – ISBN 0470164891
- [82] FENG, Jia ; LI, Xiamiao ; LIU, Haidong ; GAO, Xing ; MAO, Baohua: Optimizing the Energy-Efficient Metro Train Timetable and Control Strategy in Off-Peak Hours with Uncertain Passenger Demands. In: *Energies* 10 (2017), Nr. 4, S. 436. <http://dx.doi.org/10.3390/en10040436>. – DOI 10.3390/en10040436. – ISSN 1996–1073
- [83] FEROLDI, Diego ; CARIGNANO, Mauro: Sizing for fuel cell/supercapacitor hybrid vehicles based on stochastic driving cycles. In: *Applied Energy* 183 (2016), S. 645–658. <http://dx.doi.org/10.1016/j.apenergy.2016.09.008>. – DOI 10.1016/j.apenergy.2016.09.008. – ISSN 03062619
- [84] FLÓREZ-ORREGO, Daniel ; OLIVEIRA JUNIOR, Silvio de: Modeling and optimization of an industrial ammonia synthesis unit: An exergy approach. In: *Energy* 137 (2017), S. 234–250. <http://dx.doi.org/10.1016/j.energy.2017.06.157>. – DOI 10.1016/j.energy.2017.06.157. – ISSN 03605442
- [85] FOURER, Robert ; GAY, David M. ; KERNIGHAN, Brian W.: *AMPL: A modeling language for mathematical programming*. 2nd ed. Pacific Grove, CA : Thomson/Brooks/Cole, 2003. – ISBN 0–534–38809–4
- [86] GAO, Xiaoyong ; JIANG, Yongheng ; CHEN, Tao ; HUANG, Dexian: Optimizing scheduling of refinery operations based on piecewise linear models. In: *Computers & Chemical Engineering* 75 (2015), S. 105–119. <http://dx.doi.org/10.1016/j.compchemeng.2015.01.022>. – DOI 10.1016/j.compchemeng.2015.01.022. – ISSN 00981354
- [87] GEB, D. ; DEMOULIN, G. ; CATTON, I.: Population-based optimization for heat sink design in electronics cooling applications. In: *Twenty Ninth Annual IEEE Semiconductor Thermal Measurement and Management Symposium*. Piscataway, NJ : IEEE, 2013. – ISBN 978–1–4673–6429–4, S. 173–180
- [88] GELMAN, Vitaly: Braking energy recuperation. In: *IEEE Vehicular Technology Magazine* 4 (2009), Nr. 3, S. 82–89. <http://dx.doi.org/10.1109/MVT.2009.933480>. – DOI 10.1109/MVT.2009.933480. – ISSN 1556–6072

- [89] GELMAN, Vitaly: Energy Storage That May Be Too Good to Be True: Comparison Between Wayside Storage and Reversible Thyristor Controlled Rectifiers for Heavy Rail. In: *IEEE Vehicular Technology Magazine* 8 (2013), Nr. 4, S. 70–80. <http://dx.doi.org/10.1109/MVT.2013.2283350>. – DOI 10.1109/MVT.2013.2283350. – ISSN 1556–6072
- [90] GHADERI, Hamid ; PISHVAEE, Mir S. ; MOINI, Alireza: Biomass supply chain network design: An optimization-oriented review and analysis. In: *Industrial Crops and Products* 94 (2016), S. 972–1000. <http://dx.doi.org/10.1016/j.indcrop.2016.09.027>. – DOI 10.1016/j.indcrop.2016.09.027. – ISSN 09266690
- [91] GHAITHAN, Ahmed M. ; ATTIA, Ahmed ; DUFFUAA, Salih O.: Multi-objective Optimization Model for a Downstream Oil and Gas Supply Chain. In: *Applied Mathematical Modelling* (2017). <http://dx.doi.org/10.1016/j.apm.2017.08.007>. – DOI 10.1016/j.apm.2017.08.007
- [92] GHANBARI, Arash ; ABBASIAN-NAGHNEH, Salman ; HADAVANDI, Esmail: An intelligent load forecasting expert system by integration of ant colony optimization, genetic algorithms and fuzzy logic. In: *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, IEEE, 2011. – ISBN 978–1–4244–9927–4, S. 246–251
- [93] GILLI, Manfred ; MARINGER, Dietmar ; SCHUMANN, Enrico: Optimization Problems in Finance. Version:2011. <http://dx.doi.org/10.1016/B978-0-12-375662-6.00010-9>. In: GILLI, Manfred (Hrsg.) ; MARINGER, Dietmar (Hrsg.) ; SCHUMANN, Enrico (Hrsg.): *Numerical methods and optimization in finance*. London : Academic, 2011. – DOI 10.1016/B978–0–12–375662–6.00010–9. – ISBN 9780123756626, S. 271–285
- [94] GOMEZ, J. F. ; KHODR, H. M. ; DEOLIVEIRA, P. M. ; OCQUE, L. ; YUSTA, J. M. ; VILLASANA, R. ; URDANETA, A. J.: Ant Colony System Algorithm for the Planning of Primary Distribution Circuits. In: *IEEE Transactions on Power Systems* 19 (2004), Nr. 2, S. 996–1004. <http://dx.doi.org/10.1109/TPWRS.2004.825867>. – DOI 10.1109/TPWRS.2004.825867. – ISSN 08858950
- [95] GONZÁLEZ-GIL, Arturo ; PALACIN, Roberto ; BATTY, Paul: Sustainable urban rail systems: Strategies and technologies for optimal management of regenerative braking energy. In: *Energy Conversion and Management* 75 (2013), S. 374–388. <http://dx.doi.org/10.1016/j.enconman.2013.06.039>. – DOI 10.1016/j.enconman.2013.06.039. – ISSN 01968904

- [96] GORISSEN, Bram L. ; YANIKOĞLU, İhsan ; DEN HERTOĞ, Dick: A practical guide to robust optimization. In: *Omega* 53 (2015), S. 124–137. <http://dx.doi.org/10.1016/j.omega.2014.12.006>. – DOI 10.1016/j.omega.2014.12.006. – ISSN 03050483
- [97] GRIGANS, Linards ; LATKOVSKIS, Leonards: Estimation of the power and energy requirements fro trackside storage systems. In: *13th European Conference on Power Electronics and Applications* (2009)
- [98] GRILLO, Samuele ; PIEVATOLO, Antonio ; TIRONI, Enrico: Optimal Storage Scheduling Using Markov Decision Processes. In: *IEEE Transactions on Sustainable Energy* 7 (2016), Nr. 2, S. 755–764. <http://dx.doi.org/10.1109/TSTE.2015.2497718>. – DOI 10.1109/TSTE.2015.2497718. – ISSN 1949–3029
- [99] GUARNIERI, Massimo ; MATTAVELLI, Paolo ; PETRONE, Giovanni ; SPAGNUOLO, Giovanni: Vanadium Redox Flow Batteries: Potentials and Challenges of an Emerging Storage Technology. In: *IEEE Industrial Electronics Magazine* 10 (2016), Nr. 4, S. 20–31. <http://dx.doi.org/10.1109/MIE.2016.2611760>. – DOI 10.1109/MIE.2016.2611760. – ISSN 1932–4529
- [100] HAIFENG, Dai ; XUEYU, Chang: A Study on Lead Acid Battery and Ultra-capacitor Hybrid Energy Storage System for Hybrid City Bus. In: *2010 International Conference on Optoelectronics and Image Processing*. [Los Alamitos, Calif.] : IEEE Computer Society, 2010. – ISBN 978–1–4244–8683–0, S. 154–159
- [101] HAJIMIRAGHA, Amir H. ; CANIZARES, Claudio A. ; FOWLER, Michael W. ; MOAZENI, Somayeh ; ELKAMEL, Ali: A Robust Optimization Approach for Planning the Transition to Plug-in Hybrid Electric Vehicles. In: *IEEE Transactions on Power Systems* 26 (2011), Nr. 4, S. 2264–2274. <http://dx.doi.org/10.1109/TPWRS.2011.2108322>. – DOI 10.1109/TPWRS.2011.2108322. – ISSN 08858950
- [102] HASSAN, Saima ; KHOSRAVI, Abbas ; JAAFAR, Jafreezal: Neural network ensemble: Evaluation of aggregation algorithms in electricity demand forecasting. In: *The 2013 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2013. – ISBN 978–1–4673–6129–3, S. 1–6
- [103] HAYKIN, Simon S.: *Neural networks: A comprehensive foundation / Simon Haykin*. New York : Macmillan and London : Maxwell Macmillan International, 1994. – ISBN 0023527617
- [104] HE, Chaoming ; SUN, Haoran ; XU, Yang ; LV, Siyun: Hydrogen refueling station siting of expressway based on the optimization of hydrogen life cycle cost. In: *International Journal of Hydrogen Energy* 42 (2017), Nr. 26, S. 16313–16324. <http://dx.d>

- oi.org/10.1016/j.ijhydene.2017.05.073. – DOI 10.1016/j.ijhydene.2017.05.073. – ISSN 03603199
- [105] HERRERA, Victor ; MILO, Aitor ; GAZTAÑAGA, Haizea ; ETXEBERRIA-OTADUI, Ion ; VILLARREAL, Igor ; CAMBLONG, Haritza: Adaptive energy management strategy and optimal sizing applied on a battery-supercapacitor based tramway. In: *Applied Energy* 169 (2016), S. 831–845. <http://dx.doi.org/10.1016/j.apenergy.2016.02.079>. – DOI 10.1016/j.apenergy.2016.02.079. – ISSN 03062619
- [106] HIKSAS, Muhammad M. ; ANINDITIO, Muhammad L.: Redox Flow Batteries for small scale energy storage. In: *2016 IEEE Conference on Technologies for Sustainability (SusTech)*, IEEE, 2016. – ISBN 978-1-5090-4158-9, S. 134–139
- [107] HIPPERT, H. S. ; PEDREIRA, C. E. ; SOUZA, R. C.: Neural networks for short-term load forecasting: A review and evaluation. In: *IEEE Transactions on Power Systems* 16 (2001), Nr. 1, S. 44–55. <http://dx.doi.org/10.1109/59.910780>. – DOI 10.1109/59.910780. – ISSN 08858950
- [108] HORNIK, Kurt ; STINCHCOMBE, Maxwell ; WHITE, Halbert: Multilayer feedforward networks are universal approximators. In: *Neural Networks* 2 (1989), Nr. 5, S. 359–366. [http://dx.doi.org/10.1016/0893-6080\(89\)90020-8](http://dx.doi.org/10.1016/0893-6080(89)90020-8). – DOI 10.1016/0893-6080(89)90020-8. – ISSN 08936080
- [109] HOSEN, Mohammad A. ; KHOSRAVI, Abbas ; CREIGHTON, Douglas ; NAHAVANDI, Saeid: Prediction interval-based modelling of polymerization reactor: A new modelling strategy for chemical reactors. In: *Journal of the Taiwan Institute of Chemical Engineers* 45 (2014), Nr. 5, S. 2246–2257. <http://dx.doi.org/10.1016/j.jtice.2014.05.021>. – DOI 10.1016/j.jtice.2014.05.021. – ISSN 18761070
- [110] HOSEN, Mohammad A. ; KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Douglas ; SALAKEN, Syed M.: *Prediction interval-based neural network controller for nonlinear processes: 12 - 17 July 2015, Killarney, Ireland*. Piscataway, NJ : IEEE, 2015. <http://dx.doi.org/10.1109/IJCNN.2015.7280382>. <http://dx.doi.org/10.1109/IJCNN.2015.7280382>. – ISBN 9781479919598
- [111] [HTTPS://MATH.STACKEXCHANGE.COM/USERS/25159/MDP](https://math.stackexchange.com/users/25159/mdp): *What is the difference between linear and affine function*. URL:[https://math.stackexchange.com/q/275327\(version:2017-04-12\)](https://math.stackexchange.com/q/275327(version:2017-04-12))
- [112] HUSSAIN, Akhtar ; BUI, Van-Hai ; KIM, Hak-Man: Robust Optimization-Based Scheduling of Multi-Microgrids Considering Uncertainties. In: *Energies* 9 (2016), Nr. 4, S. 278. <http://dx.doi.org/10.3390/en9040278>. – DOI 10.3390/en9040278. – ISSN 1996-1073

- [113] IANNUZZI, Diego: Improvement of the energy recovery of traction electrical drives using supercapacitors. In: *2008 13th International Power Electronics and Motion Control Conference*, IEEE, 2008. – ISBN 978–1–4244–1741–4, S. 1469–1474
- [114] IANNUZZI, Diego ; CICCARELLI, Flavio ; LAURIA, Davide: Stationary ultracapacitors storage device for improving energy saving and voltage profile of light transportation networks. In: *Transportation Research Part C: Emerging Technologies* 21 (2012), Nr. 1, S. 321–337. <http://dx.doi.org/10.1016/j.trc.2011.11.002>. – DOI 10.1016/j.trc.2011.11.002
- [115] IBAIONDO, Harkaitz ; ROMO, Asier: Kinetic energy recovery on railway systems with feedback to the grid. In: *Proceedings of 14th International Power Electronics and Motion Control Conference EPE-PEMC 2010*, IEEE, 2010. – ISBN 978–1–4244–7856–9
- [116] INTERNATIONAL ELECTROTECHNICAL COMMISSION: *Railway applications - Supply voltages of traction systems*. 4.0. 2014
- [117] IPCC ; EDENHOFER, O. (Hrsg.) ; R. PICHES-MADRUGA (Hrsg.) ; Y. SOKONA (Hrsg.) ; E. FARAHANI (Hrsg.) ; S. KADNER (Hrsg.) ; K. SEYBOTH (Hrsg.) ; A. ADLER (Hrsg.) ; I. BAUM (Hrsg.) ; S. BRUNNER (Hrsg.) ; P. EICKEMEIER (Hrsg.) ; B. KRIEMANN (Hrsg.) ; J. SAVOLAINEN (Hrsg.) ; S. SCHLÖMER (Hrsg.) ; C. VON STECHOW (Hrsg.) ; T. ZWICKEL AND J.C. MINX (Hrsg.): *Climate Change 2014: Mitigation of Climate Change: Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. <http://www.ipcc.ch/report/ar5/wg3/>
- [118] ISE, T. ; KITA, M. ; TAGUCHI, A.: A Hybrid Energy Storage With a SMES and Secondary Battery. In: *IEEE Transactions on Applied Superconductivity* 15 (2005), Nr. 2, S. 1915–1918. <http://dx.doi.org/10.1109/TASC.2005.849333>. – DOI 10.1109/TASC.2005.849333. – ISSN 1051–8223
- [119] JABR, R. A.: Robust Transmission Network Expansion Planning With Uncertain Renewable Generation and Loads. In: *IEEE Transactions on Power Systems* 28 (2013), Nr. 4, S. 4558–4567. <http://dx.doi.org/10.1109/TPWRS.2013.2267058>. – DOI 10.1109/TPWRS.2013.2267058. – ISSN 08858950
- [120] JAIN, Amit ; BABITA JAIN, M. ; SRINIVAS, E.: A novel hybrid method for short term load forecasting using fuzzy logic and particle swarm optimization. In: *2010 International Conference on Power System Technology*, IEEE, 2010. – ISBN 978–1–4244–5938–4, S. 1–7

- [121] JALILVAND-NEJAD, Amir ; SHAFAEI, Rasoul ; SHAHRIARI, Hamid: Robust optimization under correlated polyhedral uncertainty set. In: *Computers & Industrial Engineering* 92 (2016), S. 82–94. <http://dx.doi.org/10.1016/j.cie.2015.12.006>. – DOI 10.1016/j.cie.2015.12.006. – ISSN 03608352
- [122] JENSEN, Ida G. ; MÜNSTER, Marie ; PISINGER, David: Optimizing the supply chain of biomass and biogas for a single plant considering mass and energy losses. In: *European Journal of Operational Research* 262 (2017), Nr. 2, S. 744–758. <http://dx.doi.org/10.1016/j.ejor.2017.03.071>. – DOI 10.1016/j.ejor.2017.03.071. – ISSN 03772217
- [123] JI, Bing ; SONG, Xueguan ; SCIBERRAS, Edward ; CAO, Wenping ; HU, Yihua ; PICKERT, Volker: Multiobjective Design Optimization of IGBT Power Modules Considering Power Cycling and Thermal Cycling. In: *IEEE Transactions on Power Electronics* 30 (2015), Nr. 5, S. 2493–2504. <http://dx.doi.org/10.1109/TPEL.2014.2365531>. – DOI 10.1109/TPEL.2014.2365531. – ISSN 0885–8993
- [124] JIANG, Ruiwei ; WANG, Jianhui ; GUAN, Yongpei: Robust Unit Commitment With Wind Power and Pumped Storage Hydro. In: *IEEE Transactions on Power Systems* 27 (2012), Nr. 2, S. 800–810. <http://dx.doi.org/10.1109/TPWRS.2011.2169817>. – DOI 10.1109/TPWRS.2011.2169817. – ISSN 08858950
- [125] KAVOUSI-FARD, Abdollah ; KHOSRAVI, Abbas ; NAHAVANDI, Saeid: A New Fuzzy-Based Combined Prediction Interval for Wind Power Forecasting. In: *IEEE Transactions on Power Systems* 31 (2016), Nr. 1, S. 18–26. <http://dx.doi.org/10.1109/TPWRS.2015.2393880>. – DOI 10.1109/TPWRS.2015.2393880. – ISSN 08858950
- [126] KHARGONEKAR, P. P. ; PETERSEN, I. R. ; ZHOU, K.: Robust stabilization of uncertain linear systems: Quadratic stabilizability and H_∞ / control theory. In: *IEEE Transactions on Automatic Control* 35 (1990), Nr. 3, S. 356–361. <http://dx.doi.org/10.1109/9.50357>. – DOI 10.1109/9.50357. – ISSN 00189286
- [127] KHODABAKHSH, Raheleh: *Energy management in grid-connected microgrids with on-site storage devices*. Canada, McMaster university, Master of Applied Science, August 2015
- [128] KHOSRAVI, A. ; NAHAVANDI, S. ; CREIGHTON, D.: Prediction Intervals for Short-Term Wind Farm Power Generation Forecasts. In: *IEEE Transactions on Sustainable Energy* 4 (2013), Nr. 3, S. 602–610. <http://dx.doi.org/10.1109/TSTE.2012.2232944>. – DOI 10.1109/TSTE.2012.2232944. – ISSN 1949–3029

- [129] KHOSRAVI, Abbas ; NAHAVANDI, Saeid: Closure to the Discussion of “Prediction Intervals for Short-Term Wind Farm Generation Forecasts” and “Combined Non-parametric Prediction Intervals for Wind Power Generation” and the Discussion of “Combined Nonparametric Prediction Intervals for Wind Power Generation”. In: *IEEE Transactions on Sustainable Energy* 5 (2014), Nr. 3, S. 1022–1023. <http://dx.doi.org/10.1109/TSTE.2014.2323852>. – DOI 10.1109/TSTE.2014.2323852. – ISSN 1949–3029
- [130] KHOSRAVI, Abbas ; NAHAVANDI, Saeid: Load Forecasting Using Interval Type-2 Fuzzy Logic Systems: Optimal Type Reduction. In: *IEEE Transactions on Industrial Informatics* 10 (2014), Nr. 2, S. 1055–1063. <http://dx.doi.org/10.1109/TII.2013.2285650>. – DOI 10.1109/TII.2013.2285650. – ISSN 1551–3203
- [131] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug: Constructing prediction intervals for neural network metamodels of complex systems. In: *International Joint Conference on Neural Networks, 2009*. Piscataway, NJ : IEEE, 2009. – ISBN 978–1–4244–3548–7, S. 1576–1582
- [132] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug: Construction of Optimal Prediction Intervals for Load Forecasting Problems. In: *IEEE Transactions on Power Systems* 25 (2010), Nr. 3, S. 1496–1503. <http://dx.doi.org/10.1109/TPWRS.2010.2042309>. – DOI 10.1109/TPWRS.2010.2042309. – ISSN 08858950
- [133] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug: Load Forecasting and Neural Networks: A Prediction Interval-Based Perspective. Version:2010. http://dx.doi.org/10.1007/978-3-642-14013-6_5. In: KACPRZYK, Janusz (Hrsg.) ; PANIGRAHI, Bijaya K. (Hrsg.) ; ABRAHAM, Ajith (Hrsg.) ; DAS, Swagatam (Hrsg.): *Computational Intelligence in Power Engineering* Bd. 302. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010. – DOI 10.1007/978–3–642–14013–6_5. – ISBN 978–3–642–14012–9, S. 131–150
- [134] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug: A prediction interval-based approach to determine optimal structures of neural network metamodels. In: *Expert Systems with Applications* 37 (2010), Nr. 3, S. 2377–2387. <http://dx.doi.org/10.1016/j.eswa.2009.07.059>. – DOI 10.1016/j.eswa.2009.07.059. – ISSN 09574174
- [135] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug ; ATIYA, Amir F.: Comprehensive review of neural network-based prediction intervals and new advances. In: *IEEE transactions on neural networks* 22 (2011), Nr. 9, S. 1341–1356. <http://dx.doi.org/10.1109/TNN.2011.2394444>. – DOI 10.1109/TNN.2011.2394444. – ISSN 1045-9225

- [//dx.doi.org/10.1109/TNN.2011.2162110](http://dx.doi.org/10.1109/TNN.2011.2162110). – DOI 10.1109/TNN.2011.2162110. – ISSN 1045–9227
- [136] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug ; ATIYA, Amir F.: Lower upper bound estimation method for construction of neural network-based prediction intervals. In: *IEEE transactions on neural networks* 22 (2011), Nr. 3, S. 337–346. <http://dx.doi.org/10.1109/TNN.2010.2096824>. – DOI 10.1109/TNN.2010.2096824. – ISSN 1045–9227
- [137] KHOSRAVI, Abbas ; NAHAVANDI, Saeid ; CREIGHTON, Doug ; SRINIVASAN, Dipti: Interval Type-2 Fuzzy Logic Systems for Load Forecasting: A Comparative Study. In: *IEEE Transactions on Power Systems* 27 (2012), Nr. 3, S. 1274–1282. <http://dx.doi.org/10.1109/TPWRS.2011.2181981>. – DOI 10.1109/TPWRS.2011.2181981. – ISSN 08858950
- [138] KHUNTIA, Swasti R. ; RUEDA, José L. ; VAN DER MEIJDEN, Mart A.: Forecasting the load of electrical power systems in mid- and long-term horizons: A review. In: *IET Generation, Transmission & Distribution* 10 (2016), Nr. 16, S. 3971–3977. <http://dx.doi.org/10.1049/iet-gtd.2016.0340>. – DOI 10.1049/iet-gtd.2016.0340. – ISSN 1751–8687
- [139] KIM, Seung-Jun ; GIANNAKIS, Georgios B.: Scalable and Robust Demand Response With Mixed-Integer Constraints. In: *IEEE Transactions on Smart Grid* 4 (2013), Nr. 4, S. 2089–2099. <http://dx.doi.org/10.1109/TSG.2013.2257893>. – DOI 10.1109/TSG.2013.2257893. – ISSN 1949–3053
- [140] KIRKPATRICK, S. ; GELATT, C.D. ; VECCHI, M.P.: Optimization by Simulated Annealing. Version:1987. <http://dx.doi.org/10.1016/B978-0-08-051581-6.50059-3>. In: *Readings in Computer Vision Issues, Problem, Principles, and Paradigms*. 1987. – DOI 10.1016/B978-0-08-051581-6.50059-3. – ISBN 9780080515816, S. 606–615
- [141] KOECHLI, C. ; FUSSELL, B. K. ; PRINA, S. R. ; JAMES, D. A. ; PERRIARD, Y.: Design optimization of induction motors for aerospace applications. In: *Conference record of the 2004 IEEE industry applications conference*. Piscataway, N.J. : IEEE, 2004. – ISBN 0–7803–8486–5, S. 2501–2505
- [142] KOOHI-KAMALI, Sam ; TYAGI, V. V. ; RAHIM, N. A. ; PANWAR, N. L. ; MOKHLIS, H.: Emergence of energy storage technologies as the solution for reliable operation of smart power systems: A review. In: *Renewable and Sustainable Energy Reviews* 25 (2013), S. 135–165. <http://dx.doi.org/10.1016/j.rser.2013.03.056>. – DOI 10.1016/j.rser.2013.03.056. – ISSN 13640321

- [143] KOPANOS, Georgios M. ; PISTIKOPOULOS, Efstratios N.: Reactive Scheduling by a Multiparametric Programming Rolling Horizon Framework: A Case of a Network of Combined Heat and Power Units. In: *Industrial & engineering chemistry research* 53 (2014), Nr. 11, S. 4366–4386. <http://dx.doi.org/10.1021/ie402393s>. – DOI 10.1021/ie402393s. – ISSN 0888–5885
- [144] KOSTOV, K. S. ; KYIRA, J. J.: Genetic algorithm optimization of peak current mode controlled buck conve. In: MARTIKAINEN, Jarno (Hrsg.): *SMCia05*. Piscataway New Jersey : IEEE, 2005. – ISBN 0–7803–8942–5, S. 111–116
- [145] KUZNETSOVA, Elizaveta: *Microgrid agent-based modelling and optimization under uncertainty*. France, Université de Versailles Saint-Quentin-en-Yvelines, Ph.D., March 2014
- [146] KUZNETSOVA, Elizaveta ; LI, Yan-Fu ; RUIZ, Carlos ; ZIO, Enrico: An integrated framework of agent-based modelling and robust optimization for microgrid energy management. In: *Applied Energy* 129 (2014), S. 70–88. <http://dx.doi.org/10.1016/j.apenergy.2014.04.024>. – DOI 10.1016/j.apenergy.2014.04.024. – ISSN 03062619
- [147] KUZNETSOVA, Elizaveta ; RUIZ, Carlos ; LI, Yan-Fu ; ZIO, Enrico: Analysis of robust optimization for decentralized microgrid energy management under uncertainty. In: *International Journal of Electrical Power & Energy Systems* 64 (2015), S. 815–832. <http://dx.doi.org/10.1016/j.ijepes.2014.07.064>. – DOI 10.1016/j.ijepes.2014.07.064. – ISSN 01420615
- [148] LAHAIE, Sébastien: *How to take the Dual of a Linear Program*. <http://slahaie.net/docs/lpdual.pdf>
- [149] LAMEDICA, Regina ; RUVIO, Alessandro ; GALDI, Vincenzo ; GRABER, Giuseppe ; SFORZA, Patrizia ; GUIDIBUFFARINI, Guido ; SPALVIERI, Claudio: Application of battery auxiliary substations in 3kV railway systems. In: *2015 AEIT International Annual Conference (AEIT)*, IEEE, 2015. – ISBN 978–8–8872–3728–3, S. 1–6
- [150] LATKOVSKIS, Leonards ; GRIGANS, Linards: Estimation of the untapped regenerative braking energy in urban electric transportation network. In: *2008 13th International Power Electronics and Motion Control Conference*, IEEE, 2008. – ISBN 978–1–4244–1741–4, S. 2066–2070
- [151] LEIRAS, A. ; HAMACHER, S. ; ELKAMEL, A.: Petroleum refinery operational planning using robust optimization. In: *Engineering Optimization* 42 (2010), Nr. 12, S. 1119–1131. <http://dx.doi.org/10.1080/03052151003686724>. – DOI 10.1080/03052151003686724

- [152] LEON-ALDACO, Susana E. ; CALLEJA, Hugo ; AGUAYO ALQUICIRA, Jesus: Meta-heuristic Optimization Methods Applied to Power Converters: A Review. In: *IEEE Transactions on Power Electronics* 30 (2015), Nr. 12, S. 6791–6803. <http://dx.doi.org/10.1109/TPEL.2015.2397311>. – DOI 10.1109/TPEL.2015.2397311. – ISSN 0885–8993
- [153] LESEL, Jonathan ; CLAISSE, Gautier ; DEBAY, Patrick ; ROBYNS, Benoit: Design of daily energy optimal timetables for metro lines using metaheuristics. In: *2016 18th Mediterranean Electrotechnical Conference (MELECON)*, IEEE, 2016. – ISBN 978–1–5090–0058–6, S. 1–6
- [154] LHOMME, W. ; DELARUE, P. ; BARRADE, P. ; BOUSCAYROL, A. ; RUFER, A.: Design and control of a supercapacitor storage system for traction applications. In: *Fourtieth IAS Annual Meeting. Conference Record of the 2005 Industry Applications Conference, 2005*, IEEE, 2005. – ISBN 0–7803–9208–6, S. 2013–2020
- [155] LI, Zukui ; DING, Ran ; FLOUDAS, Christodoulos A.: A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: I. Robust Linear Optimization and Robust Mixed Integer Linear Optimization. In: *Industrial & engineering chemistry research* 50 (2011), Nr. 18, S. 10567–10603. <http://dx.doi.org/10.1021/ie200150p>. – DOI 10.1021/ie200150p. – ISSN 0888–5885
- [156] LI, Zukui ; FLOUDAS, Christodoulos A.: A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: II. Probabilistic Guarantees on Constraint Satisfaction. In: *Industrial & engineering chemistry research* 51 (2012), Nr. 19, S. 6769–6788. <http://dx.doi.org/10.1021/ie201651s>. – DOI 10.1021/ie201651s. – ISSN 0888–5885
- [157] LI, Zukui ; FLOUDAS, Christodoulos A.: A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: III. Improving the Quality of Robust Solutions. In: *Industrial & engineering chemistry research* 53 (2014), Nr. 33, S. 13112–13124. <http://dx.doi.org/10.1021/ie501898n>. – DOI 10.1021/ie501898n. – ISSN 0888–5885
- [158] LIAN, Cheng ; ZENG, Zhigang ; YAO, Wei ; TANG, Huiming ; CHEN, Chun Lung P.: Landslide Displacement Prediction With Uncertainty Based on Neural Networks With Random Hidden Weights. In: *IEEE transactions on neural networks and learning systems* 27 (2016), Nr. 12, S. 2683–2695. <http://dx.doi.org/10.1109/TNNLS.2015.2512283>. – DOI 10.1109/TNNLS.2015.2512283. – ISSN 2162–2388

- [159] LIEBREICH, Michael: *London summit 2017: Breaking Clean: Keynote: State of the Industry*. <https://about.bnef.com/blog/michael-liebreich-state-industry-keynote-bnef-emea-summit-2017/>
- [160] LIN, Fei ; LIU, Shihui ; YANG, Zhihong ; ZHAO, Yingying ; YANG, Zhongping ; SUN, Hu: Multi-Train Energy Saving for Maximum Usage of Regenerative Energy by Dwell Time Optimization in Urban Rail Transit Using Genetic Algorithm. In: *Energies* 9 (2016), Nr. 3, S. 208. <http://dx.doi.org/10.3390/en9030208>. – DOI 10.3390/en9030208. – ISSN 1996–1073
- [161] LIU, Wan-Yu ; LIN, Chun-Cheng ; YEH, Tzu-Lei: Supply chain optimization of forest biomass electricity and bioethanol coproduction. In: *Energy* 139 (2017), S. 630–645. <http://dx.doi.org/10.1016/j.energy.2017.08.018>. – DOI 10.1016/j.energy.2017.08.018. – ISSN 03605442
- [162] LÓPEZ-LÓPEZ, A. J. ; PECHARROMÁN, R. R. ; GARCÍA-MATOS, J. A. ; FERNÁNDEZ-CARDADOR, A. ; CUCALA, A. P.: Optimal Deployment Of Energy Storage Systems In A DC-electrified Railway System. In: *WIT Transactions on The Built Environment* 127 (2012), 603–614. <https://www.witpress.com/elibrary/wit-transactions-on-the-built-environment/127/23895>
- [163] LORCA, Alvaro ; SUN, Xu A.: Adaptive Robust Optimization With Dynamic Uncertainty Sets for Multi-Period Economic Dispatch Under Significant Wind. In: *IEEE Transactions on Power Systems* 30 (2015), Nr. 4, S. 1702–1713. <http://dx.doi.org/10.1109/TPWRS.2014.2357714>. – DOI 10.1109/TPWRS.2014.2357714. – ISSN 08858950
- [164] LUO, Xing ; WANG, Jihong ; DOONER, Mark ; CLARKE, Jonathan: Overview of current development in electrical energy storage technologies and the application potential in power system operation. In: *Applied Energy* 137 (2015), S. 511–536. <http://dx.doi.org/10.1016/j.apenergy.2014.09.081>. – DOI 10.1016/j.apenergy.2014.09.081. – ISSN 03062619
- [165] MALAVASI, G. ; PALLESCHI, P. ; RICCI, S.: Driving and operation strategies for traction-energy saving in mass rapid transit systems. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 225 (2011), Nr. 5, S. 475–482. <http://dx.doi.org/10.1177/2041301710395077>. – DOI 10.1177/2041301710395077. – ISSN 0954–4097
- [166] MALYSZ, Pawel ; SIROUSPOUR, Shahin ; EMADI, Ali: An Optimal Energy Storage Control Strategy for Grid-connected Microgrids. In: *IEEE Transactions on Smart Grid*

- 5 (2014), Nr. 4, S. 1785–1796. <http://dx.doi.org/10.1109/TSG.2014.2302396>. – DOI 10.1109/TSG.2014.2302396. – ISSN 1949–3053
- [167] MBAMALU, G. A. N. ; HAWARY, M. E. E.: Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation. In: *IEEE Transactions on Power Systems* 8 (1993), Nr. 1, 343–348. <http://dx.doi.org/10.1109/59.221222>. – DOI 10.1109/59.221222. – ISSN 08858950
- [168] McDONALD, Alan ; SCHRATTENHOLZER, Leo: Learning rates for energy technologies. In: *Energy Policy* 29 (2001), Nr. 4, S. 255–261. [http://dx.doi.org/10.1016/S0301-4215\(00\)00122-1](http://dx.doi.org/10.1016/S0301-4215(00)00122-1). – DOI 10.1016/S0301-4215(00)00122-1. – ISSN 03014215
- [169] MEGHWANI, Suraj S. ; THAKUR, Manoj: Multi-criteria algorithms for portfolio optimization under practical constraints. In: *Swarm and Evolutionary Computation* (2017). <http://dx.doi.org/10.1016/j.swevo.2017.06.005>. – DOI 10.1016/j.swevo.2017.06.005. – ISSN 22106502
- [170] MEINERT, Michael ; RECHENBERG, Karsten ; HEIN, Gerd: *Energy efficient solutions for the complete railway system*. http://uic.org/cdrom/2008/11_wcrr2008/pdf/PS.2.32.pdf
- [171] MELLITT, B. ; MOUNEIMNE, Z. S. ; GOODMAN, C. J.: Simulation study of DC transit systems with inverting substations. In: *IEE Proceedings B Electric Power Applications* 131 (1984), Nr. 2, S. 38. <http://dx.doi.org/10.1049/ip-b.1984.0008>. – DOI 10.1049/ip-b.1984.0008. – ISSN 01437038
- [172] MOKHTARI, Hadi ; HASANI, Aliakbar: An energy-efficient multi-objective optimization for flexible job-shop scheduling problem. In: *Computers & Chemical Engineering* 104 (2017), S. 339–352. <http://dx.doi.org/10.1016/j.compchemeng.2017.05.004>. – DOI 10.1016/j.compchemeng.2017.05.004. – ISSN 00981354
- [173] MUHLETHALER, Jonas ; SCHWEIZER, Mario ; BLATTMANN, Robert ; KOLAR, Johann W. ; ECKLEBE, Andreas: Optimal Design of LCL Harmonic Filters for Three-Phase PFC Rectifiers. In: *IEEE Transactions on Power Electronics* 28 (2013), Nr. 7, S. 3114–3125. <http://dx.doi.org/10.1109/TPEL.2012.2225641>. – DOI 10.1109/TPEL.2012.2225641. – ISSN 0885–8993
- [174] MURTY, Katta G.: *Linear programming*. N Y : Wiley, 1983. – ISBN 978–0–471–09725–9
- [175] MUSOLINO, V. ; PIEVATOLO, A. ; TIRONI, E.: A statistical approach to electrical storage sizing with application to the recovery of braking energy. In: *Energy* 36 (2011),

- Nr. 11, S. 6697–6704. <http://dx.doi.org/10.1016/j.energy.2011.07.037>. – DOI 10.1016/j.energy.2011.07.037. – ISSN 03605442
- [176] NAFISI, Hamed ; AGAH, Seyed Mohammad M. ; ASKARIAN ABYANEH, Hossien ; ABEDI, Mehrdad: Two-Stage Optimization Method for Energy Loss Minimization in Microgrid Based on Smart Power Management Scheme of PHEVs. In: *IEEE Transactions on Smart Grid* 7 (2016), Nr. 3, S. 1268–1276. <http://dx.doi.org/10.1109/TSG.2015.2480999>. – DOI 10.1109/TSG.2015.2480999. – ISSN 1949–3053
- [177] NAGARAJA, Y. ; DEVARAJU, T. ; KUMAR, M. V. ; MADICHETTY, Sreedhar: *A survey on wind energy, load and price forecasting: (Forecasting methods)*. Piscataway, NJ : IEEE, 2016. <http://dx.doi.org/10.1109/ICEEOT.2016.7754792>. <http://dx.doi.org/10.1109/ICEEOT.2016.7754792>. – ISBN 9781467399401
- [178] NGUYEN, D. ; WIDROW, B.: Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In: *1990 IJCNN International Joint Conference on Neural Networks*, IEEE, 1990, S. 21–26 vol.3
- [179] NGUYEN, Duc Cong H. ; ASCOUGH, James C. ; MAIER, Holger R. ; DANDY, Graeme C. ; ANDALES, Allan A.: Optimization of irrigation scheduling using ant colony algorithms and an advanced cropping system model. In: *Environmental Modelling & Software* 97 (2017), S. 32–45. <http://dx.doi.org/10.1016/j.envsoft.2017.07.002>. – DOI 10.1016/j.envsoft.2017.07.002. – ISSN 13648152
- [180] OGURA, Koki ; NISHIMURA, Kazuya ; MATSUMURA, Takahiro ; TONDA, Chiyoharu ; YOSHIYAMA, Eiji ; ANDRIANI, Maurice ; FRANCIS, Willard ; SCHMITT, Robert A. ; VISGOTIS, Anthony ; GIANFRANCESCO, Nicholas: Test Results of a High Capacity Wayside Energy Storage System Using Ni-MH Batteries for DC Electric Railway at New York City Transit. In: *2011 IEEE Green Technologies Conference (IEEE-Green)*, IEEE, 2011. – ISBN 978–1–61284–713–9, S. 1–6
- [181] OMAR, Abdul Malek S. ; SAMAT, Ahmad Asri A. ; ISA, Siti Sarah M. ; SHAMSUDDIN, Sarah A. ; JAMALUDIN, Nur F. ; KHYASUDEEN, Muhammad F.: New model of inverting substation for DC traction with regenerative braking system, Author(s), 2017 (AIP Conference Proceedings), S. 030016
- [182] PAVELKA, A. ; PROCHAZKA, A.: Algorithms for initialization of neural network weights. In: *Sborník příspěvků 12th rocní konference MATLAB 2* (2004), S. 453–459
- [183] PEÑA-ALCARAZ, Maite ; FERNÁNDEZ, Antonio ; CUCALA, Asunción P. ; RAMOS, Andres ; PECHARROMÁN, Ramon R.: Optimal underground timetable design based on power flow for maximizing the use of regenerative-braking energy. In: *Proceedings*

- of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 226 (2011), Nr. 4, S. 397–408. <http://dx.doi.org/10.1177/0954409711429411>. – DOI 10.1177/0954409711429411. – ISSN 0954–4097
- [184] PITCHAN, Mohan K. ; BHOWMIK, Shantanu ; BALACHANDRAN, Meera ; ABRAHAM, M.: Process optimization of functionalized MWCNT/polyetherimide nanocomposites for aerospace application. In: *Materials & Design* 127 (2017), S. 193–203. <http://dx.doi.org/10.1016/j.matdes.2017.04.081>. – DOI 10.1016/j.matdes.2017.04.081. – ISSN 02641275
- [185] PSALTIS, Andreas ; SINOQUET, Delphine ; PAGOT, Alexandre: Systematic optimization methodology for heat exchanger network and simultaneous process design. In: *Computers & Chemical Engineering* 95 (2016), S. 146–160. <http://dx.doi.org/10.1016/j.compchemeng.2016.09.013>. – DOI 10.1016/j.compchemeng.2016.09.013. – ISSN 00981354
- [186] QIU, Xin ; NGUYEN, Tu A. ; CROW, Mariesa L.: Heterogeneous Energy Storage Optimization for Microgrids. In: *IEEE Transactions on Smart Grid* 7 (2016), Nr. 3, S. 1453–1461. <http://dx.doi.org/10.1109/TSG.2015.2461134>. – DOI 10.1109/TSG.2015.2461134. – ISSN 1949–3053
- [187] QUAN, Hao ; SRINIVASAN, Dipti ; KHOSRAVI, Abbas: Construction of neural network-based prediction intervals using particle swarm optimization. In: *The 2012 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2012. – ISBN 978–1–4673–1490–9, S. 1–7
- [188] QUAN, Hao ; SRINIVASAN, Dipti ; KHOSRAVI, Abbas: Short-term load and wind power forecasting using neural network-based prediction intervals. In: *IEEE transactions on neural networks and learning systems* 25 (2014), Nr. 2, S. 303–315. <http://dx.doi.org/10.1109/TNNLS.2013.2276053>. – DOI 10.1109/TNNLS.2013.2276053. – ISSN 2162–2388
- [189] QUAN, Hao ; SRINIVASAN, Dipti ; KHOSRAVI, Abbas: Incorporating Wind Power Forecast Uncertainties Into Stochastic Unit Commitment Using Neural Network-Based Prediction Intervals. In: *IEEE transactions on neural networks and learning systems* 26 (2015), Nr. 9, S. 2123–2135. <http://dx.doi.org/10.1109/TNNLS.2014.2376696>. – DOI 10.1109/TNNLS.2014.2376696. – ISSN 2162–2388
- [190] RANA, Mashud ; KOPRINSKA, Irena ; KHOSRAVI, Abbas ; AGELIDIS, Vassilios G.: Prediction intervals for electricity load forecasting using neural networks. In: *The 2013 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2013. – ISBN 978–1–4673–6129–3, S. 1–8

- [191] RAO, R. V. ; MORE, K. C.: Design optimization and analysis of selected thermal devices using self-adaptive Jaya algorithm. In: *Energy Conversion and Management* 140 (2017), S. 24–35. <http://dx.doi.org/10.1016/j.enconman.2017.02.068>. – DOI 10.1016/j.enconman.2017.02.068. – ISSN 01968904
- [192] RAO, R. V. ; RAI, Dhiraj P. ; BALIC, Joze: A multi-objective algorithm for optimization of modern machining processes. In: *Engineering Applications of Artificial Intelligence* 61 (2017), S. 103–125. <http://dx.doi.org/10.1016/j.engappai.2017.03.001>. – DOI 10.1016/j.engappai.2017.03.001. – ISSN 09521976
- [193] RAO, R. V. ; SAROJ, Ankit: Economic optimization of shell-and-tube heat exchanger using Jaya algorithm with maintenance consideration. In: *Applied Thermal Engineering* 116 (2017), S. 473–487. <http://dx.doi.org/10.1016/j.applthermaleng.2017.01.071>. – DOI 10.1016/j.applthermaleng.2017.01.071. – ISSN 13594311
- [194] RATNIYOMCHAI, Tosaphol ; TRICOLI, Pietro ; HILLMANSEN, Stuart: Recent developments and applications of energy storage devices in electrified railways. In: *IET Electrical Systems in Transportation* 4 (2014), Nr. 1, S. 9–20. <http://dx.doi.org/10.1049/iet-est.2013.0031>. – DOI 10.1049/iet-est.2013.0031. – ISSN 2042–9738
- [195] REZVAN, A. T. ; GHARNEH, N. S. ; GHAREHPETIAN, G. B.: Robust optimization of distributed generation investment in buildings. In: *Energy* 48 (2012), Nr. 1, S. 455–463. <http://dx.doi.org/10.1016/j.energy.2012.10.011>. – DOI 10.1016/j.energy.2012.10.011. – ISSN 03605442
- [196] RUFER, A. ; HOTELLIER, D. ; BARRADE, P.: A Supercapacitor-Based Energy Storage Substation for Voltage Compensation in Weak Transportation Networks. In: *IEEE Transactions on Power Delivery* 19 (2004), Nr. 2, S. 629–636. <http://dx.doi.org/10.1109/TPWRD.2004.824408>. – DOI 10.1109/TPWRD.2004.824408. – ISSN 0885–8977
- [197] RUPP, A. ; BAIER, H. ; MERTINY, P. ; SECANELL, M.: Analysis of a flywheel energy storage system for light rail transit. In: *Energy* 107 (2016), S. 625–638. <http://dx.doi.org/10.1016/j.energy.2016.04.051>. – DOI 10.1016/j.energy.2016.04.051. – ISSN 03605442
- [198] SAHA, Promotes ; KSAIBATI, Khaled: An optimization model for improving highway safety. In: *Journal of Traffic and Transportation Engineering (English Edition)* 3 (2016), Nr. 6, S. 549–558. <http://dx.doi.org/10.1016/j.jtte.2016.01.004>. – DOI 10.1016/j.jtte.2016.01.004. – ISSN 20957564

- [199] SALOMON, Shaul ; AVIGAD, Gideon ; FLEMING, Peter J. ; PURSHOUSE, Robin C.: Active robust optimization: Enhancing robustness to uncertain environments. In: *IEEE transactions on cybernetics* 44 (2014), Nr. 11, S. 2221–2231. <http://dx.doi.org/10.1109/TCYB.2014.2304475>. – DOI 10.1109/TCYB.2014.2304475. – ISSN 2168–2275
- [200] SAMPAT, Apoorva M. ; MARTIN, Edgar ; MARTIN, Mariano ; ZAVALA, Victor M.: Optimization formulations for multi-product supply chain networks. In: *Computers & Chemical Engineering* 104 (2017), S. 296–310. <http://dx.doi.org/10.1016/j.compchemeng.2017.04.021>. – DOI 10.1016/j.compchemeng.2017.04.021. – ISSN 00981354
- [201] SHABANZADEH, Morteza ; FATTAHI, Mohammad: Generation Maintenance Scheduling via robust optimization. In: *2015 23rd Iranian Conference on Electrical Engineering*, IEEE, 2015. – ISBN 978–1–4799–1972–7, S. 1504–1509
- [202] SHABANZADEH, Morteza ; SHEIKH-EL-ESLAMI, Mohammad-Kazem ; HAGHIFAM, Mahmoud-Reza: The design of a risk-hedging tool for virtual power plants via robust optimization approach. In: *Applied Energy* 155 (2015), S. 766–777. <http://dx.doi.org/10.1016/j.apenergy.2015.06.059>. – DOI 10.1016/j.apenergy.2015.06.059. – ISSN 03062619
- [203] SHAHZADEH, Abbas ; KHOSRAVI, Abbas ; NAHAVANDI, Saeid: Improving load forecast accuracy by clustering consumers using smart meter data. In: *2015 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2015. – ISBN 978–1–4799–1960–4, S. 1–7
- [204] SHEN, Yanxia ; LU, Xin ; YU, Xinyan ; ZHAO, Zhipu ; WU, Dinghui: Short-term wind power intervals prediction based on generalized morphological filter and artificial bee colony neural network. In: *2016 35th Chinese Control Conference (CCC)*, IEEE, 2016. – ISBN 978–9–8815–6391–0, S. 8501–8506
- [205] SILVENTE, Javier ; KOPANOS, Georgios M. ; PISTIKOPOULOS, Efstratios N. ; ESPUÑA, Antonio: A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids. In: *Applied Energy* 155 (2015), S. 485–501. <http://dx.doi.org/10.1016/j.apenergy.2015.05.090>. – DOI 10.1016/j.apenergy.2015.05.090. – ISSN 03062619
- [206] SIRAJ, M. M. ; VAN DEN HOF, Paul M. ; JANSEN, Jan D.: Robust optimization of water-flooding in oil reservoirs using risk management tools. In: *IFAC-PapersOnLine* 49 (2016), Nr. 7, S. 133–138. <http://dx.doi.org/10.1016/j.ifacol.2016.07.229>. – DOI 10.1016/j.ifacol.2016.07.229. – ISSN 24058963

- [207] SOYSTER, A. L.: Technical Note—Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. In: *Operations Research* 21 (1973), Nr. 5, S. 1154–1157. <http://dx.doi.org/10.1287/opre.21.5.1154>. – DOI 10.1287/opre.21.5.1154
- [208] STEINER, M. ; SCHOLTEN, J.: Energy storage on board of railway vehicles. In: *2005 European Conference on Power Electronics and Applications*, IEEE, 2005. – ISBN 90–75815–09–3, S. 10 pp–P.10
- [209] STEINER, Michael ; KLOHR, Markus ; PAGIELA, Stanislaus: Energy storage system with ultracaps on board of railway vehicles. In: *2007 European Conference on Power Electronics and Applications*, IEEE, 2007, S. 1–10
- [210] STRANO, M.: Optimization under uncertainty of sheet-metal-forming processes by the finite element method. In: *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 220 (2006), Nr. 8, S. 1305–1315. <http://dx.doi.org/10.1243/09544054JEM480>. – DOI 10.1243/09544054JEM480. – ISSN 0954–4054
- [211] SUOMINEN, Olli ; MÖRSKY, Ville ; RITALA, Risto ; VILKKO, Matti: Framework for optimization and scheduling of a copper production plant. Version:2016. <http://dx.doi.org/10.1016/B978-0-444-63428-3.50212-5>. In: KRAVANJA, Z. (Hrsg.): *26th European Symposium on Computer Aided Process Engineering* Bd. 38. Amsterdam : Elsevier, 2016. – DOI 10.1016/B978-0-444-63428-3.50212-5. – ISBN 9780444634283, S. 1243–1248
- [212] SUTANTO, D. ; CHENG, K.W.E.: Superconducting magnetic energy storage systems for power system applications. In: *2009 International Conference on Applied Superconductivity and Electromagnetic Devices*, IEEE, 2009. – ISBN 978–1–4244–3686–6, S. 377–380
- [213] SUZUKI, S. ; BABA, J. ; SHUTOH, K. ; MASADA, E.: Effective Application of Superconducting Magnetic Energy Storage (SMES) to Load Leveling for High Speed Transportation System. In: *IEEE Transactions on Applied Superconductivity* 14 (2004), Nr. 2, S. 713–716. <http://dx.doi.org/10.1109/TASC.2004.830082>. – DOI 10.1109/TASC.2004.830082. – ISSN 1051–8223
- [214] TAORMINA, Riccardo ; CHAU, Kwok-Wing: ANN-based interval forecasting of stream-flow discharges using the LUBE method and MOFIPS. In: *Engineering Applications of Artificial Intelligence* 45 (2015), S. 429–440. <http://dx.doi.org/10.1016/j.engappai.2015.07.019>. – DOI 10.1016/j.engappai.2015.07.019. – ISSN 09521976

- [215] TAYLOR, James W. ; MCSHARRY, Patrick E.: Short-Term Load Forecasting Methods: An Evaluation Based on European Data. In: *IEEE Transactions on Power Systems* 22 (2007), Nr. 4, 2213–2219. <http://dx.doi.org/10.1109/TPWRS.2007.907583>. – DOI 10.1109/TPWRS.2007.907583. – ISSN 08858950
- [216] THE "TICKET TO KYOTO" PROJECT: *WP2B ENERGY RECOVERY: Guidelines for braking energy recovery systems in urban rail networks*
- [217] THE FEDERAL MINISTRY FOR ECONOMIC AFFAIRS AND ENERGY ; BMWi (Hrsg.): *Fifth "Energy Transition" Monitoring Report: The Energy of the Future: 2015 Reporting Year*. Berlin,
- [218] TIAN, Peigen ; XIAO, Xi ; WANG, Kui ; DING, Ruoxing: A Hierarchical Energy Management System Based on Hierarchical Optimization for Microgrid Community Economic Operation. In: *IEEE Transactions on Smart Grid* 7 (2016), Nr. 5, S. 2230–2241. <http://dx.doi.org/10.1109/TSG.2015.2470551>. – DOI 10.1109/TSG.2015.2470551. – ISSN 1949–3053
- [219] TSEKOURAS, G. J. ; KANELLOS, F. D. ; MASTORAKIS, N.: Short Term Load Forecasting in Electric Power Systems with Artificial Neural Networks. Version:2015. http://dx.doi.org/10.1007/978-3-319-15765-8_{_}2. In: MASTORAKIS, Nikos (Hrsg.) ; BULUCEA, Aida (Hrsg.) ; TSEKOURAS, George (Hrsg.): *Computational Problems in Science and Engineering* Bd. 343. Cham : Springer International Publishing, 2015. – DOI 10.1007/978-3-319-15765-8_2. – ISBN 978-3-319-15764-1, S. 19–58
- [220] VERBAND DEUTSCHER VERKEHRSUNTERNEHMEN: *VDV-Schrift 507: Aufbau und Schutzmaßnahmen von elektrischen Energieanlagen an Strecken von Gleichstrom-Nahverkehrsbahnen*. 2005
- [221] WANG, Bin ; YANG, Zhongping ; LIN, Fei ; ZHAO, Wei: An Improved Genetic Algorithm for Optimal Stationary Energy Storage System Locating and Sizing. In: *Energies* 7 (2014), Nr. 10, S. 6434–6458. <http://dx.doi.org/10.3390/en7106434>. – DOI 10.3390/en7106434. – ISSN 1996–1073
- [222] WANG, J. ; GRIFFO, A.: Design optimization of passive DC filters for aerospace applications. In: ALEXANDER, Geoff (Hrsg.): *The nonprofit survival guide*. Jefferson, North Carolina : McFarland & Company, 2015. – ISBN 978, S. 314
- [223] WANG, Ran ; WANG, Ping ; XIAO, Gaoxi: A robust optimization approach for energy generation scheduling in microgrids. In: *Energy Conversion and Management* 106 (2015), S. 597–607. <http://dx.doi.org/10.1016/j.enconman.2015.09.066>. – DOI 10.1016/j.enconman.2015.09.066. – ISSN 01968904

- [224] WANG, Suqin ; JIA, Cuiling: Prediction Intervals for Short-Term Photovoltaic Generation Forecasts. In: *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, IEEE, 2015. – ISBN 978-1-4673-7723-2, S. 459–463
- [225] WANG, Zhaoyu ; CHEN, Bokan ; WANG, Jianhui ; KIM, Jinho ; BEGOVIC, Miroslav M.: Robust Optimization Based Optimal DG Placement in Microgrids. In: *IEEE Transactions on Smart Grid* 5 (2014), Nr. 5, S. 2173–2182. <http://dx.doi.org/10.1109/TSG.2014.2321748>. – DOI 10.1109/TSG.2014.2321748. – ISSN 1949–3053
- [226] WARIN, Y. ; LANSELLE, R. ; THIOUNN, M.: Active substation. In: *9th World Congress on Railway Research - WCRR 2011* (Lille, France. 2011)
- [227] WILLIS, H. L. ; NORTHCOTE-GREEN, J.E.D.: Spatial electric load forecasting: A tutorial review. In: *Proceedings of the IEEE* 71 (1983), Nr. 2, S. 232–253. <http://dx.doi.org/10.1109/PROC.1983.12562>. – DOI 10.1109/PROC.1983.12562. – ISSN 0018–9219
- [228] XIA, Huan ; CHEN, Huaixin ; YANG, Zhongping ; LIN, Fei ; WANG, Bin: Optimal Energy Management, Location and Size for Stationary Energy Storage System in a Metro Line Based on Genetic Algorithm. In: *Energies* 8 (2015), Nr. 10, S. 11618–11640. <http://dx.doi.org/10.3390/en81011618>. – DOI 10.3390/en81011618. – ISSN 1996–1073
- [229] XIDONAS, Panos ; HASSAPIS, Christis ; SOULIS, John ; SAMITAS, Aristeidis: Robust minimum variance portfolio optimization modelling under scenario uncertainty. In: *Economic Modelling* 64 (2017), S. 60–71. <http://dx.doi.org/10.1016/j.econmod.2017.03.020>. – DOI 10.1016/j.econmod.2017.03.020. – ISSN 02649993
- [230] XIDONAS, Panos ; MAVROTAS, George ; HASSAPIS, Christis ; ZOPOUNIDIS, Constantin: Robust multiobjective portfolio optimization: A minimax regret approach. In: *European Journal of Operational Research* 262 (2017), Nr. 1, S. 299–305. <http://dx.doi.org/10.1016/j.ejor.2017.03.041>. – DOI 10.1016/j.ejor.2017.03.041. – ISSN 03772217
- [231] XU, Fang ; CHEN, Hong ; GONG, Xun ; MEI, Qin: Fast Nonlinear Model Predictive Control on FPGA Using Particle Swarm Optimization. In: *IEEE Transactions on Industrial Electronics* 63 (2016), Nr. 1, S. 310–321. <http://dx.doi.org/10.1109/TIE.2015.2464171>. – DOI 10.1109/TIE.2015.2464171. – ISSN 0278–0046
- [232] YANG, Zhitian ; MA, Zhiliang ; WU, Song: Optimized flowshop scheduling of multiple production lines for precast production. In: *Automation in Construction* 72

- (2016), S. 321–329. <http://dx.doi.org/10.1016/j.autcon.2016.08.021>. – DOI 10.1016/j.autcon.2016.08.021. – ISSN 09265805
- [233] YU, Chian-Son ; LI, Han-Lin: A robust optimization model for stochastic logistic problems. In: *International Journal of Production Economics* 64 (2000), Nr. 1-3, S. 385–397. [http://dx.doi.org/10.1016/S0925-5273\(99\)00074-2](http://dx.doi.org/10.1016/S0925-5273(99)00074-2). – DOI 10.1016/S0925-5273(99)00074-2. – ISSN 09255273
- [234] YUAN, Yuan ; LI, Zukui ; HUANG, Biao: Robust optimization under correlated uncertainty: Formulations and computational study. In: *Computers & Chemical Engineering* 85 (2016), S. 58–71. <http://dx.doi.org/10.1016/j.compchemeng.2015.10.017>. – DOI 10.1016/j.compchemeng.2015.10.017. – ISSN 00981354
- [235] ZHANG, Danxu ; LUH, Peter B. ; FAN, Junqiang ; GUPTA, Shalabh: Chiller Plant Operation Optimization With Minimum Up/Down Time Constraints. In: *IEEE Robotics and Automation Letters* 3 (2018), Nr. 1, S. 9–15. <http://dx.doi.org/10.1109/LRA.2017.2723467>. – DOI 10.1109/LRA.2017.2723467. – ISSN 2377–3766
- [236] ZHANG, Hairong ; ZHOU, Jianzhong ; YE, Lei ; ZENG, Xiaofan ; CHEN, Yufan: Lower Upper Bound Estimation Method Considering Symmetry for Construction of Prediction Intervals in Flood Forecasting. In: *Water Resources Management* 29 (2015), Nr. 15, S. 5505–5519. <http://dx.doi.org/10.1007/s11269-015-1131-7>. – DOI 10.1007/s11269-015-1131-7. – ISSN 0920–4741
- [237] ZHANG, Jun ; CHUNG, H.S.-H. ; LO, A.W.-L. ; HUANG, Tao: Extended Ant Colony Optimization Algorithm for Power Electronic Circuit Design. In: *IEEE Transactions on Power Electronics* 24 (2009), Nr. 1, S. 147–162. <http://dx.doi.org/10.1109/TPEL.2008.2006175>. – DOI 10.1109/TPEL.2008.2006175. – ISSN 0885–8993
- [238] ZHANG, Yong ; JIANG, Yunjian: Robust optimization on sustainable biodiesel supply chain produced from waste cooking oil under price uncertainty. In: *Waste management (New York, N.Y.)* 60 (2017), S. 329–339. <http://dx.doi.org/10.1016/j.wasman.2016.11.004>. – DOI 10.1016/j.wasman.2016.11.004. – ISSN 1879–2456
- [239] ZOKAEE, Shiva ; BOZORGI-AMIRI, Ali ; SADJADI, Seyed J.: A robust optimization model for humanitarian relief chain design under uncertainty. In: *Applied Mathematical Modelling* 40 (2016), Nr. 17-18, S. 7996–8016. <http://dx.doi.org/10.1016/j.apm.2016.04.005>. – DOI 10.1016/j.apm.2016.04.005