



Faculty of Engineering

Thermal Engineering and Combustion Unit



VRIJE  
UNIVERSITEIT  
BRUSSEL

Department of Mechanical Engineering

Thermo and Fluid Dynamics

# Robust design optimization of hybrid renewable energy systems

---

Diederik Coppitters

## **Jury composition:**

Prof. W. De Paepe (UMONS, Supervisor)

Prof. F. Contino (VUB, Supervisor)

Prof. L. Bricteux (UMONS, President)

Prof. R. Vounckx (VUB, Vice-President)

Prof. N. Gillis (UMONS, Secretary)

Prof. H. Jeanmart (UCLouvain)

Prof. P. De Wilde (University of Plymouth)

Dr. M. Montero Carrero (3E)

Thesis submitted in  
fulfillment of the  
requirements for the award  
of the degree of Doctor of  
Engineering Sciences

Mons,

November 24, 2021



Alle rechten voorbehouden. Niets van deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook, zonder voorafgaande schriftelijke toestemming van de auteur.

All rights reserved. No part of this publication may be produced in any form by print, photoprint, microfilm, electronic or any other means without permission from the author.

Printed by

Crazy Copy Center Productions

VUB Pleinlaan 2, 1050 Brussel

Tel : +32 2 629 33 44

[crazycopy@vub.ac.be](mailto:crazycopy@vub.ac.be)

[www.crazycopy.be](http://www.crazycopy.be)

NUR CODE: 961



# Acknowledgments

First and foremost, I am profoundly grateful to Francesco and Ward, my mentors. Thank you for your excellent guidance in conducting research, the interesting discussions, the trust and freedom, on helping me to focus on what is most important, on learning to care about time management, on delivering presentations, and much more. I am truly grateful for the personal and professional development I gained under your supervision.

I was privileged to work in three dynamic research groups, resulting in diverse interactions with many researchers. I thank the professors and researchers for the warm, open-minded environments, which enabled me to flourish in every way. For the good times shared and making me enjoy every working day, I would like to specifically thank Kevin, Roeland, Thibault, François, Charles, Jarl, Xavier, Gauthier, Cedric and Alessio. A special reference to Svend, for conveying his passion for thermodynamics and teaching, as only he can. Finally, thank you Simon and Panos, for introducing me to the world of uncertainty quantification and optimization.

For the fruitful collaborations, I would like to thank Kevin, Simone, Jarl, Xavier and Gauthier. Aside from the efficient working environment you created, you gave me invaluable help with exploring different aspects of my research.

In my first year, I worked on the GenComm project. It was an excellent opportunity to meet European researchers and industrials with a passion for hydrogen. Thank you Paul, Rory, Arya, and the other members, for appreciating the input from the youngest member, for the unforgettable evenings, for the Irish proverbs and for making me understand the Irish accent.

## Acknowledgments

---

In 2019, I was privileged to conduct a brief research stay at the impressive Stanford University. I would like to thank Prof. Catherine Gorié for the opportunity and Sergio Tarantino for making this stay unforgettable.

Having a discussion on my work for multiple hours with experts in closely-related fields is a true luxury. Thank you, jury members, for reviewing the thesis, the interesting discussion, and valuable suggestions to improve the work.

I thank the Fonds de la Recherche Scientifique for believing in this project and for financing the thesis.

Finally, my research would have been impossible without the support of my family and friends. In your unique way, you ensured that I could focus on, and disconnect from, work whenever needed. A specific mention to Leenie, *mon amour pour toujours*, for your endless support and love.

# Abstract

Driven by carbon-neutrality, the deployment of photovoltaic arrays and wind turbines increases rapidly in the power, heating and mobility sectors. To comply with the needs of each sector, these renewable energy systems are coupled with different energy storage technologies and energy conversion technologies, resulting in a diverse set of hybrid renewable energy systems. Designing such a hybrid renewable energy system requires information on the technical, economic and environmental performance of each component, as well as information on the climate and energy demand. These parameters are likely to vary during the system lifetime (i.e., aleatory uncertainty), and data resources on these variations are usually limited (i.e., epistemic uncertainty). Considering these uncertainties in the design of hybrid renewable energy systems is still an exception rather than the norm. Although, disregarding uncertainty can result in a drastic mismatch between simulated and actual performance, and thus lead to a *kill-by-randomness* of the system. In other fields, such as structural mechanics and aerospace engineering, robust design optimization has already resulted in improved product quality, by providing designs that are less sensitive to the random environment. Despite its potential, applying robust design optimization on hybrid renewable energy systems is not yet studied. Therefore, the research question of this thesis reads:

*What is the added value of robust design optimization  
to hybrid renewable energy systems?*

To answer this question, this thesis followed three steps. First, a surrogate-assisted robust design optimization framework has been developed, using state-of-the-art optimization and uncertainty quantification algorithms. Despite being limited to problems with a low stochastic dimension (i.e., less than 15 uncertainties), this framework allows defining robust designs for two-component renewable energy systems, optimized for a single quantity of interest. However, hybrid renewable energy systems are typically multi-component systems,

## Abstract

---

with multiple, cross-field objectives (i.e., technical, economic and environmental objectives). Hence, in the second step of this thesis, the uncertainty quantification algorithm has been modified. This modification allowed to handle a large stochastic dimension, and thus to define robust designs for complex, multi-component hybrid renewable energy systems in a holistic context. In the third and final step, an imprecise probability method is proposed, to distinguish between epistemic and aleatory uncertainty on a parameter. In this new formulation, the robust design is optimized for the irreducible, aleatory uncertainty, and the global sensitivity analysis is reserved for the reducible, epistemic uncertainty.

The robust design optimization algorithm has been applied on three specific hybrid renewable energy systems: a photovoltaic-battery-hydrogen system, a renewable-powered hydrogen refueling station and a photovoltaic-battery-heat pump system with thermal storage. The results indicate that the robust designs are characterized by a higher penetration of renewable energy systems and by considering energy storage: Coupling battery storage and hydrogen storage to a grid-connected photovoltaic array reduces the standard deviation of the levelized cost of electricity by 42 %; A photovoltaic-battery-heat pump with thermal storage system reduces the standard deviation of the levelized cost of exergy by 36 %, as opposed to the photovoltaic-battery-gas boiler system; Shifting towards a bus fleet that partly consists of hydrogen-fueled buses (54 % of the fleet) reduces the standard deviation of the levelized cost of driving (36 %), the mean of the carbon intensity (46 %) and the standard deviation of the carbon intensity (51 %), at the expense of a limited increase in the mean of the levelized cost of driving (11 %).

As a conclusion, robust design optimization provides an added value in the design of hybrid renewable energy systems, the method complies with the computational burden of holistic design expectations, and it is adaptable to more advanced uncertainty characterization techniques.

# Contents

<b>Acknowledgments</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>List of figures</b>	<b>ix</b>
<b>List of tables</b>	<b>xiii</b>
<b>List of publications</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Hybrid renewable energy systems: modelling, uncertainties and optimization	3
1.1.1 HRES model classification . . . . .	4
1.1.2 Deterministic optimization . . . . .	5
1.1.3 Uncertainty characterization . . . . .	8
1.1.4 Stochastic optimization . . . . .	10
1.2 Problem statement and thesis outline . . . . .	11
<b>2 Surrogate-assisted robust design optimization</b>	<b>15</b>
2.1 Contributions . . . . .	17
2.2 Surrogate-assisted robust design optimization methodology . . . . .	18
2.2.1 The inputs: Uncertainty characterization . . . . .	18
2.2.2 The inner loop: Uncertainty quantification . . . . .	19
2.2.3 The outer loop: Design optimization . . . . .	27
2.2.4 Surrogate-assisted robust design optimization main procedure . . . . .	31
2.3 Implementation . . . . .	36
2.4 Application on a directly-coupled photovoltaic-electrolyzer system . . . . .	40
2.4.1 Introduction . . . . .	40

## Contents

---

2.4.2	System modelling . . . . .	42
2.4.3	Design variables and quantity of interest . . . . .	46
2.4.4	Uncertainty characterization . . . . .	49
2.4.5	Polynomial order and stochastic dimension reduction . . . . .	49
2.4.6	Results and discussion . . . . .	51
2.4.7	Conclusion . . . . .	60
<b>3</b>	<b>Sparse polynomial chaos expansion: handling large stochastic dimensions</b>	<b>63</b>
3.1	Contributions . . . . .	65
3.2	Stepwise regression Polynomial Chaos Expansion . . . . .	66
3.3	Power-to-power: photovoltaic-battery-hydrogen system . . . . .	68
3.3.1	Introduction . . . . .	68
3.3.2	System modelling . . . . .	70
3.3.3	Design variables and quantities of interest . . . . .	77
3.3.4	Uncertainty characterization . . . . .	80
3.3.5	Stepwise regression PCE . . . . .	82
3.3.6	Results and discussion . . . . .	83
3.3.7	Conclusion . . . . .	91
3.4	Power-to-mobility: renewable-powered hydrogen refueling station . . . . .	92
3.4.1	Introduction . . . . .	93
3.4.2	System modelling . . . . .	94
3.4.3	Design variables and quantities of interest . . . . .	98
3.4.4	Uncertainty characterization . . . . .	99
3.4.5	Stepwise regression PCE . . . . .	101
3.4.6	Results and discussion . . . . .	101
3.4.7	Conclusion . . . . .	113
3.5	Decision making from the Pareto front . . . . .	114
<b>4</b>	<b>Imprecise probabilities: handling uncertainty about uncertainty</b>	<b>119</b>
4.1	Contributions . . . . .	120
4.2	Polynomial chaos expansion for parametric probability boxes . . . . .	121
4.3	Power-to-heat: photovoltaic-battery-heat pump with thermal storage . . . . .	129
4.3.1	Introduction . . . . .	129
4.3.2	System modelling . . . . .	130

4.3.3	Design variables and quantity of interest . . . . .	136
4.3.4	Uncertainty characterization . . . . .	137
4.3.5	Results and discussion . . . . .	139
4.3.6	Conclusion . . . . .	152
<b>5</b>	<b>Conclusions and perspectives</b>	<b>155</b>
5.1	Main conclusions . . . . .	155
5.2	Recommendation and guidance . . . . .	157
5.3	Future perspectives . . . . .	160
<b>A</b>	<b>Application of the methods on an illustrative example</b>	<b>163</b>
A.1	Four-bar truss case . . . . .	164
A.2	Uncertainty characterization with precise probabilities . . . . .	164
A.3	Determination of the polynomial order . . . . .	165
A.4	Characterization of the sparse PCE . . . . .	167
A.5	Robust design characteristics . . . . .	168
A.6	Uncertainty quantification with imprecise probabilities . . . . .	170
<b>B</b>	<b>Physical behavior of system components</b>	<b>175</b>
B.1	Wind turbine array . . . . .	175
B.2	Photovoltaic array . . . . .	178
B.3	PEM electrolyzer . . . . .	179
B.4	PEM fuel cell . . . . .	180
B.5	Air Source Heat Pump . . . . .	182



# List of Figures

1.1	An illustrative Hybrid Renewable Energy System (HRES).	4
2.1	The Robust Design Optimization (RDO) process	16
2.2	One iteration of the NSGA-II procedure.	30
2.3	Illustration of Latin Hypercube Sampling.	32
2.4	The hypervolume metric.	32
2.5	The surrogate-assisted robust design optimization algorithm.	33
2.6	Number of training samples in function of the number of uncertainties.	36
2.7	The effect of variability in the climate conditions.	47
2.8	The power transferred from the photovoltaic array to the electrolyzer.	48
2.9	The Leave-One-Out error for the photovoltaic-electrolyzer system.	51
2.10	The hypervolume for the techno-economic design optimization.	52
2.11	Trade-off between hydrogen production and Levelized Cost Of Hydrogen.	54
2.12	Sensitivity analysis on the operating temperature of the cost-efficient design.	55
2.13	Sensitivity analysis on the discount rate.	56
2.14	Robust Design Optimization on the hydrogen production.	58
2.15	Probability density functions on the optimized LCOH.	59
2.16	Sensitivity analysis on the discount rate.	61
3.1	The considered Hybrid Renewable Energy System (HRES).	72
3.2	The climate data and demand data.	78
3.3	Statistical moments of sparse Polynomial Chaos Expansion screening.	83
3.4	Robust Design Optimization on the Levelized Cost Of Electricity for dwelling.	85
3.5	Robust Design Optimization on the Levelized Cost Of Electricity for community.	86
3.6	The Cumulative Density Functions of three optimized designs for dwelling.	89
3.7	The Cumulative Density Functions of three optimized designs for community.	89

## List of Figures

---

3.8	Probability Density Functions of the Self-Sufficiency Ratio for the dwelling. . .	90
3.9	Probability Density Functions of the Self-Sufficiency Ratio for the community. . .	90
3.10	The refueling profile indicates the daily fuel demand. . . . .	95
3.11	The considered refueling station. . . . .	96
3.12	The considered capacity factor for the photovoltaic array and wind turbine. . .	97
3.13	Evaluation of the accuracy of the sparse PCE. . . . .	102
3.14	Robust Design Optimization on the Levelized Cost Of Driving. . . . .	104
3.15	Robust Design Optimization on the Carbon Intensity. . . . .	106
3.16	Overall performance comparison. . . . .	107
3.17	The hypervolume for the holistic robust design optimization. . . . .	108
3.18	Holistic Robust Design Optimization. . . . .	109
3.19	2-dimensional and 4-dimensional Robust Design Optimization overlap. . . .	109
3.20	The results from the holistic Robust Design Optimization (RDO). . . . .	110
3.21	The Sobol' indices on the Levelized Cost Of Driving and on the Carbon Intensity.	112
3.22	The knee point on a Pareto front. . . . .	116
4.1	Illustration of a parametric probability box. . . . .	122
4.2	The reference system consists of photovoltaics, batteries and gas boiler. . . .	132
4.3	The photovoltaic-battery-heat pump system. . . . .	132
4.4	The hourly solar irradiance, ambient temperature and energy demand profiles.	135
4.5	The parametric probability box of the annual space heating demand. . . . .	138
4.6	The Robust Design Optimization on the Levelized Cost Of Exergy. . . . .	142
4.7	The Pareto fronts for the PV-battery-ASHP configurations. . . . .	144
4.8	The PV self-consumption, grid electricity consumption and system capacities.	146
4.9	The evolution of the Sobol' indices for the optimized designs. . . . .	147
4.10	The probability boxes for the optimized designs. . . . .	149
4.11	Top-5 total order Sobol' indices for the PV-battery-gas boiler design. . . . .	150
4.12	Top-5 total order Sobol' indices for the optimized PV-battery-ASHP designs. .	151
A.1	The four-bar truss structure. . . . .	164
A.2	The worst-case Leave-One-Out error For a set of design samples. . . . .	165
A.3	Leave-One-Out error (LOO) for Polynomial Chaos Expansions. . . . .	166
A.4	Normalized statistical moments of the displacement over the design space. . .	167
A.5	Statistical moments of sparse Polynomial Chaos Expansion screening. . . . .	168

A.6	Leave-One-Out error of sparse Polynomial Chaos Expansion screening. . . . .	168
A.7	Trade-off between minimizing the displacement statistical moments. . . . .	169
A.8	Pareto front for normalized objectives . . . . .	170
A.9	The probability density functions of the designs. . . . .	170
A.10	The cumulative distribution functions of the designs. . . . .	171
A.11	Probability boxes on the four-bar truss displacement. . . . .	172
A.12	Sobol' indices on epistemic uncertainty of the displacement. . . . .	172
B.1	Power curve of Vestas V90 2000 . . . . .	177
B.2	Current-voltage characteristics for the photovoltaic array. . . . .	178
B.3	The energy efficiency of the considered photovoltaic technology. . . . .	179
B.4	The current-voltage characteristic of an electrolyzer . . . . .	180
B.5	Voltage, power density and efficiency of electrolyzer. . . . .	181
B.6	The current-voltage characteristic of a fuel cell . . . . .	182
B.7	Voltage, power density and efficiency of a fuel cell. . . . .	183
B.8	Isentropic efficiency of a scroll compressor . . . . .	184
B.9	The Coefficient Of Performance of an Air-Source Heat Pump . . . . .	185



# List of Tables

2.1	The multivariate polynomials for the example case. . . . .	23
2.2	The climate data for the considered locations. . . . .	46
2.3	Techno-economic uncertain parameters for the PV-electrolyzer system. . . . .	50
2.4	The Sobol' index above the threshold for each parameter. . . . .	51
2.5	Optimized PV-electrolyzer deterministic designs. . . . .	55
2.6	The optimized stochastic PV-electrolyzer designs. . . . .	57
2.7	The Sobol' indices for the robust design. . . . .	60
3.1	The ranges for the stochastic parameters. . . . .	81
3.2	The characteristics of the three evaluated optimized designs for the dwelling. . . . .	87
3.3	The characteristics of the three evaluated optimized designs for the community. . . . .	88
3.4	The cost of hydrogen production. . . . .	94
3.5	Parameters affected by uncertainty. . . . .	100
4.1	Parameters characterized by a probability box. . . . .	139
4.2	Parameters affected by aleatory uncertainty. . . . .	139
4.3	Parameters affected by epistemic uncertainty. . . . .	140
4.4	The characteristics of the three representative designs. . . . .	145
A.1	The imprecise Sobol' indices for the displacement of both designs. . . . .	173



# Acronyms and abbreviations

ANOVA	ANalysis Of VAriance
ASHP	Air Source Heat Pump
CDF	Cumulative Distribution Function
CI	Carbon Intensity
CRF	Capital Recovery Factor
DDO	Deterministic Design Optimization
FC	Fuel Cell
FCEB	Fuel Cell Electric Buses
HRES	Hybrid Renewable Energy System
HRS	Hydrogen Refueling Station
LAR	Least Angle Regression
LCOD	Levelized Cost Of Driving
LCOE	Levelized Cost Of Electricity
LCOH	Levelized Cost Of Hydrogen
LCOX	Levelized Cost Of eXergy
LHS	Latin Hypercube Sampling
MCS	Monte Carlo Simulation
MPP	Maximum Power Point
NSGA	Nondominated Sorting Genetic Algorithm
PCE	Polynomial Chaos Expansion
PDF	Probability Density Function
PEM	Proton Exchange Membrane
PV	PhotoVoltaic
RBRDO	Reliability-Based Robust Design Optimization

## List of Tables

---

RDO	Robust Design Optimization
RHEIA	Robust design optimization of renewable Hydrogen and dErIved energy cArrier systems
SOC	State Of Charge
SSR	Self-Sufficiency Ratio
UQ	Uncertainty Quantification

# List of publications

The work resulted in 9 international, peer-reviewed journal articles (4 as the first author, 3 as the second author), 1 article under review in the international, peer-reviewed journal Applied Energy, and 12 international, peer-reviewed conference articles (6 as the first author, 5 as the second author). These papers resulted in the following number of citations and h-index:

- citations: 86 (Scopus), 97 (Google Scholar);
- h-index: 5 (Scopus and Google Scholar).

Data were retrieved on November 16, 2021.

## International journal article with peer-review

1. **Coppitters Diederik**, Verleysen Kevin, De Paepe Ward, Contino Francesco, "How can renewable hydrogen compete with diesel in mobility? Holistic robust design optimization of a bus fleet under techno-economic and environmental uncertainty" in Applied Energy, (2021), **under review**
2. Rixhon Xavier, Limpens Gauthier, **Coppitters Diederik**, Jeanmart Hervé, Contino Francesco, "The Role of Electrofuels under Uncertainties for the Belgian Energy Transition" in Energies, 14, 4027 (2021)
3. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Robust design optimization of a photovoltaic-battery-heat pump system with thermal storage under aleatory and epistemic uncertainty" in Energy, 229, 120692 (2021)
4. De Paepe Ward, Pappa Alessio, **Coppitters Diederik**, Montero Carrero Marina, Tsirikoglou Panagiotis, Contino Francesco, "Recuperator performance assessment in humidified micro Gas Turbine applications using experimental data extended with preliminary support vector regression model analysis" in Journal of Engineering for Gas Turbines and Power, GTP-20-1547 (2021)

## List of Tables

---

5. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Robust Design optimization and stochastic performance analysis of a grid-connected photovoltaic system with battery storage and hydrogen storage" in *Energy*, 213, 118798 (2020)
6. Beckers Jarl, **Coppitters Diederik**, De Paepe Ward, Contino Francesco, Van Mierlo Joeri, Verrelst Björn, "Multi-Fidelity Design Optimisation of a Solenoid-Driven Linear Compressor" in *Actuators*, 9, 2, 38 (2020)
7. Verleysen Kevin, **Coppitters Diederik**, Parente Alessandro, De Paepe Ward, Contino Francesco, "How can power-to-ammonia be robust? Optimization of an ammonia synthesis plant powered by a wind turbine considering operational uncertainties" in *Fuel*, 266, 117049 (2020)
8. Giorgetti Simone, **Coppitters Diederik**, Contino Francesco, De Paepe Ward, Bricteux Laurent, Aversano Gianmarco, Parente Alessandro, "Surrogate-Assisted Modeling and Robust Optimization of a micro Gas Turbine Plant with Carbon Capture" in *Journal of Engineering for Gas Turbines and Power*, 142, 1, 011010 , GTP-19-1326 (2019)
9. **Coppitters Diederik**, Contino Francesco, El-Baz Ahmed, Breuhaus Peter, De Paepe Ward, "Techno-economic feasibility study of a solar-powered distributed cogeneration system producing power and distillate water: Sensitivity and exergy analysis" in *Renewable Energy*, 150, 1089-1097 (2019)
10. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Surrogate-assisted robust design optimization and global sensitivity analysis of a directly coupled photovoltaic-electrolyzer system under techno-economic uncertainty" in *Applied Energy*, 248, 310-320 (2019)

## International conference article with peer-review

1. De Paepe Ward, Pappa Alessio, **Coppitters Diederik**, Montero Carrero Marina, Tsirikoglou Panagiotis, Contino Francesco, "Control strategy development for optimized operational flexibility from humidified micro gas turbine: Saturation tower performance assessment" in "ASME Turbo Expo 2021: Turbomachinery Technical Conference and Exposition" , GT2021-60209, Virtual, Online (2021)
2. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Robust design optimization of a renewable-powered demand with energy storage using imprecise probabilities" in "100RES 2020 – Applied Energy Symposium (ICAE), Virtual, Online (2020)
3. Lhuillier Charles, **Coppitters Diederik**, Verleysen Kevin, Brequigny Pierre, Mounaïm-

- Rousselle Christine, Contino Francesco, "Power-to-ammonia-to-X: cost assessment of an integrated renewable ammonia-based system providing heat and power for residential use" in "33rd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems" , Osaka, Japan (2020)
4. Limpens Gauthier, **Coppitters Diederik**, Rixhon Xavier, Contino Francesco, Jeanmart Hervé, "The impact of uncertainties on the Belgian energy system: application of the Polynomial Chaos Expansion to the EnergyScope model" in "33rd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems" , Osaka, Japan (2020)
  5. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Thermo-economic performance optimization of a grid-connected, solar-powered building with hydrogen-based energy storage, including epistemic and aleatory uncertainty quantification" in "33rd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems" , Osaka, Japan (2020)
  6. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Robust design optimization of a renewable-powered demand with energy storage using imprecise probabilities" in "100RES 2020 - Applied Energy Symposium (ICAE)" , Pisa, Italy (2020)
  7. Giorgetti Simone, **Coppitters Diederik**, Contino Francesco, De Paepe Ward, Bricteux Laurent, Aversano Gianmarco, Parente Alessandro, "Surrogate-Assisted Modeling and Robust Optimization of a Micro Gas Turbine Plant With Carbon Capture" in "ASME Turbo Expo 2019: Turbomachinery Technical Conference and Exposition" , GT2019-91400, V003T06A017, Phoenix, AZ, USA (2019)
  8. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Does the inclusion of a hydrogen-based energy system improves the robustness of the levelized cost of electricity for an urban, grid connected household?" in "11th International Conference on Applied Energy", ICAE2019, Paper ID: 272, Västerås, Sweden (2019)
  9. Verleysen Kevin, **Coppitters Diederik**, Parente Alessandro, De Paepe Ward, Contino Francesco, "Alternative Storage Of Hydrogen: Robust Design Optimization Of A Standalone Wind-powered Ammonia Plant" in "1st International Conference on Smart Energy Carriers" , Napels, Italy (2019)
  10. De Paepe Ward, **Coppitters Diederik**, Abraham Simon, Tsirikoglou Panagiotis, Ghorbaniasl Ghader, Contino Francesco, "Robust Operational Optimization of a Typical micro Gas Turbine" in Energy Procedia, 158, 5795, 5803 (2019)

## List of Tables

---

11. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Techno-economic uncertainty quantification and robust design optimization of a directly coupled photovoltaic-electrolyzer system" in Energy Procedia, 158, 1750, 1756 (2019)
12. **Coppitters Diederik**, Contino Francesco, El-Baz Ahmed, Breuhaus Peter, De Paepe Ward, "Is a distributed system based on a micro gas turbine integrated with renewable energy sources suitable for effective desalination? Sensitivity and exergy analysis." in "The 31th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2018)" , Guimaraes, Portugal (2018)

## Conference abstract

1. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Epistemic and aleatory uncertainty quantification of a grid-connected photovoltaic system with battery storage and hydrogen storage" in "Smart Energy Systems International conference", Aalborg, Denmark (2020)
2. **Coppitters Diederik**, De Paepe Ward, Contino Francesco, "Robust design optimization of a hydrogen-based solar energy storage system using an efficient stepwise regression method for sparse polynomial chaos expansions" in "Uncertainty Quantification and Optimization conference" , Sorbonne University, Paris, France (2019)

# 1 Introduction

## Chapter overview

- The need for hybrid renewable energy systems;
- The role of uncertainty in hybrid renewable energy system design;
- State of the art: hybrid renewable energy system modelling, optimization and uncertainty;
- Problem statement and thesis outline.

Energy is a primary element of life. To supply this energy, we mostly rely on fossil fuels: 80 % came from the combustion of coal, oil and natural gas in 2018 [1]. Fossil fuel combustion releases CO<sub>2</sub>, a GreenHouse Gas (GHG) that drives global warming. To avoid these emissions, well-known alternatives are solar energy and wind energy. Both energy sources are carbon-free and inexhaustible, and are therefore referred to as *renewable* energy sources. These renewable energy sources provide a tremendous potential to cover the global energy demand: capturing 2 % of the harvestable annual solar energy ( $5 \times 10^{22}$  J [2]) is sufficient to cover the global primary energy demand [3]. Harvesting solar and wind energy is typically done through PhotoVoltaic (PV) arrays and wind turbines, respectively. The deployment of these technologies increases rapidly: tremendous global capacity additions for PV arrays (107 GW) and wind turbines (65 GW) were realised in 2020 [4]. However, due to their intermittent behavior (e.g., the sun does not shine at night), these energy sources can not comply with the energy demand at any given time. To store excess energy when solar and wind are abundant, and to reuse this energy when the PV array and wind turbines fail to comply with the given demand, energy storage technologies are needed.

## Chapter 1. Introduction

---

While applying PV arrays and wind turbines in the power sector is evident, applying them to decarbonize other sectors, such as heating and mobility, is less evident [5]. As these sectors typically rely on fossil fuels (e.g., diesel in cars and natural gas in boilers), either electrically-driven technologies are needed in these sectors, or alternative fuels need to be produced with renewable electricity (i.e., electrofuels [6]). The use of electrically-driven technologies is gaining interest, such as the use of heat pumps for heating and the use of electric cars for mobility. Alternatively, electricity can be used to produce hydrogen, which can be used in combustion processes or fuel cells, for heating and mobility purposes. As a conclusion, to comply with the global energy demand in a renewable way, PV arrays and wind turbines need to be coupled with other technologies, to store and reuse the electricity when needed, and to generate heat and electrofuels in other sectors. Such a multi-component system, with two or more energy generators among which at least one is a renewable energy source, is defined as a Hybrid Renewable Energy System (HRES) [7].

Like any other system, an HRES should be designed to achieve an optimized performance (e.g., high efficiency, low costs). To do so, the designer needs information on a large set of technical, economic and environmental parameters, such as the efficiency of a PV array, the investment cost of a wind turbine, and the GHG emissions during the construction of a battery stack. Determining the true values for these parameters requires a significant amount of reliable data, an asset that is often not at hand. In addition to the characteristics of the HRES components, information is needed on the solar irradiance, wind speed and energy demand, preferably with a high spatial and temporal resolution. Thus, as data is not always at hand and predictions on the future are always wrong to some extent, these HRES parameters are subject to uncertainty. When a deterministic value (i.e., fixed and free from inherent variation) is assigned to these parameters during HRES design, the uncertainty can lead to a drastic mismatch between the simulated performance and the actual performance of the system. Such issues occurred recently in Europe: several industrial Combined Heat and Power units were taken out of operation, following highly fluctuating electricity prices, high gas prices and overcapacity on the electricity market [8]. Particularly, the 1.3 GW *Claus C* Combined Cycle Gas Turbine in the Netherlands, in operation since 2012, was shut down in 2014 because of poor market conditions (rising natural gas price and low electricity price on the market) [9]. As a conclusion, uncertainties can drastically affect the performance of an energy system and ultimately result in unfavorable operating conditions.

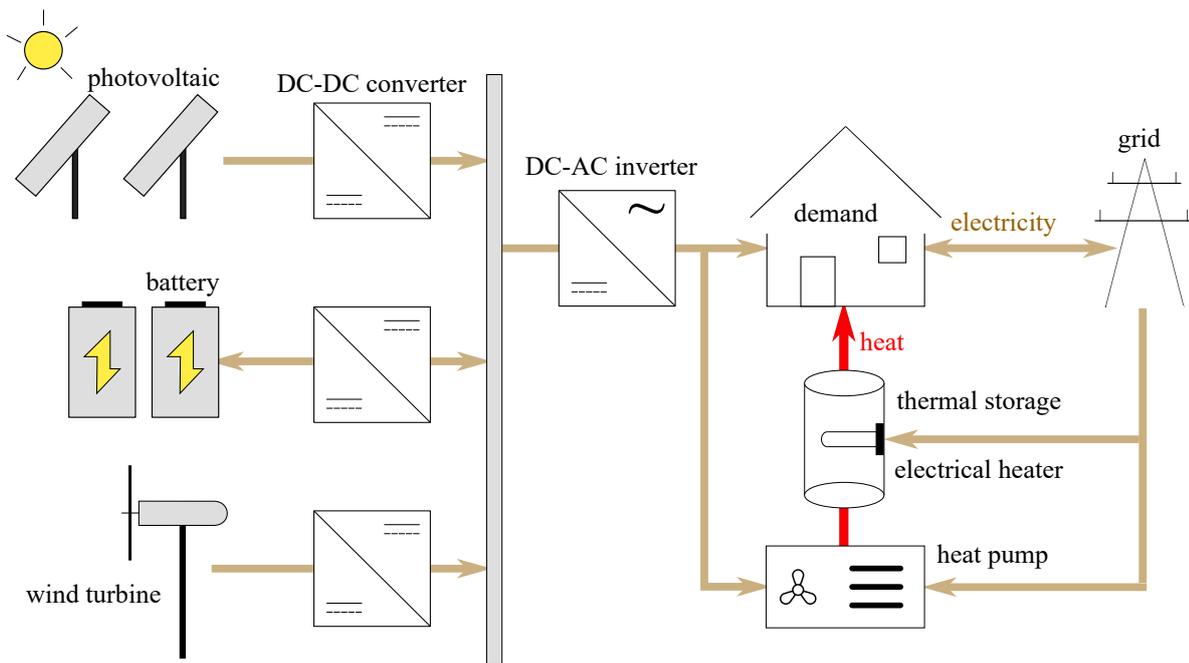
## **1.1 Hybrid renewable energy systems: modelling, uncertainties and optimization**

---

While HRES are crucial in the transition towards a carbon-neutral energy mix, these HRES need to be properly designed, considering the large set of uncertainties that affect the system performance. In this first chapter, an overview is given on how the HRES are modelled, optimized, and on how these parametric uncertainties are considered in the existing literature. After listing some limitations on the current approaches, the research question is defined, followed by a summary of the contributions of this thesis.

### **1.1 Hybrid renewable energy systems: modelling, uncertainties and optimization**

A Hybrid Renewable Energy System, or HRES, is defined as an energy system with at least two energy generators, among which at least one is a renewable energy source (e.g., PV array, wind turbine, battery and heat pump, Figure 1.1) [7]. As opposed to a single intermittent renewable energy source, an HRES achieves increased reliability in power supply. While the number of different HRES configurations is high (Rekioua et al. [10] mentioned 74 different configurations), the HRES that receive the most attention are combinations of a PV array, wind turbines, batteries, diesel generators and fuel cells [10]. To evaluate the performance of these HRES, a large number of HRES models exist, with different spatial, technical and temporal resolutions. This section presents the classification of the HRES models first (Subsection 1.1.1). These HRES models provide the system performance, depending on the fixed model parameters (e.g., PV panel efficiency) and user-defined design variables (e.g., the size of the PV array). Typically, these user-defined design variables are optimized through deterministic design optimization methods. The common deterministic design optimization approaches for HRES are discussed in Subsection 1.1.2. In these optimization approaches, the model parameters are considered deterministic (i.e., fixed and perfectly known, with no uncertainty). However, many parameters are subject to uncertainty. How this uncertainty is addressed in current HRES literature is discussed in Subsection 1.1.3. This parametric uncertainty can be considered in design optimization as well, resulting in alternative designs. The discussion on such stochastic optimization approaches in Subsection 1.1.4 closes this section.



**Figure 1.1:** In this illustrative Hybrid Renewable Energy System (HRES), a photovoltaic array and wind turbine produce electricity, which can be stored in, and recuperated from, a battery. The electricity is used to cover the electricity demand of the dwelling, as well as to generate heat in the heat pump. When the electricity from renewable energy sources and battery is insufficient, grid electricity is used to cover the electricity and heat demand.

### 1.1.1 HRES model classification

To model an HRES, a wide range of model types exist, with different temporal, spatial and technical resolutions [11]. Broadly, these models can be classified into two groups: models intended for system operation and models intended for system design [12].

System operation models are constructed for four different applications, with timescales from seconds to months: system control, dispatch, unit commitment and maintenance scheduling [13]. When the aim is to control the technical operation of an HRES (i.e., stable performance during nominal conditions), a high level of temporal detail (i.e., milliseconds) and technical detail are required [14]. He et al. [15] evaluated the active and reactive power of a PV-wind HRES under a low-voltage-ride-through control scheme, considering a 20 kHz switching frequency and an evaluation period of 1 s. The control scheme improved the energy efficiency and accelerated the post-fault recovery process. Dispatch models determine the load distribution among the generators, considering ramp-up times and transmission constraints. While the aim remains to ensure a reliable energy supply, economic dispatch

## 1.1 Hybrid renewable energy systems: modelling, uncertainties and optimization

---

models include the objective to find the least-cost combination of running the generators. Surender Reddy and Bijwe [16] applied an economic dispatch model on a 30 bus and 118 bus test system, including solar and wind energy and considering minute-to-minute variability, voltage, reactive power and line flow constraints. The economic dispatch model reduced the total cost of the system up to 5 %, when compared to three conventional economic dispatch approaches. In a broader time frame (hours to days), unit commitment models determine the commitment of generation units, based on an extensive, but coarser, set of technical constraints (e.g., simplified start-up times) and operation costs [13]. Lee et al. [17] applied a contextual learning algorithm for a unit commitment problem: minimizing the average total operating costs for a microgrid with wind turbines and thermal power generation units in a one-day time period. Finally, to ensure a reliable, cost-efficient operation, maintenance should be scheduled without harming the reliability of the power supply [11]. This subject is addressed in maintenance optimization models, which cover a timeframe of approximately one month. Yang et al. [18] proposed a cost-effective opportunistic maintenance for a wind farm, considering the wind conditions and turbine aging conditions. The framework proved superior over conventional health-centred maintenance models when failure-induced losses are high.

While operation models evaluate the operation of an existing system, design models create new systems, either by replacing an existing component with a new one (expansion planning), or by designing a new system from scratch (system planning) [13]. Expansion planning covers, among others, the integration of renewable energy systems in the current energy mix. Dagoumas et al. [19] provided an overview of models that evaluated the integration of renewable energy systems in generation expansion planning. The design of a complete system is considered in capacity planning models, which evaluate the performance over a multi-year time frame. These design models evaluate the technical, economic, environmental, and/or socio-political performance over the lifetime of the system. This topic is actively researched, with a multitude of models available [20]. An overview of HRES models for sizing purposes is provided by Anoune et al. [21].

### 1.1.2 Deterministic optimization

An HRES optimization problem is defined by three main allocations: the *model inputs*, the *objectives* and the *constraints*. Iqbal et al. [22] provided an overview of common inputs,

objectives and constraints in HRES optimization. The model inputs correspond to both the design variables, i.e., the variables that can be controlled by the designer within a certain range, and the model parameters, i.e., the parameters which are out of the control of the designer. Among others, these model inputs can correspond to the technical (e.g., efficiency, size) and economic (e.g., investment cost) parameters for each technology, the climate data and the energy demand data. Constraints limit the number of feasible designs. Common constraints include the need to comply with the energy demand, budget limitations, CO<sub>2</sub>-emissions and regulatory constraints [22]. The objectives depend on the type of assessment (e.g., technical, economic, environmental, socio-political). Commonly selected technical objectives include system reliability and minimization of loss of load probability [20, 22, 23], minimization of levelized cost and total annual cost are frequently used economic objectives [20, 22], minimization of pollutants [20, 22] and life cycle assessment [20] are typical environmental objectives, while human development index and job creation [20, 24] are commonly adopted social objectives. These objectives can be conflicting, meaning that no single design results in the optimal values for all the considered objectives. Iqbal et al. [22] illustrated the conflicting relationship between common HRES objectives. Following these conflicting relationships, a holistic approach is often recommended [23], in which several objectives are considered simultaneously during a design procedure. Eriksson et al. [25] advocated for a holistic approach in the transition towards a renewable energy economy, considering also environmental and social criteria, in addition to the more commonly adopted technical and economic objectives.

To obtain the optimized design for an HRES, an optimization algorithm is required. Given a set of constraints, such an optimization algorithm solves the problem of finding values for the design variables that lead to the optimized values for the objectives. Briefly, optimization approaches in HRES can be divided into two segments: traditional and heuristic [26]. Traditional approaches are based on differential calculus to acquire an optimized solution. These approaches include, among others, graphical construction, probabilistic method and iterative methods [23]. Despite being easy to implement, these methods adopt strong assumptions, which generally leads to over-conservative designs [27]. Moreover, due to better-performing alternatives, these traditional algorithms receive limited interest nowadays [26]. Among these better-performing alternatives, linear programming is a widely used technique for the optimization of HRES [22]. The method proves to be more flexible than the previously mentioned methods, as a wide range of complex problems

## 1.1 Hybrid renewable energy systems: modelling, uncertainties and optimization

---

can easily be evaluated [20]. However, a linear input-output relation is an unrealistic assumption for HRES [26]. To address this issue, piecewise linear approximations were developed, for which the approximation gains complexity with the number of degrees of freedom considered in the model [28].

Heuristic optimization represents algorithms that need little information on the system model being optimized. Despite being more difficult to implement, these algorithms provide efficient performance in finding the global optimum for nonlinear, complex problems, subject to intermittent energy sources like solar and wind [27, 29]. However, heuristic optimization algorithms do not ensure mathematical optimality, nor provide information on the distance between the optimized design and the truly optimal solution. Well-established heuristic optimization algorithms include the evolutionary algorithms Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). Fadaee et al. [30] provided a review on multi-objective optimization of HRES via evolutionary algorithms. GA, the most-used heuristic algorithm [27], is based on the natural evolution of mutation and selection to optimize system models [31]. Mayer et al. [32] used GA to optimize an HRES in economic and environmental objectives and concluded that GA is a suitable way to solve the complex HRES design optimization problem in multiple locations. The evolutionary PSO algorithm is based on predator behavior of swarms [33]. PSO is easier to configure than GA (i.e., no crossover and mutation operators) but suffers from premature convergence in local optima [27]. Maleki et al. [34] evaluated different PSO variants for economic size optimization of a PV-wind-battery HRES, and concluded that PSO with constriction factor finds the optimized solution at the lowest computational cost.

Several of these optimization algorithms have been included in software packages to perform design optimization of HRES. An overview of these software packages is provided by Erdinc et al. [29]. Among these software packages, Hybrid Optimization Model for Electric Renewables (HOMER) is widely used. HOMER is a software that allows sizing HRES to obtain a minimized Net Present Value, based on hourly solar irradiance, wind speed and load data [27]. In addition to optimization, HOMER provides a tool for sensitivity analysis as well. As a disadvantage, the system components are not modifiable by the user and the code is based on first-degree linear equations. Nevertheless, despite the linear representation, HOMER allows an easy interpretation of the sizing procedures for HRES [29].

### 1.1.3 Uncertainty characterization

HRES models are characterized by many model parameters. Depending on the interest, these parameters enable quantifying the technical, economic, environmental and socio-political performance of an HRES. These parameters are often subject to uncertainty. The uncertainty on a model parameter can be divided into two categories: *epistemic* and *aleatory* [35]. Epistemic uncertainty is related to the lack of knowledge on the parameter, but this knowledge can be improved by gaining more information on the parameter (e.g., measuring the temperature of water by holding your hands above the water results in a significant uncertainty, which can be reduced by using a thermometer instead). The aleatory uncertainty relates to the unknown evolution of the parameter value and is therefore irreducible. To illustrate, the aleatory uncertainty on wholesale electricity price comes from the unknown evolution of this parameter in the coming years.

Generally, the characterization of uncertainty can be performed with probabilistic or non-probabilistic approaches [36]. In a probabilistic approach, the uncertain parameter follows a specific Probability Density Function (PDF). Such a PDF is typically estimated by statistical inference when a large amount of data is available (e.g., measurement data), by expert judgement when no data is available, or by Bayesian inference when the dataset is limited, but expert knowledge is at hand [37]. Non-probabilistic approaches include interval analysis, fuzzy set theory and stochastic processes, like Markov chains [38]. For model parameters in HRES sizing, the number of data is generally rather limited (e.g., only a handful of predictions exists on the price of natural gas in the coming years), with exceptions on measurement data for, e.g., solar irradiance and wind speed. Therefore, whenever PDFs are assigned to model parameters in HRES, the distributions are usually assumed to be normally or uniformly distributed [36].

Whenever uncertainty is considered in HRES, the uncertainty is generally considered on the intermittent renewable energy supply (i.e., solar and wind), on the energy demand and on the economic parameters related to grid energy (e.g., grid electricity price, gas price, feed-in tariffs) [36]. The uncertainty on the intermittent renewable energy supply relates to the total availability (e.g., annual) and temporal availability (e.g., hourly variation). The uncertainty on the total availability corresponds to interannual variability and can be addressed by an interval or by PDFs. The temporal variation for wind and solar profiles

## 1.1 Hybrid renewable energy systems: modelling, uncertainties and optimization

---

are typically represented by PDFs (e.g., Weibull and Rayleigh distribution for wind [39], Normal and Beta distribution for solar [36]). Generally, these distributions are fitted on the available time series. Consequently, a single distribution can be sampled to generate a time series, or several distributions can be constructed to represent seasonal or diurnal variations in the time series [40, 41]. Despite the satisfying match between the intermittent renewable energy profile and the corresponding PDF, sampling from a distribution to attain hourly wind speed or solar irradiance data results in a time series with independent values. Hence, the autocorrelation in consecutive hours is lost. To avoid this independent relation, long-term measurement data (i.e., years) can be adopted, while Markov chains or synthetic time series can be constructed based on that dataset [42, 43]. As a conclusion, when limited data or nominal profiles are available, defining a PDF might be most appropriate, while a large set of data enables the construction of Markov chains or synthetic time series [36]. Similar to wind and solar irradiance, energy demand profiles exist with an underlying autocorrelation between consecutive hours. Therefore, applying PDFs on nominal energy demand profiles is the most popular approach [36]. When significant information is available on the demand (e.g., a building), the energy demand profile can be generated using a building performance simulation model. However, these models are characterized by a large set of parameters (e.g., material properties, occupant density, lighting, electrical equipment), which are inherently uncertain as well [44]. Finally, parameters related to grid energy (e.g., energy carrier price, feed-in tariff, carbon tax) evolve over the system lifetime, following the changes in the energy mix, economy, policy and regulation [45]. However, for these parameters, only a handful of predictions are available. Therefore, discrete non-probabilistic approaches are most common, while normal and uniform distributions are typically assigned to these parameters whenever a probabilistic approach is considered [36]. In the case of a time-dependent representation for these parameters, to illustrate the uncertainty of the long-term evolution, Geometric Brownian Motion [46] and the AutoRegressive Integrated Moving Average model [47] are commonly deployed. However, such methods can only be deployed when sufficient historical data is available. Moreover, they can not possibly predict unforeseen events (e.g., due to policy changes, or *black swans*) [48].

The technical and economic parameters related to the components of an HRES, such as the energy efficiency and operating costs, are generally considered deterministic [36]. However, technical parameters are subject to uncertainty on the commissioning, ageing and maintenance quality. Whenever uncertainties are considered in a technical framework, it is

usually deployed on the energy conversion efficiency by considering PDFs. Maheri et al. [49] considered a uniform distribution on the efficiency of a PV array and on the rotor power coefficient of a wind turbine. The uncertainty allocation on economic parameters for components, such as the investment cost, depends on the change in market conditions between the design and investment stage. Nevertheless, for mature technologies, these changes might be small, while novel technologies might be subject to larger uncertainty. In addition, the epistemic uncertainty related to the knowledge of the market can affect the uncertainty characterization for these parameters. Typically, these uncertainties are represented by normal distribution (e.g., [47]), uniform distributions (e.g., [50]) or discrete scenarios (e.g., [51]).

### 1.1.4 Stochastic optimization

In stochastic optimization, an objective function is optimized, while considering randomness in the model parameters. Liu et al. [52] provided an overview of optimization methodologies under uncertainty for energy systems and mentioned *stochastic programming* and *robust optimization* as the most common approaches. Stochastic programming is a scenario-based approach in which the expectation of the system performance is optimized, subject to probabilistic input uncertainties [53]. Zakaria et al. [54] provided an overview of stochastic programming applications on HRES. While stochastic programming is rather easy to implement, a large number of model evaluations are required to quantify the expectation during the multi-stage stochastic programming procedure, leading to a computationally expensive approach [55]. Alternatively, robust optimization provides robust designs, found under the worst-case combination of the input parameters values, which are defined by interval uncertainty [56, 57]. As a nested anti-optimization routine needs to be performed to find the worst-case combination of the model parameters, the method might become computationally intractable when many uncertainties are considered. Moreover, as this approach generally leads to over-conservative designs [38, 55], modifications were suggested to the algorithm, to allow controlling the conservativeness of the robust counterpart [58, 59]. The applications of robust optimization in energy systems exist [60–62], but are rather limited [63].

In other fields, such as aeronautics, an alternative method is considered when performing robust optimization. This alternative way of performing robust optimization, often referred

to as *Robust Design Optimization* (RDO), considers probabilistic uncertainties at the model inputs. RDO aims at optimizing both the expected performance (i.e., the mean) and the variability in that performance (i.e., the standard deviation) [35]. Hence, by considering the mean and standard deviation of the quantity of interest as design objectives, RDO provides alternative designs which are less sensitive to the random environment. Note that the reduction in variability of the quantity of interest is achieved by modifying the design variables, instead of reducing the uncertainty in the model parameters. RDO has provided designs with improved quality in, among others, ship design [64], structural mechanics [65, 66], electrical machine design [67] and aerospace engineering [68, 69]. Ghisu et al. [70] demonstrated a robust design for the efficiency of a gas turbine compression system, for which the improvements were focused on off-design performance, as opposed to the deterministic optimized design which focused on optimized performance in nominal conditions. Bilel et al. [71] performed a deterministic design optimization and robust design optimization of a mechatronic system in a sewing machine, aiming to optimize the motor current and the motor current fluctuations. The authors showed that the robust design can reduce the variance on the motor current and its fluctuation by 67 % and 83 %, respectively, at the expense of a slight increase up to 3 % in the corresponding mean. Fang et al. [72] performed an RDO on the fatigue life of a truck cab, concluding that by reducing the thickness of two segments in the cab, the robust design achieves a fatigue life standard deviation that drops by 40 %, at the expense of a decrease in mean fatigue life by 20 %, when compared to an optimized design in fatigue life mean. Despite its proven potential in other fields, RDO applications of HRES or similar systems are nearly non-existent. To illustrate, Tian et al. [73] highlighted that RDO applications in building energy analysis are very limited.

## 1.2 Problem statement and thesis outline

While it is clear that many HRES model parameters are subject to uncertainty, considering uncertainty in HRES decision-making processes is still an exception rather than the norm: Keirstead et al. [74] highlighted that only 3 out of 219 reviewed studies on urban energy system models mentioned uncertainty or sensitivity analysis, out of which 2 performed an optimization under uncertainty. In the limited application of design optimization of HRES under uncertainty, the studies are limited to stochastic programming and robust

optimization on linear models, which generally leads to over-conservative designs. In other fields, such as structural mechanics and aerospace engineering, RDO has already gained significant interest and resulted in improved product quality. By modifying the designs (e.g., increasing the PV array size) as opposed to reducing the uncertainties at the model input, RDO provides designs that are less sensitive to the random environment. Literature on applying RDO on HRES and illustrating its advantages as opposed to other design optimization approaches is nearly non-existent. Therefore, the research question of this thesis reads:

*What is the added value of robust design optimization  
to hybrid renewable energy systems?*

To answer this research question, an RDO framework has been developed and applied to different HRES. Out of the spectrum of HRES, a power-to-power, power-to-mobility and power-to-heat application are selected, to cover the main types of energy demand with HRES [75]. In each application, PV arrays and/or wind turbines are combined with energy storage and energy conversion technologies, and the HRES operate under conventional power management strategies. The resulting novel designs will illustrate the advantages of using RDO in decision-making processes for HRES. To achieve these results, the different steps undertaken are outlined in three main chapters:

- **Chapter 2** introduces surrogate-assisted RDO, the stochastic optimization approach adopted in this thesis. Thereafter, the open-source software is presented, which allows performing surrogate-assisted RDO on open-source and closed-source models. Finally, this chapter concludes with an application of surrogate-assisted RDO on a directly-coupled PV-electrolyzer system, a system with moderate complexity (i.e., only two components). To reduce the number of uncertainties considered in this application, and thus to reduce the computational cost, a stochastic dimension procedure is applied. The results illustrate the potential of RDO and the benefits from the analytically-derived global sensitivity indices, namely, formulating guidelines on how to effectively reduce the uncertainty on the quantity of interest.
- Typically, HRES are complex, multi-component systems for which the design should be formulated based on multiple, cross-field objectives (i.e., technical, economic, environmental and socio-political objectives). Considering multiple objectives and

multiple components in an HRES inevitably leads to a large set of uncertain parameters. In surrogate-assisted RDO, the computational cost grows dramatically with the number of uncertainties considered. **Chapter 3** addresses this issue by taking advantage of the *sparsity-of-effects* principle, a principle that states that most engineering models are mainly driven by main effects and low-order interactions [76]. With a modified algorithm, RDO has been applied on a complex, multi-component HRES (i.e., a PV-battery-hydrogen system) and on a renewable-powered hydrogen refueling station with cross-field objectives (i.e., considering techno-economic and environmental objectives). The results illustrate that RDO is suitable for the design optimization of complex, multi-component HRES in a holistic framework. In addition, the novel robust designs provide significant advantages as opposed to the deterministic alternative.

- Next to the computational efficiency, surrogate-assisted RDO suffers from a simplified probabilistic representation of uncertainty on model parameters. **Chapter 4** addresses the limitations of precise probabilistic uncertainty characterization in HRES. An imprecise probability method is proposed, which elegantly distinguishes between epistemic uncertainty and aleatory uncertainty. To use this imprecise probability method in the surrogate-assisted RDO, the RDO framework is modified. This novel method is applied on a PV-battery-heat pump system with thermal storage, providing novel robust designs based on the aleatory, irreducible uncertainty and a sensitivity analysis on the epistemic, reducible uncertainty.

Finally, the concluding remarks and future perspectives are provided in **Chapter 5**.



## 2 Surrogate-assisted robust design optimization

### Chapter overview

- Details on the adopted surrogate-assisted robust design optimization algorithm;
- Presentation of the open-source robust design optimization tool;
- Application on a directly-coupled photovoltaic-electrolyzer system.

*This chapter describes the surrogate-assisted robust design optimization algorithm, which is implemented in an open-source Python framework [77]. The directly-coupled photovoltaic-electrolyzer application is published in Applied Energy [78].*

RDO is a design optimization approach that considers the effect of the uncertain environment on the performance of the system [38]. By considering this effect, RDO provides designs that are less sensitive to variations induced by the random environment, as opposed to deterministic optimization schemes (Figure 2.1) [79]. Instead of eliminating the causes of uncertainty, this reduction of sensitivity towards the random environment is achieved by changing the design. Thus, the robust design corresponds to the design that achieves the highest stability in performance [35, 80].

A general formulation of RDO is:

$$\begin{aligned} & \text{find } \mathbf{d} \\ & \text{minimize } f(\mathbf{d}, \mathbf{X}) \\ & \text{subject to } g_j(\mathbf{d}, \mathbf{X}) \geq 0, \quad j = 1, 2, \dots, c \\ & \quad \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U \end{aligned} \tag{2.1}$$

## Chapter 2. Surrogate-assisted robust design optimization

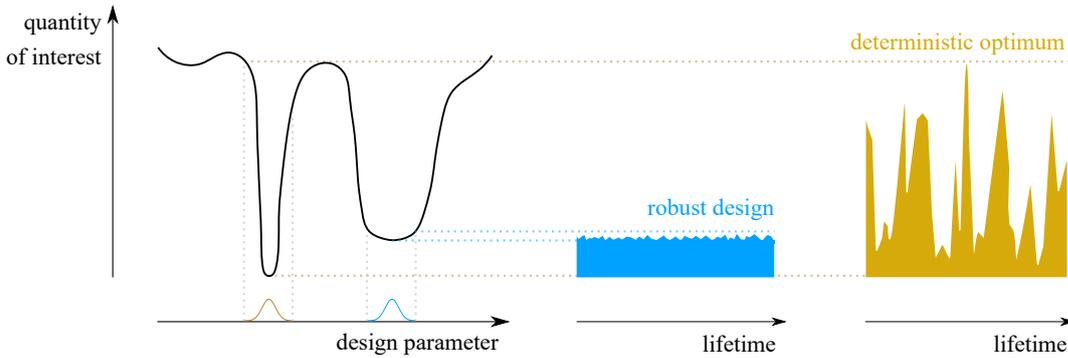
$\mathbf{d}$  is a vector in the design space. The vector components can correspond to deterministic design variables (e.g., the nominal capacity of a PV array), or the distribution parameters for an uncertain design variable (e.g., the operating temperature in an electrolyzer).  $\mathbf{d}^L$  and  $\mathbf{d}^U$  are the lower bounds and upper bounds for the design variables, respectively,  $g_j(\mathbf{d}, \mathbf{X})$  represents an uncertain constraint (e.g., the loss of load probability) and  $\mathbf{f}(\mathbf{d}, \mathbf{X})$  is the objective function (e.g., the total cost of the system).  $\mathbf{X}$  is the vector of distribution parameters for the uncertain model parameters (e.g., the ambient temperature). In RDO, the objective function can be expressed as a vector, containing the mean  $\mu_f$  and standard deviation  $\sigma_f$  for the quantities of interest, leading to a multi-objective optimization problem [64, 81]:

$$\mathbf{f}(\mathbf{d}, \mathbf{X}) = [\mu_f(\mathbf{d}, \mathbf{X}), \sigma_f(\mathbf{d}, \mathbf{X})]. \quad (2.2)$$

Alternatively, these objectives can be combined and normalized, leading to a scalar that represents the contribution from both statistical moments [35, 82]:

$$f(\mathbf{d}, \mathbf{X}) = w_1 \frac{\mu_f(\mathbf{d}, \mathbf{X})}{\mu^*} + w_2 \frac{\sigma_f(\mathbf{d}, \mathbf{X})}{\sigma^*}, \quad (2.3)$$

where  $w_1$  and  $w_2$  are positive weight factors, satisfying  $w_1 + w_2 = 1$ , and  $\mu^*$  and  $\sigma^*$  are the normalization factors. This corresponds to a single-objective optimization framework, for which the result is significantly affected by the weight factors.



**Figure 2.1:** The deterministic optimum ignores the variance in the quantity of interest. In this illustration, the deterministic optimum is subject to a significant variance, while the robust design results in a significantly smaller uncertainty on the actual outcome of the quantity of interest.

## 2.1 Contributions

During this thesis, a Python-based, open-source framework has been developed, that allows to conveniently perform deterministic design optimization, Uncertainty Quantification (UQ) and RDO on open-source and closed-source models (Section 2.3). The framework is extensively documented on Read The Docs [77], including a tutorial, examples and Application Programming Interface. To illustrate the working procedure, the framework is applied to a typical four-bar truss optimization problem in Appendix A. In addition to these algorithms, the framework comes with several hydrogen-based HRES models, for which the relevant technical, economic and environmental uncertainties are characterized. Combined with an included method to gather climate and load data for any location and demand type, these features enable to use the framework as a decision support tool for optimized, robust designs.

To evaluate the potential of using RDO on energy systems, the RDO framework was applied on four energy systems during this thesis, resulting in three journal papers [78, 83, 84] and one conference paper [85]. Indeed, as the framework considers the models as a black box, it is convenient to apply RDO on existing system models, for which the parametric uncertainties are characterized. A robust design was characterized for a micro gas turbine, by minimizing the relative standard deviation of the electrical efficiency and power output [85]. Despite that the micro gas turbine is rather insensitive to the technical uncertainties, the robustness can be improved by reducing the rotational speed and operating near the turbine outlet temperature upper limit. A similar approach was applied to a micro gas turbine with carbon capture [83]. For that system, the robust design deviates from the deterministic solution by not operating at the highest exhaust gas recirculation ratio. In the application of surrogate-assisted RDO on a wind power-to-ammonia plant [84], the sensitivity indices indicated that the uncertainty on the wind speed drives the uncertainty on the ammonia production of the deterministic design, while the uncertain operating temperature of the Haber-Bosch process dominates the uncertainty on the ammonia production of the robust design. Finally, a robust design was found for a directly-coupled photovoltaic-electrolyzer system [78]. The results of this work indicated that bulk manufacturing of the technology and more demonstration projects are the main actions to further improve the robustness. In this thesis, only the directly-coupled photovoltaic-electrolyzer case is detailed (Section 2.4), as it best fits with the methodology applied to answer the research question (i.e., it is the

only renewable energy case, subject to a significant stochastic dimension).

## 2.2 Surrogate-assisted robust design optimization methodology

The surrogate-assisted RDO aims at optimizing (i.e., minimizing or maximizing) the mean and minimizing the standard deviation of each quantity of interest. To define the random environment, parametric uncertainties are considered (Subsection 2.2.1). To quantify the mean and standard deviation of a quantity of interest for a design, a UQ method is required (Subsection 2.2.2). When quantified, these statistical moments are used as objective values. To find the optimized designs concerning these objectives, an optimization algorithm is coupled to the UQ algorithm (Subsection 2.2.3). Finally, the combined operation of the UQ and optimization methods, leading to the surrogate-assisted RDO algorithm, is discussed in Subsection 2.2.4.

### 2.2.1 The inputs: Uncertainty characterization

Uncertainty can be present due to inaccuracy of the system model (i.e., model form uncertainty) and due to the uncertainty on the parameter values. In this thesis, no model form uncertainty is considered, but instead, it is indirectly minimized by mainly adopting experimentally-validated component models for the HRES. Thus, the random environment is characterized by independent parameter uncertainties. This uncertainty is divided into two categories: epistemic uncertainty and aleatory uncertainty (Subsection 1.1.3). This classification is relevant in theory, but in practice, these types of uncertainty can blend when treated in general terms [86]. Therefore, these definitions are adapted to the context of HRES design, based on the classification of Moret et al. [87]. In this formulation, epistemic uncertainties are handled as *present uncertainties*, referring to uncertainties that can be addressed by increasing the model detail and data collection (e.g., the cost of buying a PV panel today, which is uncertain in an early design stage, but can be specified in a detailed design stage). Aleatory uncertainties are handled as *future uncertainties*, referring to uncertainties that are irreducible due to the future evolution of the parameter (e.g., the price of natural gas over 20 years). These formulations show similarities with other

## 2.2 Surrogate-assisted robust design optimization methodology

---

definitions proposed in literature, such as design uncertainty and scenario uncertainty, respectively [73]. However, in this thesis, epistemic and aleatory uncertainty are used to indicate present and future uncertainties, respectively, to be consistent with most literature.

The parametric uncertainties are characterized by distributions. Distributions are typically defined based on large datasets, which are not always at hand for each parameter. Therefore, when no meaningful information can be extracted on the distribution from a limited dataset, a uniform distribution is typically assumed, which assigns an equal probability to each value within a range. This approach is similar to assigning an interval to a parameter, but takes advantage of the central tendency when uniform distributions are propagated through a system model (i.e., central limit theorem).

In this thesis, the data for uncertainty characterization is typically adopted from ranges defined in literature, such as review papers on a specific technology (e.g., on hydrogen electrolyzers [88]), or from a set of predictions on the future value of a parameter (e.g., six scenarios on the future evolution of the wholesale electricity price for Belgium [89]), or from historic data (e.g., the historic interannual variability of the solar irradiance [90]). For each application, the main philosophy of the uncertainty characterization is elaborated in a specific section.

### 2.2.2 The inner loop: Uncertainty quantification

Following the formulation of the objective function in Equation 2.2, the mean and standard deviation of the quantity of interest should be quantified for each design sample. The state-of-the-art method that propagates parameter uncertainties through a system model and quantifies the statistical moments of the model output is crude Monte Carlo Simulation (MCS) [91, 92]. This method is an always-converging and easy-to-implement technique: Through random sampling of the probabilistic uncertain parameters at the model input, a set of deterministic values for the quantity of interest is obtained, out of which the mean and standard deviation can be derived [93]. Despite the advantage of considering the model as a black box, MCS suffers from a low computational efficiency to reach an acceptable level of convergence, i.e.,  $10^3 - 10^4$  model evaluations are minimally required [94]. The computational efficiency of MCS can be improved by adopting low-discrepancy sampling methods such as Latin Hypercube Sampling (LHS) [95], quasi-random Sobol sampling [96]

and importance sampling [97]. Computationally-efficient alternatives to MCS are surrogate model construction methods [98, 99]. Typical surrogate models include Kriging [100, 101], support vector machines [102], ANalysis Of VAriance (ANOVA) [103] and Polynomial Chaos Expansion (PCE) [104]. While these surrogate models are often considered in deterministic optimization, their application in stochastic optimization remains limited [105, 106]. Chatterjee et al. [107] assessed different surrogate modelling techniques in RDO and concluded that ANOVA decomposition and PCE are the most promising surrogates. In the post-processing of the surrogate model (i.e., the quantification of the statistical moments), PCE provides significant advantages, such as the analytical quantification of the statistical moments and the sensitivity indices (i.e., Sobol' indices) [94]. Therefore, in this thesis, PCE was adopted as a UQ method. In the remainder of this subsection, the PCE method is presented.

### 2.2.2.1 Polynomial Chaos Expansion

Consider a generic function  $y$ , which represents the input/output relation of a system model as an infinite series:

$$y = \sum_{i=0}^{\infty} u_i \Psi_i(\mathbf{X}), \quad (2.4)$$

where  $\{u_i\}_{i=0}^{\infty}$  are coefficients and  $\{\Psi_i\}_{i=0}^{\infty}$  corresponds to a set of basis functions, characterized by the input vector  $\mathbf{X}$  [94].

**Construction of the basis** To construct the basis functions, the input random vector  $\mathbf{X}$  should consist of independent components ( $\mathbf{X} = \{X_i, i = 1, \dots, M\}$ , where  $M$  denotes the number of random parameters considered). When the random components, each characterized by a distribution  $f_{X_i}$ , are considered independent, the joint distribution for  $\mathbf{X}$  corresponds to the product of the  $M$  distributions:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^M f_{X_i}(x_i), \quad (2.5)$$

where  $x_i \in D_{X_i}$  and  $D_{X_i}$  is the support of  $X_i$  (i.e., the subset of  $X_i$  for which the points evaluated in  $f_{X_i}$  are different from zero,  $D_{X_i} = \{x \in X_i | f_{X_i}(x) \neq 0\}$ ). For each single random

## 2.2 Surrogate-assisted robust design optimization methodology

parameter  $X_i$ , a functional inner product can be defined, with two functions  $(\phi_1, \phi_2)$  that map the points in  $D_{X_i}$  to real values:

$$\langle \phi_1, \phi_2 \rangle_i = \int_{D_{X_i}} \phi_1(x) \phi_2(x) f_{X_i}(x) dx. \quad (2.6)$$

This inner product corresponds to the expectation  $\mathbb{E}[\phi_1(X_i) \phi_2(X_i)]$  to the distribution  $f_{X_i}$  (i.e., the expected value for the average if the random process occurred infinite times). When these two functions are orthogonal to the probability measure  $f_{X_i}(x) dx$ , the expectation becomes equal to zero. Hence, introducing a family of orthogonal polynomials  $\{\pi_k^{(i)}, k \in \mathbb{N}\}$  in the inner product leads to:

$$\int_{D_{X_i}} \pi_j^{(i)}(x) \pi_k^{(i)}(x) f_{X_i}(x) dx = \langle \pi_j^{(i)}, \pi_k^{(i)} \rangle_i \delta_{jk}, \quad (2.7)$$

with  $\delta_{jk}$  denoting the Kronecker delta (i.e., equal to 1 when  $j = k$  and equal to 0 otherwise). Finally, the orthogonal polynomial family is normalized, leading to an orthonormal family  $\{\psi_j^{(i)}\}_{j=0}^{\infty}$ :

$$\psi_j^{(i)} = \frac{\pi_j^{(i)}}{\sqrt{\langle \pi_j^{(i)}, \pi_j^{(i)} \rangle_i}} \quad i = 1, \dots, M, \quad j \in \mathbb{N}. \quad (2.8)$$

To obtain the orthogonal family, the Gram-Schmidt orthogonalization procedure can be applied. However, for standard distributions, the corresponding families are known [94]. To illustrate, the Legendre polynomial family corresponds to a random parameter characterized by a uniform distribution over  $[-1, 1]$  (i.e.,  $X_i \sim \mathcal{U}(-1, 1)$ ).

Following the univariate orthonormal polynomials, the basis functions for the series in Equation 2.4 can be constructed. These basis functions are multivariate polynomials, which correspond to tensor products of the univariate orthonormal polynomials:

$$\Psi_{\alpha}(\mathbf{x}) = \prod_{i=1}^M \psi_{\alpha_i}^{(i)}(x_i). \quad (2.9)$$

In this formulation,  $\alpha \in \mathbb{N}^M$  corresponds to a multi-index, which denotes a list of ordered

integers:

$$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_M). \quad (2.10)$$

Hence, the multivariate orthogonal polynomial  $\Psi_{\boldsymbol{\alpha}}$  corresponds to the tensor product of univariate polynomials for which the order is defined by the corresponding index in  $\boldsymbol{\alpha}$ . Following Equation 2.7, these multivariate polynomials in the random input vector  $\mathbf{X}$  are also orthonormal:

$$\int_{D_{X_i}} \Psi_{\boldsymbol{\alpha}}(\mathbf{x}) \Psi_{\boldsymbol{\beta}}(\mathbf{x}) f_X(\mathbf{x}) d\mathbf{x} = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}} \quad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^M, \quad (2.11)$$

where the squared norm is removed due to the orthonormality of the multivariate polynomials. Finally, the set of all the multivariate polynomials in the input random vector  $\mathbf{X}$  forms a basis for an infinite series, that provides an exact representation of the input/output relation of a system model  $\mathcal{M}$ :

$$\mathcal{M}(\mathbf{X}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} u_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{X}). \quad (2.12)$$

**Truncation scheme** The series presented in Equation 2.12 is exact when infinite elements are considered. However, quantifying the coefficients for an infinite series is computationally intractable. Hence, the series should be truncated. A typical truncation scheme limits the number of multivariate polynomials in the series based on the polynomial order of the multivariate polynomial. The total degree of the multivariate polynomial  $|\boldsymbol{\alpha}|$  corresponds to the sum of the order for each univariate polynomial in the multivariate polynomial. Hence, the multi-indices related to the multivariate polynomials for which the total degree is below or equal to the limiting degree  $p$  can be stored in the truncated series  $\mathcal{A}^{M,p}$ :

$$\mathcal{A}^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : |\boldsymbol{\alpha}| \leq p\}. \quad (2.13)$$

The number of terms in this truncated series equals:

$$\text{card}(\mathcal{A}^{M,p}) = \binom{p+M}{p} = \frac{(M+p)!}{M!p!}. \quad (2.14)$$

## 2.2 Surrogate-assisted robust design optimization methodology

Hence, following this truncation scheme, the PCE representation  $\mathcal{M}^{\text{PCE}}$  of the input/output relation of the actual system model  $\mathcal{M}$  is an approximation, containing only the multi-indices in  $\mathcal{A}^{M,p}$ :

$$\mathcal{M}^{\text{PCE}}(\mathbf{X}) = \sum_{\alpha \in \mathcal{A}^{M,p}} u_{\alpha} \Psi_{\alpha}(\mathbf{X}) \approx \mathcal{M}(\mathbf{X}). \quad (2.15)$$

**Example** Consider a model  $y = \mathcal{M}(x_1, x_2)$  for which the two input parameters are characterized by a standard normal distribution (i.e., mean and standard deviation equal to 0 and 1, respectively). The PCE representation for this model reads:

$$\mathcal{M}^{\text{PCE}}(X_1, X_2) = \sum_{\alpha \in \mathbb{N}^2} u_{\alpha} \Psi_{\alpha}(X_1, X_2). \quad (2.16)$$

As standard normal distributions are considered, the Hermite Polynomials are adopted. Assuming a maximum polynomial degree of 3, the truncation set  $\mathcal{A}^{2,3}$  contains 10 multi-indices (Equation 2.14). The 10 corresponding multivariate polynomials and the normalized Hermite polynomials up to order 3 are presented in Table 2.1.

**Table 2.1:** The multivariate polynomials for the example case.

$\Psi$	$\alpha$	$\Psi_{\alpha}$
$\Psi_0(x) = 1$	(0,0)	1
$\Psi_1(x) = x$	(1,0)	$X_1$
$\Psi_2(x) = \frac{x^2 - 1}{\sqrt{2}}$	(0,1)	$X_2$
$\Psi_3(x) = \frac{x^3 - 3x}{\sqrt{6}}$	(2,0)	$\frac{X_1^2 - 1}{\sqrt{2}}$
	(1,1)	$X_1 X_2$
	(0,2)	$\frac{X_2^2 - 1}{\sqrt{2}}$
	(3,0)	$\frac{X_1^3 - 3X_1}{\sqrt{6}}$
	(2,1)	$\frac{(X_1^2 - 1) X_2}{\sqrt{2}}$
	(1,2)	$\frac{(X_2^2 - 1) X_1}{\sqrt{2}}$
	(0,3)	$\frac{X_2^3 - 3X_2}{\sqrt{6}}$

## Chapter 2. Surrogate-assisted robust design optimization

---

Hence, the PCE for the model  $\mathcal{M}(X_1, X_2)$  is presented as follows:

$$\begin{aligned} \mathcal{M}^{\text{PCE}}(X_1, X_2) = & u_0 + u_1 X_1 + u_2 X_2 + u_3 \frac{X_1^2 - 1}{\sqrt{2}} + u_4 X_1 X_2 + \\ & u_5 \frac{X_2^2 - 1}{\sqrt{2}} + u_6 \frac{X_1^3 - 3X_1}{\sqrt{6}} + u_7 \frac{(X_1^2 - 1) X_2}{\sqrt{2}} + \\ & u_8 \frac{(X_2^2 - 1) X_1}{\sqrt{2}} + u_9 \frac{X_2^3 - 3X_2}{\sqrt{6}}. \end{aligned} \quad (2.17)$$

**Quantifying the coefficients** To calculate the coefficients, the regression approach is used, based on least-square minimization. In this approach, a set of equations is constructed. Each equation corresponds to the PCE, evaluated by a set of random samples  $\{\mathbf{x}^{(i)}, i = 1, \dots, n\}$ , where each sample contains  $M$  values, one for each of the  $M$  random parameters. As the coefficients should be the same for each equation, these coefficients are stored in the matrix  $\mathbf{u}$ . The multivariate polynomials, each evaluated by a random sample, are stored in the information matrix  $\mathbf{A}$ :

$$\mathbf{A} = \left\{ \mathbf{A}_{ij} = \Psi_j(\mathbf{x}^{(i)}), i = 1, \dots, n, j = 1, \dots, \text{card}(\mathcal{A}^{M,p}) \right\}. \quad (2.18)$$

Hence, in the information matrix  $\mathbf{A}$ , the number of multivariate polynomials depends on the number of random parameters considered and the maximum total degree selected in the truncation scheme (Equation 2.14) and  $n$  corresponds to the number of samples considered. Finally, the deterministic result of evaluating  $\mathbf{x}^{(i)}$  on the actual model  $\mathcal{M}(\mathbf{x}^{(i)})$  is stored in a matrix  $\mathbf{y}$ . The solution to the least-square minimization problem reads:

$$\hat{\mathbf{u}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}. \quad (2.19)$$

The Moore-Penrose inverse  $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$  is computed using singular value decomposition. The quality of the estimated values for the coefficients  $\hat{\mathbf{u}}$  depends on the number of random samples  $n$  evaluated on the model. The empirical rule of thumb to acquire a well-posed least-square minimization is to have a number of random samples equal to at least 2 times the number of coefficients ( $\text{card}(\mathcal{A})$ ).

**Retrieve the statistical moments** The statistical moments can be derived from the PCE coefficients analytically, i.e., no more model evaluations are required. To illustrate, the

## 2.2 Surrogate-assisted robust design optimization methodology

---

mean  $\mu^{\text{PCE}}$  and standard deviation  $\sigma^{\text{PCE}}$  are derived as follows (higher moments could also be obtained):

$$\mu^{\text{PCE}} = u_0, \quad (2.20)$$

$$\sigma^{2 \text{ PCE}} = \sum_{\alpha \in \mathcal{A}, \alpha \neq \mathbf{0}} u_{\alpha}^2. \quad (2.21)$$

**The Sobol' indices** The Sobol' indices represent the decomposition into fractions of the variance of the quantity of interest, for which each fraction is allocated to a random input parameter (first-order) or a set of random input parameters (higher-order). Hence, these indices correspond to the variance on the model output induced by one or more random parameters, divided by the total variance of the model output. To illustrate, a first-order Sobol' index  $S_i$  represents the ratio of the variance induced by the random parameter  $i$  to the total variance:

$$S_i = \frac{D_i}{D} = \frac{\text{Var}[\mathcal{M}_i(X_i)]}{\text{Var}[\mathbf{y}]}. \quad (2.22)$$

The second-order Sobol' indices correspond to:

$$S_{\mathbf{v}} = \frac{D_{\mathbf{v}}}{D} = \frac{\text{Var}[\mathcal{M}_{\mathbf{v}}(X_i, X_j)]}{\text{Var}[\mathbf{y}]}, \quad (2.23)$$

where  $\mathbf{v}$  corresponds to  $\{i, j\}$ . This set of indices  $\mathbf{v}$  can be defined to quantify the Sobol' indices of any order:

$$\mathbf{v} = i_1, \dots, i_s \subset 1, \dots, M. \quad (2.24)$$

The total Sobol' index  $S_i^T$  quantifies the total impact of a random parameter  $X_i$  and corresponds to the sum of the Sobol' indices of any order that contain index  $i$  in the considered set of indices  $\mathbf{v}$ :

$$S_i^T = \sum_{i \in \mathbf{v}} S_{\mathbf{v}}. \quad (2.25)$$

## Chapter 2. Surrogate-assisted robust design optimization

---

By using the PCE coefficients  $\mathbf{u}$ , the Sobol' indices can be quantified analytically. First, consider the set of multivariate polynomials  $\Psi_{\alpha}$  which depend only on  $\mathbf{u}$ :

$$\mathcal{A}_{\mathbf{u}} = \{\alpha \in \mathcal{A} : \alpha_k \neq 0 \text{ if and only if } k \in \mathbf{u}\}. \quad (2.26)$$

Hence, the truncated PCE for this set corresponds to:

$$\mathcal{M}_{\mathbf{u}}^{\text{PCE}}(\mathbf{x}_{\mathbf{u}}) = \sum_{\alpha \in \mathcal{A}_{\mathbf{u}}} u_{\alpha} \Psi_{\alpha}(\mathbf{x}). \quad (2.27)$$

Following the orthogonality of the multivariate polynomials, the partial variances depend solely on the corresponding PCE coefficients:

$$D_{\mathbf{u}} = \text{Var}[\mathcal{M}_{\mathbf{u}}^{\text{PCE}}(\mathbf{x}_{\mathbf{u}})] = \sum_{\alpha \in \mathcal{A}_{\mathbf{u}}} u_{\alpha}^2. \quad (2.28)$$

Hence, the Sobol' indices derived from a PCE are quantified by a mere combination of the squares of the PCE coefficients. To illustrate, the first-order Sobol' indices derived from a PCE are defined by the coefficients with a multi-index for which the value at index  $i$  is larger than zero, and the values at the other indices are equal to zero:

$$S_i^{\text{PCE}} = \sum_{\alpha \in \mathcal{A}_i} u_{\alpha}^2 / D \quad \mathcal{A}_i = \{\alpha \in \mathcal{A} : \alpha_i > 0, \alpha_{j \neq i} = 0\}, \quad (2.29)$$

and the total-order Sobol' indices derived from a PCE are defined by the coefficients with a multi-index for which the value at index  $i$  is larger than zero. As the total-order Sobol' index computes all mutual interactions, the values for the other indices can be any Natural number:

$$S_i^{T,\text{PCE}} = \sum_{\alpha \in \mathcal{A}_i^T} u_{\alpha}^2 / D \quad \mathcal{A}_i^T = \{\alpha \in \mathcal{A} : \alpha_i > 0\}. \quad (2.30)$$

**The Leave-One-Out error** To estimate the error of the PCE, Leave-One-Out (LOO) cross-validation is considered. First, one sample  $\mathbf{x}^{(i)}$  is left out of the set of  $n$  random samples and the PCE is constructed without the information from this sample, i.e.,  $\mathcal{M}^{\text{PCE} \setminus i}$ . The residual error at that point is quantified as the difference between the result from the actual

## 2.2 Surrogate-assisted robust design optimization methodology

model  $\mathcal{M}$  and the result from the PCE without information on that point  $\mathcal{M}^{\text{PCE}\setminus i}$ :

$$\Delta_i = \mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{\text{PCE}\setminus i}(\mathbf{x}^{(i)}). \quad (2.31)$$

The residual error can be quantified by leaving out each sample, one at a time, out of the set of samples to construct the PCE. As  $n$  samples are present in the set of samples, the LOO error corresponds to the sum of the  $n$  residual errors:

$$E_{\text{LOO}} = \frac{1}{n} \sum_{i=1}^n \Delta_i^2. \quad (2.32)$$

This approach requires constructing  $n$  PCEs, which can become computationally demanding in a large stochastic dimension. However, following algebraic derivations [94], the LOO error can be quantified based on the  $i$ -th diagonal terms in the matrix  $\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ . Eventually, the LOO error corresponds to:

$$E_{\text{LOO}} = \frac{1}{n} \sum_{i=1}^n \left( \frac{\mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{\text{PCE}}(\mathbf{x}^{(i)})}{1 - h_i} \right)^2. \quad (2.33)$$

Note that in this formulation,  $\mathcal{M}^{\text{PCE}}$  is constructed based on the full experimental design (i.e.,  $n$  samples). The LOO error can be normalized by dividing the LOO error by the variance of the model outputs  $\text{Var}[\mathbf{y}]$ .

### 2.2.3 The outer loop: Design optimization

As for each quantity of interest, two objectives are considered (i.e., the mean and standard deviation), a multi-objective optimization algorithm should be adopted [108]. In case that the objectives are conflicting, the method should provide design compromises between the conflicting objectives (i.e., a Pareto front). Multi-objective Genetic Algorithms are frequently used metaheuristic optimizers in surrogate-assisted RDO, due to an extensive investigation of their performance [109]. Deb et al. [110] established a Genetic Algorithm optimization scheme for RDO, while Gaspar et al. [111] and Saha et al. [112] further developed the robustness measure and improved the optimization strategy. Following its validated use, a Genetic Algorithm was adopted for the surrogate-assisted RDO algorithm, namely, Nondominated Sorting Genetic Algorithm (NSGA-II). Lee et al. [113] used NSGA-II in an RDO formulation to characterize a robust design on the power output of a thermo-electric generator. The

robust design decreased the power output standard deviation by 49% when compared to the standard deviation achieved by the deterministic optimized design. NSGA-II is a gradient-free, population-based optimizer that uses heuristics to converge to an optimized solution [31]. The population contains the design samples found and is iteratively updated after each iteration. The new population contains design samples with improved performance as opposed to the previous population. The way the NSGA-II determines if one design sample has improved performance over another design sample in a multi-objective context is explained first. Thereafter, the main procedure of NSGA-II is described.

### 2.2.3.1 Selection procedure

In a multi-objective context, determining which design sample achieves the highest performance in the objectives is not straightforward. To illustrate, a design sample  $\mathbf{d}_i$  with objective values  $\mathbf{f}_i$  might achieve a value on the first objective that is better than the value on the first objective achieved by a design sample  $\mathbf{d}_j$  with objective values  $\mathbf{f}_j$  (e.g., when minimization is aimed at,  $f_{i,1} < f_{j,1}$ ), but  $\mathbf{d}_i$  might achieve a value on the second objective that is worse than the value on the second objective from  $\mathbf{d}_j$  (e.g., when minimization is aimed at,  $f_{i,2} > f_{j,2}$ ). To determine the design sample with the *best performance*, NSGA-II defines two performance indicators: the rank, achieved by non-dominated sorting, and the crowding distance.

**Non-dominated sorting** In non-dominated sorting, the aim is to sort the design samples in a population into non-dominated fronts ( $F_i$ ) [31]. Consider a population with  $n$  design samples  $\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n\}$ . The corresponding fitness values, one value for each objective, are stored in  $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ . In a minimization problem,  $\mathbf{d}_i$  dominates  $\mathbf{d}_j$  when no value of  $\mathbf{f}_j$  is less than the corresponding value in  $\mathbf{f}_i$  and at least one value in  $\mathbf{f}_j$  is greater than the corresponding value in  $\mathbf{f}_i$  [114]. To sort the design samples in non-dominating fronts, first, all design samples in the population are stored in a pool. The design samples that are not dominated by any other design sample in the population are stored in the first non-dominated front  $F_1$ . When all design samples in the first non-dominated front are determined, those design samples are removed from the pool of design samples. Then, the procedure is repeated, determining the design samples which are not dominated by any other design sample that remains in the pool. These design samples are stored in

## 2.2 Surrogate-assisted robust design optimization methodology

---

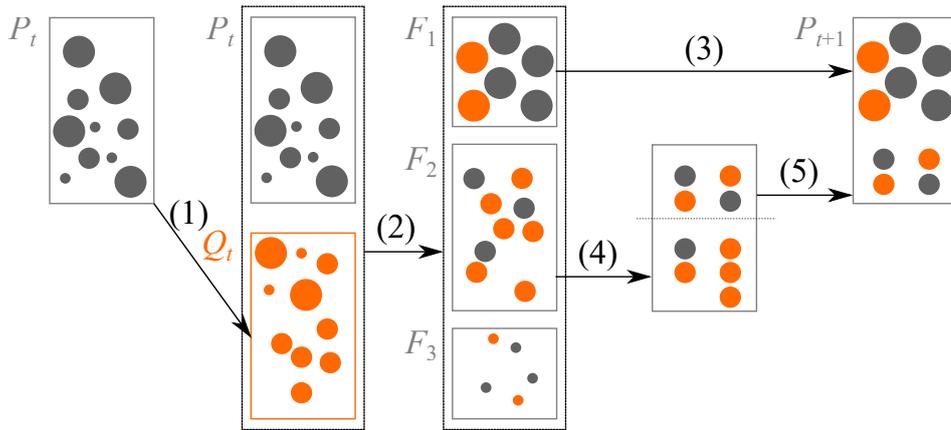
the second front  $F_2$ . This procedure is repeated until all design samples are stored in a non-dominated front. Finally, each design sample is assigned with a rank, equal to the number of the non-dominated front, e.g., the design samples in  $F_1$  are assigned with rank 1. Note that this is a naive description, for the sake of clarity. The actual implementation is, however, slightly different, to reduce the number of comparisons. The implementation of a fast non-dominated sorting approach is described in pseudo-code by Deb et al. [31].

**Crowding distance** In addition to the rank, a second performance indicator is allocated to each design sample, namely the crowding distance. The crowding distance corresponds to the euclidian distance between each individual in the objective space [31]. Hence, when the crowding distance of a design sample is small, other design samples with similar values for the objectives are present in the population. Instead, a design sample with a large crowding distance has no design samples in a near distance and might contain unique information in the design space. Therefore, design samples with a higher crowding distance are favoured, as these design samples improve the diversity of the population. In that sense, the design samples at the boundaries of the objective space, i.e., the design samples with the minimum or maximum value for at least one objective, are assigned with an infinite crowding distance.

**Selection** Based on the rank and crowding distance of each design sample, the crowded-comparison-operator  $<_n$  is used to carry out the selection procedure of the best design samples [31].  $\mathbf{d}_i <_n \mathbf{d}_j$  if the rank of  $\mathbf{d}_i$  is lower than the rank of  $\mathbf{d}_j$ . In the case of an equal rank, the crowding distance of  $\mathbf{d}_i$  should be larger than the crowding distance of  $\mathbf{d}_j$ .

### 2.2.3.2 NSGA-II main procedure

Based on the previously described selection procedure, the NSGA-II main procedure iteratively updates the population of design samples (Figure 2.2). The starting population determines the global exploration quality. Hence, the initial population should sufficiently cover the design space, characterized by the bounds of the different design variables. A typical method to generate an initial population is LHS [95]. In LHS, a Latin Hypercube is constructed for  $N$  design samples. Such a Latin Hypercube is a multi-dimensional generalization of the two-dimensional Latin square, in which the design variables are di-



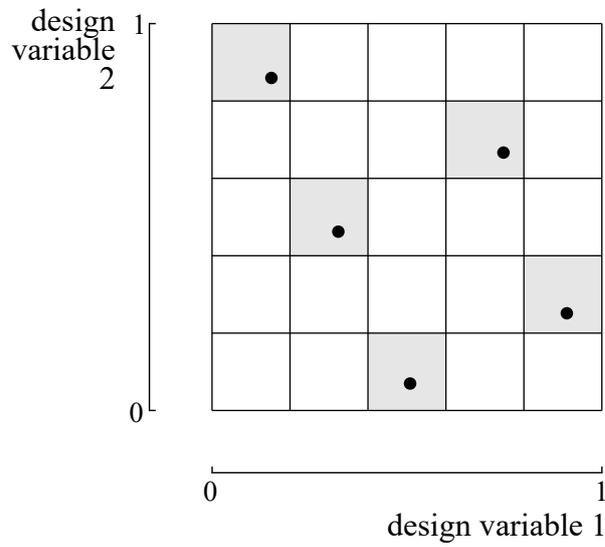
**Figure 2.2:** One iteration of the NSGA-II procedure. First, an offspring  $Q_t$  is created from the population  $P_t$  (1). The design samples from the population and offspring are sorted in non-dominated fronts, based on their dominance in the objectives (2). The design samples from the first front  $F_1$  are all stored in the new population  $P_{t+1}$  (3). The design samples in the second front  $F_2$  are sorted based on crowding distance (4) and the remaining places in the new population  $P_{t+1}$  are filled with the design samples with the highest crowding distance among the samples in  $F_2$  (5). The remaining design samples in  $F_2$  and  $F_3$  are rejected.

vided into  $N$  equal intervals and each row and each column contains only one design sample (Figure 2.3). Starting from a population  $P_t$  with  $N$  design samples, an offspring  $Q_t$  should be created ((1) in Figure 2.2). The offsprings are created based on a binary tournament selection, crossover and mutation. First, in a binary tournament selection, two design samples are chosen randomly out of the population and based on the crowding-comparison-operator  $<_n$ , the best design sample is chosen (i.e., parent 1). A second parent is chosen analogously. These parents generate two offsprings, based on crossover and mutation [115]. Crossover corresponds to swapping part of the values of one design sample with part of the values from another design sample, aiming to mix the solutions and reach convergence in a design subspace. Mutation changes part of the design values randomly, aiming to increase diversity in the population and to escape from a local optimum. A probability is assigned to each operator, providing the possibility to control the exploitation and exploration capabilities of the algorithm. Typically, the probability for crossover is at least 85 %, while the probability for mutations remains below 10 %. This process is repeated until the number of offsprings is equal to the number of design samples in the initial population. The initial population and offsprings are stored in a combined population, with size  $2N$ . For this combined population, the samples are sorted into the non-dominated fronts ((2) in Figure 2.2). To create the new population  $P_{t+1}$ , first, the samples in the non-dominated

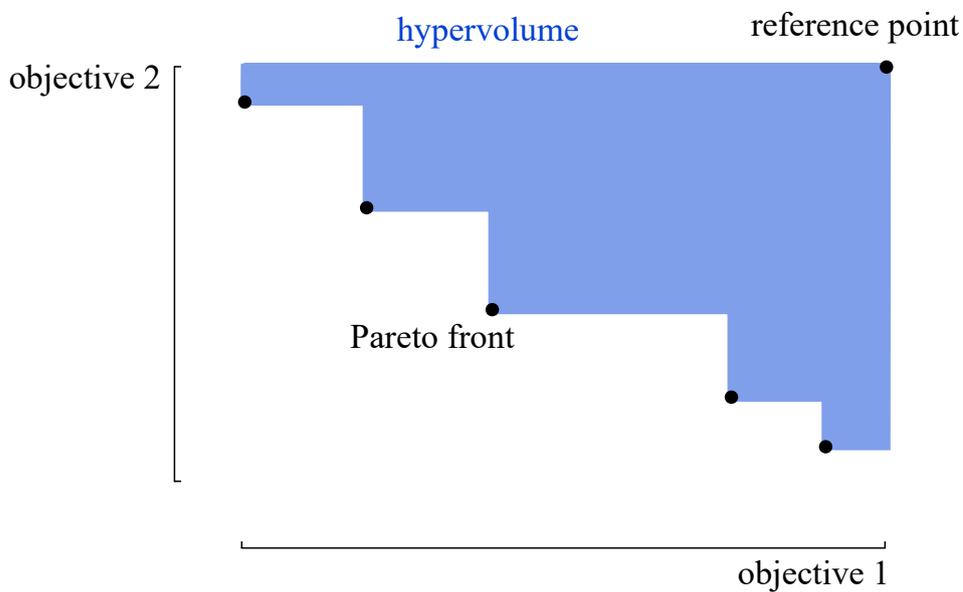
front  $F_1$  are considered, as these design samples correspond to the best solutions ((3) in Figure 2.2). If  $F_1$  contains less than  $N$  design samples, all the design samples in  $F_1$  are stored in the new population  $P_{t+1}$ . The remaining samples for the new population are retrieved from the subsequent non-dominated fronts, until adding the design samples from a specific front  $F_i$  leads to a new population with more than  $N$  design samples (this is inevitable, as the combined population includes  $2N$  design samples). In that case, the design samples in front  $F_i$  are sorted based on the crowding distance ((4) in Figure 2.2), and the remaining places in the new population are filled by the samples with the highest crowding distance among the samples in front  $F_i$  ((5) in Figure 2.2). From population  $P_{t+1}$ , again an offspring  $Q_{t+1}$  is created, and so on. The number of iterations for this process depends on the stopping criterion. Typical stopping criteria include direct termination criteria (e.g., limiting the time budget), derived termination criteria (e.g., stop when the difference between the best objective value and the mean of the objective values in the last generation is below a given threshold), and performance indicator termination criteria (e.g., the hypervolume metric, which represents the volume defined by the set of design values and a reference sample in the objective space, Figure 2.4) [116]. The direct termination criteria are conservative and realistic, meaning that the algorithm terminates within a finite number of iterations. However, the optimal values for the direct termination criteria require a good understanding of the behavior of the algorithm in the specific problem. Derived termination criteria and performance indicator criteria ensure a given performance of the solution when terminating the algorithm, but require information on the global optimum and do not ensure the termination of the algorithm within a finite number of iterations. Ghoreishi et al. [116] provided a comprehensive survey of the common stopping criteria in evolutionary algorithms.

### 2.2.4 Surrogate-assisted robust design optimization main procedure

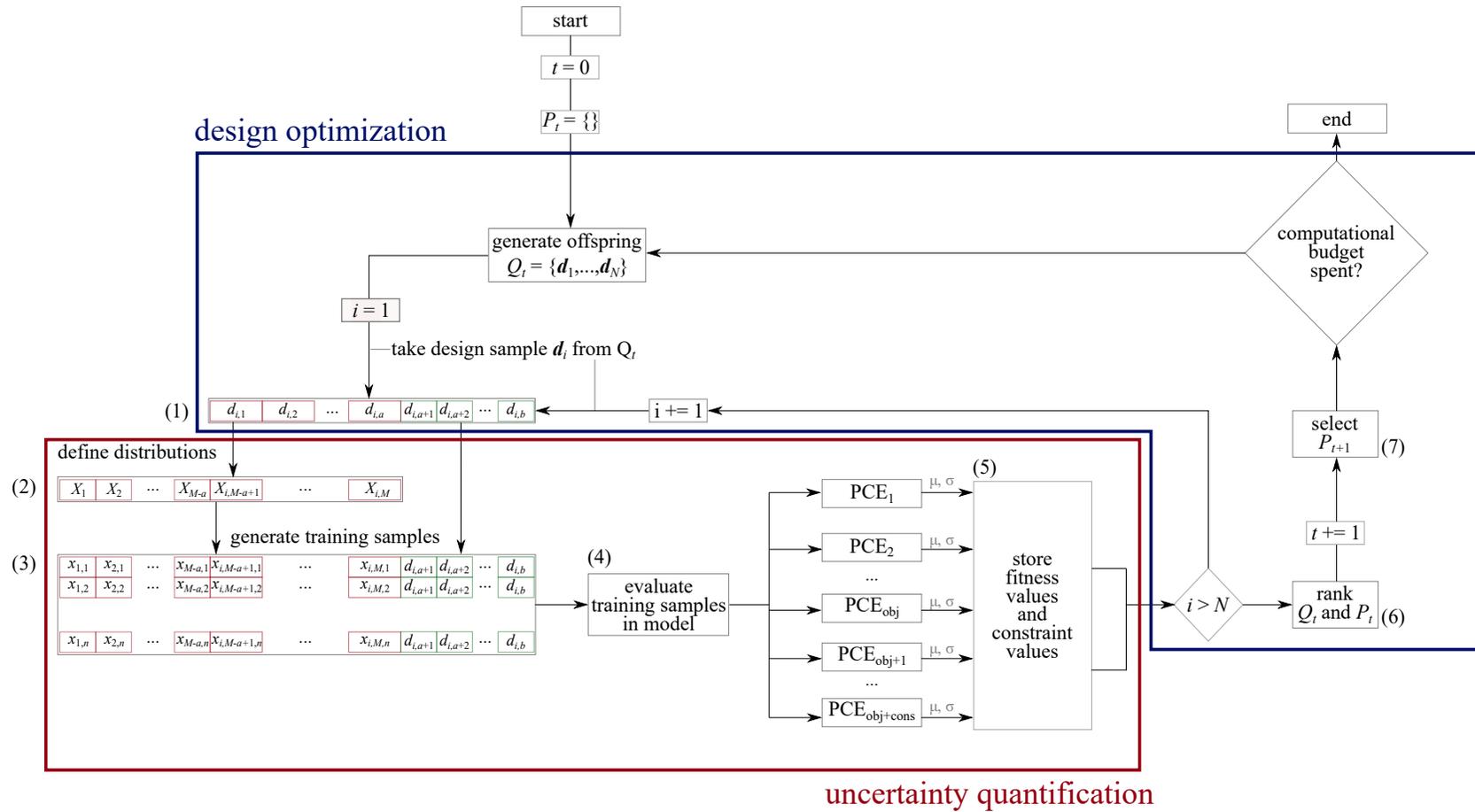
The surrogate-assisted RDO algorithm consists of NSGA-II to perform the optimization, while PCE is applied on each design sample to provide the statistical moments as fitness values (Figure 2.5). Like in NSGA-II, the first step consists of generating a first population of design samples via LHS. Each design sample  $\mathbf{d}_i$  is constructed by ( $a$ ) stochastic design variables and ( $b - a$ ) deterministic design variables, i.e.  $\mathbf{d}_i = (d_{i,1}, d_{i,2}, \dots, d_{i,a}, d_{i,a+1}, \dots, d_{i,b})$  ((1) in Figure 2.5). The values for the stochastic design variables ( $d_{i,1}, d_{i,2}, \dots, d_{i,a}$ ) corre-



**Figure 2.3:** An illustration of Latin Hypercube Sampling for 5 design samples, from two design variables with range  $[0, 1]$ . The design space is divided into 5 rows and columns, where each row and each column contains a single design sample.



**Figure 2.4:** In an optimization problem with two objectives, the hypervolume corresponds to the area surrounded by the design points on the Pareto front and the reference point.



**Figure 2.5:** The surrogate-assisted robust design optimization algorithm. After the generation of the offspring, a design sample  $\mathbf{d}_i$  is selected out of the offspring, which contains  $a + b$  values (1). Design variables  $d_1, \dots, d_a$  are considered uncertain, which means that the values in  $d_{i,1}, \dots, d_{i,a}$  are used as mean value for the distributions  $X_{i,M-a+1}, \dots, X_{i,M}$  (2). From the  $M$  distributions, the training samples are generated and each training sample is appended with the values from the deterministic design variables  $d_{i,a+1}, \dots, d_{i,b}$  (3). After evaluating the training samples in the deterministic model (4), the fitness values and constraint values (i.e., mean  $\mu$  and standard deviation  $\sigma$  from each PCE) are stored (5). This process is repeated until all  $N$  design samples in  $Q_t$  are processed. The offspring  $Q_t$  is combined with the population  $P_t$  (6) and the new population  $P_{t+1}$  is generated using the NSGA-II selection procedure (7). As long as the user-defined computational budget is not spent, a new offspring is generated and the entire process is repeated. Note that for the first iteration,  $P_0$  is empty and the offspring  $Q_0$  is generated through Latin Hypercube Sampling. After going through the uncertainty quantification loop,  $P_1 = P_0$ .

respond to the mean in the corresponding distributions  $(X_{i,M-a+1}, X_{i,M-a+2}, \dots, X_{i,M})$ . Combined with the distributions for the stochastic model parameters  $(X_1, X_2, \dots, X_{M-a})$ , the stochastic space is defined by a dimension  $M$  ((2) in Figure 2.5). This stochastic dimension  $M$ , and the polynomial order  $p$  (see Subsection 2.2.4.1 for the determination of the order), determine the size of the training samples required to construct the PCEs, according to the truncation scheme in Equation 2.14. These training samples are generated based on a Sobol quasirandom sequence [96] ((3) in Figure 2.5), which ensures an equal exploration of the stochastic space for each design sample. The training samples are evaluated in the deterministic model ((4) in Figure 2.5), the corresponding deterministic outputs are stored for each quantity of interest<sup>1</sup> and the PCE coefficients are quantified based on regression (Equation 2.19). Out of these coefficients, the mean and standard deviation are derived (Equation 2.21) and stored as fitness values or constraint values ((5) in Figure 2.5). This procedure is repeated for each design sample in the population. Finally, for each design sample, the mean and standard deviation on the quantities of interest are stored.

Out of the initial population, an offspring is created (Subsection 2.2.3.2), for which the same procedure is applied: a PCE is created for each quantity of interest and the corresponding mean and standard deviation are stored as fitness values. When the fitness values are quantified for each design sample in the uncertainty quantification inner loop, the design samples in the population and offspring are ranked based on their dominance in the objectives (Subsection 2.2.3.1, (6) in Figure 2.5). The top half of the design samples leads to the next generation of design samples ((7) in Figure 2.5). Out of this new generation, offsprings are again created and the new population and offspring are ranked, leading to the next generation. This process is repeated until the termination criterion is reached. In this procedure, a direct termination criterion is adopted, as it ensures reliability in the procedure (i.e., the algorithm terminates after a finite number of iterations). To evaluate the convergence of the final generation, the hypervolume metric is evaluated.

### 2.2.4.1 Determine the polynomial order

The maximum polynomial degree for the multivariate polynomials needs to be determined and its value should ensure accurate statistical moments on the quantity of interest, over

---

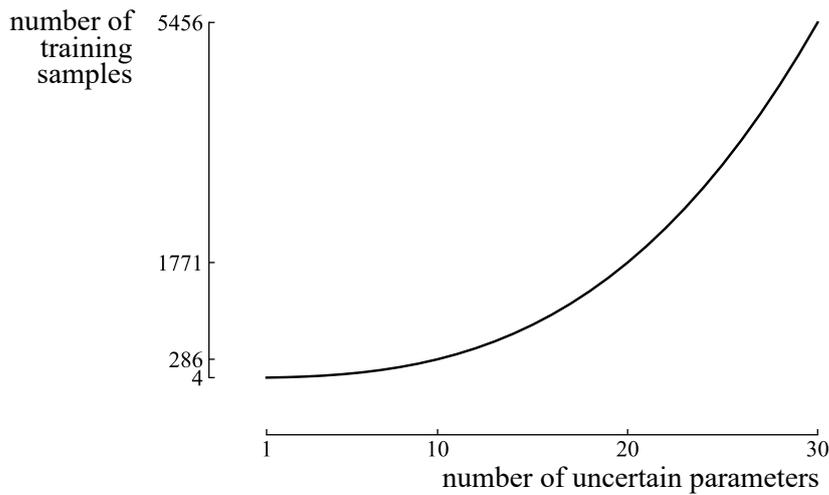
<sup>1</sup>Only one set of training samples is required to generate all the required PCEs for a design sample, as each sample generates a value for each quantity of interest.

## 2.2 Surrogate-assisted robust design optimization methodology

the entire design space [117, 118]. In other words, for each design sample provided by the optimization algorithm, the polynomial order should be sufficient to construct an accurate PCE. A fixed polynomial order is defined during the entire RDO process [117–119]. This polynomial order is defined based on a first screening of the design space. To screen the design space, LHS is used to construct a representative set of design samples. For each design sample, a PCE is constructed. If the worst-case LOO error among the design samples is below a certain threshold, then can the corresponding polynomial order be considered sufficient to generate accurate statistical moments during the RDO procedure. With this method, the polynomial order can be decided by the user. A fixed order ensures the finalization of the algorithm and avoids a large computational cost for a specific design sample when the LOO error is slightly higher than a fixed threshold. Nevertheless, a sufficiently large set of design samples should be selected, to approximate the worst-case LOO error over the design space and thus to make a well-founded decision on the polynomial order. The application procedure for this method is presented in an illustrative example in the Appendix, Section A.3.

### 2.2.4.2 Stochastic dimension reduction

The dramatic increase in model evaluations with the number of uncertain inputs in PCE is a computational burden (Figure 2.6). To address this issue, a stochastic dimension reduction is suggested, based on the methods of He et al. [120] and Turati et al. [121]. The method aims to determine the uncertain parameters with a negligible contribution to the standard deviation of the quantity of interest, and thus to consider these parameters as deterministic in the RDO algorithm. This contribution is determined by evaluating the total-order Sobol' indices for each stochastic parameter (Equation 2.30). As the contribution of a specific stochastic parameter might vary in different areas of the design space, a screening of the design space is performed by evaluating a set of design samples with size  $n_{\text{pop}}$ , generated via LHS. This results in  $n_{\text{pop}}$  values for the total-order Sobol' index  $S_j^{T,\text{PCE}}$  of each stochastic parameter  $j$  (i.e.,  $\mathbf{S}_j^{T,\text{PCE}} = (S_{j,1}^{T,\text{PCE}}, S_{j,2}^{T,\text{PCE}}, \dots, S_{j,n_{\text{pop}}}^{T,\text{PCE}})$ , where  $j = 1, 2, \dots, M$ ). If  $\max(\mathbf{S}_j^{T,\text{PCE}}) < S_{\text{thres}}$ , this parameter  $j$  can be discarded from the stochastic design space and, consequently, be considered deterministic.



**Figure 2.6:** Under the conventional truncation scheme for Polynomial Chaos Expansion, the number of training samples increases dramatically with the number of uncertainties considered. In this example, the polynomial order is fixed at 3.

### 2.3 Implementation

A surrogate-assisted RDO algorithm has been developed during this thesis, which has been released as an open-source framework, named Robust design optimization of renewable Hydrogen and dErIved energy cARrier systems (RHEIA) [77]. RHEIA includes the algorithms for multi-objective deterministic optimization (NSGA-II), RDO (NSGA-II + PCE) and uncertainty quantification (PCE). Moreover, algorithms are present to determine a sufficient polynomial order over the entire design space (Subsection 2.2.4.1) and to reduce the stochastic design space (Subsection 2.2.4.2). These algorithms can be applied to hydrogen-based energy systems, which are included in RHEIA. In addition, RHEIA allows connecting your own models to the algorithms as well. The application of these different methods is described in an extensive documentation [77], and the Python source code is located on Github<sup>2</sup>. The RHEIA framework includes several sub-packages:

```
RHEIA
+---CASES
+---OPT
+---POST_PROCESS
+---RESULTS
```

---

<sup>2</sup><https://github.com/rheia-framework/RHEIA>

```
\---UQ
```

The different cases can be characterized in *CASES*. The optimization and uncertainty quantification algorithms are coded in the sub-packages *OPT* and *UQ*. The results of applying one of these algorithms to the cases are stored in *RESULTS*, while the sub-package *POST\_PROCESS* provides a module to post-process the results (e.g., plot Pareto front, Sobol' indices, LOO error).

The PCE and NSGA-II algorithms are programmed in the *pce* and *genetic\_algorithms* modules, respectively:

```
RHEIA
+---OPT
|   |   genetic_algorithms.py
|   |   optimization.py
|   |   __init__.py
|
\---UQ
|   |   pce.py
|   |   uncertainty_quantification.py
|   |   __init__.py
```

These algorithms can be applied to predefined cases present in RHEIA, as well as to cases defined by the users. The following cases are present in RHEIA:

```
RHEIA
+---CASES
|   +---ENERGYPLAN
|   +---FOUR_BAR_TRUSS
|   +---H2_FUEL
|   +---H2_MOBILITY
|   +---H2_POWER
```

*H2\_FUEL*, *H2\_POWER* and *H2\_MOBILITY* are Python-based system models which represent three main valorization pathways of hydrogen as an energy carrier. In each model,

## Chapter 2. Surrogate-assisted robust design optimization

---

the system component capacities can be considered as design variables (e.g., photovoltaic array capacity, electrolyzer array capacity), while techno-economic and environmental objectives are quantified (e.g., levelized cost of electricity, carbon intensity). In addition to the system models, a method is described to generate climate data and demand data for the location of interest, as well as stochastic data on the model parameters (e.g., uncertainty on the grid electricity price, operating costs of the components and the energy consumption of a hydrogen-fueled bus). The *FOUR\_BAR\_TRUSS* case is present to illustrate how to connect your own Python-based system model. As RHEIA allows the evaluation of closed-source system models as well, connecting an EnergyPLAN [122] model is described as well in the documentation.

In each case folder, files to characterize and evaluate the case are required, e.g., for the *H2\_FUEL* case:

```
RHEIA
+---CASES
|   +---H2_FUEL
|   |   |   case_description.py
|   |   |   h2_fuel.py
|   |   |   design_space
|   |   |   stochastic_space
|   |   |   __init__.py
```

Either the system model is present (Python-based model, e.g. *h2\_fuel.py*), or a Python wrapper is present to run the system model, which is then called by e.g., a closed-source executable file. The *design\_space* file includes the bounds for the design variables and the mean values for the model parameters. The *stochastic\_space* file includes the stochastic characterization of the random parameters (i.e., the distribution type and variance). Finally, the *case\_description* Python module operates as a Python wrapper, which enables the evaluation of the system model with the samples provided by the optimization or uncertainty quantification algorithms.

The surrogate-assisted RDO can be initiated by characterizing a dictionary and providing this dictionary as an argument to the *run\_opt()* function. To illustrate for the *H2\_FUEL* case:

---

```
1 import rheia.OPT.optimization as rheia_opt
2 import multiprocessing as mp
3
4 dict_opt = {'case':          'H2_FUEL',
5            'objectives':    {'ROB': (-1, -1)},
6            'objective names': ['LCOH', 'mh2'],
7            'objective of interest': ['LCOH'],
8            'stop':          ('BUDGET', 72000),
9            'n jobs':         int(mp.cpu_count() / 2),
10           'population size': 20,
11           'results dir':    'run_1',
12           'pol order':      2,
13           }
14
15 rheia_opt.run_opt(dict_opt)
```

---

In the dictionary, the case folder name *'H2\_FUEL'* is provided, followed by the optimization type *'ROB'* and the weights for both objectives, i.e., minimization for both the mean and standard deviation of the Levelized Cost Of Hydrogen (LCOH). The default settings for the crossover and mutation probability are adopted. A computational budget of 72000 model evaluations is selected as a stopping criterion and the number of available physical cores are used to parallelize the evaluations. The population contains 20 samples and the total polynomial order for the multivariate polynomials in the PCE is set to two. Similar to the crossover and mutation probability setting, Sobol' sampling is considered for the random samples to construct each PCE, which is a default setting. To run the optimization, the *run\_opt()* function is called. In a similar matter, deterministic design optimization and uncertainty quantification can be performed. The population and fitness values for each generation are saved in the folder *'run\_1'*. The result files include the population for each generation (*population*), fitness values for each design sample in each generation (*fitness*) and a status file that logs the computational budget spent and the number of generations (*STATUS*):

RHEIA

```
+---RESULTS
|   +---H2_FUEL
|   |   +---ROB
|   |   |   \---run_1
|   |   |   |       fitness
|   |   |   |       population
|   |   |   |       STATUS
```

The results can be plotted using the methods provided in the post-processing classes, defined in the post-processing module.

```
RHEIA
+---POST_PROCESS
|   |   lib_post_process.py
|   |   __init__.py
```

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

To illustrate the potential of RDO in renewable energy systems, the surrogate-assisted RDO framework has been applied on a directly-coupled photovoltaic-electrolyzer system. As mentioned in the chapter overview, this application is published in Applied Energy [78]. After a short introduction on the reasons for directly-coupled photovoltaic-electrolyzer systems in Subsection 2.4.1, the system model (Subsection 2.4.2), quantity of interest (Subsection 2.4.3), uncertainty characterization (Subsection 2.4.4) and stochastic dimension reduction (Subsection 2.4.5) are discussed. The RDO results follow in Subsection 2.4.6, and the section concludes with a summary of the main messages from RDO on this application (Subsection 2.4.7).

### 2.4.1 Introduction

Hydrogen provides a viable alternative as an energy carrier for long-term solar energy storage [123]. Among different renewable hydrogen production methods, water electrolysis

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

---

is the most developed and provides high-purity hydrogen [124]. In the framework of solar-powered water electrolysis, the direct coupling of PV energy systems with electrolyzers is widely studied [125–127]. When considering direct coupling, the inclusion of DC-DC converters with Maximum Power Point (MPP) tracking is avoided, which reduces the system complexity, cost and transmission loss. As a result, this approach results in the highest economic performance for low-power applications [126].

When designing a directly coupled PV-electrolyzer system, multiple objectives can be defined to determine the system performance. In Khalilnejad et al. [125], a hybrid PV and wind turbine system is coupled with a 10 kW electrolyzer. In their work, the system is designed such that the system excess power production, divided by the hydrogen production, is minimized. In Maroufmashat et al. [126], the electrolyzer stack is designed to minimize the annual energy loss and maximize the hydrogen production. In García-Valverde et al. [127], the number of electrolyzers in series is defined to maximize the energy transfer from 72 series-connected solar panels. From previous work, it is clear that the hydrogen production and energy loss are common PV-electrolyzer system objectives. While the hydrogen production is a technical performance indicator, the energy transfer loss is only an indirect economic performance indicator. Instead, the Levelized Cost Of Hydrogen (LCOH) is generally used to indicate the techno-economic performance of hydrogen-based energy systems [128]. In a recent work of Sayedin et al. [129], a techno-economic design optimization of a directly coupled PV-electrolyzer system is performed, aiming to minimize the energy transfer loss and LCOH by configuring the number of electrolyzers in series and parallel. Despite its clear significance (cost estimated around 11 €/kWh [130]), the hydrogen storage system is rarely addressed in directly coupled PV-electrolyzer optimization studies, as these studies focus on the feasibility of producing renewable hydrogen in this way. Moreover, water treatment (energy consumption ranging between 1.5 kWh/m<sup>3</sup> and 2.5 kWh/m<sup>3</sup>, cost ranging between 0.40 €/m<sup>3</sup> and 1.52 €/m<sup>3</sup> [131]) and potential oxygen revenues (59 €/Nm<sup>3</sup> for medical grade oxygen from retail level vendors, 0.069 €/Nm<sup>3</sup> for large industrial-scale producers [132]) are not addressed.

The characteristics of a techno-economic robust design of a directly coupled PV-electrolyzer system are not yet demonstrated. Next to that, the operating temperature is rarely considered as a design variable and no specific techno-economic uncertainties are applied. As the optimality of the direct coupling strongly depends on the electrolyzer stack configuration

and therefore on the assumed techno-economic parameter values, a surrogate-assisted RDO has been applied for various climate conditions to ensure an optimized, robust direct coupling of the PV-electrolyzer system. Next to the number of electrolyzers in series and parallel, which are common design variables in literature, the operating temperature is selected as an additional design variable. All technical and economic model parameters are considered uncertain and the ranges are based on literature rather than using a general assumption. Thereafter, a stochastic dimension reduction is applied to determine the parameters that contribute significantly to the objective variation, instead of relying on the vaguely defined modeller judgement. When the robust designs are characterized, their advantages are illustrated compared to the respective deterministic designs. For each robust design, the individual contribution of each uncertain parameter to the objective variation is quantified. The determination of these individual contributions allows the formulation of technical and economic guidelines to further enhance the robustness of the performance during real-life operations.

### 2.4.2 System modelling

The adopted models for the PV system and the electrolyzer stack are described, followed by the climate data of the three locations considered in this study.

#### 2.4.2.1 Photovoltaic array

To convert the solar energy into electric power, a PV array of 72 series-connected MSX-60 panels has been adopted in this thesis [127, 133]. A single PV cell is characterized based on a single diode model without parallel resistance, following the experimentally validated structure presented by González-Longatt [134].

The produced PV cell current  $I_{PV}$  depends on the photocurrent  $I_L$ , the diode current  $I_0$  and the series resistance  $R_s$ :

$$I_{PV} = I_L - I_0 \left( \exp \left( \frac{q(U_{PV} + I_{PV}R_s)}{n_d k T_{amb}} \right) - 1 \right), \quad (2.34)$$

where  $q$  is the electron charge,  $n_d$  the diode ideality factor and  $k$  is the Boltzmann constant. To solve this non-linear equation for  $I_{PV}$ , the Newton-Raphson numerical method is selected

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

for its fast convergence and accuracy [134, 135]. The photocurrent  $I_L$  depends on the solar irradiance  $G$  and ambient temperature  $T_{amb}$ :

$$I_L = I_{SC}(T_{1,nom}) \frac{G}{1000} + K_0(T_{amb} - T_1). \quad (2.35)$$

The coefficient  $K_0$  depends on 2 reference temperatures and the corresponding short-circuit currents  $I_{SC}$ :

$$K_0 = \frac{I_{SC}(T_2) - I_{SC}(T_1)}{T_2 - T_1}. \quad (2.36)$$

The diode current  $I_0$  is represented according to the following equation:

$$I_0 = I_0(T_1) \left( \frac{T_{amb}}{T_1} \right)^{\frac{3}{n}} \exp \left( \frac{qU_q(T_1)}{nk \left( \frac{1}{T_{amb}} - \frac{1}{T_1} \right)} \right), \quad (2.37)$$

where  $I_0(T_1)$  is given by:

$$I_0(T_1) = \frac{I_{SC}(T_1)}{\left( \exp \left( \frac{qU_{OC}(T_1)}{nkT_1} \right) - 1 \right)}. \quad (2.38)$$

The series resistance represents the internal losses:

$$R_s = -\frac{dU}{dI_{U_{OC}}} - \frac{1}{X_U}, \quad (2.39)$$

where the term  $dU/dI_{U_{OC}}$  is equal to  $1.15/2/N_{PV,s}$ , based on the PV panel current-voltage characteristic provided by the manufacturer [134].  $X_U$  is given by:

$$X_U = I_0(T_1) \frac{q}{nkT_1} \exp \left( \frac{qU_{OC}(T_1)}{nkT_1} \right) - \frac{1}{X_U}. \quad (2.40)$$

The shunt resistance, which corresponds to the leakage current to the ground, is commonly neglected and therefore it is not considered in this work [134, 136].

### 2.4.2.2 Electrolyzer stack

To produce hydrogen from an intermittent electricity supply, a Proton Exchange Membrane (PEM) electrolyzer has been selected, based on its fast response time ( $<1$  s) and full operational flexibility [88]. To characterize the PEM electrolyzer, the experimentally validated model of Garcia-Valverde et al. [137] has been adopted. To induce an electric current in the PEM electrolyzer, an operating voltage  $U_{\text{PEMEL}}$  has to be applied:

$$U_{\text{PEMEL}} = U_{\text{rev}} + U_{\text{electrodes}} + U_{\text{ohm}}. \quad (2.41)$$

$U_{\text{rev}}$  is reversible potential,  $U_{\text{electrodes}}$  the overpotential at the electrodes and  $U_{\text{ohm}}$  the ohmic overpotential. At atmospheric pressure, the reversible potential  $U_{\text{rev}}$  is found empirically:

$$U_{\text{rev}} = 1.5184 - 1.5421 \cdot 10^{-3} T_{\text{PEMEL}} + 9.523 \cdot 10^{-5} T_{\text{PEMEL}} \ln(T_{\text{PEMEL}}) + 9.84 \cdot 10^{-8} T_{\text{PEMEL}}^2. \quad (2.42)$$

While the reversible potential represents only the open-circuit voltage, an overpotential has to be applied to induce the electric charge transfer and to overcome the mass transport phenomena. This required overpotential is represented by the activation overpotential and diffusion overpotential respectively:

$$U_{\text{electrodes}} = U_{\text{act}} + U_{\text{diff}}. \quad (2.43)$$

However, the mass transport phenomena are relatively small at low current densities [137]. Therefore, this effect is negligible in commercial PEM electrolyzers and is not considered in this work. The activation overpotential is defined by the overpotential that is distributed over the anode and cathode, where respectively the reduction and oxidation take place:

$$U_{\text{act}} = U_{\text{act,a}} + |U_{\text{act,c}}|, \quad (2.44)$$

$$U_{\text{act,a}} = \frac{RT_{\text{PEMEL}}}{\alpha_a z F} \ln\left(\frac{i_a}{i_{0,a}}\right), \quad (2.45)$$

$$U_{\text{act,c}} = \frac{RT_{\text{PEMEL}}}{\alpha_c z F} \ln\left(\frac{i_c}{i_{0,c}}\right). \quad (2.46)$$

The activation overpotential depends on the ideal gas constant  $R$ , the operating temperature of the electrolyzer  $T_{\text{PEMEL}}$ , the coefficient of transferred electrons  $z$  ( $= 2$  for water electrolysis), the Faraday constant  $F$  ( $= 96485$  C/mol) and the current density  $i$ . The

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

---

electrode contribution to the overpotential is unequally divided. The anode exchange current density is generally much lower ( $\approx 1 \times 10^{-7}$  A/cm<sup>2</sup>) than at the cathode electrode ( $\approx 1 \times 10^{-3}$  A/cm<sup>2</sup>) [137]. Therefore, the cathode contribution is neglected. By using the common assumption of a symmetric reaction, the charge transfer coefficients  $\alpha$  are equal to 0.5. To define the exchange current density, an Arrhenius expression is applied, which characterizes the effect of the operating temperature on the current density:

$$i_0 = i_{0,\text{ref}} \exp\left(-\frac{E_{\text{exc}}}{R} \left(\frac{1}{T_{\text{PEMEL}}} - \frac{1}{T_{\text{ref}}}\right)\right). \quad (2.47)$$

Where  $E_{\text{exc}}$  is the activation energy for the electrode reaction. The final overpotential that defines the operating voltage is the Ohmic overpotential, which depends on the resistive losses:

$$U_{\text{ohm}} = iR_{\text{l}} = i(R_{\text{ele}} + R_{\text{ion}}). \quad (2.48)$$

The electronic contribution to the resistive losses is limited compared to the ionic contribution, since the membrane conductivity is significantly lower than the conductivity in the other components [137]. The ionic membrane resistance depends on the thickness and conductivity:

$$R_{\text{ion}} = \frac{t_{\text{m}}}{\sigma_{\text{m}}}. \quad (2.49)$$

Assuming a fully hydrated membrane, the membrane conductivity can be modelled using an Arrhenius expression at a reference temperature:

$$\sigma_{\text{m}} = \sigma_{\text{m,ref}} \exp\left(-\frac{E_{\text{pro}}}{R} \left(\frac{1}{T_{\text{ref}}} - \frac{1}{T_{\text{PEMEL}}}\right)\right), \quad (2.50)$$

where  $E_{\text{pro}}$  is the activation energy for the proton transport in the membrane. When combining the electrolyzers in parallel  $N_{\text{p}}$  and series  $N_{\text{s}}$ , the electrolyzer stack voltage and current are expressed as:

$$U_{\text{stack}} = N_{\text{s}}U_{\text{PEMEL}}, \quad (2.51)$$

$$I_{\text{stack}} = N_{\text{p}}I_{\text{PEMEL}}. \quad (2.52)$$

The hydrogen production is proportional to the current:

$$\dot{m}_{\text{H}_2} = \frac{N_s I_{\text{stack}}}{2000F} \eta_F, \quad (2.53)$$

where  $\eta_F$  is the Faraday efficiency that represents the loss in hydrogen collection due to leakage and electrons that do not participate in the decomposition reaction [138].

### 2.4.2.3 Climate data

To analyze the effect of different climate conditions on the system, this thesis considered hourly solar irradiance and ambient temperature data for one year for Bern in Switzerland (46.9°N, 7.4°E), San Francisco in the United States of America (37.8°N, -122.5°E) and Johannesburg in South-Africa (-26.2°N, 28.0°E). The hourly climate data is obtained from Meteonorm [139]. These locations are subject to different climate conditions, where Johannesburg achieves the highest total yearly solar irradiance and average yearly ambient temperature, followed by San Francisco and Bern (Table 2.2). Due to the variability in the climate conditions, the PV system in each location produces different PV system characteristics and MPPs over the year. The MPPs in Johannesburg are concentrated in the top-right region of the current-voltage graph, resulting in the highest yearly power production, while in the other locations the MPPs are more widespread (Figure 2.7).

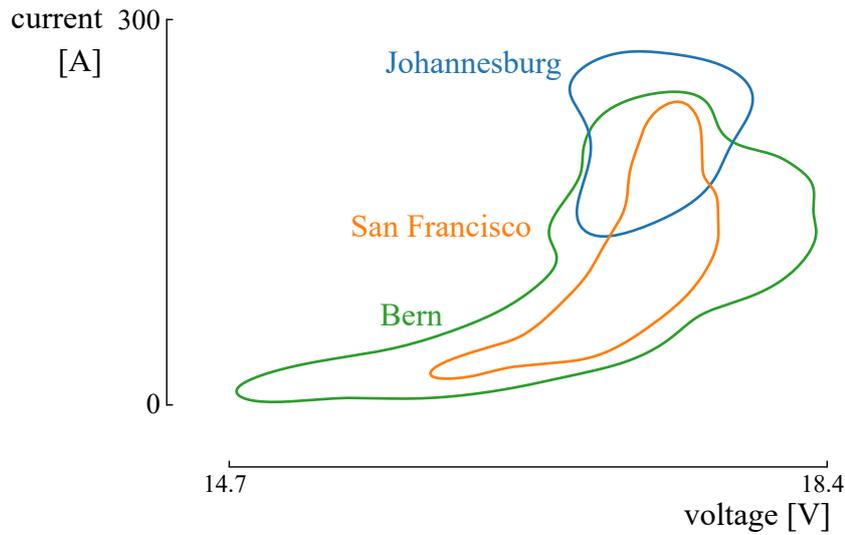
**Table 2.2:** The climate data for the considered locations illustrate that Johannesburg receives the highest total yearly solar irradiance and highest average ambient temperature.

Location	total solar irradiance $G_{\text{total}}$ kWh/m <sup>2</sup> /year	mean $\mu_G$ W/m <sup>2</sup>	standard deviation $\sigma_G$ W/m <sup>2</sup>	mean $\mu_{T_{\text{amb}}}$ °C	standard deviation $\sigma_{T_{\text{amb}}}$ °C
Bern	1243	143	262	10.7	8.0
San Francisco	1842	210	303	14.2	4.1
Johannesburg	2302	263	373	16.0	5.8

### 2.4.3 Design variables and quantity of interest

The solar irradiance is converted into electrical energy by the PV array. To convert the electrical energy into hydrogen, an electrolyzer stack is directly coupled to the PV array. Due to this direct coupling, the electrolyzer stack needs to be configured such that the

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

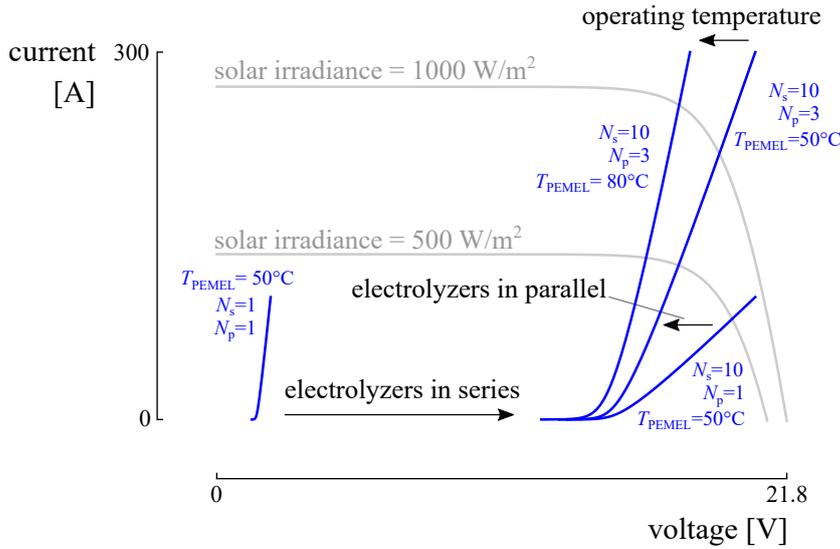


**Figure 2.7:** To illustrate the effect of the variability of the climate conditions, the region is marked where 50 % of the hourly produced Maximum Power Points (MPPs) over the year are situated. Clearly, during the year, the climate conditions in Johannesburg induce that 50 % of the MPPs produced by the PV array are located in the high-current, high-voltage region. The MPPs produced in San Francisco and Bern are more widespread towards lower currents and voltages, resulting in a lower overall power production during the year.

optimized energy transfer occurs from the PV panels to the electrolyzer stack. To find an optimized configuration for the electrolyzer stack for each considered location, the number of electrolyzers in series [126], the number of parallel strings of series electrolyzers [126] and the operating temperature [88] are selected as design variables:

$$\begin{aligned}
 1 &\leq N_s \leq 30, \\
 1 &\leq N_p \leq 30, \\
 323 \text{ K} &\leq T_{\text{PEMEL}} \leq 353 \text{ K}.
 \end{aligned}
 \tag{2.54}$$

By adapting these design variables, the electrolyzer stack characteristic is modified, resulting in a different intersection point for the PV system characteristic and electrolyzer stack characteristic and therefore a change in transferred power between both systems (Figure 2.8). To illustrate, increasing the operating temperature decreases the operating voltage for the same current density. Therefore, less electrical power is required for the same current flowing through the stack. At a higher operating temperature, the electrochemical reaction rate is improved, resulting in higher exchange current density and decreased voltage loss [140].



**Figure 2.8:** The voltage and current, transferred from the PhotoVoltaic (PV) array to the electrolyzer stack, are defined at the intersection point of the PV and electrolyzer stack voltage-current characteristic. The electrolyzer stack characteristic can be adapted by changing the stack configuration: The operating voltage is proportional to the number of electrolyzers in series ( $N_s$ ), while adding electrolyzers in parallel ( $N_p$ ) increases the limiting current of the electrolyzer stack and therefore the range of the characteristic. Finally, the stack characteristic is steepened by increasing the number of electrolyzers in parallel and increasing the operating temperature ( $T_{PEMEL}$ ).

To characterize the PV-electrolyzer system techno-economic performance for each design sample, the LCOH is quantified. The LCOH depends on the hydrogen production, CAPEX and OPEX of the installed PV panels and electrolyzers [141]:

$$LCOH = \frac{CRF(N_{PV}CAPEX_{PV} + N_{PEMEL}CAPEX_{PEMEL}) + N_{PV}OPEX_{PV} + N_{PEMEL}OPEX_{PEMEL}}{\sum_{i=1}^{8760} \dot{m}_{H_2,i}}, \quad (2.55)$$

where  $N_{PV}$  and  $N_{PEMEL}$  correspond to the PV array and electrolyzer array capacity, in  $kW_p$  and kW, respectively. CRF is the Capital Recovery Factor, which is characterized by the discount rate  $r$  and system lifetime  $n$ :

$$CRF = \frac{(r(1+r)^n)}{(1+r)^n - 1}. \quad (2.56)$$

The discount rate is adopted from literature assessing similar energy systems and includes the risk perception. The considered values for the different parameters are listed in Table 2.3.

### 2.4.4 Uncertainty characterization

To represent the random environment, the technical and economic parameters of the PV-electrolyzer system are characterized as uncertain (Table 2.3). For the system economics, the CAPEX and OPEX are considered uncertain. While the design optimization on system models employs at an early stage to evaluate the feasibility, the final investment stage usually occurs at a significantly later time, as it is preceded by a thorough planning study. Therefore, the proposed designs are vulnerable to changing market conditions, leading to uncertainty in the initial investment cost. Next to that, the OPEX can be considered uncertain, as unexpected operating costs might emerge. Considering uncertainty on the discount rate takes into account changes in finance types and policies [36]. The technical parameters of the PV-electrolyzer system are subject to commissioning and maintenance quality during operation. Moreover, inherent parameter variations are present, as well as measurement inaccuracy [36]. The limiting current  $i_{lim}$ , which forms a constraint on the operating range of the electrolyzer, is considered at its worst-case value. This will ensure, with a reliability of 3 times the standard deviation, that the proposed design will not violate the constraint due to the uncertain behavior of the model parameters. The climate data (i.e., solar irradiance and ambient temperature) can be considered uncertain due to measurement inaccuracy, inter-annual variability and spatial variability [36]. In this work, all uncertainties follow a Gaussian distribution.

### 2.4.5 Polynomial order and stochastic dimension reduction

A fixed polynomial order is selected for the construction of each PCE in the surrogate-assisted RDO algorithm (Subsection 2.2.4.1). A screening of the design space with 24 samples resulted in a LOO error below 1% for a polynomial order of 3 (Figure 2.9). When considering the full stochastic design space and polynomial order of 3, 2660 model evaluations are required to construct the PCE (Equation 2.14). To increase the computational efficiency, the stochastic dimension reduction method is applied (Subsection 2.2.4.2). The stochastic dimension reduction is applied for both the hydrogen production and LCOH separately, since the model parameter uncertainties contribute differently to each objective. For the LCOH, most economic parameters, the solar irradiance, membrane area and short-circuit current are the parameters contributing more than 2 % to the standard deviation in at least one considered design sample (Table 2.4). The CAPEX<sub>PV</sub> is a significant contributor

## Chapter 2. Surrogate-assisted robust design optimization

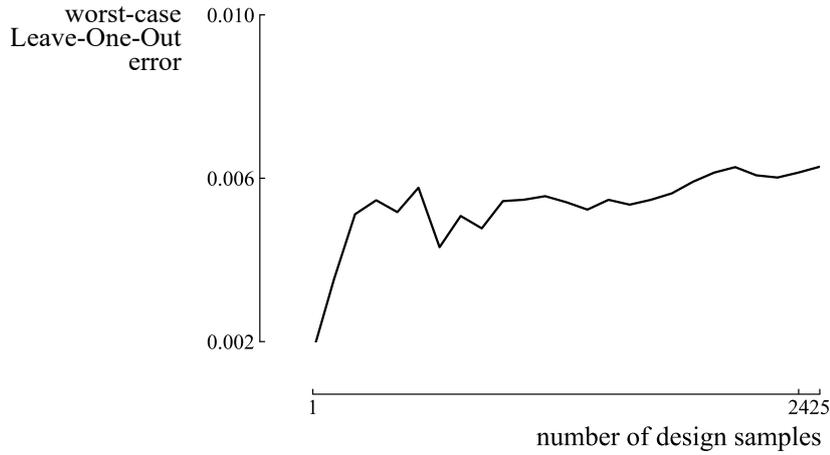
to the LCOH variance in PV-electrolyzer system designs with a relatively small amount of electrolyzers installed. In such a design, the LCOH depends mainly on the cost of installed PV panels, and therefore the LCOH variation depends mainly on the CAPEX<sub>PV</sub> variation. When the hydrogen production is selected as an objective, evidently the economic parameters do not contribute to the variation. Instead, the solar irradiance is the most dominant parameter (Table 2.4).

The uncertainty of only 8 and 7 parameters remain significant for the LCOH and hydrogen production, respectively. As a result, the total required number of model evaluations is reduced to 330 for 8 stochastic parameters and 240 for 7 parameters, as opposed to 2660 for the full stochastic space. Consequently, the computational cost is reduced by 88 % and 91 %, respectively, to achieve statistical moments for each considered objective.

**Table 2.3:** For the PV-electrolyzer system model, 18 techno-economic model parameters are considered uncertain.

Parameter	mean	3 standard deviation	unit	Ref.
CAPEX <sub>PV</sub>	780	520	€/kW <sub>p</sub>	[142]
OPEX <sub>PV</sub>	17.5	1.5	€/kW <sub>p</sub> /year	[142]
lifetime <sub>PV</sub>	25	5	year	[142, 143]
temperature coefficient short-circuit current $\mu_{I_{SC}}$	0.065	0.015	A/K	[133]
temperature coefficient open-circuit voltage $\mu_{U_{OC}}$	0.08	0.01	V/K	[133]
short-circuit current $I_{SC}$	3.8	0.1	A	[133]
open-circuit voltage $U_{OC}$	21.1	0.1	V	[133]
CAPEX <sub>PEMEL</sub>	1750	350	€/kW	[88]
OPEX <sub>PEMEL</sub>	4	1	% of CAPEX <sub>PEMEL</sub>	[88]
lifetime <sub>PEMEL</sub>	80	20	kh	[88]
discount rate $r$	6	4	% <sub>abs</sub>	[144, 145]
membrane area $A_m$	50	1	cm <sup>2</sup>	[137]
membrane thickness $t_m$	0.0051	0.0001	cm	[126]
Faraday efficiency $\eta_F$	99.5	0.5	%	[137]
temperature <sub>PEMEL</sub> $T_{PEMEL}$	variable	1	K	[146]
degradation voltage	6	2	μV/h	[88]
solar irradiance	hourly data	7	%	[139]
ambient temperature	hourly data	0.5	K	[139]

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system



**Figure 2.9:** For each set of design samples, the worst-case Leave-One-Out (LOO) error is quantified for a Polynomial Chaos Expansion of order 3. The worst-case LOO error remains below 0.01.

### 2.4.6 Results and discussion

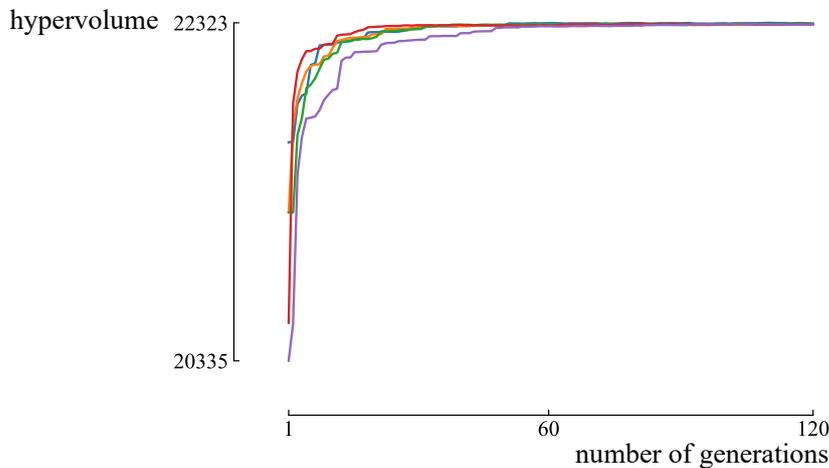
First, the results of the Deterministic Design Optimization (DDO) are presented and discussed, followed by a sensitivity analysis of the discount rate. Thereafter, the surrogate-assisted RDO is performed. To conclude, a sensitivity analysis of the discount rate is performed, considering only the operational uncertainties.

**Table 2.4:** The maximum Sobol' index above the threshold for each parameter over the considered design samples are presented. Therefore, these parameters are considered uncertain in the surrogate-assisted Robust Design Optimization (RDO). The CAPEX parameters, discount rate  $r$  and total annual solar irradiance  $G$  are the most contributing parameters to the Levelized Cost Of Hydrogen (LCOH) variation in at least one of the considered design samples. For the hydrogen production, the variation is mainly dominated by the total annual solar irradiance  $G$ .

LCOH		hydrogen production	
parameter	Sobol' index	parameter	Sobol' index
CAPEX <sub>PV</sub>	0.80	solar irradiance	0.91
discount rate $r$	0.51	membrane area $A_m$	0.18
CAPEX <sub>PEMEL</sub>	0.40	short-circuit current $I_{SC}$	0.12
solar irradiance	0.22	temperature <sub>PEMEL</sub> $T_{PEMEL}$	0.12
OPEX <sub>PEMEL</sub>	0.05	membrane thickness $t_m$	0.06
lifetime <sub>PEMEL</sub>	0.04	ambient temperature	0.05
membrane area $A_m$	0.03	lifetime <sub>PEMEL</sub>	0.02
short-circuit current $I_{SC}$	0.02		

### 2.4.6.1 Techno-economic Deterministic Design Optimization

The optimization algorithm is applied to the system, resulting in Pareto design samples that make a trade-off between minimizing the LCOH and maximizing the hydrogen production (Figure 2.11). The algorithm is configured with a crossover and mutation probability of 0.9 and 0.1, respectively. As the NSGA-II algorithm does not ensure mathematical optimality, several independent optimization runs were performed, and the corresponding hypervolumes were quantified (Figure 2.10). The optimization converged after 60 generations (i.e., the hypervolume stabilizes after 60 generations), which is comparable to the computational efficiency reported by Sayedin et al. [147] (71 generations for a population size of 20).



**Figure 2.10:** The hypervolume, presented for five independent optimization runs, stabilizes after 60 generations. The reference point is defined as [100 €/kg, 0 kg/year].

Despite the clear link between the power applied to the electrolyzer stack and the hydrogen production (Equation 2.53), a trade-off exists near the MPPs between maximizing the energy transfer efficiency from the PV array to the electrolyzer stack ( $=E_{\text{elec,year}}/E_{\text{MPP,year}}$ ) and maximizing the hydrogen production (Table 2.5) [126]. Consequently, operating exactly in the MPP does not induce the highest hydrogen production.

For the 3 considered locations, a configuration of 2 parallel-connected strings of electrolyzers, where each string consists of 10 electrolyzers in series, results in the minimum LCOH for the system. The optimized LCOH ranges from 6.3 €/kg (Johannesburg) to 10.5 €/kg (Bern), which is comparable to recently published LCOH for directly coupled PV-electrolyzer system designs in various cities in Iran (between 7.32 €/kg and 9.46 €/kg) [129]. With 2 parallel strings of electrolyzers, the electrolyzer stack can handle up to 200 A, which is sufficient to

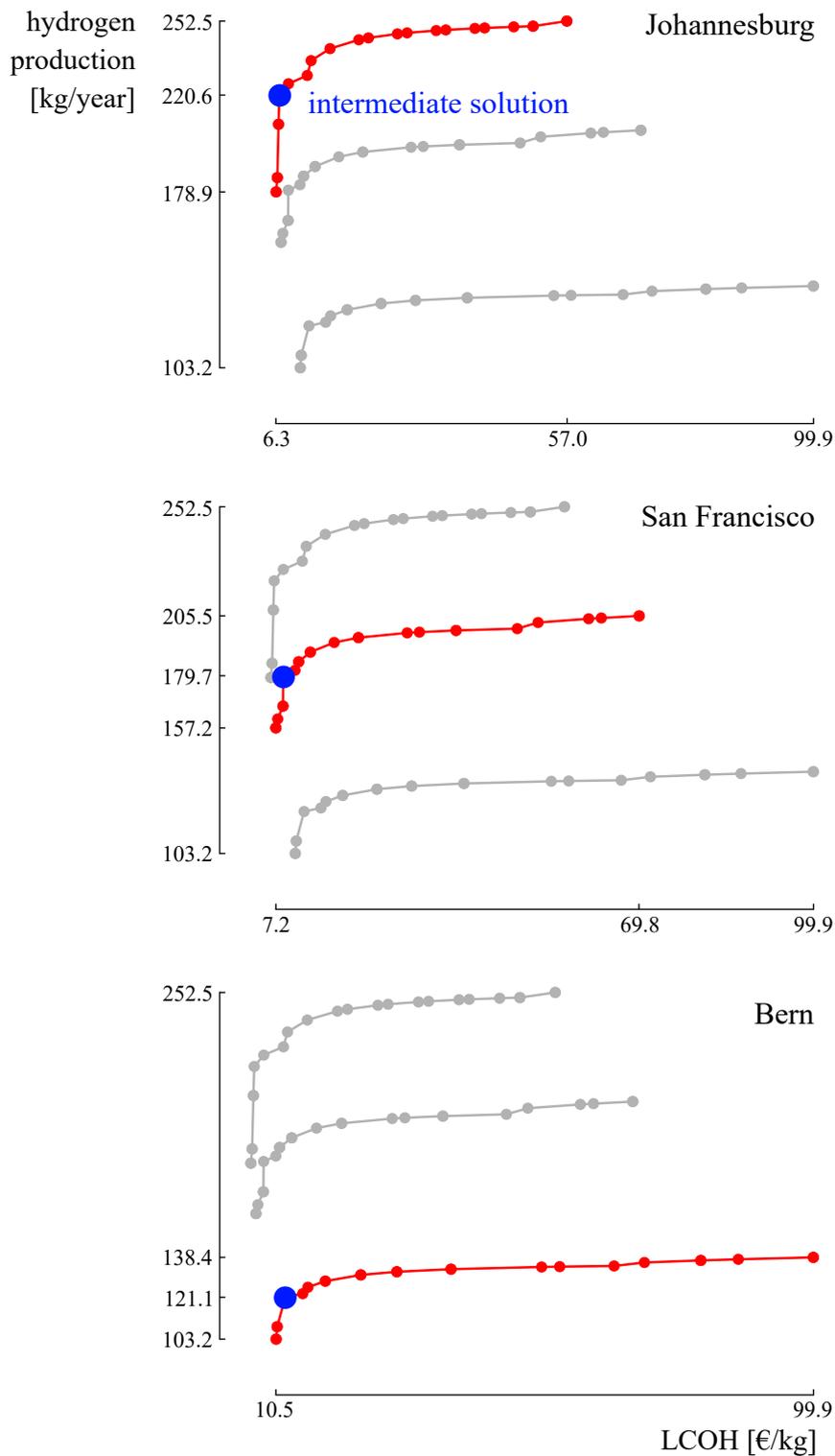
## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

---

achieve an optimized amount of operating hours for each location. To illustrate, for this configuration, in only 0.7 % of the time in Johannesburg, the produced PV power is situated outside the operating range of the electrolyzer stack. When 1 electrolyzer string in parallel is considered, this downtime is equal to 40.8 %. Next to the regulation of the limiting current of the electrolyzer stack by the number of electrolyzer strings in parallel, the operating voltage of the electrolyzer stack is mainly controlled by the number of cells in series. Since the operating voltage range near the MPPs is narrow for all locations (Figure 2.7), the optimized designs in each location consist of the same number of electrolyzers in series. The final design variable, the electrolyzer operating temperature, is slightly different:  $T_{\text{PEMEL}} = 333 \text{ K}$  in Johannesburg,  $T_{\text{PEMEL}} = 336 \text{ K}$  in San Francisco and  $T_{\text{PEMEL}} = 331 \text{ K}$  in Bern. Remarkably, the operating temperature in the optimized designs is selected below the highest temperature, where the minimum voltage losses would have occurred. By operating at these temperatures, the intersection points between the PV array and electrolyzer stack characteristic is optimized, resulting in the optimized hydrogen production, even if a higher operating voltage is required compared to operating at the maximum temperature. To illustrate, for this electrolyzer stack configuration, operating out of the optimized operating temperature can increase the LCOH up to 78 % in Johannesburg (Figure 2.12).

The PV-electrolyzer system in Johannesburg achieves the lowest LCOH out of all locations, leading thus to the highest hydrogen production, since the investment cost of the electrolyzer stack is equal in all 3 locations. Indeed, the hydrogen production ranges between 103.2 kg/year in Bern and 178.9 kg/year in Johannesburg. Therefore, the hydrogen production is proportional to the total yearly solar irradiance, while the LCOH is inversely proportional to the total yearly solar irradiance. The hydrogen production can be maximized in each location by adapting the system design into a configuration of 30 parallel-connected strings of 12 electrolyzers in series and an operating temperature of 353 K in all locations. This configuration consists of 18 times more electrolyzers than the minimum LCOH design (360 electrolyzers instead of 20), while the resulting increase in hydrogen production ranges from 30.7 % (San Francisco) to 40.9 % (Johannesburg). Consequently, the LCOH is drastically increased ( $\approx 9$  times) when installing this configuration. As a trade-off between both extreme designs, an intermediate design configuration is presented, consisting of 3 parallel strings of 10 electrolyzers operating at 353 K (Table 2.5). This design significantly improves the hydrogen production in each location compared to the cost-efficient design, at the expense of an acceptable increase in LCOH (Figure 2.11).

## Chapter 2. Surrogate-assisted robust design optimization

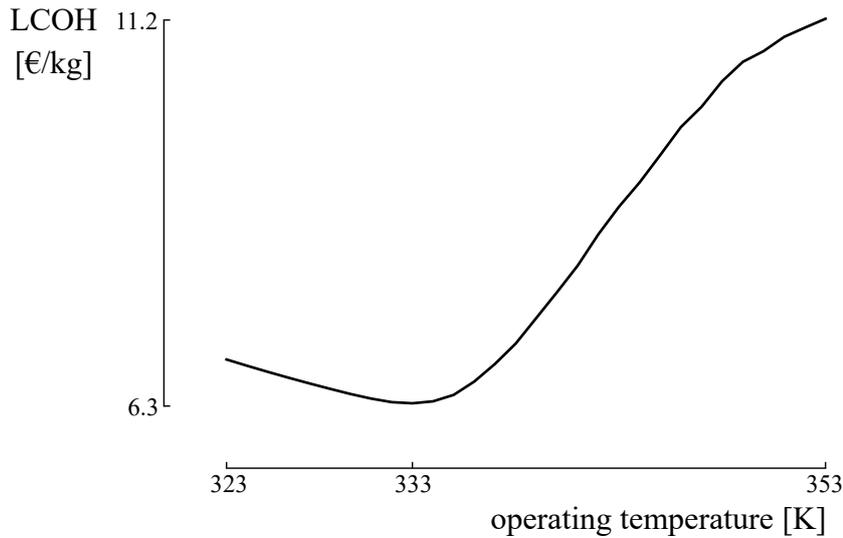


**Figure 2.11:** A trade-off exists between maximizing the hydrogen production and minimizing the Levelized Cost of Hydrogen (LCOH) in each considered location. Next to the 2 extreme designs, an intermediate solution is appointed as well. In Johannesburg, the optimized LCOH and hydrogen production is achieved.

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

**Table 2.5:** For each location, the same electrolyzer stack design leads to a maximum hydrogen production. To acquire the minimum Levelized Cost Of Hydrogen (LCOH), the number of electrolyzers is significantly lower, with slight differences in operating temperature for each location. As a trade-off between both extremes, an intermediate design is presented, which drastically increases the hydrogen production compared to the cost efficient design, at the expense of a slight increase in LCOH.

Location	$N_s$	$N_p$	$T_{PEMEL}$ K	LCOH €/kg	$\dot{m}_{H_2}$ kg/year	$E_{elec,year}/E_{MPP,year}$ %
Bern	10	2	331	10.5	103.2	93.5
	12	30	353	99.9	138.4	98.7
	10	3	353	12.1	121.2	97.9
San Francisco	10	2	336	7.2	157.2	95.8
	12	30	353	69.8	205.5	99.3
	10	3	353	8.5	179.7	98.4
Johannesburg	10	2	333	6.3	178.9	88.9
	12	30	353	57.0	252.5	98.4
	10	3	353	6.9	220.6	99.1

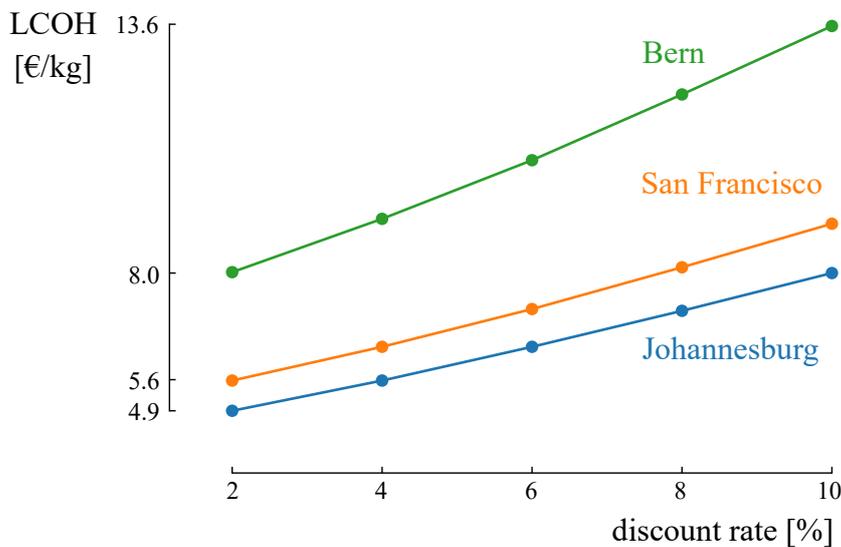


**Figure 2.12:** When performing a sensitivity analysis on the operating temperature of the cost-efficient design in Johannesburg, the LCOH can increase up to 78% when deviating from the optimized operating temperature of 333 K.

### 2.4.6.2 Sensitivity analysis of the discount rate

The discount rate of the PV-electrolyzer system is a highly uncertain parameter that has a significant effect on the LCOH [145]. Therefore, in this section, the effect of the discount rate on the LCOH is quantified and the potential LCOH reduction by de-risking the technology

is illustrated. For the design leading to the minimum LCOH in each location (Table 2.5), the discount rate is varied between 2 % and 10 %. A decrease in LCOH of 63 % is achieved in Johannesburg, 64 % in San Francisco and 69 % in Bern (Figure 2.13). Hence, this result demonstrates the importance of low-interest loan policies and technology de-risking for the economic maturity of PV-electrolyzer systems. Moreover, this result shows the increasing importance of reducing the discount rate in locations with a lower total yearly solar irradiance, since the LCOH reduction effect is the largest in Bern, followed by San Francisco and Johannesburg.



**Figure 2.13:** Increasing the discount rate from 2% to 10% increases the Levelized Cost Of Hydrogen (LCOH) of the PV-electrolyzer system by 63% in Johannesburg, 64% in San Francisco and 69% in Bern.

### 2.4.6.3 Surrogate-assisted Robust Design Optimization on the hydrogen production

The surrogate-assisted RDO algorithm is applied to the PV-electrolyzer system, to find a Pareto set of design samples that makes a trade-off between maximizing the mean hydrogen production and minimizing the standard deviation for each considered location (Figure 2.14). The converged Pareto front is achieved around 80 generations for the considered locations. As the most robust design corresponds to a negligible mean hydrogen production, this design is not considered. Instead, an intermediate design is presented for each location, which achieves a significant drop in standard deviation (from 43 % in Johannesburg up to 56 % in Bern), at the expense of a tolerable decrease in mean hydrogen

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

production (from 23 % in Johannesburg up to 34 % in Bern, Table 2.6). Clearly, the effect of a more severe constraint on the limiting current does not imply a reduction of the maximum mean hydrogen production, compared to the predicted deterministic hydrogen production. With 30 electrolyzer strings in parallel, this design is clearly oversized, reaching a limiting current far above the highest current produced by the PV system. Therefore, the energy from the PV system at each hour is transferred to the electrolyzer stack, even if the limiting current is a bit more severe.

**Table 2.6:** For each location, the same electrolyzer stack design leads to a maximum hydrogen production and a minimal Levelized Cost Of Hydrogen (LCOH), with slight differences in the operating temperature. An intermediate design is found, which drastically increases the hydrogen production compared to the cost-efficient design, at the expense of a slight increase in LCOH.

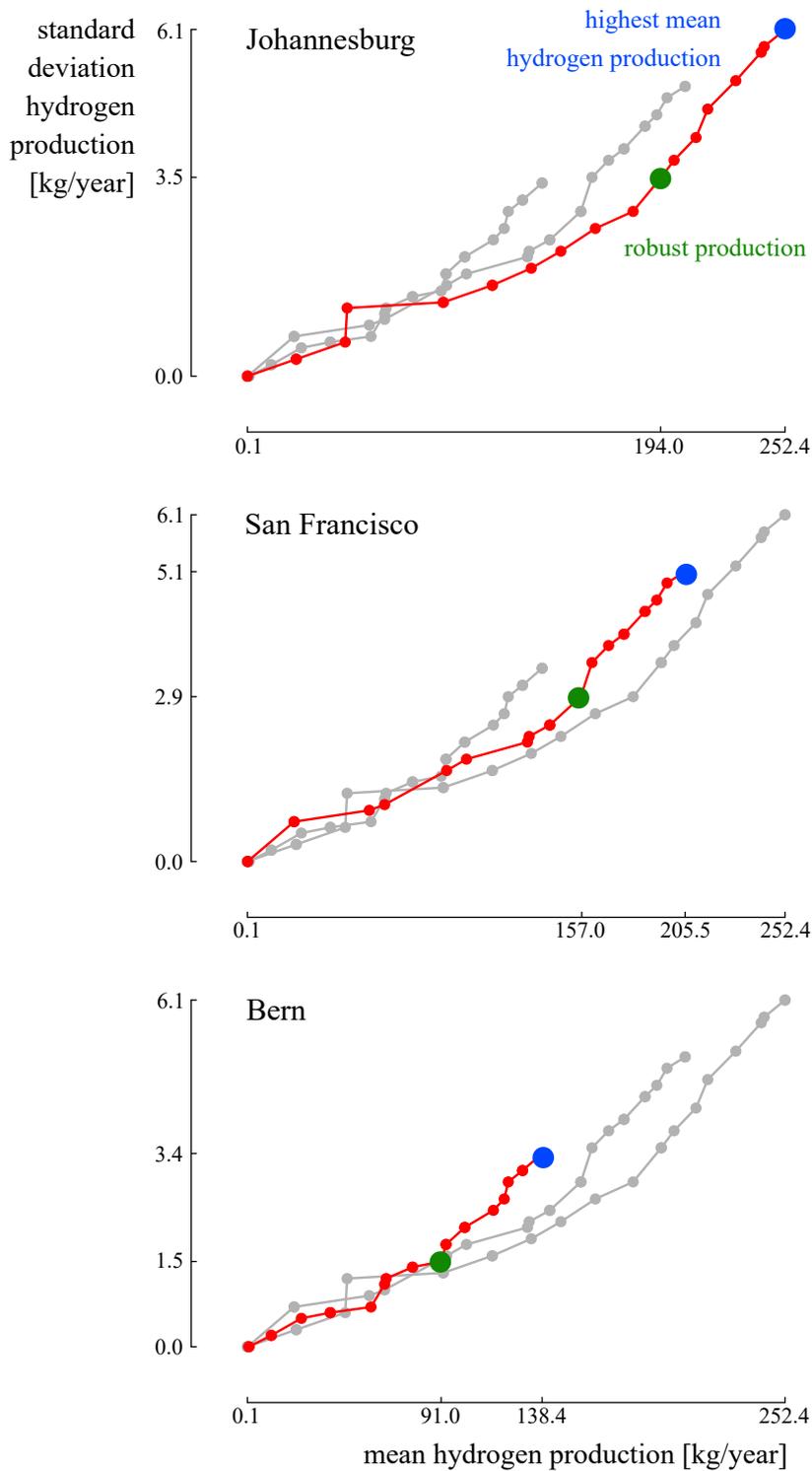
Location	$N_s$	$N_p$	$T_{PEMEL}$ K	$\sigma \dot{m}_{H_2}$ kg/year	$\mu \dot{m}_{H_2}$ kg/year
Bern	11	2	331	1.5	91.0
	30	12	353	3.4	138.4
San Francisco	11	2	349	2.9	157.0
	30	12	353	5.1	205.5
Johannesburg	11	2	351	3.5	194.0
	30	12	353	6.1	252.4

When quantifying the contribution of the stochastic input parameters to the standard deviation of the hydrogen production in the intermediate design, the solar irradiance is the dominating parameter (Table 2.7). Hence, increased accuracy in solar irradiance data and predictions is the most significant action to improve the quality and robustness of the hydrogen production.

### 2.4.6.4 Techno-economic surrogate-assisted Robust Design Optimization

Next to the hydrogen production, the surrogate-assisted RDO algorithm is applied for the LCOH. In this approach, no trade-off exists between minimizing the LCOH standard deviation and minimizing the mean LCOH. Hence, the robust design achieves the optimized mean LCOH as well. Clearly, the standard deviation of the LCOH is the lowest for the PV-electrolyzer system installed in the location with the highest total yearly solar irradiance (i.e., Johannesburg) (Figure 2.15). An optimized LCOH standard deviation of 0.74 €/kg is achieved in Johannesburg, followed by 0.84 €/kg in San Francisco and 1.26 €/kg in Bern.

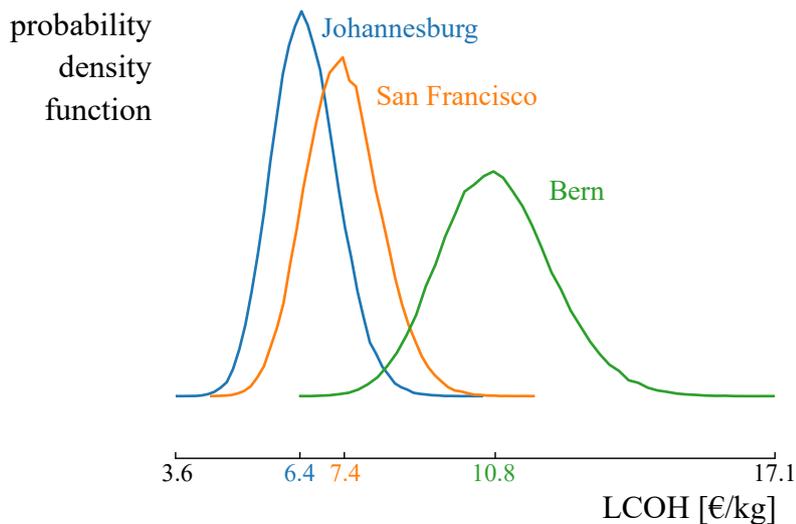
## Chapter 2. Surrogate-assisted robust design optimization



**Figure 2.14:** A clear trade-off exists between maximizing the mean hydrogen production and minimizing the standard deviation of the hydrogen production in all locations. The minimum standard deviation is found at a negligible production. Therefore, an intermediate, practical solution is presented as a robust design instead.

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system

Therefore, installing PV-electrolyzer systems in locations with high yearly total solar irradiance is not only beneficial for the mean LCOH, but also ensures optimized robustness. Due to the more severe constraint on the limiting current, the mean LCOH for each location is slightly increased compared to the LCOH achieved in the corresponding deterministic design (Table 2.5). Due to the reduced limiting current, the operating range of the electrolyzer stack is reduced, leading to a decrease in total transferred current from the PV system to the electrolyzer stack over the year, eventually leading to a decreased yearly hydrogen production. The uncertainty on the discount rate and CAPEX parameters dominates the LCOH standard deviation by 91 % (Table 2.7). Therefore, bulk manufacturing of these systems to reduce the CAPEX variation and more demonstration projects to stabilize the discount rates are the main actions to improve the LCOH robustness. During system operation (i.e., CAPEX and discount rate fixed), the electrolyzer lifetime and OPEX are the dominating parameters. Therefore, ensuring high-quality maintenance is the main action to ensure the LCOH robustness.



**Figure 2.15:** The mean LCOH is optimized in the location with the highest total yearly solar irradiance, Johannesburg. Moreover, the robust design in Johannesburg achieves the lowest standard deviation out of the robust designs in all considered locations.

### 2.4.6.5 Sensitivity analysis of the discount rate during system operation

When the PV-electrolyzer system is designed for an optimized LCOH, the variation is mainly dominated by the CAPEX parameters and the discount rate (Subsection 2.4.6.4). These

## Chapter 2. Surrogate-assisted robust design optimization

**Table 2.7:** The Sobol' indices are quantified for the robust design for the Levelized Cost Of Hydrogen (LCOH) and hydrogen production. The LCOH variation is dominated by the discount rate  $r$  and the CAPEX parameters of the PV-electrolyzer system. The uncertainty on the solar irradiance dominates the variation of the hydrogen production.

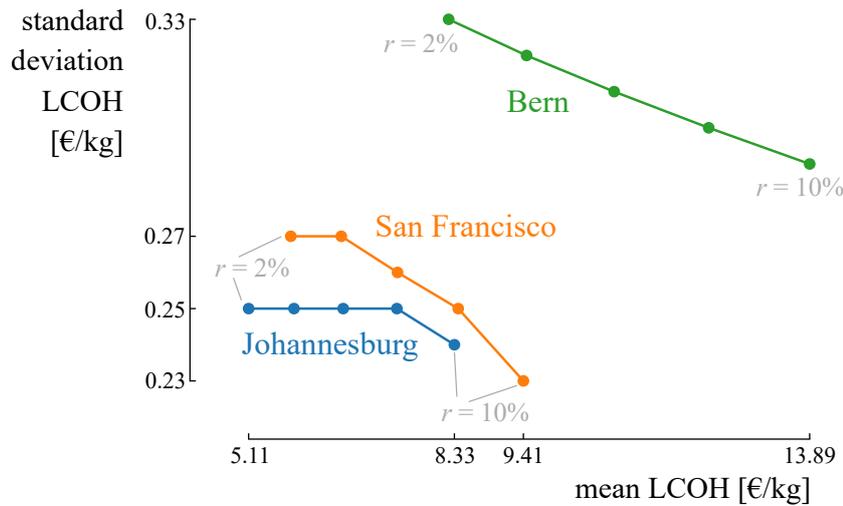
LCOH		hydrogen production	
parameter	Sobol' index	parameter	Sobol' index
discount rate $r$	0.52	solar irradiance	0.85
CAPEX <sub>PV</sub>	0.23	short-circuit current $I_{SC}$	0.04
CAPEX <sub>PEMEL</sub>	0.16	membrane area $A_m$	0.03
lifetime <sub>PEMEL</sub>	0.05	temperature <sub>PEMEL</sub> $T_{PEMEL}$	0.03
OPEX <sub>PEMEL</sub>	0.03	ambient temperature	0.02
solar irradiance	0.01	membrane thickness $t_m$	0.02
membrane area $A_m$	0.01		

parameters can be considered uncertain in the planning stage, but are fixed during system operation. Therefore, to illustrate the effect of the discount rate on the LCOH during system operation under uncertainty, the CAPEX parameters are considered deterministic, while different discrete non-probabilistic scenarios for the discount rate are evaluated. For each location, the robust design is presented for variable discount rates between 2 % and 10 % (Figure 2.16). As expected, an increasing discount rate escalates the mean LCOH. However, the standard deviation drops for an increasing discount rate. Moreover, this stabilizing effect increases when considering less optimal locations. Since the effect on the mean LCOH is significantly large compared to the gain in robustness, financing the system at the lowest discount rate remains the most beneficial situation.

### 2.4.7 Conclusion

Directly coupled photovoltaic-electrolyzer systems deliver a viable solution to store solar energy in the form of hydrogen. From a techno-economic perspective, such a system should achieve a minimum levelized cost of hydrogen. A configuration of 2 parallel strings of 10 electrolyzers in series results in a minimum levelized cost of hydrogen ranging between 6.3 €/kg (Johannesburg) and 10.5 €/kg (Bern). Despite the equal number of electrolyzers in each design, the operating temperature is different, ranging between 331 K and 336 K for the different locations. A sensitivity analysis of the discount rate from 10 % to 2 % illustrates that the levelized cost of hydrogen can be decreased by 63 % in the optimal location and 69 % in the least optimal location. Therefore, low-interest loan policies and de-risking of the

## 2.4 Application on a directly-coupled photovoltaic-electrolyzer system



**Figure 2.16:** Despite that the discount rate is proportional to the mean LCOH, it reduces the standard deviation and therefore provides a stabilizing effect for the LCOH of the PV-electrolyzer system. The discount rate is varied in steps of 2 %.

technology are crucial to improving the economic maturity of the photovoltaic-electrolyzer system, especially in suboptimal climate conditions.

When considering uncertainty on the techno-economic model parameters, a robust design for the hydrogen production is proposed, where a decrease in standard deviation is acquired (from 43 % in Johannesburg up to 56 % in Bern), at the expense of a reduction in mean hydrogen production (from 23 % in Johannesburg up to 34 % in Bern). The solar irradiance is the main contributor to the hydrogen production variation in the robust design, representing 85 % of the total variation. Consequently, increased solar irradiance data accuracy and more accurate predictions are the most significant improvements to increase the hydrogen production quality during real-life system operation.

No trade-off is found between minimizing the mean levelized cost of hydrogen and minimizing the standard deviation. The levelized cost of hydrogen statistical moments are optimal in the location with the highest yearly total solar irradiance (mean levelized cost of hydrogen of 6.4 €/kg at a standard deviation of 0.74 €/kg). Therefore, installing a photovoltaic-electrolyzer system in locations with a high total yearly solar irradiance is favourable for both the mean and robustness of the levelized cost of hydrogen. The main contributors to the levelized cost of hydrogen variation are the capital expenditure parameters and the discount rate, representing 52 % and 39 % of the total variation, respectively. During system operation, the most significant contributing parameters to the variation

## **Chapter 2. Surrogate-assisted robust design optimization**

---

are the operational expenditure and lifetime of the electrolyzer stack. Therefore, bulk manufacturing of the technology, promoting more demonstration projects and continuous high-quality maintenance are the most significant actions to improve the levelized cost of hydrogen robustness. An increased discount rate has a stabilizing effect on the levelized cost of hydrogen variation, resulting in a more robust performance. Since the stabilizing effect is small compared to the increase in mean levelized cost of hydrogen, financing the photovoltaic-electrolyzer system with the lowest discount rate is still considered optimal.

## 3 Sparse polynomial chaos expansion: handling large stochastic dimensions

### Chapter overview

- Implementation of the stepwise regression PCE algorithm in RDO;
- Application on a photovoltaic-battery-hydrogen system;
- Application of holistic RDO on a renewable-powered hydrogen refueling station.

*The application of RDO on the photovoltaic-battery-hydrogen system is published in Energy [148]. The application of holistic RDO is under review in Applied Energy.*

In design optimization of HRES, a holistic approach is often recommended [25, 30, 149], to deliver optimized solutions in a technical, economic, environmental and socio-political context. To realize such a holistic optimization, the effect of each HRES component on each objective should be captured (e.g., the energy efficiency, investment cost, CO<sub>2</sub> emissions during construction and job creation related to a PV array affect the technical, economic, environmental and socio-political performance, respectively). Consequently, the RDO of a multi-component HRES inevitably leads to a large stochastic dimension. The surrogate-assisted RDO, as presented in the previous chapter, suffers from the *curse-of-dimensionality*: with the conventional truncation scheme (subsection 2.2.2.1), the number of model evaluations increases drastically with the stochastic dimension. Hence, the computational cost becomes burdensome when the stochastic dimension is significant (i.e., more than 15 uncertain parameters). A stochastic dimension reduction method was suggested (Subsection 2.2.4.2), which identifies the stochastic parameters with a negligible contribution to the variance of the quantity of interest, and transforms these stochastic parameters

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

---

into deterministic parameters. The accuracy of this method is sensitive to user-defined constants (i.e., the number of design samples and the threshold). Therefore, in this chapter, an alternative UQ method is integrated into the surrogate-assisted RDO, to handle a large stochastic dimension, and thus, to perform RDO on multi-component HRES, in a holistic context.

In the conventional PCE, the truncation set  $\mathcal{A}^{M,p}$  contains multivariate polynomials, for which a large amount of these polynomials represent the interactions between input variables. However, in many engineering cases, the coefficients of the high-interaction terms (i.e., terms that represent the effect of interaction between two or more uncertain parameters) have a negligible value. This follows out of the heuristic statement that physical phenomena are mainly driven by main effects and low-order interactions [150]. Hence, performing deterministic model evaluations to determine these high-interaction coefficients through regression is inefficient. In the truncation scheme, a large number of coefficients correspond to high-interaction terms in the PCE. To illustrate, for a case with a stochastic dimension of 50 and a maximum polynomial degree of 3, the number of terms in the set  $\mathcal{A}^{50,3}$  corresponds to 23 426. In this set, only 150 polynomials depend on a single random parameter (i.e., univariate polynomials), which is less than 1 % of the total number of terms. Thus, a large number of coefficients may be negligible and should, therefore, not drive the computational cost of constructing the PCE (i.e., sparsity-of-effect principle [151]). Based on this principle, sparse PCE methods have been developed, trying to capture the significant coefficients a priori [76].

Initially, sparse PCE algorithms were developed in the context of compressive sensing [152]. Compressive sensing methods allowed to retrieve a sparse solution with fewer design samples than basis functions  $P$  [76]. Regression-based sparse PCE were first introduced with forward-backward selection procedure [153]. Blatman et al. [154] proposed an alternative truncation scheme in which only the multivariate polynomials which depend on a limited number of input parameters are included. The number of input parameters on which the multivariate polynomials may depend is a user-defined constant, defined a priori. A more relevant method is Least Angle Regression (LAR) [155]. In LAR, a hyperbolic truncation scheme is proposed in which terms with high-interaction order are excluded. LAR achieves promising performance in surrogate-assisted RDO [107]. Recently, Abraham et al. [156] proposed a stepwise regression method that adaptively determines the significant terms

from the pool of basis functions. The algorithm promises improved sparsity and improved accuracy in the sparse PCE than LAR. Despite its potential, the method has not been used in surrogate-assisted RDO. Therefore, the stepwise regression method is considered in the surrogate-assisted RDO in this chapter.

## 3.1 Contributions

The stepwise regression sparse PCE algorithm has been integrated in the surrogate-assisted RDO (Section 3.2). This updated surrogate-assisted RDO enables to handle a large stochastic dimension, as it tackles the curse-of-dimensionality related to the typical truncation scheme in PCE. Hence, with this updated method, complex, multi-component HRES can be designed with RDO in a holistic context. This updated method was applied in this thesis on two different HRES, to explore the advantages of robust designs for these HRES. In the first application, RDO is performed on the Levelized Cost Of Electricity (LCOE) of a PV-battery-hydrogen system (Section 3.3) [148]. The results illustrate that the economically preferred solution is a photovoltaic array (LCOE mean of 269 €/MWh), when the self-sufficiency ratio is irrelevant ( $\leq 30\%$ ). When a higher self-sufficiency ratio threshold is of interest, i.e. up to 59 %, photovoltaic-battery designs and photovoltaic-battery-hydrogen designs provide the cost-competitive alternatives which are least-sensitive to real-world uncertainty (LCOE standard deviation reduction up to 42 %, when compared to the optimized photovoltaic array). In the second application, an RDO was performed on the Levelized Cost Of Driving (LCOD) and Carbon Intensity (CI) for solar and wind-powered hydrogen refueling station, subject to technical, economic and environmental uncertainty (Section 3.4). By considering both objectives simultaneously, a holistic RDO is performed, enabling to provide designs that consider the stochastic techno-economic and environmental performance of the system. The results illustrate that a diesel-fueled bus fleet achieves an optimized LCOD mean. However, shifting towards a bus fleet that partly consists of hydrogen-fueled buses (54 % of the fleet) reduces the LCOD standard deviation by 36 %, the CI mean by 46 % and CI standard deviation by 51 %, at the expense of an increase in LCOD mean by only 11 %. In conclusion, the results illustrate that the surrogate-assisted RDO method enables performing RDO on HRES and to identify the advantages of the robust designs. These robust designs are characterized by an increased renewable penetration, e.g., switching from diesel to hydrogen as a fuel for mobility, combining battery and hydrogen storage with

intermittent renewable energy sources.

## 3.2 Stepwise regression Polynomial Chaos Expansion

The philosophy of sparse PCE is to integrate only the significant basis functions into the expansion, i.e., the basis functions that contribute significantly to the accuracy of the surrogate model. However, these significant basis functions are not known a priori. The stepwise regression method to determine the significant basis functions (and corresponding coefficients) is adopted from Abraham et al. [156].

The available basis functions, defined by the truncation scheme in subsection 2.2.2.1, are present in a pool of basis functions  $[\Psi_0, \Psi_1, \dots, \Psi_P]$ . The algorithm first creates an initial design of experiment  $\chi = [\xi^1, \xi^2, \dots, \xi^n]$ , where the size  $n$  is a user-defined constant. The design of experiment is evaluated in the deterministic model and the output for the quantity of interest is stored in  $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$ . In the first forward step of the algorithm, using the basis functions individually,  $P + 1$  independent one-predictor regression models are created and assessed. This assessment is performed by solving the least-squares problem:

$$\hat{u} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}, \quad (3.1)$$

where  $\mathbf{A}$  is the information matrix of size  $n \times 1$ , containing the response of a one-predictor regression model to each sample in the design of experiments. Following the dimension of  $\mathbf{A}$  (i.e. information on a single predictor is quantified), a single coefficient  $\hat{u}$  is estimated.

Out of the  $P + 1$  one-predictor regression models, the preferred basis function  $\Psi_{j^*}$  is selected as follows:

$$j^* = \operatorname{argmax}_j \left\{ \frac{|\hat{u}_j|}{\sqrt{\operatorname{Var}[\hat{u}_j]}}, j = 0, \dots, P \right\}, \quad (3.2)$$

where  $\hat{u}_j$  and  $\operatorname{Var}[\hat{u}_j]$  correspond to the estimate and the variance of the coefficient, respectively. Hence, the criterion favors a coefficient with a high weight (i.e., a high estimate) in most of the samples (i.e., a low variance). Hence, this corresponds to the coefficient that carries the most significant information. This criterion results in an effective and robust selection procedure. The best performing one-predictor model is stored in the final regres-

### 3.2 Stepwise regression Polynomial Chaos Expansion

sion model and it is removed from the pool of basis functions. The residual  $\hat{\boldsymbol{e}}$  is updated, based on the difference between the deterministic model results  $\boldsymbol{y}$  and the results from the final regression model in the design of experiment points  $\hat{\boldsymbol{y}}$ :

$$\hat{\boldsymbol{e}} = \boldsymbol{y} - \hat{\boldsymbol{y}}. \quad (3.3)$$

After the first forward step, the stepwise regression algorithm enters a forward-backward loop. The forward step is the same as the first forward step, except for the fact that the one-predictor regression models are fitted on the residual  $\hat{\boldsymbol{e}}_i$  quantified at iteration  $i$ . This fitting on the residual  $\hat{\boldsymbol{e}}_i$  enables to fit the response from the one-predictor regression models on the effects that are still missing in the final regression model. After each forward step, a backward step is initiated. In this backward step, the confidence intervals of the regression coefficients entered in the regression model are evaluated, to determine if the importance of the coefficients that already entered the final regression model is affected by the entry of a new coefficient. The confidence interval is evaluated as:

$$u_i \in \left[ \hat{u}_i \pm z_{[1-\alpha/2]} \sqrt{\text{Var}(\hat{u}_i)} \right], \quad (3.4)$$

where  $\text{Var}(\hat{u}_i)$  is the  $i$ -th diagonal term of  $\text{Var}(\hat{\boldsymbol{u}})$  and  $z_{[1-\alpha/2]}$  is the  $1 - \alpha/2$  quantile of the standard normal distribution and  $\alpha$  corresponds to 0.05. When a confidence interval includes 0, the corresponding coefficient is removed from the model and the residual is updated. This process is repeated until the maximum number of iterations is reached. Additional details on this sparse PCE method are described by Abraham et al. [156]. The post-processing of the final PCE, the determination of the statistical moments, LOO error and Sobol' indices, follows the same procedure as for a conventional PCE (Subsection 2.2.2).

The sparse PCE algorithm should be characterized in such a way that, for each design configured during the RDO, accurate statistical moments are provided for each quantity of interest. The two main parameters for the characterization of the sparse PCE are the maximum total order  $p$  and the number of training samples. The characterization of these two parameters is performed based on a screening method. First, a set of design samples is created using LHS, to ensure the exploration of the design space. For each design sample, a full PCE is constructed with a polynomial order equal to 1, and the LOO error is evaluated. The worst LOO error among the design samples is captured and

compared with a user-defined threshold value. If the worst-case LOO error is below the threshold value, the maximum total degree  $p$  is fixed at 1. If not, the process is repeated with  $p = 2$ . This goes on until the threshold value is reached. To determine the required number of training samples, the sample size is gradually increased, and for each size, a PCE is constructed with the stepwise regression method. For each sparse PCE, the mean and standard deviation are compared with the reference statistical moments, in this case retrieved by a conventional PCE method. Similar to the determination of the polynomial order, the design sample with the highest amount of training samples needed to achieve a converged sparse PCE determines the number of training samples adopted for each design sample during RDO. The application procedure for this method is presented in an illustrative example in Appendix, Section A.4.

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

In this application, RDO is performed on a PV-battery-hydrogen system. As this is a multi-component system, a large number of uncertainties are considered, which are handled with the updated surrogate-assisted RDO algorithm. As mentioned in the chapter overview, the results from this application are published in Energy [148]. After a brief introduction on PV-battery-hydrogen HRES (Subsection 3.3.1), Subsection 3.3.2 provides the lay-out of the HRES and the adopted models for each component, followed by the quantities of interest (Subsection 3.3.3), the characterization of the uncertainties (Subsection 3.3.4) and the characterization of the stepwise regression PCE (Subsection 3.3.5). The optimized designs, following the application of RDO on the system, are discussed in Subsection 3.3.6. The main messages of this section are provided in Subsection 3.3.7.

#### 3.3.1 Introduction

For intermittent balancing, battery energy storage and hydrogen energy storage provide a flexible, adequate solution [123]. Battery storage achieves a fast response time and a high round-trip efficiency ( $\approx 80\%$ , as opposed to  $\approx 35\%$  for hydrogen storage), which makes it suitable for short-term energy storage (i.e., from seconds to days) [123]. Nevertheless, due to their self-discharge rate and low energy density (30 Wh/kg), batteries are suboptimal for long-term energy storage (i.e., from weeks to months) [157]. Instead, hydrogen storage

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

---

is suggested for seasonal energy storage, as hydrogen storage operates with a negligible discharge rate, achieves a higher energy density (up to 10 kWh/kg) and provides a flexible capacity setting of power charging, energy storage and power discharging [123, 157]. In this framework, power charging is generally performed by splitting water into hydrogen and oxygen in a PEM electrolyzer [158], because of its fast response time ( $<1$  s) and full load flexibility [88]. The reverse reaction is generally performed in a PEM fuel cell to discharge the power [159]. The PEM fuel cell technology is widespread commercially, it operates at low temperature ( $70^{\circ}\text{C}$ -  $100^{\circ}\text{C}$ ) and achieves high power densities (up to  $2\text{ W}/\text{cm}^2$ ) [160].

Design optimization studies on HRES including battery storage and hydrogen storage are well-established [161]. Han et al. [162] described a mode-triggered droop scheme for an islanded PV-battery-hydrogen DC microgrid and showed that combining hydrogen storage and battery storage provides stronger environmental adaptability and higher PV utilization. Khiareddine et al. [163] performed a design optimization of a stand-alone PV-wind-battery-hydrogen system and concluded that an adequately sized fuel cell array allows a significant reduction in battery capacity and thus reduction of the total cost. Zhang et al. [164] evaluated different configurations of a stand-alone PV-wind-battery-hydrogen system and concluded that battery storage is the most cost-effective solution, while hydrogen storage is a reliable and non-polluting alternative, which can become economically advantageous after cost and efficiency improvements. Recently, also design optimization studies on grid-connected HRES including battery storage and hydrogen storage received attention. Parra et al. [165] illustrated the increase in on-site energy production of a PV-powered, grid-connected dwelling by considering battery and hydrogen storage, resulting in an additional annual income of £112 and £102, respectively. Pellow et al. [166] compared battery storage and hydrogen storage for grid-connected systems and concluded that hydrogen storage achieves a higher electrical Energy Stored On Invested than battery storage. Zhang et al. [157] optimized a grid-connected PV-battery-hydrogen system, considering several operation strategies and two different cost scenarios. Under the optimistic cost scenario, hydrogen storage induces a higher Net Present Value. Despite the clear advantage of this type of HRES, Eriksson et al. [25] highlighted that incorporating hydrogen in HRES design optimization was still an anomaly in 2017: only 5 out of 30 surveyed HRES studies incorporated hydrogen-based energy systems. Recently, Yue et al. [167] evaluated the publications related to hydrogen-powered energy systems and illustrated that the number of publications are doubled between 2015 and 2019, which indicates the increasing trend in

## **Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions**

---

the interest of hydrogen-based energy systems.

In design optimization studies of HRES, the optimal integration of battery systems and hydrogen-based energy systems in grid-connected applications received limited attention. To fill the research gap on design optimization under uncertainty of grid-connected, HRES including hydrogen storage and battery storage, this section provides the following main contributions: the significant techno-economic uncertain parameters are characterized by their uncertainty as described in literature; to handle this large stochastic dimension, the advantages of the stepwise regression sparse PCE algorithm (Section 3.2) are exploited for the first time in a surrogate-assisted RDO algorithm; the Cumulative Distribution Functions (CDF) of the optimized designs are used to compare their respective stochastic performances, which provides new insights into the probability of attaining an affordable LCOE when combining battery storage and hydrogen storage with a PV array.

### **3.3.2 System modelling**

The considered system is a grid-connected load, supported by an HRES (Figure 3.1). The HRES consists of a PV array, which is coupled to a DC bus bar through a DC-DC converter with Maximum Power Point Tracking. A battery stack and electrolyzer array with storage tank are integrated to store the excess of PV array electricity. A fuel cell array is implemented to produce electricity from the stored hydrogen. To transfer the DC electricity from the battery system and fuel cell to the AC load, a DC-AC converter is connected.

To set the hierarchy between the subsystems, a typical power management strategy is implemented, which primarily aims to satisfy the demand [168]. In this strategy, excess PV power is supplied to the battery stack. When the determined charge current violates the maximum charge current, the nominal charge current is considered instead. Then, or when the battery array reaches its maximum State Of Charge (SOC), the surplus energy is used to power the electrolyzer array. When the hydrogen tank is full, or when the surplus power lies outside the electrolyzer array operating range, the surplus energy is sold to the grid at the wholesale electricity price. In the opposite case, when the PV array does not cover the demand, the remaining demand is covered by the battery array, if the discharge current does not violate the maximum discharge current and the SOC remains above the SOC lower limit. If not, the nominal discharge current is extracted and the additional power

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

is supplied by the fuel cell array. When insufficient, the grid covers the remaining demand.

#### 3.3.2.1 Photovoltaic array

The PV array model proposed in Subsection 2.4.2.1 is experimentally validated for a 60 W<sub>p</sub>, MSX-60 PV panel. This PV panel achieves a rather low peak power, when compared to PV panels nowadays (around 200 W<sub>p</sub> - 300 W<sub>p</sub> per PV panel). Therefore, in this application, another PV array model is adopted, to model the performance of state-of-the-art PV panels. In this thesis, the single-diode model was imported from the open-source *PVlib* Python package [169]. The *PVlib* package is validated with experimental data and with results from existing commercial software [170]. The single-diode model provides the PV current and voltage:

$$I_{PV} = I_L - I_0 \left( \exp \left( \frac{U + IR_s}{n_{diode} N_s U_{th}} \right) - 1 \right) - \frac{U + IR_s}{R_{sh}}. \quad (3.5)$$

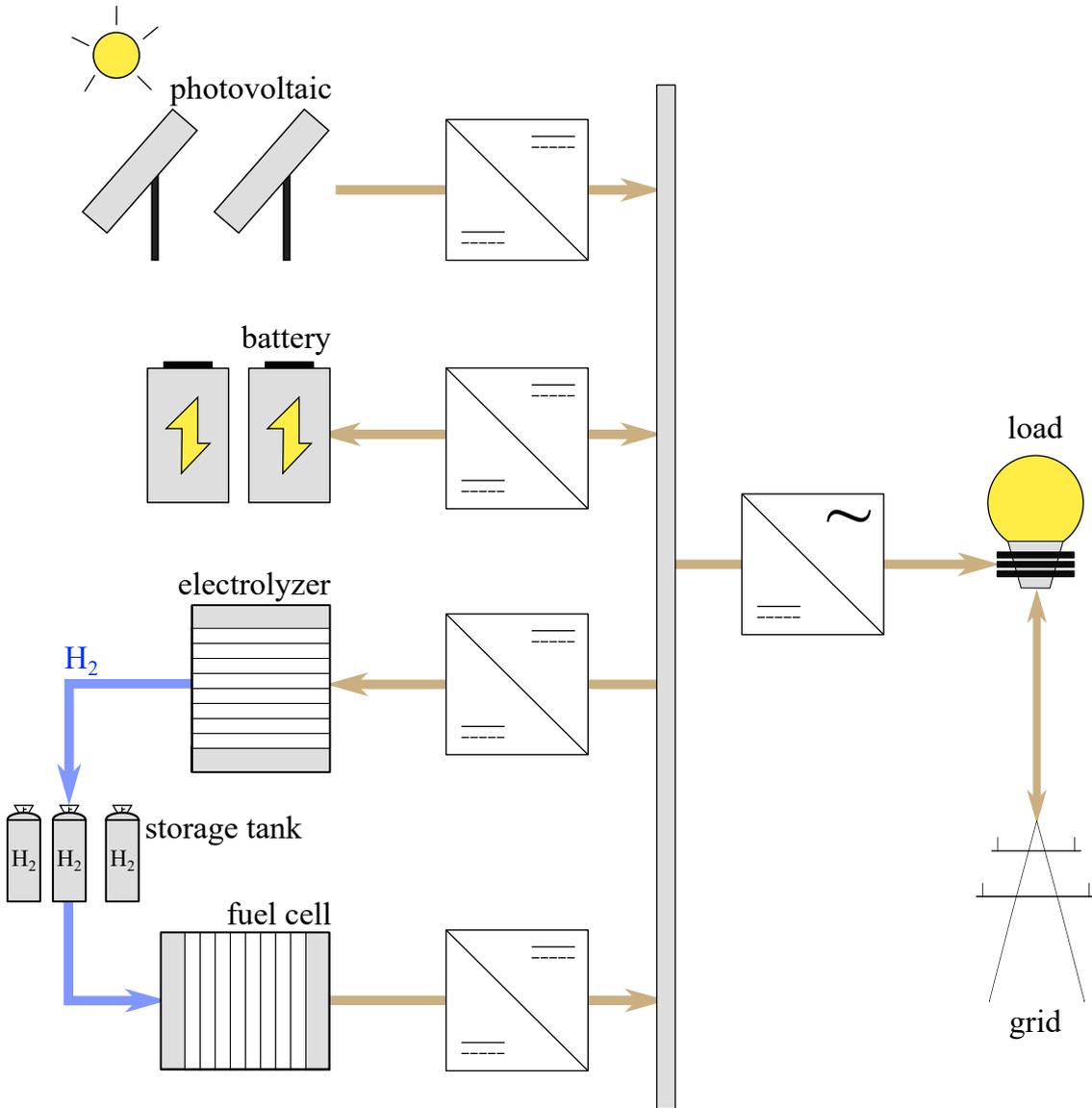
In this model, the parameters are determined based on manufacturer data, through the method developed by De Soto et al. [171]. Additional details on the physical behavior of the PV array are present in Section B.2.

#### 3.3.2.2 Battery stack

Lead-acid battery technology is selected in this work, as it is the most widespread technology and a mature lead-acid battery industry exists with significant learning rates during the last decades [172]. To characterize the performance of a lead-acid battery stack, the experimentally-validated model from Blaifi et al. [173] was adopted. The general voltage-current relation for a lead-acid battery is defined as:

$$U_{bat} = U_{bat,oc} + I_{bat} R_{bat}, \quad (3.6)$$

where the current  $I_{bat}$  is positive during charge and negative during discharge. The resistance component  $R_{bat}$  is different during charge and discharge and depends on the temperature, the current magnitude and the capacity. Hence, the characterization of the



**Figure 3.1:** The considered system includes the load connected to the grid and supported by a Hybrid Renewable Energy System (HRES). This HRES consists of a PhotoVoltaic (PV) array which converts solar energy into electricity. The battery stack and hydrogen-based energy system (electrolyzer, fuel cell and storage tank) enable the system to store an excess of PV array electricity and to comply, up to the available energy, with the demand when the solar irradiance is insufficient.

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

voltage during charge and discharge is given by:

$$U_{\text{ch}} = (2.085 - 0.12(1 - \text{SOC})) - \frac{I}{C_{\text{nom}}} \left( \frac{4}{1 + I^{1.3}} + \frac{0.27}{\text{SOC}^{1.5}} + 0.02 \right) (1 - 0.007\Delta T), \quad (3.7)$$

$$U_{\text{dch}} = (2 - 0.16 \text{SOC}) + \frac{I}{C_{\text{nom}}} \left( \frac{6}{1 + I^{0.86}} + \frac{0.48}{(1 - \text{SOC})^{1.2}} + 0.036 \right) (1 - 0.025\Delta T). \quad (3.8)$$

The State Of Charge (SOC) is the fraction of the total capacity stored in the battery:

$$\text{SOC}(t) = \text{SOC}_0 + \frac{1}{C(t)} \int_0^t \eta_{\text{bat}}(t) I(t) dt, \quad (3.9)$$

In this work, the minimum SOC is set at 20 % [130, 174].

I estimated the lifetime based on SOC variations through the commonly implemented Rainflow cycles counting method [175]. To avoid excessive reduction of the lifetime, over-charging and over-discharging are avoided, while the maximum charge current and maximum discharge current are limited to  $C_{\text{nom}}/10$  and  $C_{\text{nom}}/3.3$ , respectively [176]. Because of the constraints limiting the operation in the optimal operating zone to prolong battery life, a yearly capacity degradation rate [177] and a fixed energy efficiency of 80 % [178] was assumed.

#### 3.3.2.3 Electrolyzer array

The PEM electrolyzer model presented in Subsection 2.4.2.2 was validated under atmospheric pressure. In this application, a pressurized PEM electrolyzer is considered, as the pressurized hydrogen increases the energy density of the hydrogen storage tank. To determine the voltage-current characteristic and hydrogen flow rate of the pressurized PEM electrolyzer array, the experimentally-validated model from Abdin et al. [179] was selected, with an operating pressure of 30 bar [180]. The operating voltage is characterized according to the following equation:

$$U_{\text{PEMEL}} = U_{\text{PEMEL,oc}} + U_{\text{PEMEL,act}} + U_{\text{PEMEL,ohm}} + U_{\text{PEMEL,con}}. \quad (3.10)$$

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

In the remainder of this subsection, the subscript *PEMEL*, which refers to *PEM electrolyzer*, is left out for ease of reading. The open-circuit voltage follows out of the Nernst equation for electrolysis:

$$U_{oc} = (1.229 - 0.9 \times 10^{-3} (T - 298)) + \frac{R_u T}{2F} \left( \ln \left( \frac{p_{H_2} \sqrt{p_{O_2}}}{a_{H_2O}} \right) \right), \quad (3.11)$$

where  $R_u$ ,  $F$  and  $a$  represent the universal gas constant, Faraday constant and water activity between electrode and membrane, respectively. The activation overpotential  $U_{act}$  represents the voltage used to transfer electrons between the electrodes. By inverting the Butler-Volmer equation for the reactions at the electrode surface, the activation voltage can be quantified:

$$U_{act} = \frac{R_u T}{\alpha_{an} F} \operatorname{arcsinh} \left( \frac{i}{2i_{0,an}} \right) + \frac{R_u T}{\alpha_{cat} F} \operatorname{arcsinh} \left( \frac{i}{2i_{0,cat}} \right). \quad (3.12)$$

At high currents, a concentration overpotential  $U_{con}$  is created due to an excess of reactants (e.g., oxygen bubbles slowing down the reaction). To quantify this overpotential, a combination of the Nernst equation and Fick's law is adopted:

$$U_{con} = \frac{R_u T}{4F} \ln \frac{C_{O_2}^{mem}}{C_{O_2,0}^{mem}} + \frac{R_u T}{2F} \ln \frac{C_{H_2}^{mem}}{C_{H_2,0}^{mem}}, \quad (3.13)$$

where  $C^{mem}$  is the concentration at the membrane-electrode interface, and the subscript 0 refers to the reference working condition. The final overpotential that occurs in the PEM electrolyzer is the ohmic overpotential  $U_{ohm}$ , which is driven by the electric resistance of the electrodes  $R_{el}$ , bipolar plates  $R_{pl}$  and the membrane  $R_{mem}$ :

$$U_{ohm} = I (R_{el} + R_{pl} + R_{mem}). \quad (3.14)$$

The electric resistance of the electrodes and the flow plates can be quantified by applying Ohms law. I refer to Abdin et al. [179] for further details on the determination of the overpotentials. Finally, following the working point of the electrolyzer array, the hydrogen molar flow rate  $n_{H_2}$  is formulated as:

$$n_{H_2} = \frac{I}{2F}. \quad (3.15)$$

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

Despite the effect of intermittent loading on degradation and lifetime is not yet well quantified, this effect is reported to be negligible by several manufacturers [88]. Therefore, this effect is not considered in this study. Additional details on the physical behavior of this component is present in Section B.3.

#### 3.3.2.4 Hydrogen storage tank

The pressurized hydrogen produced in the electrolyzer is stored in a storage tank. Filling up the storage tank increases the tank pressure, until the electrolyzer outlet pressure is reached, according to the ideal gas law [181]:

$$p_t - p_{t,\text{init}} = Z \frac{N_{\text{H}_2} R_u T_t}{M_{\text{H}_2} V_t}. \quad (3.16)$$

The compressibility factor  $Z$  for  $\text{H}_2$  is equal to 1 at room temperature and moderate pressure (<100 bar) [181].

#### 3.3.2.5 Fuel cell array

To represent the conversion of hydrogen and oxygen into water in a PEM fuel cell array, the model from Murugesan et al. [182] was adopted, which is experimentally validated on the Ballard-Mark-V PEM fuel cell. The operating current depends on the converted hydrogen molar flow rate:

$$I_{\text{PEMFC}} = 2F n_{\text{PEMFC},\text{H}_2}. \quad (3.17)$$

The electric potential produced during water composition out of hydrogen and oxygen is equal to the Nernst potential minus the losses:

$$U_{\text{PEMFC}} = U_{\text{PEMFC},\text{Nernst}} - U_{\text{PEMFC},\text{act}} - U_{\text{PEMFC},\text{ohm}} - U_{\text{PEMFC},\text{con}}. \quad (3.18)$$

In the remainder of this subsection, the subscript *PEMFC* is left out for ease of reading. The Nernst equation determines the maximum fuel cell voltage and considers the operating temperature and reactant pressures:

$$U_{\text{Nernst}} = 1.229 - 0.85 \times 10^{-3} (T - 298.15) + 4.31 \times 10^{-5} T (\ln(p_{\text{H}_2}) + 0.5 \ln(p_{\text{O}_2})). \quad (3.19)$$

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

The activation losses  $U_{\text{act}}$  occur due to a low rate of charge transfer at lower current densities. This activation loss corresponds to:

$$U_{\text{act}} = -0.948 + 0.00354T + 7.6 \times 10^{-5} T \ln(C_{\text{O}_2}) + -1.93 \times 10^{-4} T \ln(I). \quad (3.20)$$

The ohmic losses  $U_{\text{ohm}}$  result from electrolyte resistance, contact resistance at the collector plates and the graphite electrodes. This loss is linearly dependent on the load:

$$U_{\text{ohm}} = iR_{\text{ohm}} = i \frac{\delta_{\text{mem}}}{\sigma_{\text{mem}}}, \quad (3.21)$$

where  $R_{\text{ohm}}$  is the electrical resistance that depends on the membrane thickness  $\delta_{\text{mem}}$  and membrane conductivity  $\sigma_{\text{mem}}$  (Nafion 117). The concentration loss occurs due to changes in concentration of reactants at higher current density regions:

$$U_{\text{con}} = -0.016 \ln \left( 1 - \frac{i}{i_{\text{lim}}} \right). \quad (3.22)$$

I refer to the work of Murugesan et al. [182] for the detailed quantification of these losses. Similar to the PEM electrolyzer, the effect of intermittent operation on the degradation and lifetime is not yet fully understood [183]. Additional details on the physical behavior of this component is present in Section B.4.

#### 3.3.2.6 Power conversion

The components are connected to a DC bus bar through DC-DC converters. To provide the power to the load or the grid, the DC bus bar is connected through a DC-AC inverter. The conversion efficiency of a DC-DC converter  $\eta_{\text{conv}}$  depends on the component output power:

$$P_{\text{DC,out}} = \eta_{\text{conv}}(P_{\text{DC,out}}) P_{\text{DC,in}}. \quad (3.23)$$

The conversion efficiency depends on the converter type. the experimentally buck-boost converter efficiency profile, presented by Taghvaei et al. [184], was considered for the PV array, PEM electrolyzer array and fuel cell array and a bidirectional buck-boost converter efficiency profile for the battery stack. Similar to the DC-DC converter, the DC-AC inverter

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

---

efficiency profile  $\eta_{\text{inv}}$  depends on the AC output power  $P_{\text{AC,out}}$ :

$$P_{\text{AC,out}} = \eta_{\text{inv}}(P_{\text{AC,out}}) P_{\text{DC,in}}. \quad (3.24)$$

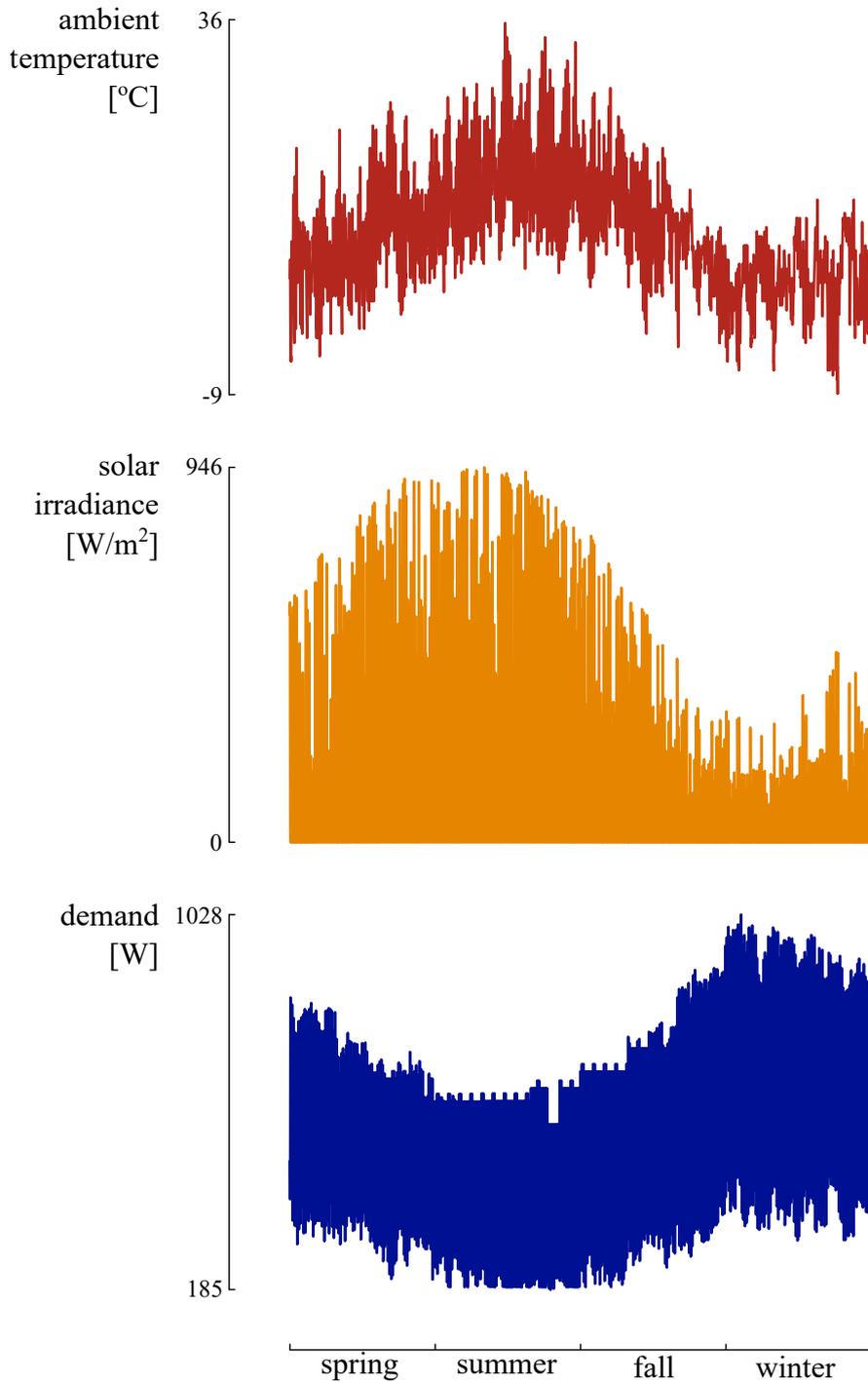
In this work, the inverter efficiency profile is determined by the experimentally-validated method described by Rampinelli et al. [185].

#### 3.3.2.7 Climate and demand data

The HRES performance depends on the climate and demand characteristics. In this application, a dwelling and a community (i.e., 2500 dwellings) are considered for Brussels, Belgium (Figure 3.2). When adopting climate data and demand data, it is important to adopt the climate data that was used to generate the demand data, as the solar irradiance and ambient temperature affect how the demand behaves. Therefore, Typical Meteorological Year data and hourly demand data were adopted from the National Renewable Energy Laboratory, as the former is used to construct the latter [90, 186]. As the data only exists for locations in the United States of America, the method presented by Montero Carrero et al. [187] was applied to adapt the climate and demand profiles to Belgium.

#### 3.3.3 Design variables and quantities of interest

To be able to find the designs that minimize these objectives, the capacity of the PV array, battery stack, PEM electrolyzer array, PEM fuel cell array and storage tank are selected as design variables. By considering the capacities as independent design variables, the optimization algorithm can exclude any technology from the HRES. Hence, the lower bound for the design variables is set at zero. The upper bounds for the design variables are 15 kW<sub>p</sub> for the PV array, 15 kWh for the battery stack, 10 kW for the PEM electrolyzer and PEM fuel cell, and 1 MWh for the hydrogen storage tank, per dwelling. In this thesis, no spatial constraints are considered for installing the proposed systems, as a generic dwelling is considered (e.g., a 15 kW<sub>p</sub> PV array needs around 75 m<sup>2</sup> of suitable area, which might not be available for every dwelling). This approach allows evaluating the raw potential of improving the LCOE robustness by installing an HRES. However, to avoid unreasonable designs and to avoid that the computational budget is spent on evaluating designs with unreasonably large capacities, the design variable upper bounds are selected based on the



**Figure 3.2:** The climate data (solar irradiance and ambient temperature) and demand data are inversely proportional, resulting in a reduced demand during spring and summer and a peak demand during fall and winter.

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

expected order of magnitude for the optimized designs. However, note that for all cases in this thesis, none of the optimized designs is constrained by the upper bound of the design variables.

To evaluate the techno-economic performance of the HRES, the LCOE is selected. The LCOE of the hybrid system reflects the system cost per unit of electricity covered [130]:

$$\text{LCOE} = \frac{\text{CAPEX}_a + \text{OPEX}_a + C_{\text{repl},a} + C_{\text{grid},a}}{\sum_{i=1}^{8760} P_{\text{load}}}. \quad (3.25)$$

To determine the system cost, the annualized investment cost  $\text{CAPEX}_a$ , operational cost  $\text{OPEX}_a$ , replacement cost  $C_{\text{repl},a}$ , difference between grid electricity cost and the gain from selling excess electricity  $C_{\text{grid},a}$  are evaluated.  $\text{CAPEX}_a$  represents the annualized capital expenses for every system component [130]:

$$\text{CAPEX}_a = \text{CRF} \sum_{k=0}^c N_k \text{CAPEX}_k, \quad (3.26)$$

where  $c$  refers to the different components (i.e., PV array, battery stack, PEM electrolyzer array, storage tank, PEM fuel cell array, DC-DC converters and DC-AC inverter) and  $N$  represents the corresponding installed capacity. The Capital Recovery Factor (CRF) is determined by the real discount rate  $i$  and the system lifetime  $L$  (20 years):

$$\text{CRF} = \frac{i(1+i)^L}{(1+i)^L - 1}. \quad (3.27)$$

The real discount rate considers the effect of inflation  $f$  on the nominal discount rate at the moment of loan  $i'$ :

$$i = \frac{i' - f}{1 + f}. \quad (3.28)$$

In addition to the investment cost, the system is subject to costs related to the component operation and maintenance. The annualized operating expenses are quantified through [130]:

$$\text{OPEX}_a = \sum_{k=0}^c N_k \text{OPEX}_k. \quad (3.29)$$

## Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

The system components come with different lifetimes. Consequently, some components might need replacement during the system lifetime, which induces additional cost. The annualized replacement cost is described as [130]:

$$C_{\text{repl,a}} = \text{CRF} \sum_{k=0}^c \left( N_k R_{c,k} \sum_{l=0}^{r_k} (1+i)^{-lt_k} \right), \quad (3.30)$$

where  $r_k$  is the number of replacements during the system lifetime for every component and  $t_k$  is the replacement period. Finally, when the system is unable to cover part of the demand, electricity is extracted from the grid. As a result, additional costs are included which relate to the electricity price. In this work, a fixed annual electricity tariff is adopted, which corresponds to the contract of nearly 75% of the households in Belgium [188]. The specific cost parameters values are listed in Table 3.1.

To indicate the fraction of the load that is covered by the HRES, the Self-Sufficiency Ratio (SSR) is quantified as a secondary performance indicator [157]:

$$\text{SSR} = 1 - \frac{E_{\text{grid}}}{E_{\text{load}}}, \quad (3.31)$$

where  $E_{\text{load}}$  is the total energy demand over the system lifetime. The SSR is an important factor for adopters of HRES, as it illustrates the resilience against large electricity price increases and the protection against power cuts, which are more likely in the future [189]. Moreover, reaching a significant SSR threshold is beneficial for grid operators, as it reduces the simultaneous power extraction from the grid and therefore reduces the risk of black outs.

### 3.3.4 Uncertainty characterization

To represent the uncertainty during design and operation of the HRES, the model parameters are characterized by a uniform distribution based on literature (Table 3.1). During operation, the system is subject to a natural variability of the technical parameters, e.g., fluctuating operating temperature and pressure, varying degradation speed and an uncertain component lifetime. Additionally, inter-annual variability is present in the electricity demand, solar irradiance and ambient temperature [36]. From an economic point of view, the system is subject to commissioning and maintenance quality, which affects the operat-

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

**Table 3.1:** The ranges for the stochastic parameters, used in the Robust Design Optimization (RDO) process.

parameter	min	max	unit	Ref.
CAPEX <sub>PV</sub>	430	780	€/kW <sub>p</sub>	[190]
OPEX <sub>PV</sub>	16	19	€/kW <sub>p</sub> /year	[142]
power tolerance <sub>PV</sub>	0	5	%	[191]
CAPEX <sub>PEMEL</sub>	1400	2100	€/kW	[88]
OPEX <sub>PEMEL</sub>	3	5	% of CAPEX <sub>PEMEL</sub>	[88]
replacement cost <sub>PEMEL</sub>	15	20	% of CAPEX <sub>PEMEL</sub>	[192, 193]
lifetime <sub>PEMEL</sub>	60	100	kh	[88]
temperature <sub>PEMEL</sub>	347	359	K	[194]
pressure <sub>PEMEL</sub>	29.9	30.1	bar	[195]
degradation <sub>PEMEL</sub>	4	8	μV/h	[88]
CAPEX <sub>tank</sub>	10.4	14.4	€/kWh	[196]
OPEX <sub>tank</sub>	1	2	% of CAPEX <sub>tank</sub>	[157, 192]
CAPEX <sub>PEMFC</sub>	1500	2400	€/kW	[197–199]
OPEX <sub>PEMFC</sub>	0.045	0.135	€/h	[197, 198]
replacement cost <sub>PEMFC</sub>	25	30	% of CAPEX <sub>PEMFC</sub>	[197, 198]
lifetime <sub>PEMFC</sub>	20	30	kh	[157, 174]
temperature <sub>PEMFC</sub>	347	359	K	[194]
pressure <sub>PEMFC</sub>	0.9	1.1	bar	[195]
degradation <sub>PEMFC</sub>	2	10	μV/h	[200]
CAPEX <sub>bat</sub>	102	354	€/kWh	[174]
OPEX <sub>bat</sub>	15	28	€/kWh/year	[130, 174]
replacement cost <sub>bat</sub>	61	141	€/kWh	[130, 174]
lifetime <sub>bat</sub>	500	2000	cycles	[174]
self discharge <sub>bat</sub>	0.1	0.3	%/day	[130]
degradation <sub>bat</sub>	3.5	4.0	%/year	[177]
CAPEX <sub>DCDC</sub>	100	200	€/kW	[201, 202]
OPEX <sub>DCDC</sub>	1	5	% of CAPEX <sub>DCDC</sub>	[203, 204]
CAPEX <sub>DCAC</sub>	342	519	€/kW	[190]
OPEX <sub>DCAC</sub>	1	5	% of CAPEX <sub>DCAC</sub>	[203, 204]
total solar irradiance	1082	1266	kWh/m <sup>2</sup> /year	[205]
average ambient temperature	9.4	11.2	°C	[205]
electricity demand	4.0	4.3	MWh/year	[89]
wholesale electricity price	46	97	€/MWh	[89]
fraction wholesale/retail price	20	40	%	[206]
inflation rate	1	3	%	[207]
nominal discount rate	4	8	%	[163, 208, 209]

## **Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions**

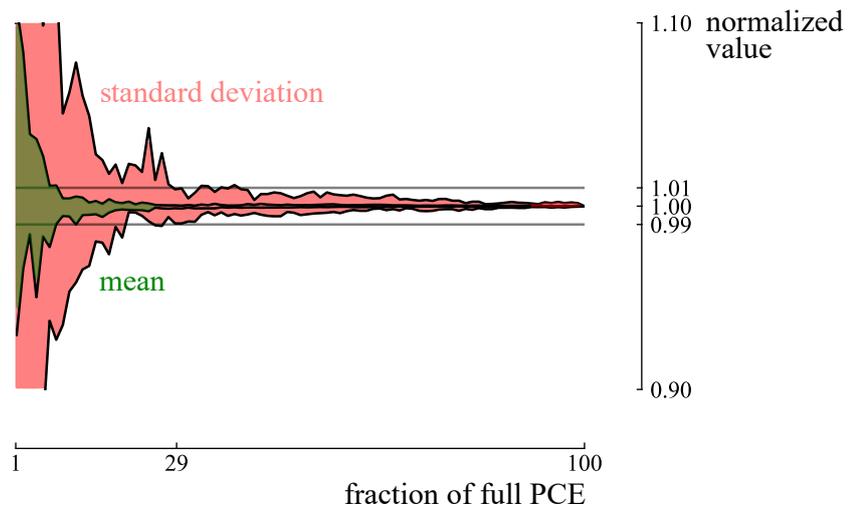
---

ing and maintenance cost, uncertain replacement cost due to evolving market conditions and a highly uncertain wholesale electricity price due to an evolving energy mix, improved energy efficiency and increased electrification of fossil-based energy sectors [89]. Otherwise, the discount rate and the investment cost can be considered deterministic or uncertain, depending on the actual step of the design process and the finance type considered [36]. Most studies assume a single-stage investment at the project start, which implies a deterministic characterization of these parameters. However, a change in the finance type of the project and a significant timeframe between the design stage and investment stage, which increases the possibility for the market conditions to change between the stages, implies uncertainty on the discount rate and investment costs. Therefore, in this work, two scenarios are assumed. The first scenario considers the discount rate and investment costs as deterministic (i.e., fixed market conditions), while the second scenario treats these parameters as uncertain (i.e., varying market conditions). As the results are significantly influenced by the design step and investment type, the handling of two scenarios enables comparing with both commonly adopted scenarios in scientific literature.

### **3.3.5 Stepwise regression PCE**

Following the large number of uncertainties considered (36 in the variable market conditions scenario and 28 in the fixed market conditions scenario), the sparse PCE method, presented in Section 3.2, was adopted in this thesis. To use the sparse PCE algorithm in the RDO algorithm, the polynomial order and the number of training samples need to be defined. Following a screening of the design space with 50 samples and a worst-case LOO error of approximately 1% among the design samples (Subsection 2.2.4.1), a polynomial order of 3 is selected. From these PCEs of order 3 (one for each design sample), the statistical moments are stored as reference values. Then, for each design sample, the number of training samples is gradually increased, to determine the number of training samples required to reach a similar accuracy on the statistical moments with the sparse PCE. The sparse PCE method achieves a similar accuracy ( $\pm 1\%$ ) on the statistical moments, for all the design samples, with only 29% of the required evaluations for PCE with the conventional truncation scheme (Figure 3.3).

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system



**Figure 3.3:** With 29 % of the training samples required for a conventional Polynomial Chaos Expansion (PCE), the sparse PCE achieves the statistical moments for a set of 50 design samples with an error below 1 %, when compared to the statistical moments from a conventional PCE. The areas for the mean and standard deviation are bounded by the highest and lowest (normalized) mean and standard deviation found among the set of design samples, respectively.

#### 3.3.6 Results and discussion

The RDO method is applied to the HRES to determine the optimized designs and to evaluate their performance under the variable market conditions scenario and fixed market conditions scenario. First, the Pareto set of optimized designs is presented, supplemented by a global sensitivity analysis to capture the driving uncertainties. Additionally, the stochastic design performance of these optimized designs is compared based on the CDF.

##### 3.3.6.1 Robust Design Optimization and global sensitivity analysis

The RDO method is applied four times independently on the HRES: for the dwelling and community in the fixed market conditions scenario and the variable market conditions scenario. In both scenarios, a trade-off exists between minimizing the LCOE mean and minimizing the LCOE standard deviation for the dwelling (Figure 3.4) and for the community (Figure 3.5), which is illustrated by the Pareto set of optimized designs. For each optimized design, the total Sobol' indices of the LCOE and the statistical moments of the corresponding SSR are quantified.

The minimum LCOE mean is achieved by a PV array (e.g., 2.7 kW<sub>p</sub> for the dwelling), which

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

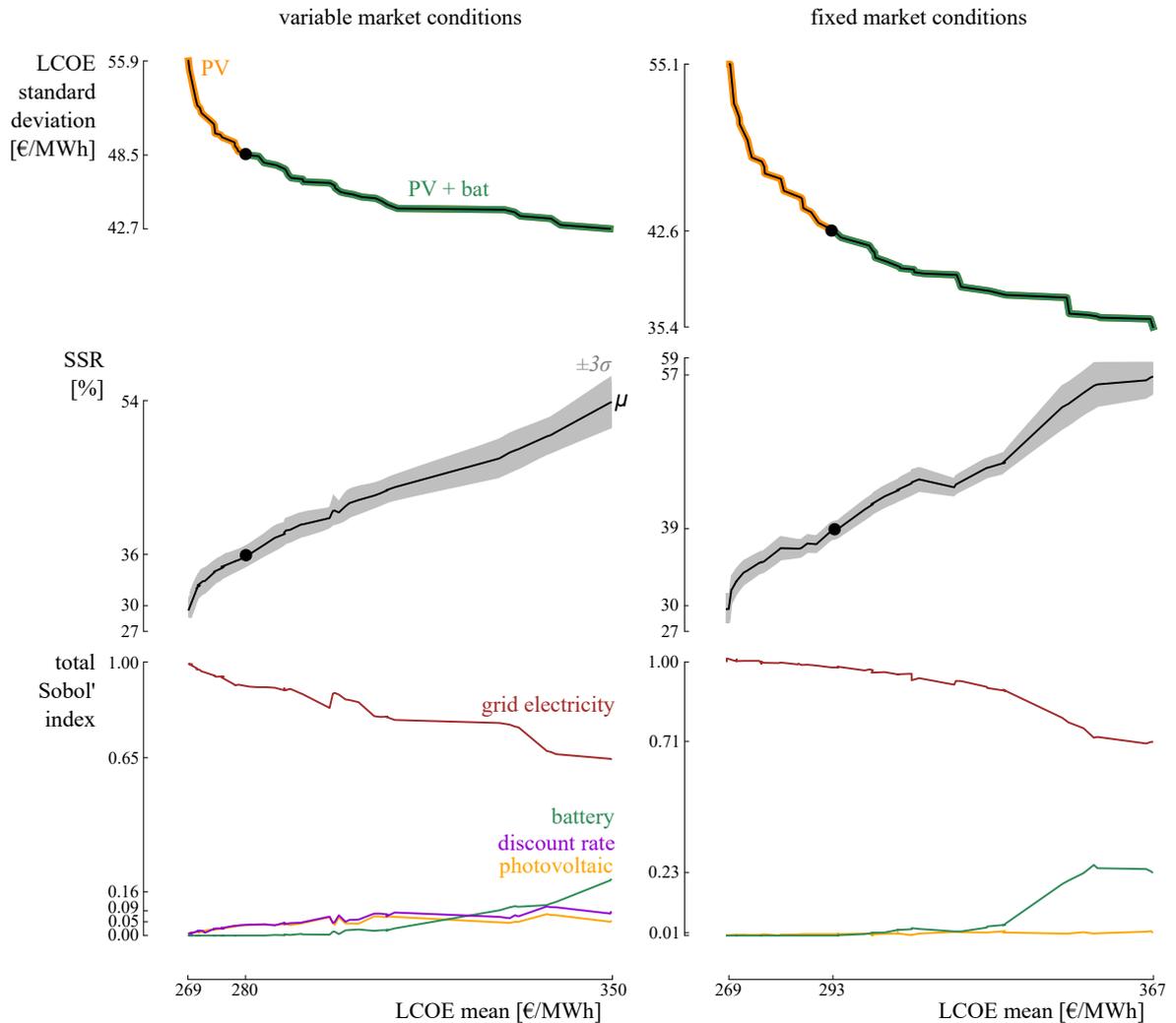
---

in the variable market conditions scenario induces a slightly larger LCOE standard deviation (e.g., 55.9 €/MWh for the dwelling) than in the fixed market conditions scenario (e.g., 55.1 €/MWh for the dwelling), due to the additional uncertainty present on the PV array investment cost and discount rate (Figure 3.4 and Figure 3.5). For this design, the uncertainty related to the grid electricity price (i.e., wholesale electricity price and proportion of the wholesale electricity price in the total charged cost per MWh consumed) dominates the LCOE standard deviation, due to the significant dependency on the grid to comply with the electricity demand (SSR mean = 30%). To reduce the LCOE standard deviation, at the expense of a minimal increase in LCOE mean, the Pareto set of optimized designs implies to increase the PV array capacity, which consequently increases the SSR mean and therefore decreases the Sobol' indices related to grid electricity. To illustrate for the dwelling, increasing the PV array capacity proves to be a cost-efficient approach (i.e., with minimal increase in LCOE mean) to decrease the LCOE standard deviation down to 48.5 €/MWh in the variable market conditions scenario and down to 42.6 €/MWh in the fixed market conditions scenario. Despite the LCOE standard deviation can be reduced modestly by increasing the PV array capacity, the proposed PV designs remain primarily dependent on grid electricity (SSR mean < 40%) and the envisioned LCOE is therefore subject to a significant standard deviation.

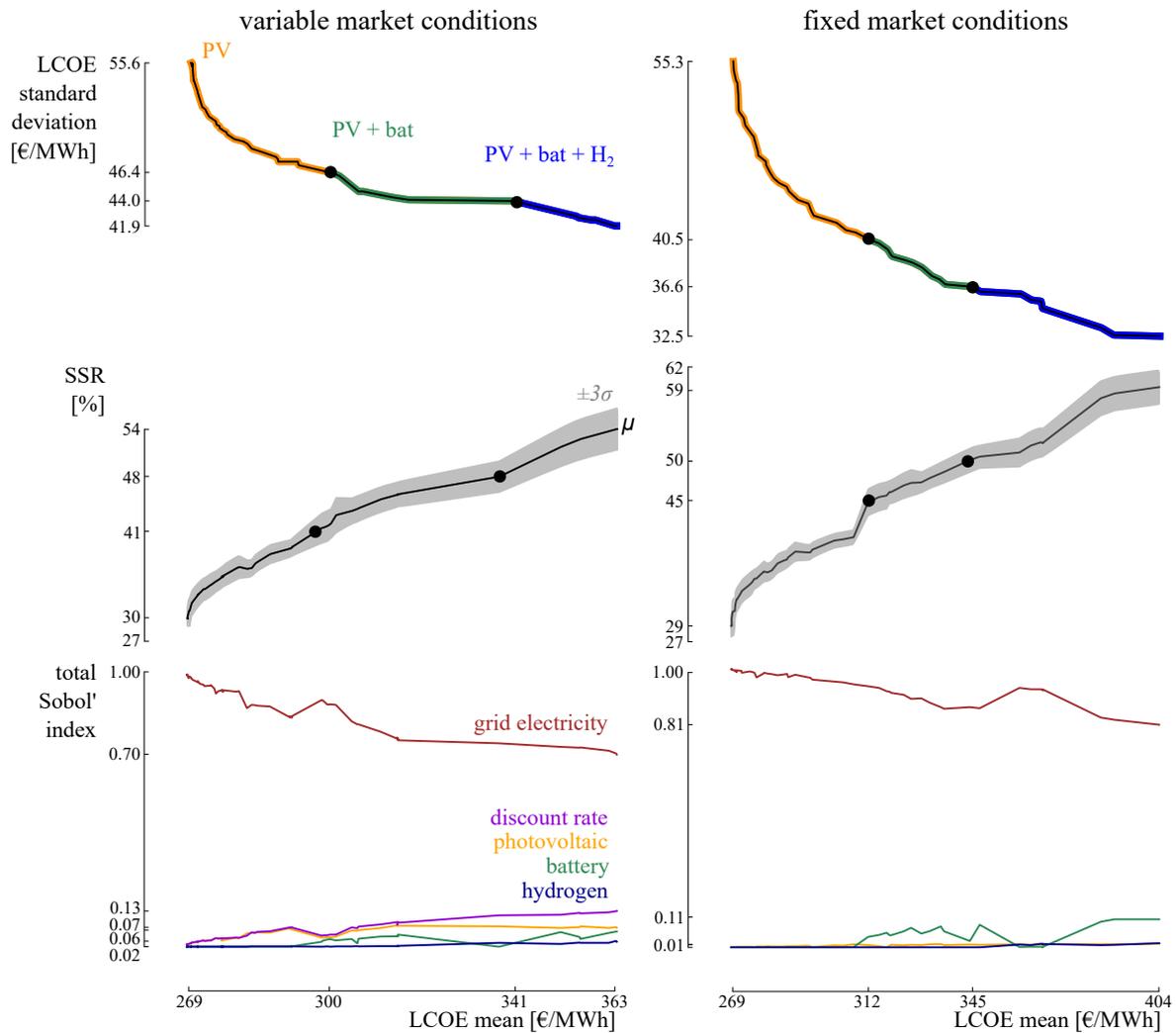
To further decrease the LCOE standard deviation over a higher SSR mean threshold (e.g., SSR mean > 40% for the community, Figure 3.5), PV-battery designs are configured by the optimization algorithm, instead of further increasing solely the PV array capacity. This follows from the fact that the SSR mean stagnates over a certain PV capacity threshold, while the inclusion of a battery stack enables to cover part of the demand when insufficient solar irradiance is available. Hence, the optimization algorithm suggests designs that subsequently increase in both PV array capacity and battery stack capacity to further reduce the LCOE standard deviation (and thus increase in SSR mean). For these designs, the LCOE standard deviation is significantly characterized by the uncertainty related to the battery (e.g., total Sobol' index up to 0.16 in the variable market conditions scenario for the dwelling, Figure 3.4). Therefore, improving the battery lifetime estimation is considered an effective approach to reducing the LCOE standard deviation of these designs with an external measure.

For the dwelling, a PV-battery configuration defines the robust design which achieves an

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system



**Figure 3.4:** For the dwelling, the Pareto front illustrates that a trade-off exists between minimizing the Levelized Cost Of Electricity (LCOE) mean and LCOE standard deviation. Reducing the LCOE standard deviation cost-efficiently is suggested by subsequently expanding a PV array and battery stack. This system capacity evolution improves the mean Self-Sufficiency Ratio (SSR) and consequently reduces the importance of the grid price uncertainty. Instead, the uncertainty related to the discount rate (in the variable market conditions scenario) and the battery stack gradually gain importance in the LCOE variation.



**Figure 3.5:** For the community, the Pareto front illustrates that a trade-off exists between minimizing the Levelized Cost of Electricity (LCOE) mean and LCOE standard deviation. Reducing the LCOE standard deviation cost-efficiently is suggested by subsequently expanding a PV array, battery stack and the hydrogen-based energy system. This system capacity evolution improves the mean Self-Sufficiency Ratio (SSR) and consequently reduces the importance of the grid price uncertainty. Instead, the uncertainty related to the discount rate (in the variable market conditions scenario) and the battery stack gradually gain importance in the LCOE variation.

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

SSR mean of 54 % in the variable market conditions scenario and an SSR mean of 57 % in the fixed market conditions scenario. Consequently, from this point, improving the SSR mean leads to an increase in LCOE standard deviation due to the significant increase in system capacity, from which the corresponding uncertainty overcompensates the reduction in the uncertainty related to grid electricity. As lower uncertainty is related to the system capacity in the fixed market conditions scenario (i.e., deterministic investment costs), the SSR mean for the robust design in this scenario is higher than for the robust design in the variable market scenario. Therefore, the LCOE standard deviation for the robust design in the fixed market conditions scenario (35.4 €/MWh) is lower than for the robust design characterized in the variable market scenario (42.7 €/MWh). For the community, the robust design includes both a battery stack and a hydrogen-based energy system (Figure 3.5), despite the large uncertainty related to the specific cost of the electrolyzer and fuel cell (e.g.,  $\sigma_{\text{CAPEX}_{\text{PEMEL}}} = 202 \text{ €/kW}$  as opposed to  $\sigma_{\text{CAPEX}_{\text{bat}}} = 73 \text{ €/kWh}$ ). The decoupling of power (i.e., electrolyzer and fuel cell) and energy (i.e., hydrogen tank,  $\sigma_{\text{CAPEX}_{\text{tank}}} = 1.2 \text{ €/kWh}$ ) enables to curb the uncertainty related to the hydrogen-based energy system present in the LCOE standard deviation. Moreover, this flexible capacity setting of hydrogen storage results in the fact that the significant amount of energy stored over a longer period can be achieved by solely increasing the capacity of the hydrogen tank, as opposed to sizing the entire battery stack for long-term energy storage. In conclusion, a PV-battery-hydrogen design provides a cost-efficient alternative for a large battery stack at moderate SSR and is therefore configured as a robust alternative in both scenarios for the community.

**Table 3.2:** The characteristics of the three evaluated optimized designs for the dwelling: a PV design and a PV-battery design that achieve the lowest LCOE mean among their corresponding category and the PV-battery robust design that achieves the lowest LCOE standard deviation.

	$N_{\text{PV}}$ kW <sub>p</sub>	$N_{\text{bat}}$ kWh	$\mu_{\text{LCOE}}$ €/MWh	$\sigma_{\text{LCOE}}$ €/MWh	$\mu_{\text{SSR}}$ %
<b>variable market scenario</b>					
PV	2.7		269	55.9	30
PV + bat	5.1	1.0	280	48.5	36
PV + bat robust	5.2	6.9	350	42.7	54
<b>fixed market scenario</b>					
PV	2.7		269	55.1	30
PV + bat	6.9	1.1	293	42.6	39
PV + bat robust	9.1	5.8	367	35.4	57

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

**Table 3.3:** The characteristics of the three evaluated optimized designs for the community: a PV design and a PV-battery design that achieve the lowest LCOE mean among their corresponding category and the PV-battery-hydrogen robust design that achieves the lowest LCOE standard deviation.

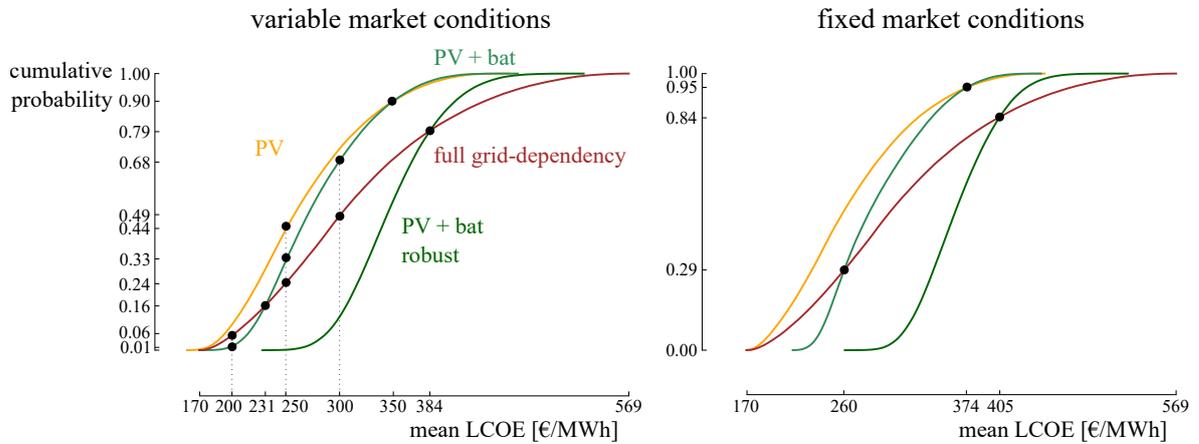
	$N_{PV}$ MW <sub>p</sub>	$N_{bat}$ MWh	$N_{PEMEL}$ MW	$N_{PEMFC}$ MW	$N_{tank}$ MWh	$\mu_{LCOE}$ €/MWh	$\sigma_{LCOE}$ €/MWh	$\mu_{SSR}$ %
<b>variable market scenario</b>								
PV	6.9					269	55.6	30
PV + bat	11.6	6.1				300	46.4	41
PV + bat + H <sub>2</sub>	16.3	7.1	1.7	0.5	16.7	363	41.9	54
<b>fixed market scenario</b>								
PV	6.5					269	55.3	29
PV + bat	15.0	6.9				312	40.5	45
PV + bat + H <sub>2</sub>	21.2	8.4	2.3	0.5	16.7	404	32.5	59

#### 3.3.6.2 Comparison of stochastic performance

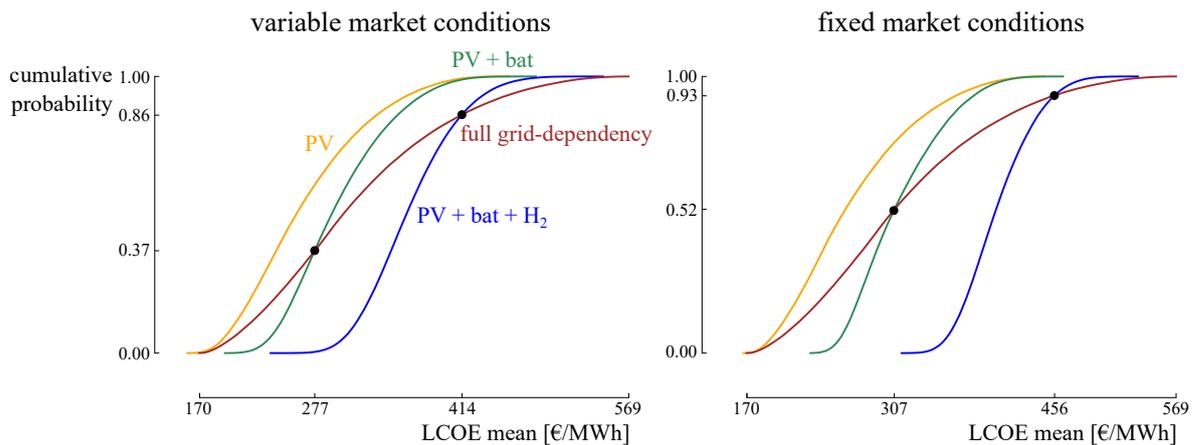
Due to the trade-off between minimizing the LCOE mean and minimizing the LCOE standard deviation, each design out of the Pareto set of optimized designs carries an advantage in either average performance or robustness. To evaluate the overall stochastic performance, the CDF is constructed for three representative optimized designs: the PV design that achieves the lowest LCOE mean, the PV-battery design that achieves the lowest LCOE mean among the proposed PV-battery designs and a robust design, which corresponds to the lowest LCOE standard deviation for the dwelling (Table 3.2) and for the community (Table 3.3). Among these designs, the PV design achieves the highest probability that, in reality, the resulting LCOE over the lifetime will be lower than any predefined LCOE upper limit for the dwelling and community (yellow CDF in Figure 3.6 and Figure 3.7, respectively). To illustrate in a variable market conditions scenario, if the dwelling owner predefines an LCOE upper limit of 250 €/MWh to ensure an affordable cost of electricity, the PV design provides a probability of 44 % that the LCOE, in reality, will be below or equal to that upper limit. Alternatively, a lower probability that the real LCOE will be below this upper limit is achieved by the PV-battery design (33 %) and by full grid-dependency (24 %). However, when the upper limit value is determined higher or equal to 350 €/MWh in the variable market conditions scenario (374 €/MWh in the fixed market conditions scenario), the PV-battery design achieves a similar probability of resulting in an LCOE below the upper limit as the PV design. This observation is explained by the lower LCOE standard deviation of the PV-battery design, which steepens the corresponding CDF and therefore realises the intersection between the PV and PV-battery CDF, despite the larger LCOE mean of the

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

PV-battery design.



**Figure 3.6:** The Cumulative Density Functions (CDF) of three optimized designs and full grid-dependency for the dwelling. The intersection point between two CDFs illustrates the minimum LCOE upper limit that has to be defined, for the optimized design to achieve a higher probability than full grid-dependency to achieve an LCOE below this upper limit in reality.

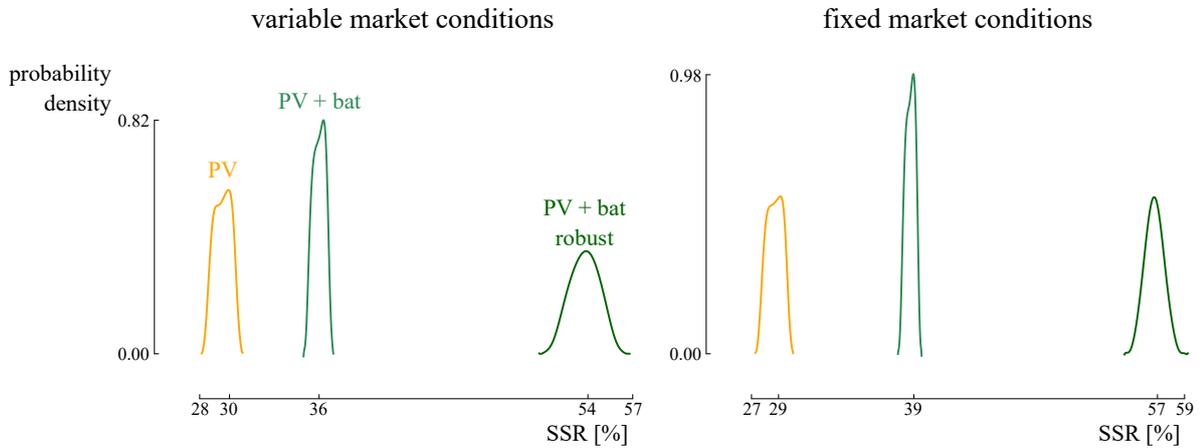


**Figure 3.7:** The Cumulative Density Functions (CDF) of three optimized designs and full grid-dependency for the community. The intersection point between two CDFs illustrates the minimum LCOE upper limit that has to be defined, for the optimized design to achieve a higher probability than full grid-dependency to achieve an LCOE below this upper limit in reality.

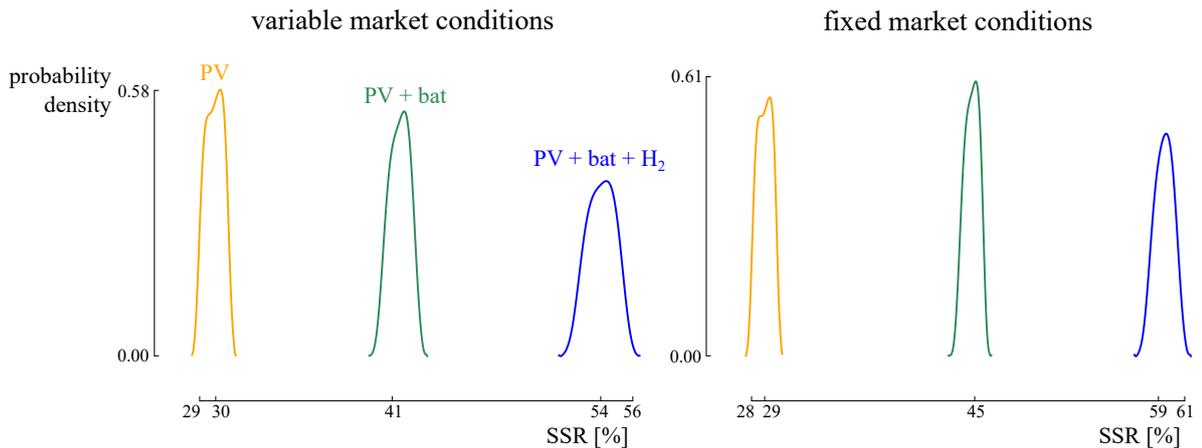
The PV design achieves a rather low SSR mean (30 %), which makes this design vulnerable to grid behavior, fluctuating electricity prices and potential power cuts. Moreover, operating over a higher SSR threshold reduces the risk of black-out, as it avoids the simultaneous power extraction from the grid of different demand types. Therefore, the PV-battery design and robust design are of interest when operating over a larger SSR threshold is preferred (Figure 3.8 for the dwelling and Figure 3.9 for the community). To illustrate for the dwelling

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

in a variable market scenario, the PV-battery design with the lowest LCOE mean achieves an SSR mean of 36 %, while the robust PV-battery design achieves an SSR mean of 54 %.



**Figure 3.8:** For the dwelling, the probability density functions of the SSR for three optimized designs illustrate the different SSR thresholds over which the designs operate.



**Figure 3.9:** For the community, the probability density functions of the SSR for three optimized designs illustrate the different SSR thresholds over which the designs operate.

When operating over a higher SSR threshold is considered, the performance of the PV-battery design (green CDF in Figure 3.6 and Figure 3.7) is of interest and can be compared with full grid-dependency (red CDF in Figure 3.6 and Figure 3.7). In this comparison, the preferred design depends on the predefined LCOE upper limit (and thus on the financial flexibility of the owner). To illustrate for the dwelling in the variable market conditions scenario, both the PV-battery design and full grid-dependency result in a probability of 16 % that the LCOE, in reality, will be below or equal to 231 €/MWh. If the LCOE upper limit is estimated even lower by the building owner to ensure an affordable cost of electricity, full

### 3.3 Power-to-power: photovoltaic-battery-hydrogen system

---

grid-dependency results in a higher probability of ending up below the upper limit than the PV-battery design: e.g., full grid-dependency achieves a probability of 6 % that the LCOE, in reality, will be below 200 €/MWh, while the PV-battery design ensures a probability of only 1 %. Instead, if the LCOE upper limit is set higher than 231 €/MWh, the PV-battery design ensures a higher probability: e.g., full grid-dependency results in a probability of 49 % that the LCOE, in reality, will be below or equal to 300 €/MWh, while the PV-battery design ensures a probability of 68 %. Thus, when the household projects an LCOE of at least 231 €/MWh as a maximum to ensure an affordable cost of electricity, the PV-battery design provides a higher probability of complying with this upper limit than full grid-dependency. Similar results are presented for the community.

Despite the lowest LCOE standard deviation, the CDF of the robust designs do not intersect with the other evaluated optimized designs and is therefore only of interest when operating over a larger SSR threshold. If so, then the robust PV-battery design for the dwelling (dark green CDF in Figure 3.6) proves to be beneficial over full grid-dependency when the LCOE upper limit is set above 384 €/MWh in the variable market conditions scenario and above 405 €/MWh in the fixed market conditions scenario. For the community, a PV-battery-hydrogen configuration is characterized as a robust design (blue CDF in Figure 3.7). This design is beneficial over full grid-dependency when the LCOE upper limit is defined higher than 414 €/MWh in the variable market conditions scenario and higher than 456 €/MWh in the fixed market conditions scenario. Consequently, a PV-battery-hydrogen design presents a cost-competitive alternative over full grid-dependency and operates over a significant SSR threshold ( $\geq 54\%$ ), which reduces the risk of black-out and power cuts.

#### 3.3.7 Conclusion

The robust design optimization method illustrates a trade-off between minimizing the levelized cost of electricity mean and minimizing the standard deviation for a dwelling and a community. A photovoltaic array achieves the lowest levelized cost of electricity mean (i.e., 269 €/MWh). Additionally, such a system ensures the highest probability that the levelized cost of electricity, in reality, will be below an upper limit, which has been predicted by the system owner to ensure an affordable cost of electricity. Nevertheless, the proposed photovoltaic arrays remain primarily dependent on grid electricity (self – sufficiency ratio mean  $\approx 30\%$ ) and therefore subject to a significant standard devia-

### **Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions**

---

tion, risk of black out and power cuts.

Photovoltaic-battery designs reduce this grid-dependency by increasing the self-sufficiency ratio and therefore decreasing the levelized cost of electricity standard deviation. When operating over a higher self-sufficiency ratio is of interest, these photovoltaic-battery designs can ensure a higher probability than full grid-dependency to operate below the maximum affordable levelized cost of electricity predicted (e.g. 231 €/MWh for a dwelling in a variable market scenario).

For a community, the robust design achieves a significant self-sufficiency ratio ( $\geq 54\%$ ) and includes both a battery stack and a hydrogen-based energy system. This design is least-sensitive to real-life uncertainty, as the decoupling of hydrogen power and hydrogen energy enables to curb the contribution of the uncertainty related to the electrolyzer and fuel cell in the levelized cost of electricity uncertainty. Moreover, these designs prove beneficial over full grid-dependency when the maximum allowed levelized cost of electricity is determined above 414 €/MWh in the fixed market conditions scenario and above 456 €/MWh in the variable market conditions scenario. In conclusion, depending on the financial flexibility of the system owner, battery and hydrogen storage provide a higher probability of realizing a levelized cost of electricity below the maximum allowed levelized cost of electricity to ensure affordability than full-grid dependency. Future work will focus on the integration of different locations. Moreover, a heat demand will be included, to fully exploit the fuel cell capability.

#### **3.4 Power-to-mobility: renewable-powered hydrogen refueling station**

In this second application, an RDO is performed on a renewable-powered Hydrogen Refueling Station (HRS). A holistic approach is adopted, considering an economic and environmental quantity of interest, which results in a multi-objective optimization problem with four objectives. Moreover, considering two quantities of interest results in a significant set of uncertain parameters, which are handled with the updated surrogate-assisted RDO algorithm. As mentioned in the chapter overview, this application is under review in Applied Energy. After a brief introduction on renewable-powered HRS (Subsection 3.4.1),

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

---

Subsection 3.4.2 provides the lay-out of the HRS and the adopted models for each component, followed by the quantities of interest (Subsection 3.4.3), the characterization of the uncertainties (Subsection 3.4.4) and the characterization of the stepwise regression PCE (Subsection 3.4.5). The optimized design characteristics for each quantity of interest and the results for a holistic RDO (i.e., considering all quantities of interest in the optimization problem) are discussed in Subsection 3.4.6. The main messages of this section are provided in Subsection 3.4.7.

#### 3.4.1 Introduction

Heavy-duty vehicles are significant emitters of CO<sub>2</sub> (6% of the global energy-related emissions) [210]. For that reason, the reduction of emissions from public buses is actively researched [211]. Significant emission reductions can be achieved by replacing diesel-powered buses with battery-electric buses [212] or Fuel Cell Electric Buses (FCEB), fueled by renewable hydrogen [213, 214]. With a refueling time similar to traditional diesel-powered buses (10 min, as opposed to 3 h - 5 h for electric buses) and promising performance over extended driving ranges (i.e., 250 km) [215], FCEB achieve significant interest across Europe, with over 90 FCEB deployed in 2018 [216]. Despite that FCEB emit no GreenHouse Gases (GHG) while driving, the production of hydrogen in a renewable-powered HRS is not entirely carbon-free [217]: Spath et al. [218] and Burkhardt et al. [217] quantified the GHG emission for a wind-powered HRS and they present values of 0.97 kg<sub>CO<sub>2</sub>,eq</sub>/kg<sub>H<sub>2</sub></sub> and 1.92 kg<sub>CO<sub>2</sub>,eq</sub>/kg<sub>H<sub>2</sub></sub> (i.e., 29 g<sub>CO<sub>2</sub>,eq</sub>/kWh and 58 g<sub>CO<sub>2</sub>,eq</sub>/kWh), respectively. Nevertheless, the GHG emissions are significantly lower compared to the well-to-wheel emission factor for diesel, quantified at 326 g<sub>CO<sub>2</sub>,eq</sub>/kWh [219]. Despite the clear environmental advantage, the cost of renewable hydrogen production using PV arrays and wind turbines remains relatively high (Table 3.4). In comparison, the CO<sub>2</sub>-emitting hydrogen production alternatives are more beneficial in terms of economic performance [220].

In summary, feasibility studies on renewable HRS indicate excellent environmental performance, at the expense of a higher cost. However, these studies assumed fixed parameters, which can lead to designs and conclusions sensitive to economic and environmental uncertainty. Therefore, RDO was applied to a wind- and solar-powered HRS, considering a large set of technical, economic, and environmental uncertainties (39 in total). In addition to sizing the HRS component capacities, the bus fleet composition is an additional design

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

**Table 3.4:** The cost of hydrogen production using solar photovoltaic and wind turbines is higher than for the traditional CO<sub>2</sub>-emitting alternatives.

type	cost range [€/kg <sub>H<sub>2</sub></sub> ]	Ref.
solar photovoltaic hydrogen	4.80 - 19.31	[220]
	9.29 - 12.48	[221]
	15.5 - 22.5	[222]
wind hydrogen	4.89 - 5.00	[220]
	4.40 - 4.81	[223]
	1.16 - 32.83	[224]
steam methane reforming	1.73 - 1.88	[220]
methane pyrolysis	1.32 - 1.41	[220]
coal gasification	1.11 - 1.35	[220]

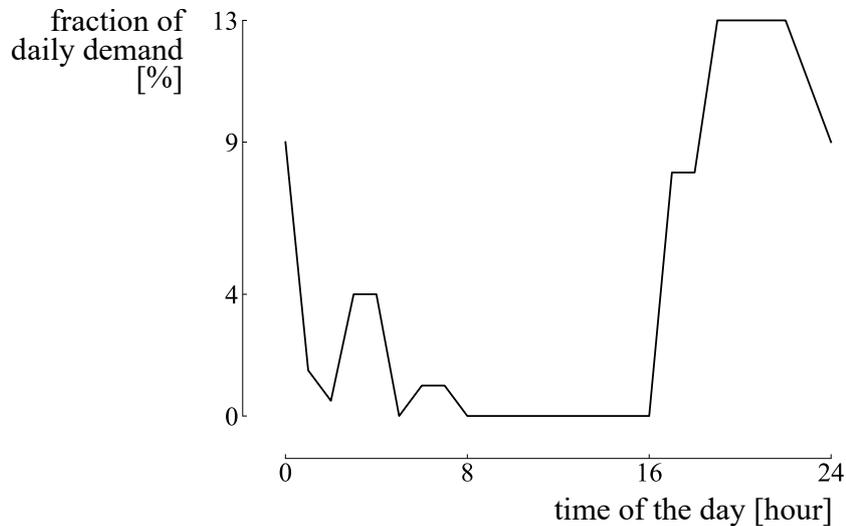
variable. Through this design variable, the optimization algorithm configures the fraction of the bus fleet powered by hydrogen, while the remainder of the bus fleet is powered by conventional diesel. Consequently, the optimized mix between hydrogen-powered and diesel-powered buses is determined based on the expected performance and the variance of the economic and environmental quantity of interest. By considering these statistical moments as optimization objectives, this application provides novel designs and highlights additional benefits in terms of robustness for hydrogen-powered heavy-duty transport.

#### 3.4.2 System modelling

The mobility demand corresponds to a bus depot in Brussels, Belgium. At the bus depot, a typical amount of 40 buses [225] are stationed with an average utilisation rate of 250 km/bus/day [216]. The bus fleet can consist of hydrogen-powered buses (i.e., FCEB), diesel-powered buses or a mix of both. To determine the period in which these buses are fueled, the European daily bus refueling profile is adopted (Figure 3.10) [226]. The dispensation pressure for FCEB is fixed at 350 bar, as opposed to 700 bar for hydrogen-powered cars, as buses allow carrying more volume than passenger cars. This reduced dispensing pressure lowers the HRS complexity (i.e., no cooling unit before dispensation), compression energy, system cost and improves reliability [226]. In a future where hydrogen-powered passenger vehicles are massively deployed, converting the dispensing pressure for FCEB to 700 bar can result in advantages related to the uniformity in distribution systems. However,

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

today and in the near future, 350 bar remains the state-of-the-art dispensation pressure for FCEB [226].

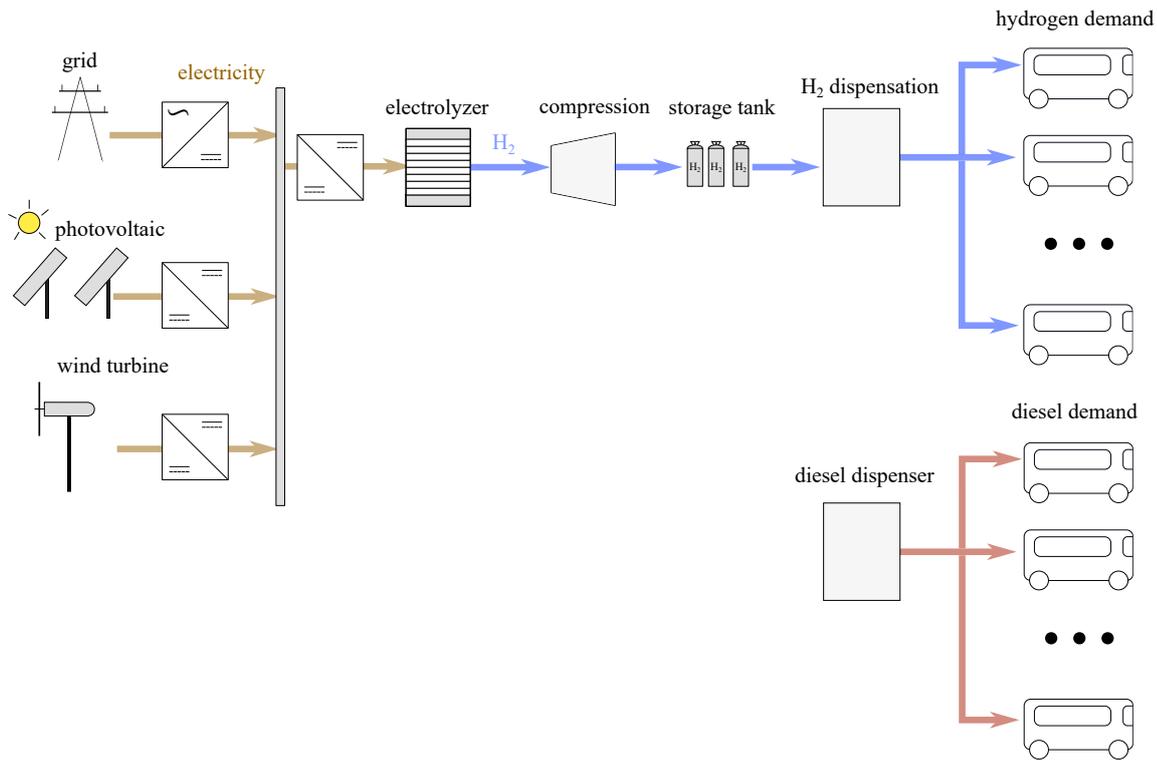


**Figure 3.10:** The refueling profile indicates the daily fuel demand. The refueling is performed in the evening and at night, between 4PM and 8AM.

To fuel the FCEB, a renewable-powered HRS is considered. The power management strategy for the HRS aims to continuously comply with the hydrogen demand. During the refueling period, hydrogen is extracted from the storage tank. When the stored hydrogen is insufficient to meet the demand, the remaining hydrogen is produced instantaneously (i.e., cold start-up time: 5 min - 10 min; warm start-up time: below 10 s [88]) in the PEM electrolyzer array, compressed in the compressor and dispensed. To power the PEM electrolyzer array and compressor, in this situation, the available renewable power is called upon first. If the instantaneous solar and wind power does not comply with the power demand from the electrolyzer and compressor, grid electricity is extracted to cover the remaining demand. The grid is considered permanently available, which means that the electricity demand to cover the hydrogen demand is always complied with (by the wind turbines, PV array and/or the grid). Alternatively, when the storage tank is able to comply with the instantaneous hydrogen demand, the PV electricity and wind electricity are used to produce and compress hydrogen. No grid reinjection is considered, as it remains highly uncertain how the grid tariff methodology will adapt after the current integration of digital meters, which ends the current net metering regime [227]. Therefore, this approach mitigates the risk of changing metering regimes, it provides globally applicable conclusions and it illustrates the raw potential of local renewable energy storage.

## Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

The climate data for Brussels is generated based on the method described in Subsection 3.3.2.7. In addition, the component models for the PV array (Subsection 3.3.2.1) and the PEM electrolyzer array (Subsection 3.3.2.3) are described in the previous application. In the remainder of this section, the component models for the wind turbine and compressor array are presented.



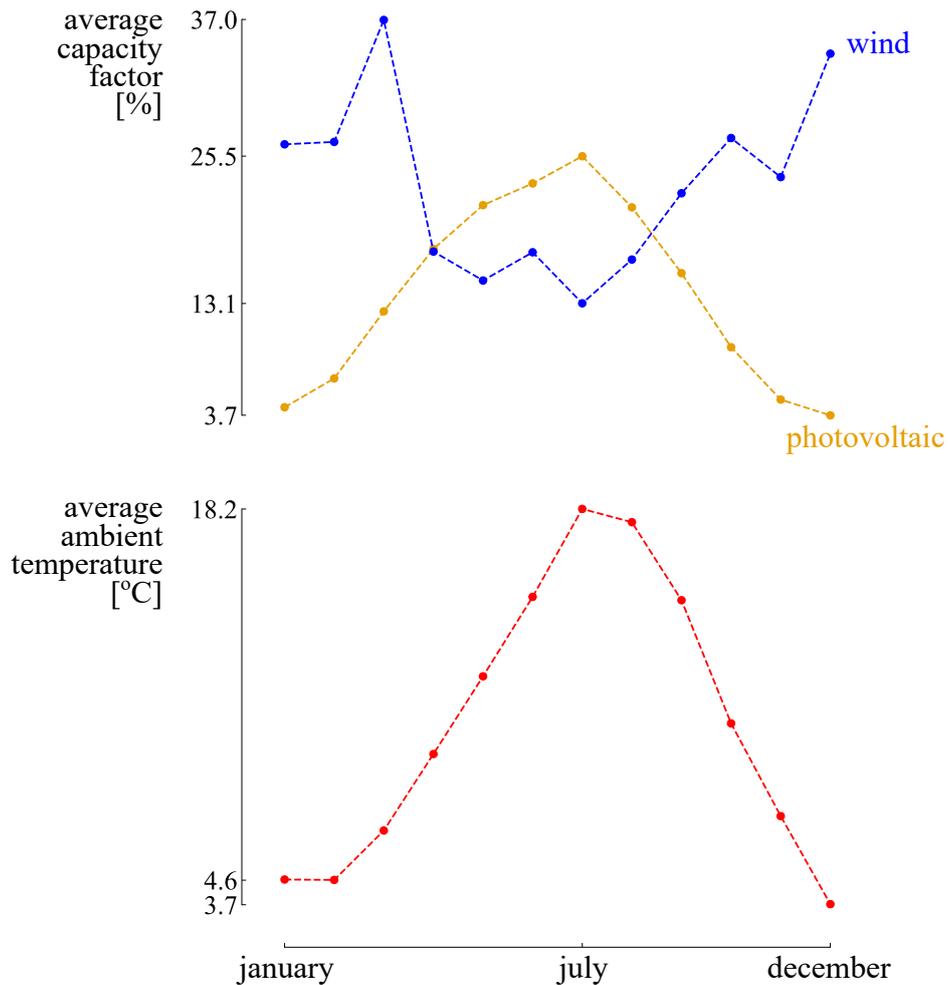
**Figure 3.11:** The bus fleet can consist of hydrogen-powered buses, diesel-powered buses or a mix of both. For the hydrogen-powered buses, a grid-connected hydrogen refueling station is considered. In the hydrogen refueling station, a PV array and wind turbine array produce electricity to generate hydrogen in the electrolyzer. The produced hydrogen is compressed and stored in storage tanks, before dispensation.

### 3.4.2.1 Wind turbine

I adopted the hourly wind power profile from renewables.ninja [228], a framework where wind turbine models and power output data are adopted from the work of Staffell et al. [229]. This data is corrected with a bias factor to obtain realistic power data via a virtual wind farm model for any location in Europe with a high spatial and temporal accuracy. The hourly capacity factor profile was based on a wind turbine model (Enercon E70/2000), which was integrated into the framework of Staffell et al. [229]. This turbine model and location

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

in Brussels provided an annual mean wind capacity factor of 23.1%, which corresponds to the approximate average wind capacity factor of Belgium for onshore wind turbine farms [229, 230]. Additional information on the physical behavior of the wind turbine array is present in Section B.1. Generally, the average monthly wind capacity factor shows a higher capacity factor for wind than for PV at the same location, with a dip between May and August (Figure 3.12).



**Figure 3.12:** For Brussels (Belgium), the monthly average capacity factor for the photovoltaic array and the monthly average ambient temperature show a similar pattern. However, the wind turbine capacity factor is opposite to the monthly average ambient temperature and monthly average photovoltaic capacity factor.

### 3.4.2.2 Compressor

In this application, high-pressure hydrogen (350 bar) is required, to compensate for the low energy density of hydrogen (around 3 kWh/Nm<sup>3</sup>), and thus to avoid large tanks in buses. A compressor pressurizes the produced hydrogen gas from the electrolyzer array up to the required pressure in the hydrogen storage tank. The hydrogen storage tank requires a pressure of 440 bar to create an overpressure when fueling the H<sub>2</sub> busses according to the National Institute of Standards [231]. An isentropic compression process was considered, to calculate the necessary power for storing the hydrogen gas in the tank at 440 bar with the following equation:

$$P_{\text{compressor}} = \frac{\dot{m}_{\text{H}_2}}{\eta_c} \frac{k}{k-1} R_{\text{H}_2} T_{\text{in}} \left[ \left( \frac{p_{\text{out}}}{p_{\text{in}}} \right)^{\frac{k-1}{k}} - 1 \right], \quad (3.32)$$

where  $\dot{m}_{\text{H}_2}$  is the mass flow rate,  $\eta_c$  is the compression efficiency (85%),  $k$  is the ratio of specific heats (1.4),  $R_{\text{H}_2}$  the gas constant of hydrogen (4.12 kJ/kg K),  $T_{\text{in}}$  is the inlet temperature (353 K) and  $p_{\text{in}}/p_{\text{out}}$  is the inlet/outlet pressure (20 bar and 440 bar, respectively). These parameters provide a specific compression energy of 2.36 kWh/kg<sub>H<sub>2</sub></sub>, which remains in the range of the real energy consumption of compressors in refueling stations (in reality, such installation requires a specific compression energy between 2 kWh/kg<sub>H<sub>2</sub></sub> and 4 kWh/kg<sub>H<sub>2</sub></sub> [231]).

### 3.4.3 Design variables and quantities of interest

The system is designed by sizing the components of the system, i.e., the capacity of the PV array, wind turbine array, electrolyzer array, storage tank and the fraction of the bus fleet fueled by hydrogen. These design variables are considered independent, which means that the optimization algorithm can exclude a technology from the HRS by setting its capacity equal to zero (i.e., the lower bounds). The upper bounds for the design variables are 10 MW<sub>p</sub> for the PV array, 20 MW for the wind turbine and 10 MW for the PEM electrolyzer.

To evaluate the system performance, an economic and an environmental quantity of interest are defined for this system. For the economic quantity of interest, the Levelized Cost Of Driving (LCOD) is selected, which corresponds to the annualized system cost per unit of distance covered by the bus fleet. The system cost depends on the CAPEX, OPEX and

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

replacement cost for the components (i.e., PV array, wind turbine array, electrolyzer array, compressor, storage tank, dispenser, DC-DC converters, DC-AC inverters, diesel-powered buses and FCEB). The adopted values for these specific costs are presented in Table 3.5. In addition to the component costs, the price for the annual grid electricity bought and for the annual diesel bought is considered in the total cost. The wholesale electricity price corresponds to 40 % of the total electricity price for buying electricity, while the remaining 60 % depends on transmission cost, distribution cost, taxes and surcharges. In summary, the LCOD is presented as the annualized system cost over the annual distance travelled by the fleet ( $D$ ):

$$\text{LCOD} = \frac{\text{CAPEX}_a + \text{OPEX}_a + C_{\text{repl},a} + C_{\text{grid},a} + C_{\text{diesel},a}}{D}. \quad (3.33)$$

The cost parameters are defined analogously to the cost parameters in Subsection 3.3.3. Similar to the LCOD, the environmental quantity of interest represents the annualized GHG emission of the system per unit of distance travelled by the bus fleet (i.e., Carbon Intensity (CI)). The system GHG emissions represent the sum of the annualized GHG emissions during the construction of the components  $\text{GHG}_{\text{comp},a}$ , the GHG emissions from grid electricity consumption  $\text{GHG}_{\text{grid},a}$  and the well-to-wheel GHG emissions from diesel consumption  $\text{GHG}_{\text{diesel},a}$ :

$$\text{CI} = \frac{\text{GHG}_{\text{comp},a} + \text{GHG}_{\text{grid},a} + \text{GHG}_{\text{diesel},a}}{D}. \quad (3.34)$$

#### 3.4.4 Uncertainty characterization

In this work, technical, economic, and environmental parameters are subject to uncertainty. As only a handful of data points exist for these parameters, determining the underlying distribution is not feasible. Therefore, a uniform distribution was assumed for all uncertainties, specified by the lower bound and upper bound value (Table 3.5).

The uncertainty on the evolution of the grid electricity price is adopted from 6 prediction scenarios [89]. Similarly, the distribution of the diesel price integrates over low, middle and high diesel cost scenarios [249]. The uncertainty on the energy consumption of the hydrogen-powered and diesel-powered buses is adopted from the quantified difference be-

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

**Table 3.5:** Parameters affected by uncertainty.

parameter	min	max	unit	Ref.
CAPEX <sub>PV</sub>	350	600	€/kW <sub>p</sub>	[232]
OPEX <sub>PV</sub>	16	19	€/kW <sub>p</sub> /year	[142]
GHG <sub>PV</sub>	520	1550	kg <sub>CO<sub>2</sub>,eq</sub> /kW <sub>p</sub>	[233]
CAPEX <sub>wind</sub>	620	800	€/kW	[234]
OPEX <sub>wind</sub>	18	36	€/kW/year	[234]
GHG <sub>wind</sub>	242	469	kg <sub>CO<sub>2</sub>,eq</sub> /kW	[235]
CAPEX <sub>PEMEL</sub>	1400	2100	€/kW	[88]
OPEX <sub>PEMEL</sub>	3	5	% of CAPEX	[88]
replacement cost <sub>PEMEL</sub>	15	20	% of CAPEX	[192, 193]
lifetime <sub>PEMEL</sub>	60	100	kh	[88]
GHG <sub>PEMEL</sub>	190	235	kg <sub>CO<sub>2</sub>,eq</sub> /kW	[236]
CAPEX <sub>tank</sub>	11	14	€/kWh	[237]
OPEX <sub>tank</sub>	1	2	% of CAPEX	[157, 192]
GHG <sub>tank</sub>	6	12	kg <sub>CO<sub>2</sub>,eq</sub> /kWh	[238]
CAPEX <sub>compressor</sub>	1000	1500	€/kW	[196]
OPEX <sub>compressor</sub>	1	2	% of CAPEX	[196]
GHG <sub>compressor</sub>	80	120	kg <sub>CO<sub>2</sub>,eq</sub> /kW	[239]
CAPEX <sub>dispenser</sub>	45 000	65 000	€/unit	[216, 221, 237]
OPEX <sub>dispenser</sub>	1	2	% of CAPEX	[221]
CAPEX <sub>diesel bus</sub>	220 000	250 000	€/unit	[216, 240, 241]
OPEX <sub>diesel bus</sub>	0.26	0.30	€/km	[241, 242]
GHG <sub>diesel engine</sub>	20.0	22.7	kg <sub>CO<sub>2</sub>,eq</sub> /kW	[243, 244]
CAPEX <sub>hydrogen bus</sub>	400 000	620 000	€/unit	[216, 240–242]
OPEX <sub>hydrogen bus</sub>	0.26	0.40	€/km	[216, 241]
GHG <sub>fuel cell</sub>	43	61	kg <sub>CO<sub>2</sub>,eq</sub> /kW	[245]
CAPEX <sub>DCDC</sub>	40	160	€/kW	[246]
OPEX <sub>DCDC</sub>	1	5	% of CAPEX	[246]
CAPEX <sub>DCAC</sub>	50	200	€/kW	[246]
OPEX <sub>DCAC</sub>	1	5	% of CAPEX	[246]
consumption <sub>diesel bus</sub>	3.7	4.5	kWh/km	[247]
consumption <sub>hydrogen bus</sub>	3.0	3.2	kWh/km	[247]
total solar irradiance	1082	1266	kWh/m <sup>2</sup> /year	[205]
wind capacity factor	20	26	%	[229]
ambient temperature	9.9	10.7	°C	[205]
inflation rate	1	2	%	[207]
wholesale electricity price	57	91	€/MWh	[89]
GHG <sub>electricity</sub>	144	176	g <sub>CO<sub>2</sub>,eq</sub> /kWh	[248]
diesel price	1.42	2.31	€/l	[249]
discount rate	5	7	%	[250–253]

### **3.4 Power-to-mobility: renewable-powered hydrogen refueling station**

---

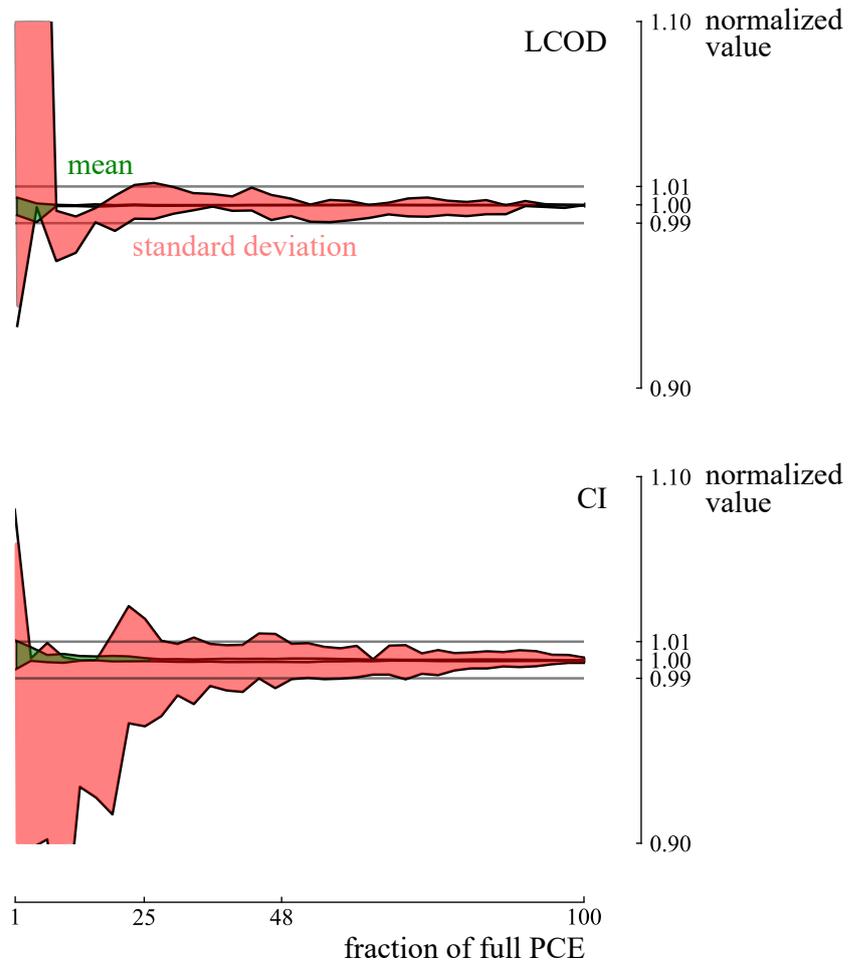
tween predicted and real-world operating conditions for these buses [247]. The uncertainty on the specific GHG emission for grid electricity consumption depends on the scenarios for the evolution of the electricity mix up to 2040 [248]. The solar irradiance and ambient temperature are characterized by a Typical Meteorological Year for which the uncertainty is characterized based on the interannual variability [90]. The annual average capacity factor for the wind turbine array is based on hourly data between 1980 and 2020 for Belgium [229]. For each component, the economic and environmental parameters are subject to uncertainty. Indeed, by considering generic models, the uncertainty lies in the lack of specifying the exact characteristics of the component. To illustrate, the CAPEX for each component is subject to the current price range on the market. Similar assumptions are made on the GHG emissions during the manufacturing of each component. To illustrate, the uncertainty on the GHG emissions during the manufacturing of the hydrogen storage tank relates to the amount of recycled steel that has been used, an uncertainty that can be addressed by specifying the hydrogen tank and its manufacturer.

#### **3.4.5 Stepwise regression PCE**

The system model is subject to 39 parametric uncertainties, out of which 31 parameters affect the LCOD, and 14 parameters affect the CI. The sparse PCE method, presented in Section 3.2, was adopted. A polynomial order of 2 results in a worst-case Leave-One-Out error of around 1% for both quantities of interest (i.e., LCOD and CI). For the LCOD and CI, 25 % and 48 % of the training samples are needed to ensure an error near  $\pm 1$  %, respectively, when compared to the statistical moments retrieved with a conventional PCE (Figure 3.13).

#### **3.4.6 Results and discussion**

The results are provided in two parts. In the first part, RDO is performed separately for each quantity of interest (i.e., LCOD and CI). These results illustrate the trade-off between minimizing the mean and standard deviation for each quantity of interest and the design trends between the corresponding optimized designs. In the second part, a holistic RDO is performed, considering the mean and standard deviation for both quantities of interest simultaneously as objectives. This strategy suggests holistic optimized designs, which consider both the stochastic performance in LCOD and CI.

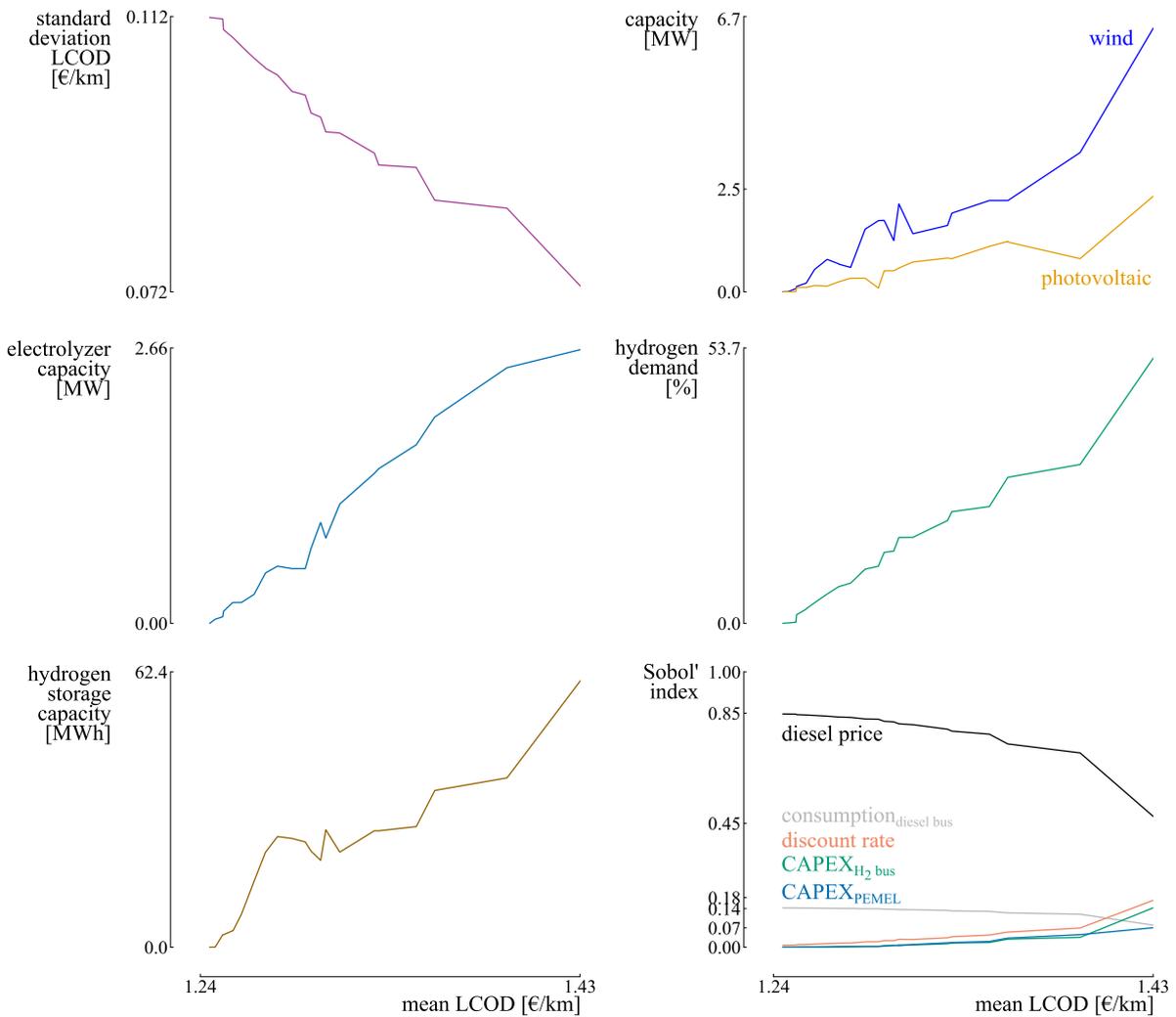


**Figure 3.13:** Based on the design sample with the worst convergence, the sparse PCE reaches statistical moments with an error below  $\pm 1\%$  on the Levelized Cost Of Driving and Carbon Intensity at 25 % and 48 % of the training samples needed for a conventional Polynomial Chaos Expansion (PCE), respectively.

#### 3.4.6.1 Individual Robust Design Optimization

Two RDOs are performed, one for each quantity of interest. The RDO on the LCOD is presented first, followed by the RDO on the CI and an overall performance comparison between the two sets of optimized designs.

**Robust Design Optimization of the Levelized Cost Of Driving** The RDO algorithm considers the LCOD mean and LCOD standard deviation as optimization objectives. The results illustrate a trade-off between minimizing the LCOD mean and minimizing the LCOD standard deviation (Figure 3.14, top-left). Hence, no single design exists for this system, subject to the aforementioned uncertainties, that simultaneously achieves the lowest LCOD mean and lowest LCOD standard deviation. The design with the best mean achieves an LCOD mean of 1.24 €/km, with a standard deviation of 0.11 €/km. Alternatively, the robust design (i.e., with the lowest LCOD standard deviation) decreases the LCOD standard deviation by 36 % (0.07 €/km), at the expense of an increase in LCOD mean by 11 % (1.43 €/km). To achieve the optimized LCOD mean, no FCEB are considered in the design, resulting in a complete diesel-powered bus fleet (Figure 3.14, middle-right). Consequently, no hydrogen production is required (Figure 3.14, middle-left) and thus no renewable capacity is considered (Figure 3.14, top-right). With only diesel-powered buses, the main drivers of the LCOD standard deviation (Figure 3.14, bottom-right) are the uncertainty on the future diesel price (Sobol' index: 0.85) and the actual fuel consumption of the diesel buses (Sobol' index: 0.14). To achieve a reduced LCOD standard deviation, a minimized increase in LCOD mean should be accepted. The optimized designs are then subsequently configured with an increasing hydrogen demand, supported by a renewable-powered HRS. Hence, converting a fraction of the diesel-powered bus fleet into FCEB, which are fueled by a renewable-powered HRS, improves the LCOD standard deviation. The robust design is configured with a 2.5 MW PV array, 6.7 MW wind turbine, 2.7 MW electrolyzer array and 62 MWh (31 m<sup>3</sup>) hydrogen storage tank (Figure 3.14), covering 54 % of the bus fleet with FCEB. As 46 % of the bus fleet is still fueled by diesel, the LCOD standard deviation for the robust design remains mainly driven by the uncertainty on the diesel price (Sobol' index: 0.45) and the uncertainty on the diesel consumption (Sobol' index: 0.07), accompanied by a significant share coming from the uncertainty on the discount rate (Sobol' index: 0.18), the CAPEX of the FCEB (Sobol' index: 0.17) and the CAPEX of the electrolyzer array (Sobol' index: 0.17).



**Figure 3.14:** The Pareto front illustrates a trade-off between minimizing the Levelized Cost Of Driving (LCOD) mean and LCOD standard deviation. To achieve the optimized LCOD mean, only diesel-powered buses are considered (i.e., no hydrogen demand). The LCOD standard deviation is reduced by subsequently converting the diesel-powered buses into Fuel Cell Electric Buses (FCEB), accompanied by a Hydrogen Refueling Station with increasing wind power, photovoltaic power and hydrogen storage.

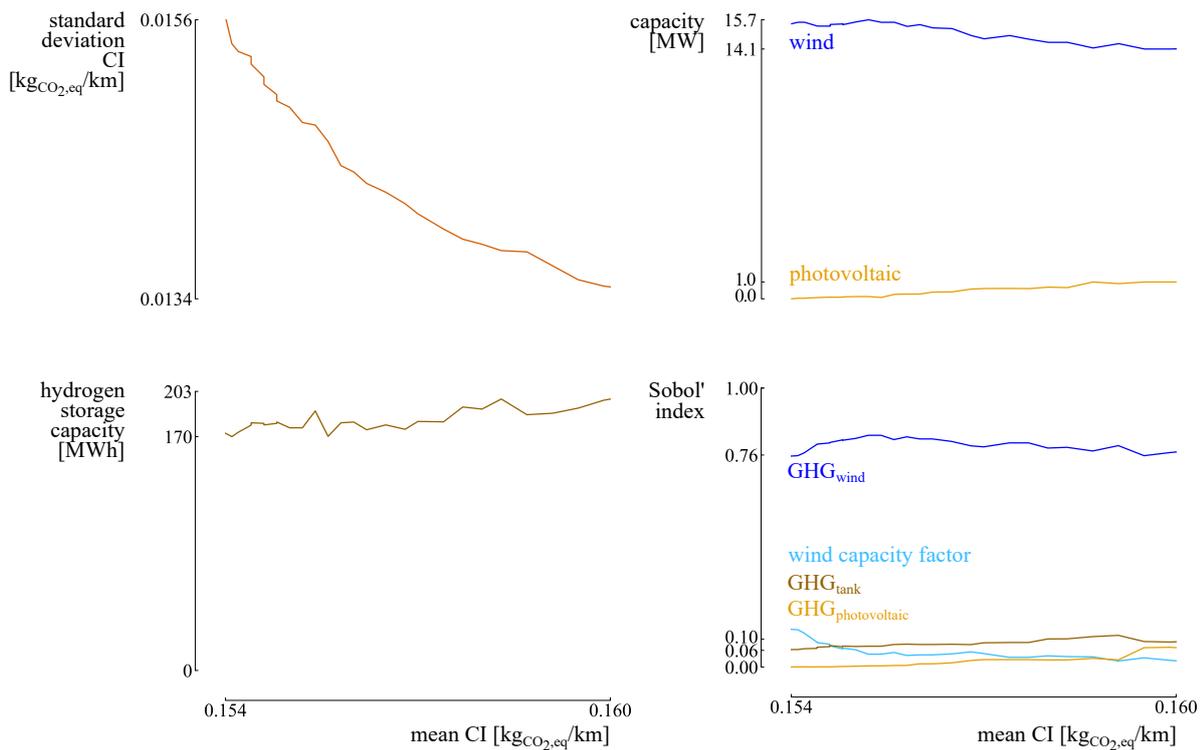
### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

---

index: 0.07). Interestingly, a higher share of the bus fleet powered by hydrogen (> 54%), accompanied by a renewable-powered HRS leads to a rise in both the LCOD mean and LCOD standard deviation. This can be explained by the rising uncertainty related to the HRS overcompensates the reducing uncertainty on the diesel price and diesel consumption in the LCOD standard deviation.

**Robust Design Optimization of the Carbon Intensity** Contrary to the RDO on the LCOD, nearly no trade-off exists between minimizing the CI mean and minimizing the CI standard deviation (Figure 3.15, top-left). Among the optimized designs, the optimized CI mean ranges only between 0.154 kg<sub>CO<sub>2</sub>,eq</sub>/km and 0.160 kg<sub>CO<sub>2</sub>,eq</sub>/km, to achieve a small reduction in CI standard deviation from 0.0156 kg<sub>CO<sub>2</sub>,eq</sub>/km to 0.0134 kg<sub>CO<sub>2</sub>,eq</sub>/km. All optimized designs consider a bus fleet entirely consisting out of FCEB, i.e., no diesel-powered buses are present in the bus fleet. As the electrolyzer capacity is typically scaled with the hydrogen demand, the capacity for the electrolyzer array is similar for all optimized designs (6.8 MW). The optimized CI mean is achieved by solely considering a wind turbine array (15.7 MW), while the robust design consists of a slightly smaller wind turbine array (14.1 MW), supported by a PV array of 1 MW (Figure 3.15, top-right). Following the reduced wind array capacity and the increased PV array capacity in the robust design, the match between the renewable power production profile and the hydrogen demand profile (during the evening, night and early morning) is worse. To store the additional hydrogen produced outside the dispensation hours, the hydrogen storage capacity is increased up to 203 MWh in the robust design (Figure 3.15, bottom-left). The Sobol' indices for the optimized CI mean design illustrate that the uncertainty on the GHG emissions related to the wind turbine array construction (Sobol' index: 0.76) and the uncertainty on the wind capacity factor (Sobol' index: 0.14) are the main drivers of the CI standard deviation (Figure 3.15, bottom-right). Furthermore, the CI for this design is significantly affected by the actual wind capacity factor, which means that the hydrogen storage tank does not cover the load during the entire year. Instead, part of the hydrogen demand is covered by the actual wind power available, which is sensitive to the wind capacity factor. By gradually increasing the hydrogen storage and replacing a fraction of the wind turbine array with a PV array, the robust design removes the contribution of the uncertainty on the wind capacity factor to the CI standard deviation by relying on the stored hydrogen, filled outside dispensation hours, to cover the hydrogen demand. Instead, the uncertainty related to the GHG emissions during

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions



**Figure 3.15:** The Pareto front illustrates a trade-off between minimizing the Carbon Intensity (CI) mean and CI standard deviation. Considering only wind power results in the optimized CI mean, while the robust CI is achieved by removing the dependency on the wind capacity factor uncertainty, through replacing part of the wind capacity with photovoltaic capacity and increasing the hydrogen storage capacity.

construction of the hydrogen storage tank (Sobol' index: 0.10) and PV array (Sobol' index: 0.06) gain importance.

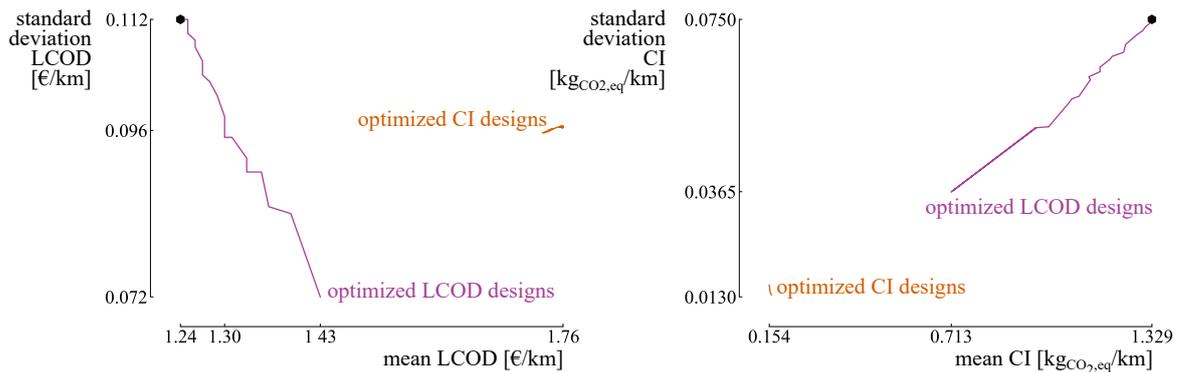
**Overall performance comparison** As illustrated in the RDO results on each quantity of interest, the optimized stochastic designs for the CI and LCOD differ in design characteristics. Therefore, the performance from the optimized designs in one quantity of interest might result in a poor performance in the other quantity of interest. Indeed, when comparing the performance in CI for the designs optimized in LCOD (and the other way around), the performance is significantly worse (Figure 3.16).

As expected from the RDO on the CI, the relatively small trade-off between minimizing the CI mean and minimizing the CI standard deviation results in a small range on the LCOD performance for these designs (Figure 3.16, optimized CI designs). The optimized CI

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

designs achieve an acceptable LCOD standard deviation, around 0.096 €/km, which ranges between the minimum and maximum LCOD standard deviation achieved by the optimized LCOD designs. However, the optimized CI designs achieve a significantly higher LCOD mean. To illustrate, when comparing with the optimized LCOD mean design that achieves a similar LCOD standard deviation of 0.096 €/km, the LCOD mean for the optimized CI design is 35 % higher (i.e., LCOD mean of 1.76 €/km, as opposed to 1.30 €/km).

Among the optimized designs, all optimized LCOD designs perform worse in both CI mean and CI standard deviation (Figure 3.16, optimized LCOD designs). The worst-performing optimized LCOD design achieves a CI mean of 1.33 kg<sub>CO<sub>2</sub>,eq</sub>/km, with a CI standard deviation of 0.075 kg<sub>CO<sub>2</sub>,eq</sub>/km. This design corresponds to the design with the optimized LCOD mean of 1.24 €/km and is characterized by a complete diesel-powered fleet. This comparison shows that, when considering an optimized LCOD design with a reduced LCOD standard deviation, this design also improves both the CI mean and CI standard deviation.



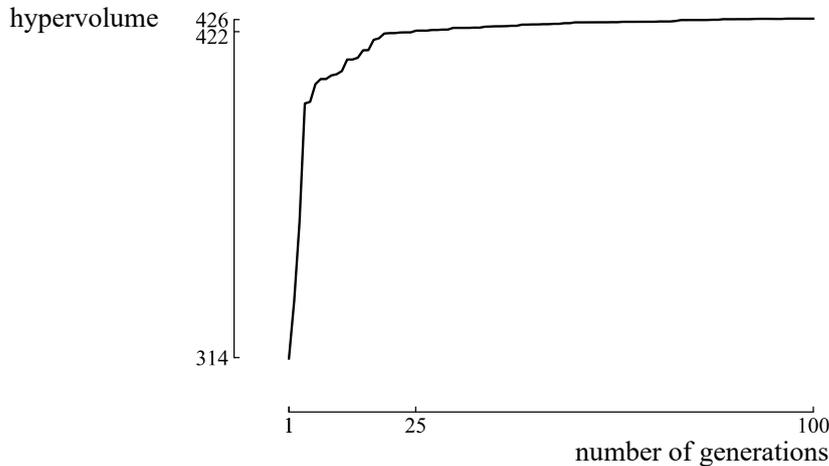
**Figure 3.16:** The performance in one quantity of interest of the designs optimized for the other quantity of interest illustrates a clear trade-off between optimizing the stochastic performance for the two quantities of interest.

#### 3.4.6.2 Holistic Robust Design Optimization

In the holistic RDO, four objectives are minimized simultaneously (i.e., minimizing the mean and standard deviation for the LCOD and CI). Hence, four objective values are quantified for each design in a generation of design samples. The hypervolume indicates a converged Pareto front after 100 generations (Figure 3.17). In each consecutive generation, the design samples converge towards the Pareto set of optimized samples in the final generation (Figure 3.18). Following the nondominated sorting principle of NSGA-II, a

### Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions

design remains in the Pareto set of optimized designs as long as it dominates every other design in at least one of the four objectives. Consequently, the optimized designs found in the individual RDO (Figure 3.16) are extended in the holistic RDO (Figure 3.19).

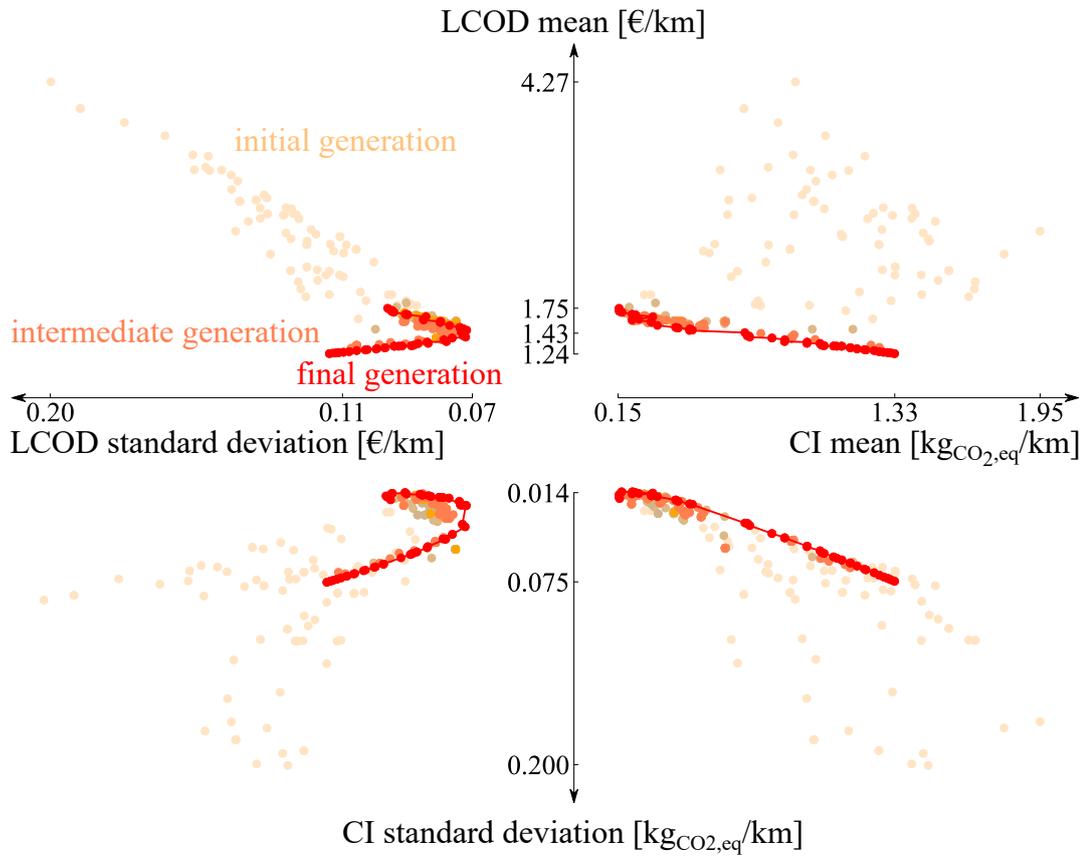


**Figure 3.17:** The hypervolume stabilizes, indicating that the Pareto front does not significantly improve over the generations. The reference point is defined as [5 €/km, 5 €/km, 5 kg<sub>CO<sub>2</sub>,eq</sub>/km, 5 kg<sub>CO<sub>2</sub>,eq</sub>/km].

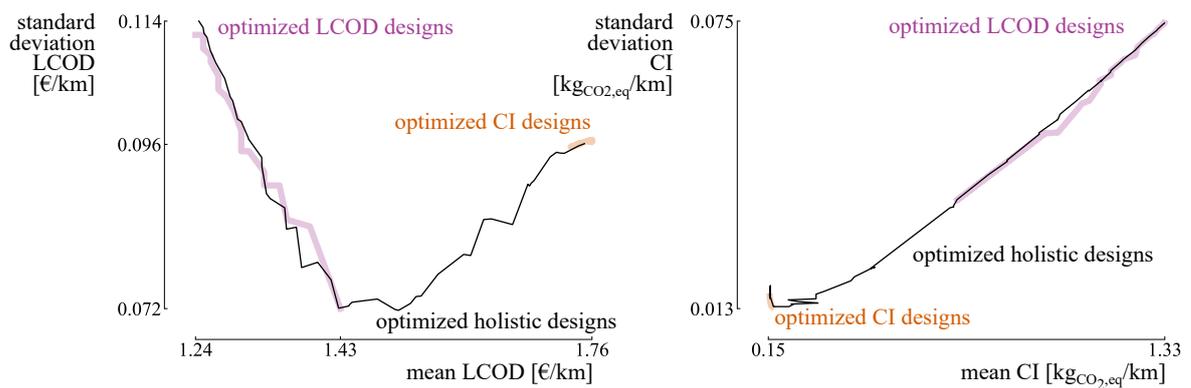
To illustrate the performance for each design and why each design is considered optimized in the holistic RDO, the achieved LCOD standard deviation, CI mean and CI standard deviation are presented in function of the achieved LCOD mean (Figure 3.20, left). For the optimized LCOD mean design (1.24 €/km), the worst LCOD standard deviation (0.11 €/km), CI mean (1.33 kg<sub>CO<sub>2</sub>,eq</sub>/km) and CI standard deviation (0.075 kg<sub>CO<sub>2</sub>,eq</sub>/km) are achieved among the optimized designs. Accepting an increase in LCOD mean up to 1.43 €/km results in alternative designs with a reduced CI mean down to 0.71 kg<sub>CO<sub>2</sub>,eq</sub>/km (46 %), reduced CI standard deviation down to 0.037 kg<sub>CO<sub>2</sub>,eq</sub>/km (51 %) and a reduced LCOD standard deviation down to 0.07 €/km (37 %), i.e., the robust LCOD (Figure 3.20, point A). Thus, this first subset of designs illustrates that switching from a diesel-powered bus fleet (i.e., the optimized LCOD mean design) into a mobility demand that is partly covered by FCEB (54 %) improves not only the LCOD robustness, but also the average expected CI and the CI robustness.

In the second part of the optimized designs, with an LCOD mean between 1.43 €/km and 1.62 €/km, the LCOD standard deviation is increasing, as the robust LCOD design was reached at an LCOD mean of 1.43 €/km. Hence, this subset of optimized designs is not considered Pareto-optimal in the LCOD statistical moments. Instead, accepting alternative

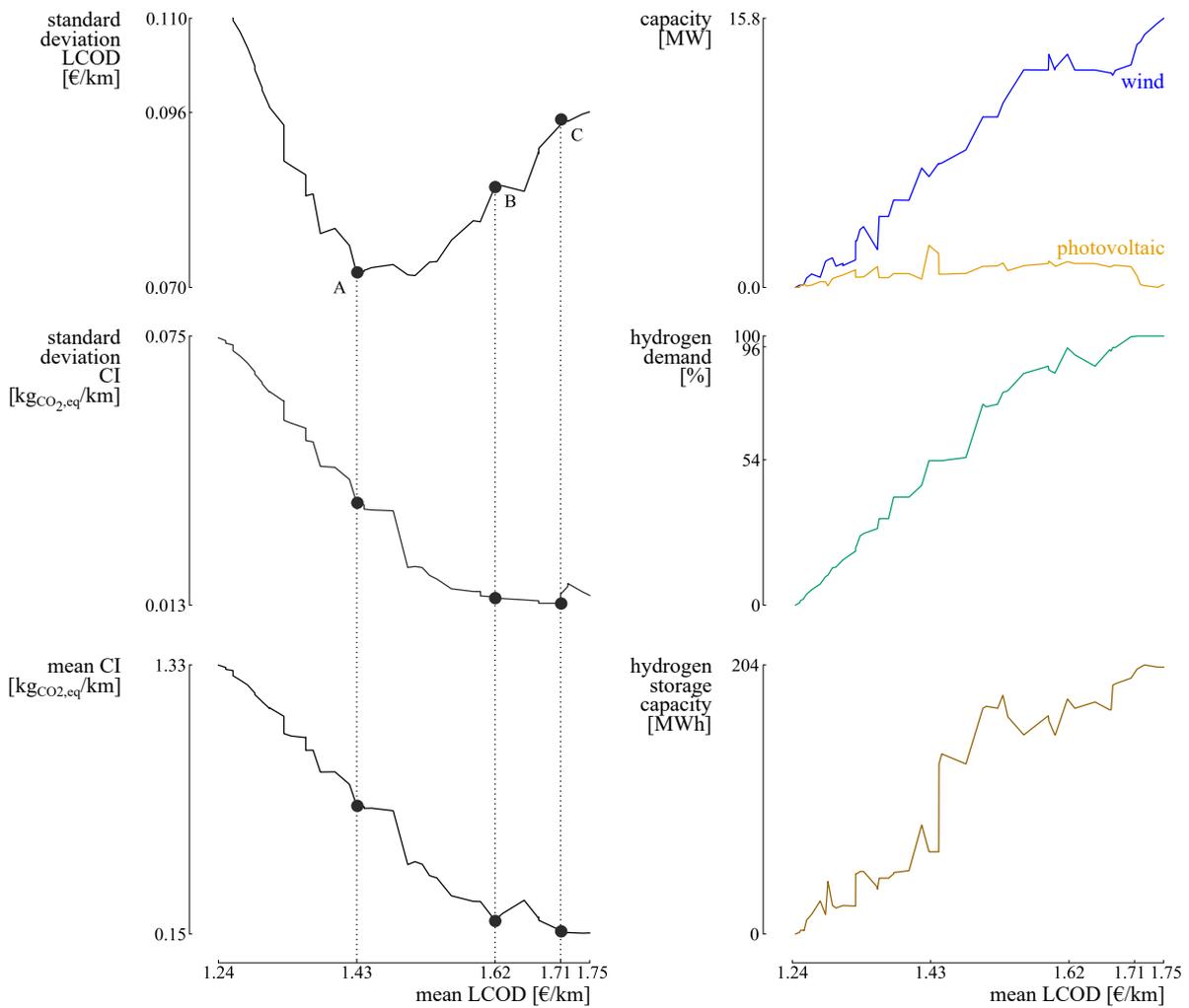
### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station



**Figure 3.18:** In the holistic Robust Design Optimization (RDO), four objectives are minimized simultaneously (mean and standard deviation for the Levelized Cost of Driving (LCOD) and Carbon Intensity (CI)), in each consecutive generation of design samples. As minimization is desired, the samples converge towards the centre of the cross-plot.



**Figure 3.19:** The optimized designs resulting from a Robust Design Optimization (RDO) on each quantity of interest separately reappear in the optimized designs resulting from the holistic RDO (optimized holistic designs).



**Figure 3.20:** The results from the holistic Robust Design Optimization (RDO) illustrate that the optimized Levelized Cost of Driving (LCOD) mean corresponds to the worst LCOD standard deviation, Carbon Intensity (CI) mean and CI standard deviation. The LCOD standard deviation, CI mean and CI standard deviation are subsequently reduced by gradually replacing the diesel-powered buses with Fuel Cell Electric Buses (FCEB), accompanied by a wind turbine array and photovoltaic array that are scaled accordingly.

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

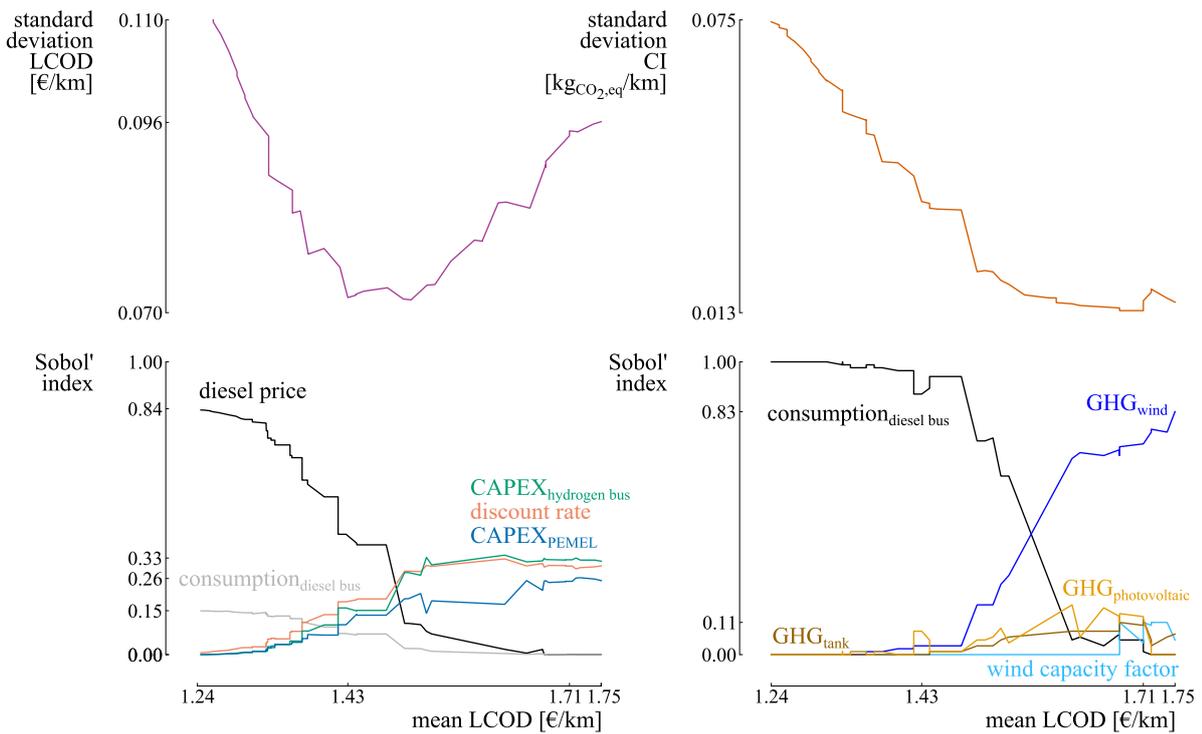
---

designs with an increase in LCOD mean up to 1.62 €/km results in a continuous decrease in CI mean and CI standard deviation (Figure 3.20, point B). In these designs, the fraction of the mobility demand converted into FCEB is increased up to 96 % (Figure 3.20, middle-right), accompanied by an increase in the size of the wind turbine array (Figure 3.20, top-right) and the hydrogen storage tank (Figure 3.20, bottom-right). Due to the significant mean and significant uncertainty on the GHG emissions during production of the PV array, the optimizer does not further increase the capacity of the PV array in this subset of optimized designs. In conclusion, this second subset of optimized designs provides intermediate solutions between the stochastic performance of LCOD and CI.

In the third subset of optimized designs, between 1.62 €/km and 1.71 €/km, the robust CI design is reached by converting the remaining diesel-powered buses into FCEB. However, the gain in CI standard deviation is marginal ( $0.013 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ , as opposed to  $0.015 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ ), while the mean CI stabilizes around  $0.21 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$  and the LCOD mean and LCOD standard deviation further increase up to 1.71 €/km and 0.094 €/km, respectively (Figure 3.20, point C).

The fourth and final subset of optimized designs, with an LCOD mean between 1.71 €/km and 1.75 €/km, reach the optimized CI mean by replacing the PV capacity with an adequate wind power capacity. This final subset corresponds to the optimized designs from the individual RDO on the CI and range between the optimized CI mean ( $0.15 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ ) and optimized CI standard deviation ( $0.013 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ ) achieved in the holistic RDO.

In addition to modifying the design, the standard deviation on the quantities of interest can be reduced by reducing the epistemic uncertainty on the input parameters. As indicated on Figure 3.20, the diesel-powered bus fleet with the optimized LCOD mean of 1.24 €/km results in the worst LCOD standard deviation and worst CI standard deviation among the optimized designs. For this design, the uncertainty on the diesel cost is the main driver for the LCOD standard deviation, while the uncertainty on the consumption of the diesel bus is the main driver for the CI standard deviation (Figure 3.21). These are aleatory uncertainties, which means that these uncertainties cannot be reduced by measures taken by the system designer. Up until the robust design for the LCOD, corresponding to an LCOD mean of 1.43 €/km, the LCOD and CI standard deviation are mostly driven by these input uncertainties. However, for the designs for which at least 54 % of the diesel-powered buses are replaced by FCEB (i.e., LCOD mean between 1.43 €/km and 1.75 €/km), the aleatory un-



**Figure 3.21:** The Sobol' indices on the Levelized Cost Of Driving (LCOD, bottom-left) and the Carbon Intensity (CI, bottom-right) indicate that the design corresponding to the best LCOD mean is subject to the diesel-related aleatory uncertainties, while the designs with higher LCOD mean, between 1.43 €/km and 1.75 €/km, are subject to epistemic uncertainties related to the renewable-powered Hydrogen Refueling Station (HRS).

### 3.4 Power-to-mobility: renewable-powered hydrogen refueling station

---

certainties related to diesel-powered buses are replaced by epistemic uncertainties. Hence, the standard deviation of the quantities of interest for the designs which are mainly characterized by FCEB can be reduced by gaining more knowledge on the important uncertainties. To illustrate, for the design corresponding to an LCOD mean of 1.71 €/km (96 % of the bus fleet are FCEB), the main drivers for the LCOD standard deviation are the uncertainty on the CAPEX of the FCEB (Sobol' index: 0.33), discount rate (Sobol' index: 0.32), CAPEX of the electrolyzer array (Sobol' index: 0.26). As the amount of diesel-powered buses in the fleet is negligible, the Sobol' indices corresponding to the diesel price and consumption of the diesel-powered buses are zero. As only epistemic uncertainties remain, the LCOD standard deviation can be reduced by gaining more knowledge on the important parameters. For this design, specifying the PEM electrolyzer and FCEB supplier and determining the finance type are the main actions to reduce the LCOD standard deviation. For the CI standard deviation, the main driver is the uncertainty on the GHG emissions during construction of the wind turbine array. Reducing this uncertainty by specifying the wind turbine technology and manufacturer is the main action to improve the CI robustness.

#### 3.4.7 Conclusion

Renewable-powered Hydrogen Refueling Stations (HRS) for a Fuel Cell Electric Bus (FCEB) fleet achieve excellent environmental performance, at the expense of a higher fuel cost as opposed to a traditional diesel-powered bus fleet. However, following the significant uncertainties associated with diesel, a renewable-powered HRS might achieve additional benefits in terms of robustness when considering techno-economic and environmental uncertainties. A diesel-powered bus fleet achieves the optimized Levelized Cost Of Driving (LCOD) mean of 1.24 €/km, subject to a standard deviation of 0.11 €/km. However, among the optimized designs, a diesel-powered bus fleet achieves the highest Carbon Intensity (CI) mean and CI standard deviation, equal to 1.33 kg<sub>CO<sub>2</sub>,eq</sub>/km and 0.075 kg<sub>CO<sub>2</sub>,eq</sub>/km, respectively. Alternative designs are proposed which reduce the LCOD standard deviation, CI mean and CI standard deviation by converting part of the diesel-powered bus fleet into FCEB and by sizing the renewable-powered HRS accordingly. The robust design for the LCOD reduces the LCOD standard deviation by 36 % (down to 0.07 €/km), at the expense of an increase in LCOD mean by 11 % (up to 1.43 €/km). This design corresponds to a bus fleet with 54 % FCEB and an HRS consisting of a 6.7 MW wind turbine array, 2.5 MW PV array

### **Chapter 3. Sparse polynomial chaos expansion: handling large stochastic dimensions**

---

and 62 MWh hydrogen storage. In addition to the optimized LCOD standard deviation, this design achieves a reduced CI mean of 46 % ( $0.71 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ ) and reduced CI standard deviation of 51 % ( $0.036 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ ) and consequently provides a suitable alternative to a diesel-powered bus fleet. The optimized CI mean and optimized CI standard deviation are achieved with a bus fleet entirely consisting out of FCEB. While the optimized CI mean of  $0.15 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$  is achieved with a wind turbine array of 15.8 MW, the robust alternative for the CI consists of a 14.1 MW wind turbine array and 1.0 MW PV array, enabling to reduce the standard deviation from  $0.015 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$  to  $0.013 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ , at the expense of an increase in CI mean up to  $0.16 \text{ kg}_{\text{CO}_2,\text{eq}}/\text{km}$ . This gain in robustness is mainly due to the reduced significance of the wind capacity factor uncertainty in the robust design by replacing part of the wind turbine array with a PV array. Nevertheless, these optimized CI designs are subject to a large LCOD mean and LCOD standard deviation, equal to  $1.75 \text{ €/km}$  and  $0.096 \text{ €/km}$ , respectively. However, these designs are subject to reducible epistemic uncertainties, while the diesel-powered bus fleet is mainly subject to irreducible aleatory uncertainties. Therefore, the robustness of the renewable-powered HRS can be improved by reducing the uncertainty on the input parameters. In this case, specifying the electrolyzer and FCEB supplier, determining the finance type and specifying the wind turbine technology and manufacturer enable to further reduce the standard deviation on the quantities of interest.

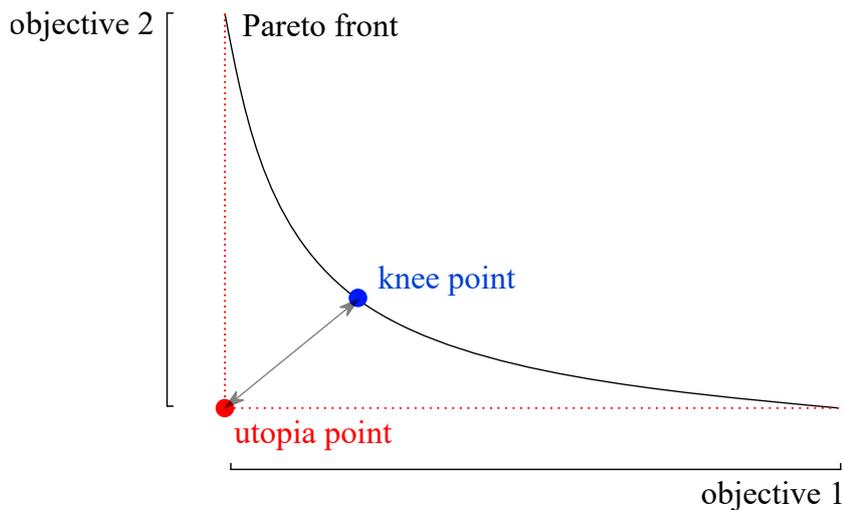
### **3.5 Decision making from the Pareto front**

For the applications, the main results are provided through a Pareto front, which illustrates the trade-off between minimizing the mean and minimizing the standard deviation for a quantity of interest. While each design on the Pareto front can be considered optimized, i.e., each design dominates all the other designs in at least one objective, a discussion on the decision-making process can guide the decision-maker towards choosing the most satisfactory design. In this section, first, a general analysis is discussed, followed by a specific evaluation for RDO. Finally, a brief decision-making process is suggested.

When analysing a Pareto front in a general context, a typical method to choose the most satisfactory design is to identify the knee point. The knee point corresponds to the design that obtains the minimum Euclidian distance towards a utopia point. The utopia point can

be identified by the user [254], or it can be defined by the best values obtained for each objective on the Pareto front (Figure 3.22). The latter approach is subject to the quality of the Pareto front, which is not optimal due to the inherent structure of NSGA-II. The range for each objective on the Pareto front will determine if the knee point is situated on the extremes, or for an intermediate design. In the cases discussed in this thesis, the absolute range on the standard deviation is significantly lower than the absolute range on the mean (e.g., for the power-to-power application, the LCOE standard deviation is reduced from 55.1 €/MWh to 35.4 €/MWh, at the expense of an increase in LCOE mean from 269 €/MWh to 367 €/MWh, Figure 3.4). Hence, the knee-point method likely favours designs with an optimized mean. Yet, one can argue on the reason for selecting the standard deviation as the objective, as the standard deviation only indicates a fraction of the possible outcomes for the quantity of interest (e.g., for a Gaussian distribution, 68 % of the possible outcomes for the quantity of interest lie within one standard deviation of the mean). Instead, a valid alternative objective can be a multitude of standard deviations, as it captures a larger share of possible outcomes for the quantity of interest (e.g., three times the standard deviation captures 99.7 % of the possible outcomes for a Gaussian distribution). Replacing the standard deviation with a multitude of standard deviations as an objective increases the weight of the robustness in the decision-making process. Alternatively, the objective values from the Pareto front can be normalized, indicating the relative difference in expected value (mean) and robustness (standard deviation) between designs. However, while the relative gain can be large, the absolute gain can be negligible (e.g., in the power-to-mobility application, a relative gain of 51 % in CI standard deviation is achieved, corresponding to a gain from 0.075 kg<sub>CO<sub>2</sub>,eq</sub>/km to 0.037 kg<sub>CO<sub>2</sub>,eq</sub>/km). Despite that the relative gain is independent from the multitude of standard deviations considered, the relative gain is significantly increased when replacing the standard deviation with the variance as an objective. As a conclusion, the choice for the objective to indicate the uncertainty on the quantity of interest significantly impacts the distance towards the utopian point on an RDO Pareto front. Therefore, the use of the knee point method on a Pareto front in RDO applications is ambiguous. To illustrate, the knee point is defined for the four-bar truss example in Section A.5, on a typical and normalized Pareto front.

Aside from generic methods to decide a satisfactory design from a Pareto front, the meaning of both objectives can be considered. The design with the optimized mean corresponds to the design with the optimized expected performance. The design with the optimized



**Figure 3.22:** The knee point is the design on the Pareto front that minimizes the distance towards the utopia point. The utopia point is configured by the best values for both objectives.

standard deviation corresponds to the design with the highest probability that the true outcome will situate near the mean. The design with the optimized mean can be adopted when the user wants the expected performance to be optimized, without much interest in the uncertainty on that performance. Alternatively, a robust alternative can be adopted when the predictability of the outcome is of high importance (e.g., the credibility of a firm can be undermined when the difference between the actual and projected costs are high). One can argue that the design with the optimized expected performance should be adopted, as long as the uncertainty on the expected performance is tolerable. Hence, absolute robustness of the performance can be achieved when the uncertainty on the expected performance is below a certain threshold. To illustrate, when evaluating the project cost, an overbudget of up to 20 % is generally accepted. Thus, a design that complies with this upper limit might be a satisfying choice.

Finally, additional information for the decision-making process can be attained by combining the two objectives: the mean and standard deviation characterize a single distribution. Such a distribution indicates the probability that a possible outcome situates below a certain value. Hence, when a reliability measure is considered on the quantity of interest, the distribution provides the probability that the design will comply with the upper limit. To illustrate, in the power-to-power application, a PV-battery configuration provides a probability of 68 % that the true outcome of the LCOE will be below 300 €/MWh, while full grid dependency achieves only a probability of 49 % (Figure 3.6). Thus, in addition to separately

considering the mean and standard deviation for a design, generating the distribution for that design provides information on the probability of not violating an upper limit on the quantity of interest. To illustrate, the CDF is defined for the optimized mean design and robust design of the four-bar truss example in Section A.5.

In summary, decision-making from a set of optimized designs, generated by RDO, is subject to several variables. When both expected performance and robustness are of interest, it is suggested to evaluate the absolute and relative trade-off between mean and standard deviation, to have an impression on the absolute and relative gain in robustness, at the expense of an absolute and relative loss in expected performance. If the absolute differences for both statistical moments is significant, the relative difference is important to decide on the adoption of the satisfactory design. Instead, when the absolute difference is negligible on e.g., the standard deviation, the design with the optimized mean might be most satisfactory and might achieve absolute robustness. Finally, if an upper limit is considered on the quantity of interest, the distribution for each design can be evaluated, to quantify the probability of violating the upper limit when installing that specific design.



## 4 Imprecise probabilities: handling uncertainty about uncertainty

### Chapter overview

- Epistemic and aleatory uncertainty characterization through imprecise probability;
- Augmented PCE to propagate imprecise probabilities;
- Application on a photovoltaic-battery-heat pump with thermal storage.

*The application on a photovoltaic-battery-heat pump with thermal storage is published in Energy [255].*

Most often, only scarce knowledge is available on the parametric uncertainty of HRES, leading to an epistemic uncertainty (i.e., lack of knowledge) alongside the aleatory uncertainty (i.e., natural variability) [86]. To illustrate, Moret et al. [87] described the parameters subject to epistemic and aleatory uncertainty in energy system planning. In the stochastic optimization of HRES, precise probabilistic uncertainty and interval uncertainty are the most commonly adopted methods for uncertainty characterization, resulting in a single range or distribution. Hence, with these methods, either no attention is given to the epistemic and aleatory uncertainty on a parameter, or only the aleatory (or epistemic) uncertainty is considered, or the epistemic uncertainty and aleatory uncertainty on a parameter are combined in a single representation. Consequently, when performing a global sensitivity analysis on a quantity of interest, and the Sobol' indices for the stochastic parameters are quantified, the meaning of this Sobol' index is ambiguous: Is the uncertainty on this parameter irreducible? Can the uncertainty be reduced by external measures, or can it be

reduced by improving the knowledge on the parameter?

Imprecise probabilities provide a parametric uncertainty characterization that elegantly distinguishes between the epistemic and aleatory uncertainty source [256]. Different imprecise probability techniques exist, such as evidence theory [257], fuzzy probability theory [258] and probability boxes (p-boxes) [259]. Among the imprecise probability characterization methods, p-boxes provide an intuitive representation of mixed epistemic and aleatory uncertainty on model parameters [260]. Nevertheless, as emphasized by Beer et al. [261] in their overview of imprecise probability techniques in engineering analyses, these methods are closely related, with strong theoretic and algorithmic interconnections: *“Whether we are using probabilistic methods, interval methods, p-boxes, or fuzzy probabilistic techniques, we are drinking the same water of truth from different sides of the same well.”* [261].

When adopting imprecise probability methods to characterize the uncertainty on the model inputs, a specific UQ method is required to propagate these imprecise probabilities through the system model and quantify their effect on the quantity of interest. UQ in an imprecise probability context is gaining interest, but is far less developed than the precise probability counterpart [262]. The main limitation is the increased computational cost, which can be already significant in a precise context [94]. UQ applications in an imprecise probability context often rely on a computationally expensive (i.e.,  $> 10^6$  evaluations) nested MCS [263]. Recently, Schöbi et al. [264] proposed an augmented PCE that enables propagating p-boxes and quantifying the imprecise Sobol’ indices in a computationally efficient manner.

### 4.1 Contributions

First, the uncertainties in an HRES are characterized using imprecise probabilities. This allows to separately characterize the epistemic and aleatory uncertainty on the stochastic parameters in an HRES. Second, the surrogate-assisted RDO algorithm was adapted, enabling to propagate the p-boxes and quantifying the imprecise probabilities in a computationally-efficient manner. In this modified surrogate-assisted RDO algorithm, the aim is to minimize the mean and standard deviation related to the aleatory uncertainty. As this uncertainty cannot be reduced by gaining more information on the input parameters, the surrogate-assisted RDO provides designs least sensitive to this aleatory uncertainty. When the optimized robust designs are determined, a global sensitivity analysis is performed on the optimized

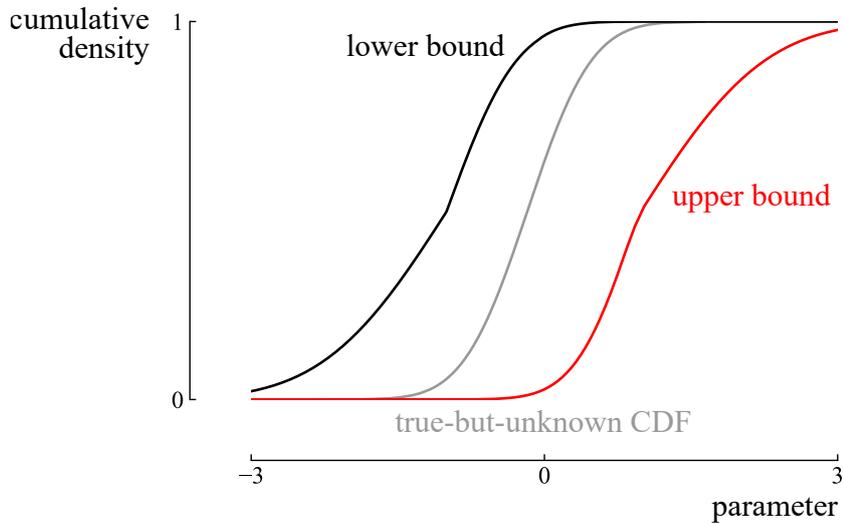
designs, to determine the main actions to reduce the remaining epistemic uncertainty on the quantity of interest. Third, this approach was applied to a PV-battery-heat pump system with thermal storage (Section 4.3) [255]. The results indicate that a PV-battery-heat pump with thermal storage system is more robust in Levelized Cost Of eXergy (LCOX) than a PV-battery system with a gas boiler. A PV-battery-heat pump with thermal storage system achieves an upper-bound LCOX standard deviation of 70.0 €/MWh, which reduces the standard deviation by 36 % as opposed to the PV-battery-gas boiler design with an upper-bound LCOX standard deviation of 109.8 €/MWh. In the optimized PV-battery-heat pump designs, the uncertainty on the occupant behavior is the main driver of the epistemic uncertainty on the LCOX. As a conclusion, the surrogate-assisted RDO enables adopting more complex uncertainty characterization methods, which allow distinguishing between the effects of reducible and irreducible uncertainties on the performance of the HRES.

## 4.2 Polynomial chaos expansion for parametric probability boxes

In this section, first, a parametric p-box is defined. Thereafter, the augmented PCE is described, followed by the methods to retrieve the statistical moments and the imprecise Sobol' indices from that augmented PCE. The section concludes with an example and the integration of the augmented PCE in the surrogate-assisted RDO algorithm. The propagation of parametric p-boxes and the quantification of imprecise Sobol' indices is performed on an illustrative example in Appendix, Section A.6.

**Parametric probability box** A precise distribution is typically represented by a PDF and CDF. Instead of being represented by a single CDF, a p-box is defined by a lower bound  $\underline{F}_X$  and upper bound  $\overline{F}_X$  (Figure 4.1). For any value of  $x \in D_X$  (i.e.,  $D_X$  is the support of  $X$ ), the true-but-unknown CDF satisfies  $\underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x)$ . P-boxes are typically categorized as *free* and *parametric* p-boxes [265]. In a free p-box, the true-but-unknown CDF can be of any shape, as long as it is a non-decreasing function, and it does not violate the p-box bounds. In a parametric p-box, each CDF realization corresponds to an underlying distribution family. Hence, the distribution family for the true-but-unknown CDF is fixed. A parametric p-box is characterized by interval-ranged statistical moments (i.e., the statistical

moments are represented by an interval, instead of a deterministic value):  $\mu \in [\mu_{\min}, \mu_{\max}]$ ,  $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ . In this formulation, the true-but-unknown mean  $\mu$  and standard deviation  $\sigma$ , which characterize the aleatory uncertainty, are situated in between the respective bounds. The range on these statistical moments represents epistemic uncertainty. Hence, due to the limited information on the parameter, the aleatory uncertainty (i.e.,  $\mu$  and  $\sigma$ ) can not be precisely determined. Parametric p-boxes are adopted in this thesis, as this approach allows to separate the epistemic and aleatory uncertainty on a parameter. A Gaussian distribution is assumed for the underlying family.



**Figure 4.1:** Illustration of a parametric probability box  $X \sim \mathcal{N}(\mu_X, \sigma_X)$ , where  $\mu_X \in [-1, 1]$  and  $\sigma_X \in [0.5, 1]$ . The true-but-unknown CDF lies between the lower bound and upper bound of the p-box, and follows a Gaussian distribution.

**Augmented PCE** To propagate p-boxes through a model and to quantify their effect on a quantity of interest, the augmented PCE representation from Schöbi et al. [264] was adopted. When parametric p-boxes are considered, the epistemic and aleatory uncertainty can be separated into two components:  $\Theta$  and  $\mathbf{X}$ , respectively [264]. Hence, an augmented input vector  $(\Theta, \mathbf{X})$  becomes the stochastic input for the system model. The components of the augmented input vector are dependent on each other, i.e., the value sampled from the distribution in  $\mathbf{X}$  depends on the statistical moments sampled from  $\Theta$ . As PCE requires independent parameters, the augmented input vector  $(\Theta, \mathbf{X})$  needs to be mapped into a vector  $\mathbf{V}$  with independent components, through isoprobabilistic transformation [264]. For a Gaussian p-box (i.e., a parametric p-box with a Gaussian underlying distribution),

## 4.2 Polynomial chaos expansion for parametric probability boxes

---

with  $\boldsymbol{\theta} = (\mu_X, \sigma_X)$  and  $\mu_X$  and  $\sigma_X$  are given in intervals, the isoprobabilistic transform corresponds to:

$$X = \mu_X + \sigma_X \zeta, \quad (4.1)$$

where  $\zeta$  is a standard Gaussian variable  $\zeta \sim \mathcal{N}(0, 1)$ . Then, the vector with independent components corresponds to  $\mathbf{V} = (\zeta, \mu_X, \sigma_X)$ , and the two independent components for epistemic uncertainty and the aleatory uncertainty are defined as  $\boldsymbol{\Theta} = (\mu_X, \sigma_X)$  and  $\mathbf{C} = (\zeta)$ , respectively. Consequently, the dimensionality  $M_{\text{aug}}$  is tripled, when compared to the number of p-boxes, i.e.,  $M_{\text{aug}} = |\mathbf{V}| = 3|\mathbf{C}|$ . Following this transformation into independent components, the PCE for parametric p-boxes can be constructed:

$$\mathcal{M}^{\text{aug,PCE}}(\mathbf{V}) = \sum_{\alpha \in \mathcal{A}} u_{\alpha} \Psi_{\alpha}(\mathbf{V}). \quad (4.2)$$

To compute the coefficients in the PCE, training samples are needed. Due to the augmented dimensionality, the conventional truncation scheme (subsection 2.2.2.1) quickly results in a computationally intractable problem. Therefore, the stepwise regression approach (Section 3.2) was adapted in this thesis to this new formulation, resulting in a computationally-efficient method to quantify the PCE coefficients.

**Retrieve statistical moments** When the PCE is constructed, the interval-ranged statistical moments on the quantity of interest can not be analytically derived from the coefficients. Instead, the interval-ranged mean and interval-ranged standard deviation for the quantity of interest are computed through a nested MCS on the PCE [266]. The nested MCS consists of an outer loop and an inner loop [267]. The outer loop samples the interval-ranged statistical moments for each p-box. Hence, each sample consists of a mean and standard deviation for each p-box. Then, for each of these outer-loop samples, a set of Gaussian distributions is constructed, one for each pair of mean and standard deviation. In the inner loop, these Gaussian distributions are sampled, to quantify the corresponding mean and standard deviation for the quantity of interest. Hence, for each outer-loop sample, a mean and standard deviation for the quantity of interest are retrieved. Following these outer-loop sampling, the intervals on the statistical moments are derived and the p-box is constructed. Despite the large number of training samples required (e.g.,  $10^3$  outer-loop samples,  $10^5$

## Chapter 4. Imprecise probabilities: handling uncertainty about uncertainty

---

inner-loop samples, resulting in  $10^8$  training samples), the computational cost remains tractable (i.e., a handful of seconds), as the training samples are evaluated directly on the PCE analytical model. To illustrate the method and philosophy, an example is presented in Section A.6, where the uncertainty characterization and UQ with parametric p-boxes are demonstrated on a typical four-bar truss case.

**Imprecise Sobol' indices** Similar to the statistical moments, imprecise Sobol' indices are defined by a range, i.e., the true-but-unknown Sobol' index  $S_j$  lies in the range  $[S_{j,\min}, S_{j,\max}]$ . The true-but-unknown Sobol' index  $S_j$  corresponds to the effect of the true aleatory uncertainty, which is unknown due to the global presence of epistemic uncertainty.

As opposed to the statistical moments, the imprecise total-order Sobol' indices can be derived analytically from the PCE. To determine these imprecise Sobol' indices, a rearrangement of the coefficients is required. First, the multi-indices and orthonormal polynomials are decomposed into the multi-indices and orthonormal polynomials related to the epistemic uncertainty  $\Theta$  and aleatory uncertainty  $\mathbf{C}$ , i.e.,  $\alpha = (\alpha_\Theta, \alpha_C)$  and  $\Psi_\alpha(\mathbf{V}) = \Psi_{\alpha_\Theta}(\Theta)\Psi_{\alpha_C}(\mathbf{C})$ , respectively. This results in the following reformulation of the augmented PCE:

$$\mathcal{M}^{\text{aug,PCE}}(\mathbf{C}, \Theta) = \sum_{\alpha \in \mathcal{A}} u_\alpha \Psi_{\alpha_C}(\mathbf{C}) \Psi_{\alpha_\Theta}(\Theta), \quad (4.3)$$

where  $\mathbf{C}$  includes the independent standard Gaussian distributions. The vector of the interval-ranged statistical moments corresponds to  $\Theta = \{\theta_1, \dots, \theta_M\}$ , with  $M$  the number of parametric p-boxes at the model input, and  $\theta_i$  represents the vector of epistemic parameters (i.e., intervals) for parameter  $i$ . Based on the idea of nested MCS, each sample from the outer loop, containing a value for the mean and standard deviation of each p-box (i.e.,  $\theta = (\boldsymbol{\mu}, \boldsymbol{\sigma})$ ,  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_M)$  and  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_M)$ ), results in a different value for the Sobol' index of parameter  $j$ . Hence, by sampling the interval on the statistical moments (i.e., the outer loop in nested MCS), the effect of epistemic uncertainty is considered on the Sobol' index, resulting in an interval-ranged imprecise Sobol' index for each stochastic parameter defined by a p-box. Thus, the PCE should be rearranged for a given value  $\theta$  as a function of

## 4.2 Polynomial chaos expansion for parametric probability boxes

$\mathbf{C}$ , such that the Sobol' indices, conditioned by  $\theta$ , can be retrieved via the PCE coefficients:

$$\mathcal{M}^{\text{aug,PCE}}(\mathbf{C}, \theta) = \sum_{\alpha \in \mathcal{A}} (u_{\alpha} \Psi_{\alpha_{\Theta}}(\theta)) \Psi_{\alpha_{\mathbf{C}}}(\mathbf{C}) = \sum_{\alpha \in \mathcal{A}} u_{\alpha, \theta}(\theta) \Psi_{\alpha_{\mathbf{C}}}(\mathbf{C}). \quad (4.4)$$

In this formulation,  $u_{\alpha, \theta}(\theta)$  is a new coefficient, depending on the sampled statistical moments, stored in  $\theta$ . However, the multi-indices  $\alpha_{\Theta}$  are not unique: several multi-indices  $\alpha_{\Theta}$  can correspond to a single  $\alpha_{\mathbf{C}}$ . To retrieve the unique indices, the PCE needs to be converted into a model with proper variance decomposition:

$$\mathcal{M}^{\text{aug,PCE}}(\mathbf{C}, \Theta) = \sum_{\alpha_{\mathbf{C}}^* \in \mathcal{A}_{\mathbf{C}}^*} \left( \sum_{i=1, \dots, n_{\mathcal{A}}} \mathbb{1}_{\{\alpha_{\mathbf{C}}^i = \alpha_{\mathbf{C}}^*\}} \left( \alpha_{\mathbf{C}}^i \right) u_{\alpha^i} \Psi_{\alpha_{\Theta}^i}(\theta) \right) \Psi_{\alpha_{\mathbf{C}}^*}(\mathbf{C}), \quad (4.5)$$

where  $\mathcal{A}_{\mathbf{C}}^*$  is the set of unique multi-indices  $\alpha_{\mathbf{C}}$  in  $\mathcal{A}$ , and  $\mathcal{A}$  consists of  $n_{\mathcal{A}}$  multi-indices.  $\mathbb{1}$  is the indicator function, which in this case equals 1 when  $\{\alpha_{\mathbf{C}}^i = \alpha_{\mathbf{C}}^*\}$ , and equals 0 otherwise. Finally,  $\alpha^i = (\alpha_{\mathbf{C}}^i, \alpha_{\Theta}^i)$ . Equation 4.5 corresponds to a PCE, defined by a set of unique indices  $\mathcal{A}_{\mathbf{C}}^*$  and the multivariate polynomials  $\Psi_{\alpha_{\mathbf{C}}^*}(\mathbf{C})$ . The coefficients for the PCE, corresponding to a unique multi-index  $\alpha_{\mathbf{C}}^*$  are identified as:

$$u_{\alpha_{\mathbf{C}}^*}(\theta) = \sum_{i=1, \dots, n_{\mathcal{A}}} \mathbb{1}_{\{\alpha_{\mathbf{C}}^i = \alpha_{\mathbf{C}}^*\}} \left( \alpha_{\mathbf{C}}^i \right) u_{\alpha^i} \Psi_{\alpha_{\Theta}^i}(\theta). \quad (4.6)$$

Note that this coefficient corresponds to the part of equation 4.5 between large brackets. Replacing this part with the symbol for the new coefficient leads to a simplified representation of the PCE:

$$\mathcal{M}^{\text{aug,PCE}}(\mathbf{C}, \Theta = \theta) = \sum_{\alpha_{\mathbf{C}}^* \in \mathcal{A}_{\mathbf{C}}^*} u_{\alpha_{\mathbf{C}}^*}(\Theta = \theta) \Psi_{\alpha_{\mathbf{C}}^*}(\mathbf{C}). \quad (4.7)$$

The coefficients presented in Equation 4.6 can be used to quantify the Sobol' indices related to  $\theta$ , similar to the approach used in PCE for precise probabilities (Equation 2.29). Hence, for each sample containing a set of statistical moments  $\theta$ , the Sobol' indices can be quantified, resulting in a set of possible outcomes for each Sobol' index. The bounds for each Sobol' index can be found through an inner optimization, where the combination of possible

statistical moments that leads to the lowest/highest Sobol' index is determined:

$$S_{\min, i_1 \dots i_s}^{\text{PCE}} = \min_{\theta \in \mathcal{D}_\Theta} \frac{D_{i_1 \dots i_s}(\theta)}{D(\theta)} = \min_{\theta \in \mathcal{D}_\Theta} \left[ \frac{\sum_{\alpha_C^* \in \mathcal{I}_{i_1 \dots i_s}} u_{\alpha_C^*}^2(\theta)}{\sum_{\alpha_C^* \in \mathcal{A}_C^*, \alpha_C^* \neq \mathbf{0}} u_{\alpha_C^*}^2(\theta)} \right], \quad (4.8)$$

$$S_{\max, i_1 \dots i_s}^{\text{PCE}} = \max_{\theta \in \mathcal{D}_\Theta} \frac{D_{i_1 \dots i_s}(\theta)}{D(\theta)} = \max_{\theta \in \mathcal{D}_\Theta} \left[ \frac{\sum_{\alpha_C^* \in \mathcal{I}_{i_1 \dots i_s}} u_{\alpha_C^*}^2(\theta)}{\sum_{\alpha_C^* \in \mathcal{A}_C^*, \alpha_C^* \neq \mathbf{0}} u_{\alpha_C^*}^2(\theta)} \right]. \quad (4.9)$$

The propagation of parametric p-boxes and the quantification of imprecise Sobol' indices is performed on an illustrative example in Appendix, Section A.6.

**Example** To illustrate the estimation of the Sobol' indices, a two-dimensional function is considered [264]:

$$f(x) = x_1 x_2, \quad (4.10)$$

where  $x_1$  and  $x_2$  are parametric p-boxes, defined by Gaussian distributions, with interval-ranged mean and standard deviation equal to  $\mu_i \in [-1, 1]$  and  $\sigma_i \in [0.5, 1.0]$ , respectively. Hence, the augmented input vector  $\mathbf{V}$  corresponds to  $(\mu_1, \sigma_1, \zeta_1, \mu_2, \sigma_2, \zeta_2)$ , and the augmented model becomes:

$$\begin{aligned} f^{\text{aug}}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\zeta}) &= (\mu_1 + \sigma_1 \zeta_1)(\mu_2 + \sigma_2 \zeta_2) \\ &= \mu_1 \mu_2 + \mu_2 \sigma_1 \zeta_1 + \mu_1 \sigma_2 \zeta_2 + \sigma_1 \sigma_2 \zeta_1 \zeta_2. \end{aligned} \quad (4.11)$$

This augmented model can be represented by a multivariate polynomial of maximum degree 4. In this example, developed by Schöbi et al. [264], the following set of multi-indices

## 4.2 Polynomial chaos expansion for parametric probability boxes

is retrieved using an adaptive Least Angle Regression [155]:

$$\mathcal{A} = \begin{pmatrix} \boldsymbol{\alpha}^{(0)} \\ \vdots \\ \boldsymbol{\alpha}^{(9)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}. \quad (4.12)$$

As the model consists of the sum of 4 monomial terms, 4 multi-indices are expected in the augmented PCE. However, an additional 6 terms are present, due to the normalization of the input distributions to a standard space. As illustrated in Equation 4.3, the matrix can be split into the contributions of the epistemic variables  $(\mu_1, \sigma_1, \mu_2, \sigma_2)$ , i.e., column 1, 2, 4 and 5, and aleatory variables  $(\zeta_1, \zeta_2)$ , i.e., column 3 and 6:

$$\mathcal{A}_{\Theta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathcal{A}_{\mathcal{C}} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}. \quad (4.13)$$

## Chapter 4. Imprecise probabilities: handling uncertainty about uncertainty

---

The row vectors in  $\mathcal{A}_C$  are not unique. Following the variance decomposition (Equation 4.5), the set of unique multi-indices  $\mathcal{A}_C^*$  contains four unique elements:

$$\mathcal{A}_C^* = \begin{pmatrix} \boldsymbol{\alpha}_C^{*(0)} \\ \boldsymbol{\alpha}_C^{*(1)} \\ \boldsymbol{\alpha}_C^{*(2)} \\ \boldsymbol{\alpha}_C^{*(3)} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (4.14)$$

Hence, the final coefficients, conditioned by the statistical moments  $\theta$ , are composed according to Equation 4.6:

$$u_{\boldsymbol{\alpha}_C^{*(0)}}(\theta) = u_{\alpha^{(0)}}\psi_{\alpha_{\Theta}^{(0)}}(\theta) + u_{\alpha^{(4)}}\psi_{\alpha_{\Theta}^{(4)}}(\theta), \quad (4.15)$$

$$u_{\boldsymbol{\alpha}_C^{*(1)}}(\theta) = u_{\alpha^{(1)}}\psi_{\alpha_{\Theta}^{(1)}}(\theta) + u_{\alpha^{(5)}}\psi_{\alpha_{\Theta}^{(5)}}(\theta) + u_{\alpha^{(6)}}\psi_{\alpha_{\Theta}^{(6)}}(\theta) + u_{\alpha^{(9)}}\psi_{\alpha_{\Theta}^{(9)}}(\theta), \quad (4.16)$$

$$u_{\boldsymbol{\alpha}_C^{*(2)}}(\theta) = u_{\alpha^{(2)}}\psi_{\alpha_{\Theta}^{(2)}}(\theta) + u_{\alpha^{(7)}}\psi_{\alpha_{\Theta}^{(7)}}(\theta), \quad (4.17)$$

$$u_{\boldsymbol{\alpha}_C^{*(3)}}(\theta) = u_{\alpha^{(3)}}\psi_{\alpha_{\Theta}^{(3)}}(\theta) + u_{\alpha^{(8)}}\psi_{\alpha_{\Theta}^{(8)}}(\theta). \quad (4.18)$$

With the final coefficients, the Sobol' indices are quantified for each combination of statistical moments  $\theta$  (Equation 4.9). Finally, the combination of statistical moments that leads to the imprecise Sobol' indices bounds should be retrieved. This can be performed through brute-force sampling of the statistical moments, or through a global optimization algorithm, such as NSGA-II. The first-order Sobol' index bounds are:

$$S_{\min,1}^{\text{PCE}} = S_{\min,2}^{\text{PCE}} = 0.0, \quad S_{\max,1}^{\text{PCE}} = S_{\max,2}^{\text{PCE}} = 0.8, \quad (4.19)$$

and the total-order Sobol' index bounds correspond to:

$$S_{\min,1}^{T,\text{PCE}} = S_{\min,2}^{T,\text{PCE}} = 0.2, \quad S_{\max,1}^{T,\text{PCE}} = S_{\max,2}^{T,\text{PCE}} = 1.0. \quad (4.20)$$

**Integration in RDO** As parametric p-boxes are considered at the model input, an isoprobabilistic transformation is required on each p-box. Hence, each p-box results in two uniform distributions (one for the interval on the mean, the other for the interval on the standard deviation), and one standard Gaussian distribution. Thus, for each p-box, three precise distributions are added to the stochastic dimension. For the construction of the augmented

PCE, the user-defined constants are the same as for the sparse PCE algorithm, namely the polynomial order and the number of training samples. Hence, the same procedure is applied as discussed in Section 3.2. Finally, with the statistical moments on the quantity of interest characterized by ranges, instead of deterministic values, the objectives of the RDO approach need to be modified. The aim is to minimize the sensitivity towards aleatory uncertainty. However, due to the epistemic uncertainty on the stochastic parameters, the true mean and true standard deviation from aleatory uncertainty are unknown. Therefore, the upper-bound mean and standard deviation are selected as design objectives, as it combines the epistemic and aleatory uncertainty in the decision-making process [268].

## 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

In this application, RDO is performed on a PV-battery-heat pump system with thermal storage. The epistemic and aleatory uncertainty is considered separately and the parameters affected by both types of uncertainty are characterized by parametric p-boxes. Due to the use of imprecise probabilities, the mean and standard deviation on the quantity of interest are characterized by a range, instead of a fixed value. Therefore, the RDO is performed on the upper-bound values of the statistical moments for the quantity of interest. As mentioned in the chapter overview, this application is published in Energy [255]. After a brief introduction on PV-battery-heat pump systems with thermal storage (Subsection 4.3.1), the system model (Subsection 4.3.2), design variables and quantity of interest (Subsection 4.3.3) and uncertainty characterization (Subsection 4.3.4) are described. The optimized designs and the global sensitivity analysis are presented in Subsection 4.3.5. The conclusion and main messages of the work are summarized in Subsection 4.3.6.

### 4.3.1 Introduction

The market expansion of electrical energy storage [269] and thermal energy storage [270] enables greater power system flexibility. To illustrate, Murugan et al. [271] reviewed polygeneration systems that combine power and heating with other energy streams, such as cooling and synthetic fuels. Kyriakarakos et al. [272] evaluated a PV capacity in a polygeneration

microgrid with a battery stack, a hydrogen-based energy system and a reverse osmosis desalination unit. Rad et al. [273] integrated a PV capacity in a hybrid power system with wind turbines, a biogas generator and a fuel cell. In the residential sector, battery technology provides a flexible, adequate technology for short-term (i.e., days) electrical energy storage to improve PV self-consumption [274]. In addition to battery energy storage, including heat pumps and thermal storage, to cover the heat demand further, improves the PV self-consumption and entails the coupling of the electricity sector and heating sector [275,276], which is anticipated to further decarbonize the heating sector [269]. Several studies evaluated the feasibility of a PV-battery-heat pump system in the residential sector [277]. Klinger evaluated the PV and battery market potential in combination with heat pumps and electric vehicles and concluded that households benefit from such a self-consuming system [278]. Angenendt et al. [279] evaluated the profitability of the participation of a residential PV-battery-heat pump system in the frequency control reserve market. Despite the increased battery ageing, this approach was proven profitable following the ability to convert negative control reserve power into residential heating. Angenendt et al. [280] evaluated the sizing of a PV array, battery stack, heat pump and thermal storage for a household under different operating strategies and concluded that small-scale energy storage is economically beneficial.

By characterizing the effects of limited information (epistemic uncertainty) and natural variability (aleatory uncertainty) through p-boxes, the work addresses the limitation of using deterministic parameters or scenarios, which can lead to biased designs and unwarranted actions to reduce performance variation. Thereafter, the effect of aleatory uncertainty on the LCOX was minimized through a surrogate-assisted RDO. Hence, this approach determines the designs least sensitive towards the irreducible aleatory uncertainty. Finally, a global sensitivity analysis was performed on the optimized designs, to quantify the significance of epistemic uncertainty on the predicted performance and robustness of the optimized designs, followed by guidelines to reduce the epistemic uncertainty effectively.

### 4.3.2 System modelling

In this application, a grid-connected dwelling was adopted for Brussels, Belgium. This household is supported by a PV array, which is primarily used to cover the electricity demand of the household (i.e., electric appliances, lighting and cooking). To cover the

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

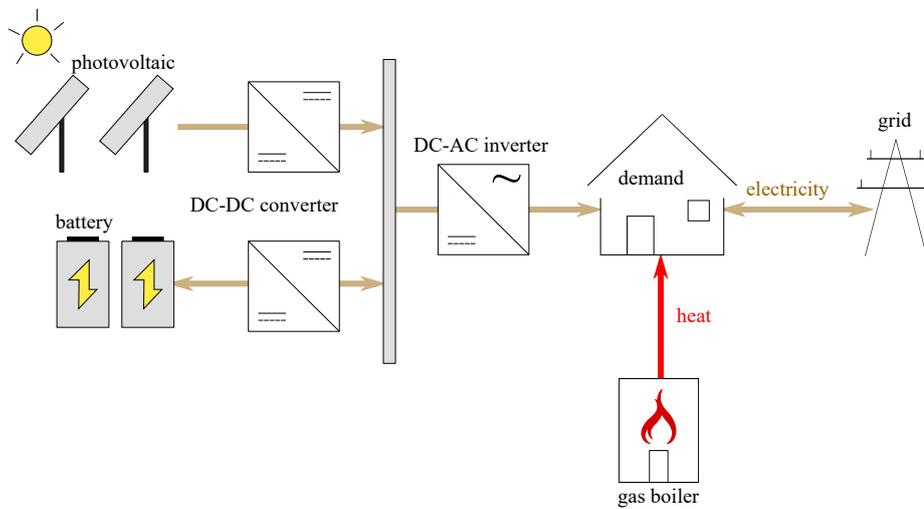
---

heat demand of the dwelling (i.e., Space Heating (SH) and Domestic Hot Water (DHW)), two different technologies are considered: a gas boiler, sized to cover the entire heat demand and connected to the natural gas grid (Figure 4.2), and an Air-Source Heat Pump (ASHP), connected to a thermal storage tank with an electric heater (Figure 4.3). The communication between the components and the monitoring of the energy system is performed by a programmable logic controller [281].

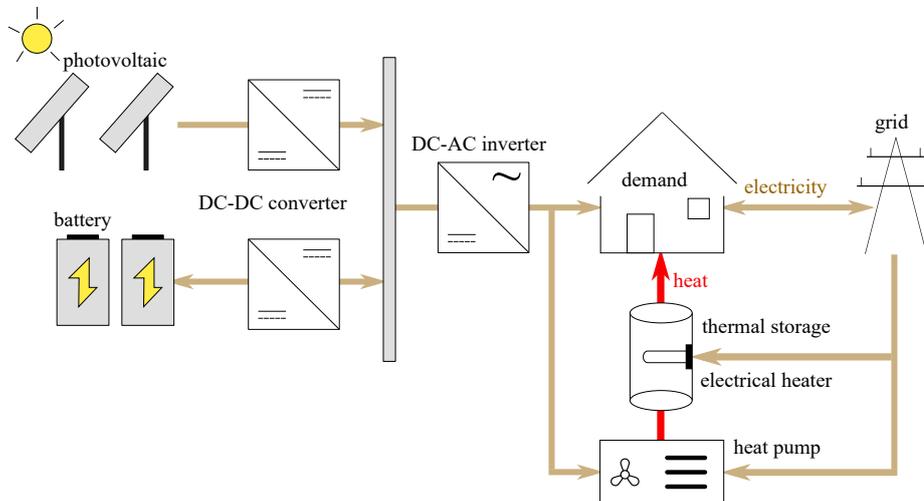
In both systems, PV electricity is used to cover the electricity demand. When excess PV electricity is available, the electricity is stored in the battery stack when considering the gas boiler, while the PV excess electricity is used to power the ASHP before storing the electricity in the battery stack in the PV-battery-ASHP configuration. Inversely, when the electricity demand exceeds the PV electricity, the battery stack is called upon first, followed by the grid whenever necessary.

To cover the DHW and SH demand, heat is supplied either by the gas boiler at an efficiency of 90 % [282] or extracted from the thermal storage tank. In the PV-battery-ASHP configuration, when the stored thermal energy is insufficient, the ASHP is used to comply with the remaining heat demand. If the PV excess electricity is insufficient to run the ASHP, two power management strategies are considered to cover the remaining required electricity for the ASHP. In the first strategy, electricity from the grid is adopted (i.e., PV-battery-ASHP without battery support). Hence, the stored excess of PV electricity in the battery is only used to cover the electricity demand, which is considered high-quality energy. In the second strategy, the electricity stored in the battery is used to run the ASHP, before covering the electricity demand (i.e., PV-battery-ASHP with battery support). This strategy enables to run the ASHP with renewable electricity when the solar irradiance is insufficient (e.g., at night) and therefore reduces the grid electricity consumption. In both power management strategies, a grid-powered, backup electric heater is considered with a fixed 99% efficiency [283]. The backup electric heater is deployed to support the ASHP and thermal storage capacity to cover demand peaks [284], which ensures thermal comfort in the household.

The component models for the PV array (Subsection 3.3.2.1), battery stack (Subsection 3.3.2.2) and power converters and inverter (Subsection 3.3.2.6) are described in the PV-battery-hydrogen application. In the remainder of this section, the component models for the ASHP and thermal storage tank are presented.



**Figure 4.2:** The reference system consists of a photovoltaic array and battery stack, which operate in support of the electricity grid to cover the electricity demand. A natural gas boiler covers the heat demand.



**Figure 4.3:** The photovoltaic-battery-heat pump system consists of a photovoltaic array that primarily complies with electricity demand. Excess photovoltaic electricity is used to power the heat pump, which is used to heat the water in the thermal storage tank to comply with the domestic hot water and space heating demand. The grid is used as a backup power generator to comply with electricity demand, to run the heat pump and to supply the electric heater.

#### 4.3.2.1 Air Source Heat Pump

The heat pump is an air-to-water, on/off-controlled heat pump, which uses ambient air as a heat source and R410A as a working fluid. The heat pump is modelled following a refrigeration cycle, in which pressure losses, evaporator superheating and condensing subcooling are neglected [285]. A pinch temperature  $T_{\text{pinch}}$  of 5 K is selected [285], which results in the following condensing temperature  $T_c$  and evaporation temperature  $T_e$ :

$$T_c = T_w + T_{\text{pinch}}, \quad (4.21)$$

$$T_e = T_{\text{amb}} - T_{\text{pinch}}. \quad (4.22)$$

To reach the condensing temperature after evaporation, the air is compressed in a compressor unit. As the compressor inlet pressure depends on the ambient temperature, the compression ratio is variable during the ASHP system lifetime. To determine the effect of the variable compression ratio on the compression isentropic efficiency, the experimentally-validated model from Underwood et al. [285] was adopted:

$$\eta_{\text{comp,is}} = \frac{a \exp(-b(r_p - c))}{1 + \exp(-d(r_p - c))}, \quad (4.23)$$

where the coefficients  $a, \dots, d$  are fitted based on manufacturer data [285]. The compression ratio  $r_p$  is determined through the pressure at which the condensation and evaporation occurs. These pressures are characterized by the respective refrigeration properties, extracted from the *CoolProp* Python library [286]. The compressor specific work is then quantified as follows:

$$\Delta h_{\text{comp}} = \frac{h(p_c, s_{\text{comp,in}}) - h_{\text{comp,in}}}{\eta_{\text{comp,is}}}. \quad (4.24)$$

Following the characterization of the enthalpy after condensation  $h_{c,\text{out}}$ , the heat generation during condensation  $Q_c$  and Coefficient Of Performance (COP) are determined:

$$Q_c = \dot{m}(h_{\text{comp,out}} - h_{c,\text{out}}), \quad (4.25)$$

$$\text{COP} = \frac{Q_c}{\dot{m}\Delta h_{\text{comp}}}. \quad (4.26)$$

At 10.4 °C ambient temperature (i.e., the average annual ambient temperature adopted for Brussels, see Table 4.2), the COP is 2.8, which approximates the experimentally-validated value presented by Angenendt et al. [280]. Additional details on the physical behavior of the ASHP is presented in Section B.5.

### 4.3.2.2 Thermal storage tank

The stratified thermal storage tank is a DHW and SH combi system with a heat exchanger that separates the drinking water from the heating water. Considering the DHW temperature of 40 °C and the heat pump limits, the lower and upper temperature are set at 45 °C and 55 °C, respectively [280]. The SH demand depends on the ambient temperature, room temperature and building losses. In this work, the SH (and DHW) demand is characterized by an energy demand profile, which represents the heat required to reach the required space temperature under specific conditions. Hence, the demand is characterized by the energy demand per hour, which is extracted from the thermal storage tank. The heat loss of the storage tank is determined through the room temperature  $T_{\text{room}}$  and the natural convection heat exchange coefficient (i.e., heat loss coefficient)  $U_{\text{loss}}$  [280]:

$$Q_{\text{loss}} = U_{\text{loss}} A_{\text{tank}} (T_{\text{w,tank}} - T_{\text{room}}), \quad (4.27)$$

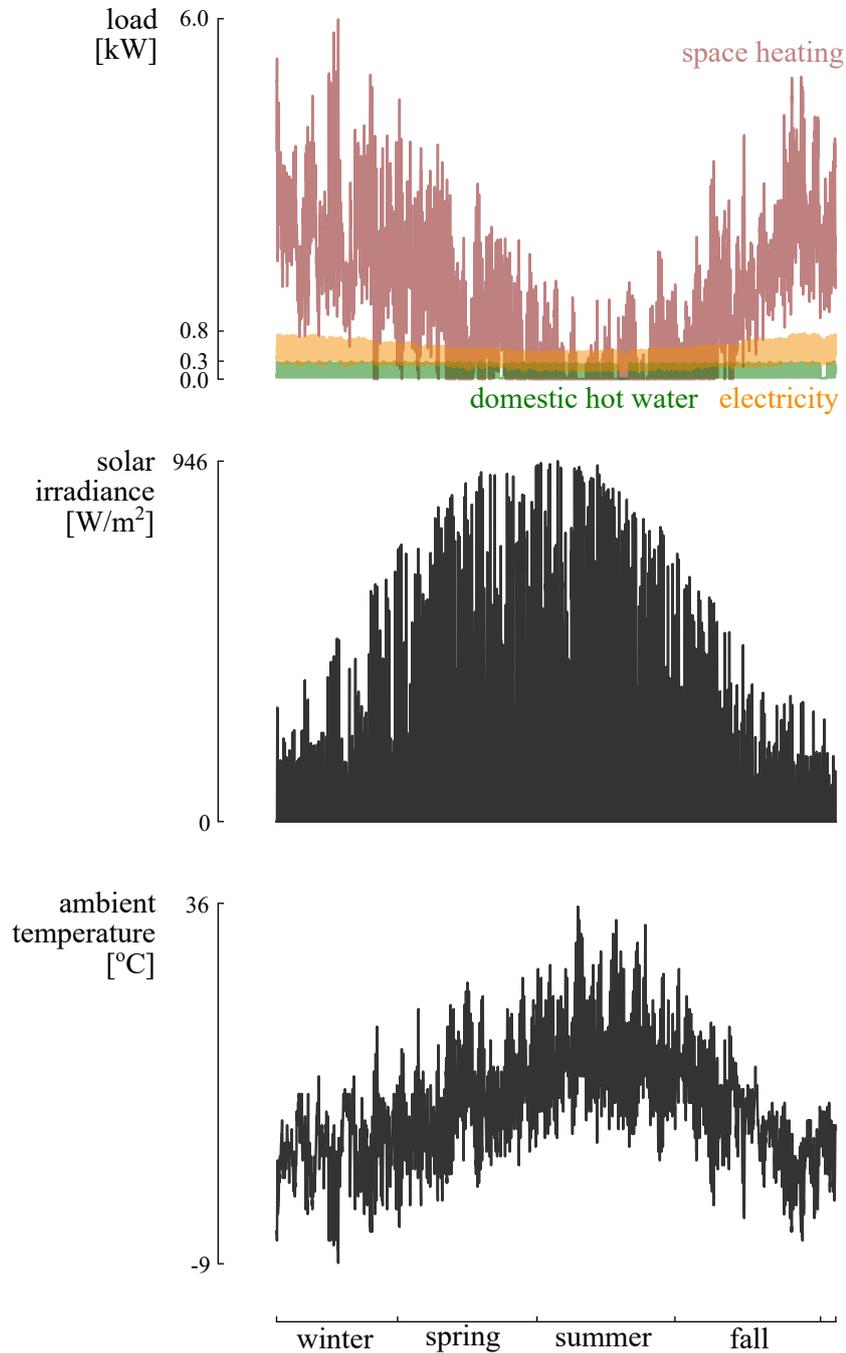
where the room temperature  $T_{\text{room}}$  is assumed constant at 20 °C and the heat loss coefficient  $U_{\text{loss}}$  is assumed at 1 W/(m<sup>2</sup>K) [280].

### 4.3.2.3 Climate and demand data

The energy consumption of the grid-connected household depends on the solar irradiance, ambient temperature, electricity demand and heating demand of the household (Figure 4.4). Therefore, these climate conditions and demand profiles are adjusted to the specific location (Brussels) and demand (dwelling). The method presented in Subsection 3.3.2.7 is adopted to generate the climate and demand data. The electricity demand profile combines the different loads of the dwelling (i.e., electric appliances, lighting and cooking). Therefore, in this work, no details are available on the contribution of each electric load to the hourly overall electricity demand. The SH period is specified based on the monthly average ambient temperature and aims to ensure thermal comfort [287]: when the monthly average

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

temperature is less than 19°C, the heating system is enabled for that month. Consequently for Brussels, the SH is considered for the entire year.



**Figure 4.4:** The hourly solar irradiance, ambient temperature and energy demand (electricity, domestic hot water and space heating) profiles for Brussels show that the energy demands drop during spring and summer and increase during fall and winter.

### 4.3.3 Design variables and quantity of interest

The system is designed by sizing the components of the system, i.e., the capacity of the PV array, battery stack, ASHP and thermal storage tank. By considering these capacities as independent design variables, the optimization algorithm can exclude any technology from the design (e.g., no thermal storage tank). The upper bounds for the design variables are 15 kW<sub>p</sub> for the PV array, 15 kWh for the battery stack, 15 kW<sub>th</sub> for the ASHP and 1000 L for the thermal storage tank. The quantity of interest is the Levelized Cost Of exergy (LCOX). In this quantity of interest, exergy is selected to account for the true quality of both energy streams in one metric [288]. Hence, the LCOX reflects the system cost per unit of exergy covered [130]:

$$\text{LCOX} = \frac{\text{CAPEX}_a + \text{OPEX}_a + C_{\text{repl},a} + C_{\text{grid},a} + C_{\text{gas},a}}{X_{\text{demand}}}. \quad (4.28)$$

To determine the annual system cost, the annualized investment cost CAPEX<sub>a</sub>, operational cost OPEX<sub>a</sub>, replacement cost C<sub>repl,a</sub>, grid electricity cost C<sub>grid,a</sub> (and gas cost C<sub>gas,a</sub> when a gas boiler is considered) are added together. The cost parameters are defined analogously to the cost parameters in Subsection 3.3.3. When the system is unable to cover part of the load (SH, DHW and electricity), electricity (and gas when considering the gas boiler) is extracted from the grid. As a result, additional costs are included which relate to the electricity price (and gas price). In this work, a fixed annual electricity tariff is adopted, which corresponds to the contract of nearly 75% of the households in Belgium [188].

The annual exergy output  $X_{\text{demand}}$  is characterized by the sum of the annual exergy required for electricity (elec), DHW and SH [289]:

$$\sum_{i=1}^{8760} \sum_{\text{type}} X^{\text{type}}, \quad \text{type} \in \{\text{elec}, \text{DHW}, \text{SH}\}. \quad (4.29)$$

While the electricity demand is covered by the highest energy quality, i.e.,  $X^{\text{elec}} = E^{\text{elec}}$ , the exergy to cover the DHW demand and SH demand are scaled by their corresponding Carnot

coefficient  $F$  [289]:

$$X^{\text{type}} = F^{\text{type}} E^{\text{type}} \quad \text{type} \in \{\text{DHW, SH}\}, \quad (4.30)$$

$$F = 1 - \frac{T_{\text{ref}}}{T_{\text{req}}}, \quad (4.31)$$

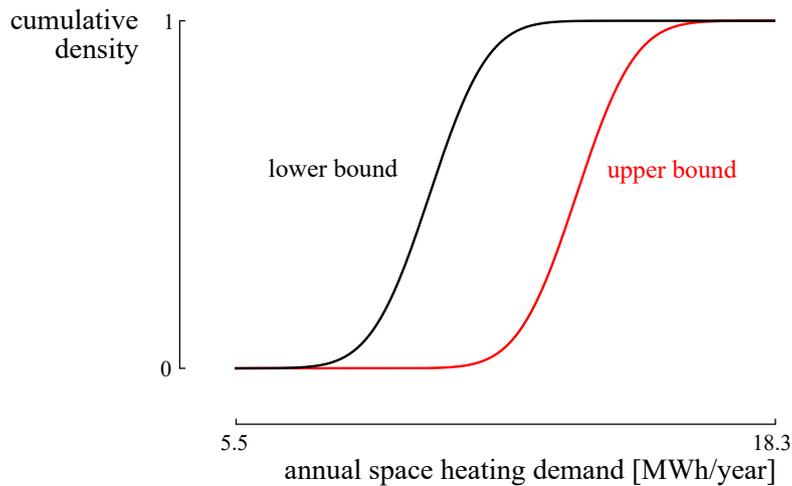
where the reference temperature  $T_{\text{ref}}$  corresponds to the ambient temperature for each specific hour of the year and  $T_{\text{req}}$  corresponds to the temperature required for DHW and SH, set at 40 °C and 35 °C, respectively [280].

Adopting the LCOX as the quantity of interest, as opposed to the Levelized Cost Of Energy (LCOEn), results in two main differences: For the same designs, the LCOX mean is higher than the LCOEn mean, due to correcting the heat demand in the denominator of the LCOX with the Carnot coefficient; The contribution of the heat demand uncertainty to the LCOX standard deviation is lower. However, as the energy demand is fixed for each design, considering the LCOX as the quantity of interest results in the same set of optimized designs as when the LCOEn is considered as a quantity of interest.

#### 4.3.4 Uncertainty characterization

The uncertainty on a model parameter can be characterized as epistemic uncertainty or aleatory uncertainty (see Subsection 1.1.3 for a brief explanation on the difference between both types of uncertainty). Both uncertainty types can affect a single model parameter simultaneously. To characterize the effect of both uncertainties on a single stochastic parameter, a parametric p-box is considered (Figure 4.5).

In this work, the grid electricity price, natural gas price, electricity demand and heat demand are subjected to both epistemic and aleatory uncertainty and thus characterized by a p-box (Table 4.1). For the electricity price, six prediction scenarios are combined to define the aleatory uncertainty, i.e., a base case scenario, a decentral scenario and a large-scale renewable energy system scenario, each determined in the coal-before-gas merit order and gas-before-coal merit order [89]. Similarly, the aleatory uncertainty on the gas cost integrates over low, middle and high gas cost scenarios [249]. The epistemic uncertainty relates to the different energy tariffs available for the household and thus to the different prices for the electricity extracted from the electricity grid (and gas from the



**Figure 4.5:** The parametric probability box of the annual space heating demand combines the effect of epistemic uncertainty (i.e., occupant behavior) and aleatory uncertainty (i.e., future evolution) on the parameter. The true-but-unknown CDF, which lies between the lower bound and upper bound of the p-box and is characterized by a mean and standard deviation, represents the aleatory uncertainty of the space heating demand. The range on the p-box represents the epistemic uncertainty on the parameter and affects the mean and standard deviation.

gas grid) [290]. Hence, the true-but-unknown mean and standard deviation represent the aleatory uncertainty on the future evolution of these prices, while not fixing the energy supplier for the household induces epistemic uncertainty, i.e., the range on the mean value. A similar philosophy is applied to the electricity demand and heat demand. The uncertainty following the different prediction scenarios for the future evolution of the energy demand represents the aleatory uncertainty [291, 292]. As a typical energy demand is assumed for average household consumption, the epistemic uncertainty is induced by the unknown occupant behavior [73]. Thus, the system performance is subjected to the uncertainty on how the energy demand will evolve (e.g., policies on increased insulation, improving efficiency of electric appliances) and how the occupants behave in their consumption (e.g., heavy consumers or light consumers).

Next to the p-boxes, several parameters are characterized only by aleatory uncertainty (Table 4.2). The solar irradiance and ambient temperature are characterized by a Typical Meteorological Year, which is assumed valid over the lifetime of the system. Hence, the epistemic uncertainty is neglected. The aleatory uncertainty corresponds to the interannual variability, which is based on percentiles from historical data [293]. Besides, the inflation rate is characterized by aleatory uncertainty following its natural variation [207].

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

Finally, the costs related to each technology are characterized by epistemic uncertainty (Table 4.3). These parameters are fixed at the beginning of the project and therefore only subject to the lack of knowledge on the market during system design, before making the actual investment.

**Table 4.1:** Parameters affected by aleatory and epistemic uncertainty and thus characterized by a parametric probability box. The true-but-unknown mean and standard deviation characterize the aleatory uncertainty. The range on the mean is induced by epistemic uncertainty.

parameter	mean	standard deviation	unit	Ref.
grid electricity price	[252,370]	68.9	€/MWh	[89, 290]
gas price	[50,86]	22.6	€/MWh	[249, 290]
electricity demand	[2.53,3.43]	0.22	MWh/year	[73, 292]
heat demand	[11.1,15.1]	1.1	MWh/year	[73, 291]

**Table 4.2:** Parameters affected by aleatory uncertainty.

parameter	mean	standard deviation	unit	Ref.
total solar irradiance	1174	92	kWh/m <sup>2</sup> /year	[205]
average ambient temperature	10.3	0.4	°C	[205]
inflation rate	1	3	%	[207]

#### 4.3.5 Results and discussion

The results and discussion section presents and discusses the trade-off between minimizing the mean and standard deviation of the aleatory uncertainty on the LCOX. Then, the p-boxes for a set of optimized designs are analysed, to see how these results are affected by epistemic uncertainty. Finally, the Sobol' indices reveal the dominating stochastic parameters to the epistemic uncertainty on the LCOX. The results and discussion are subject to the following limitations of the approach: the stochastic model parameters are considered independent, i.e., the probability of each parameter is unaffected by the probability of the other parameters; the NSGA-II does not ensure mathematical optimality, i.e., the designs are optimized, but might not represent the optimal solution; due to limited data, the aleatory uncertainty of the stochastic parameters is represented by a Gaussian distribution. Nevertheless, the robust design optimization approach provides the design trends to reduce the LCOX standard deviation induced by uncertainty which is irreducible by data collection and it illustrates the advantages of a robust design as opposed to a design

**Table 4.3:** Parameters affected by epistemic uncertainty.

parameter	min	max	unit	Ref.
CAPEX <sub>PV</sub>	350	600	€/kW <sub>p</sub>	[232]
OPEX <sub>PV</sub>	16	19	€/kW <sub>p</sub> /year	[142]
CAPEX <sub>bat</sub>	102	354	€/kWh	[174]
OPEX <sub>bat</sub>	15	28	€/kWh/year	[130, 174]
replacement cost <sub>bat</sub>	61	141	€/kWh	[130, 174]
lifetime <sub>bat</sub>	500	2000	cycles	[174]
self discharge <sub>bat</sub>	0.1	0.3	%/day	[130]
CAPEX <sub>hp</sub>	600	1100	€/kW <sub>th</sub>	[294, 295]
OPEX <sub>hp</sub>	3.7	15	€/kW <sub>th</sub> /year	[294, 295]
CAPEX <sub>thermal storage</sub>	2.8	5.4	€/l	[294, 296]
CAPEX <sub>gas boiler</sub>	110	200	€/kW <sub>th</sub>	[294, 297]
OPEX <sub>gas boiler</sub>	2.5	5	% of CAPEX <sub>gas boiler</sub>	[294, 297]
CAPEX <sub>DCDC</sub>	40	160	€/kW	[202, 246]
OPEX <sub>DCDC</sub>	1	5	% of CAPEX <sub>DCDC</sub>	[203, 204]
CAPEX <sub>DCAC</sub>	50	200	€/kW	[246]
OPEX <sub>DCAC</sub>	1	5	% of CAPEX <sub>DCAC</sub>	[203, 204]
discount rate	4	8	%	[163, 208, 209]

with an optimized LCOX mean. In addition, distinguishing between epistemic and aleatory uncertainty in the global sensitivity analysis identifies the main drivers of the epistemic uncertainty on the LCOX and thus highlights the parameters on which the knowledge should be improved, e.g., considering a specific energy demand profile as opposed to a generic one for an average household.

#### 4.3.5.1 Robust design optimization on the aleatory uncertainty

In this section, the robust design optimization results are described for the previously introduced configurations: a PV-battery configuration with a gas boiler that covers the entire heat demand; a PV-battery-ASHP configuration, where the power to run the ASHP comes from the excess PV electricity and grid electricity; a PV-battery ASHP configuration where the excess of PV electricity and stored electricity in the battery stack are used to run the

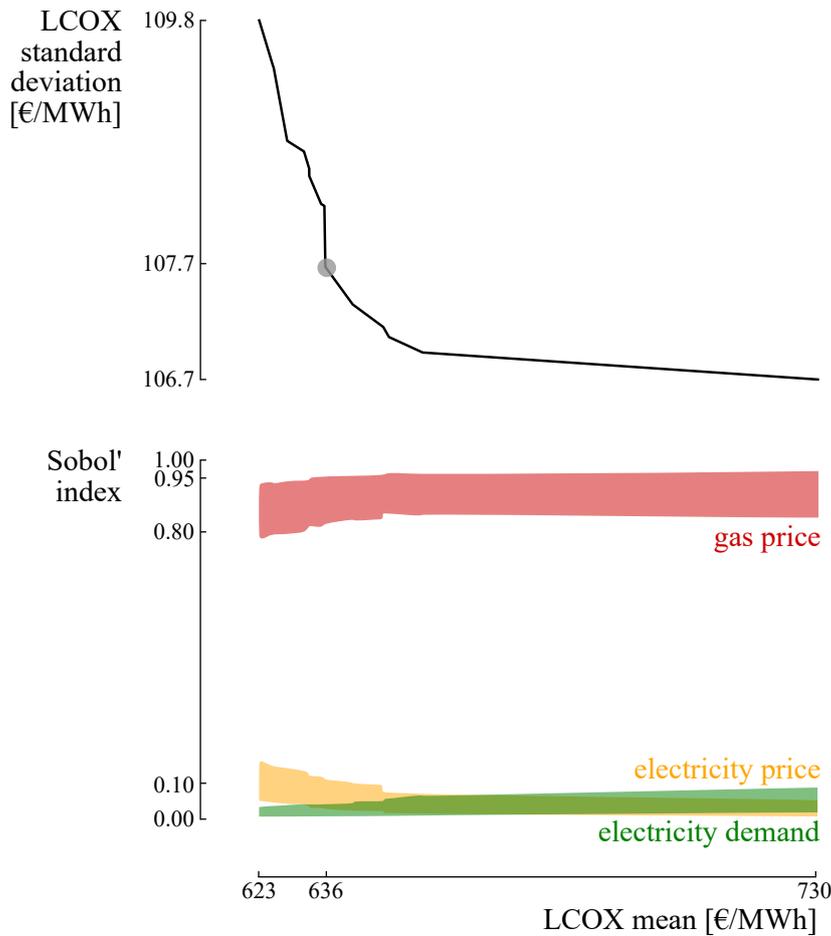
### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

---

ASHP, before the grid electricity is called upon. The optimization algorithm is characterized by a population size of 40 design samples, a crossover and mutation probability of 0.9 and 0.1, respectively.

**PV-battery-gas boiler** For the RDO of a PV-battery-gas boiler system, a trade-off exists between minimizing the LCOX mean and minimizing the LCOX standard deviation (Figure 4.6, top). Hence, the optimizer is unable to generate a single PV-battery-gas boiler design that simultaneously achieves the minimized LCOX mean and LCOX standard deviation. The Pareto front illustrates that the optimized mean design (i.e., the design that achieves the lowest upper-bound LCOX mean) achieves an LCOX mean of 623 €/MWh and an LCOX standard deviation of 109.8 €/MWh. Instead, the robust design (i.e., the design that achieves the lowest upper-bound LCOX standard deviation) achieves an LCOX mean of 730 €/MWh and an LCOX standard deviation of 106.7 €/MWh. These extreme designs on the Pareto front consist of a 1.8 kW<sub>p</sub> PV array and an 8.0 kW<sub>p</sub> PV array - 7.9 kWh battery stack, respectively. Despite the trade-off, the difference in LCOX standard deviation between the optimized mean design and the robust design is limited (3.1 €/MWh or 2.8%). Hence, the robust alternative to the optimized mean design is only slightly less sensitive to its random environment. However, an intermediate design on the Pareto front, achieving an LCOX mean of 636 €/MWh and an LCOX standard deviation of 107.7 €/MWh (grey dot in Figure 4.6), reduces the LCOX standard deviation by 2% when compared to the optimized mean design, at the expense of a 2% increase in LCOX mean. Therefore, this intermediate design, consisting of a 2.2 kW<sub>p</sub> PV array and a 3.2 kWh battery stack, provides a viable robust alternative to the optimized mean design.

The reason why the robust design is only slightly more robust than the optimized mean design comes from the complete dependency on the natural gas to comply with the heat demand (Figure 4.6, bottom). Indeed, despite the effect of epistemic uncertainty on the imprecise Sobol' indices (i.e., the epistemic uncertainty is the driver of the range of the imprecise Sobol' index), it can be concluded that the aleatory uncertainty of the natural gas price (i.e., the aleatory uncertainty is the driver of the mean of the imprecise Sobol' index) dominates the characterization of the true-but-unknown LCOX standard deviation. To illustrate, the true-but-unknown Sobol' index for the optimized mean design situates between 0.8 and 0.95, while the true-but-unknown Sobol' indices for the electricity price



**Figure 4.6:** The Pareto front between minimizing the LCOX mean and minimizing the LCOX standard deviation illustrates that the gain in robustness is limited when shifting between optimized mean and robust design (top figure). The Sobol' indices explain this limited gain in robustness by illustrating that the aleatory uncertainty on the gas price (bottom figure), which is irreducible by designing the PV array and battery capacity, is the main driver of the true-but-unknown LCOX standard deviation.

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

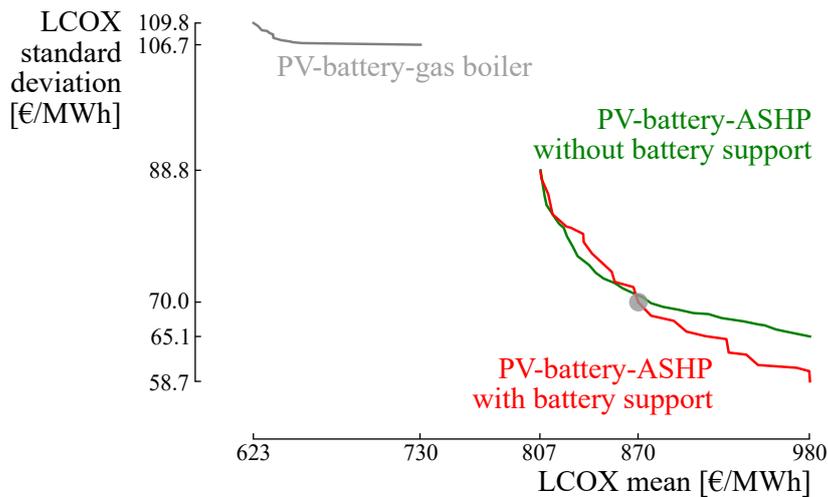
---

and electricity demand situate between 0.05 and 0.16 and between 0.00 and 0.04, respectively. Thus, considering a PV array with a battery stack is a viable robust alternative for the optimized mean design, which consists of solely a PV array. Nevertheless, due to the dominance of the gas price aleatory uncertainty on the LCOX standard deviation, the gain in robustness for the robust alternative is limited.

**PV-battery-ASHP** By considering an electrically-driven ASHP to cover the heat demand, the gas dependency is eliminated and the generated electricity by the PV array is employed to comply with both the electricity demand and heat demand.

When comparing the achieved Pareto fronts for the PV-battery-ASHP configurations with the Pareto front for the PV-battery-gas boiler configuration, the optimized LCOX standard deviation on the Pareto front is significantly lower for the PV-battery-ASHP optimized designs than for the PV-battery-gas boiler optimized designs (Figure 4.7). To illustrate, an intermediate optimized design (i.e., situated between the extreme designs on the Pareto front, indicated with a grey dot in Figure 4.7) for a PV-battery-ASHP system with a battery support power management strategy achieves an upper-bound LCOX standard deviation of 70.0 €/MWh, which reduces the standard deviation by 36 % as opposed to the PV-battery-gas boiler design with an upper-bound LCOX standard deviation of 109.8 €/MWh. However, the optimized LCOX mean values for the PV-battery-ASHP designs are higher than for the PV-battery-gas boiler designs (e.g., 39 % higher between the two previously compared designs). Similar conclusions on the fact that an ASHP with thermal storage has higher deterministic costs than a conventional gas boiler system were also presented by Renaldi et al. [298]. When comparing the optimized designs from both power management strategies in the PV-battery-ASHP configuration, the corresponding Pareto fronts illustrate that the LCOX standard deviation can be further reduced when the strategy with battery support for the ASHP is considered as opposed to when the strategy without battery support for the ASHP is considered. In conclusion, a PV-battery-ASHP configuration can provide a cost-competitive, robust alternative to a PV-battery-gas boiler configuration.

To clarify the Pareto fronts, the PV self-consumption rate (i.e., the portion of the generated PV electricity that is consumed by the electricity demand), annual grid electricity consumed and the component capacities for the optimized PV-battery-ASHP designs are presented in function of the LCOX mean in Figure 4.8. For both sets of optimized designs (i.e., the



**Figure 4.7:** The Pareto fronts for the PV-battery-ASHP configurations, which connect the LCOX mean and LCOX standard deviation for the optimized designs of each configuration, illustrate that the optimized PV-battery-ASHP designs achieve a higher LCOX mean than the optimized PV-battery-gas boiler designs. However, the optimized PV-battery-ASHP design achieves a significantly lower LCOX standard deviation.

PV-battery-ASHP with battery support and PV-battery-ASHP without battery support sets of optimized designs which resulted in the Pareto fronts in Figure 4.7), the PV array capacity evolves similarly from the optimized mean design to the robust design (i.e., the designs which achieve an upper-bound LCOX mean of 807 €/MWh and 980 €/MWh, respectively). For the optimized LCOX mean design in both power management strategies, no battery storage is considered. Hence, the difference in power management strategy does not affect the LCOX, as no battery support to run the ASHP is available. However, the trend towards the robust designs (i.e., with an LCOX mean of 980 €/MWh) illustrate that the PV-battery-ASHP configuration with battery support considers a higher capacity for the battery stack than the PV-battery-ASHP configuration without battery support. To illustrate, a design for the strategy without battery support, which achieves an LCOX mean of 870 €/MWh, consists of a 3.3 kWh battery stack, while the design for the strategy with battery support and with the same LCOX mean (870 €/MWh) consists of a 7.5 kWh battery stack (the characteristics of these designs are indicated by grey dots in Figure 4.8). The latter achieves a lower annual grid electricity consumption (5.3 MWh) than the former (5.5 MWh). Nevertheless, the annual grid electricity consumption remains significant, as the battery capacity is limited and sized for daily storage. Moreover, the evolution of the battery capacity closely relates to the PV self-consumption. As a larger share of PV excess electricity can be stored

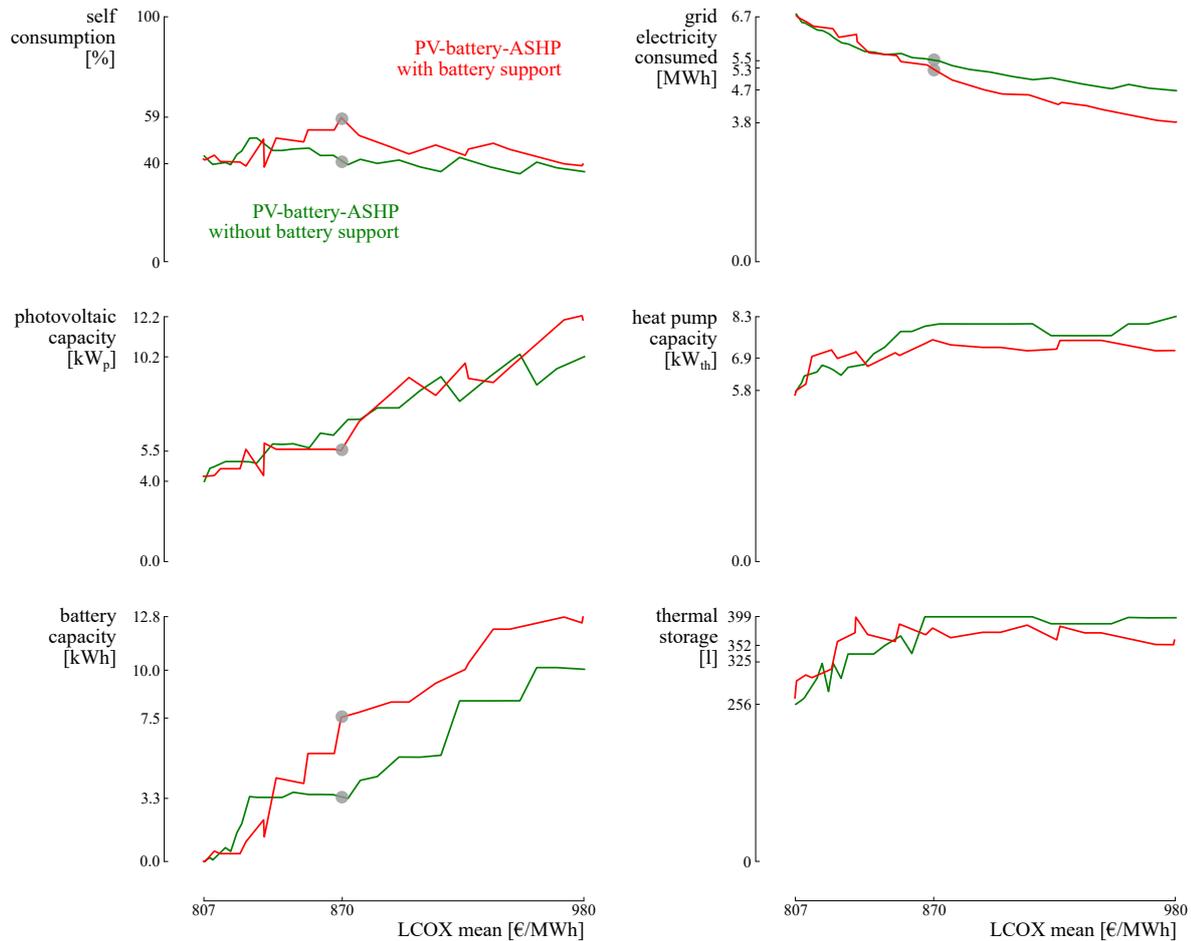
### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

in the battery stack for the designs considering battery support, this strategy achieves a higher PV self-consumption than the strategy without battery support. To illustrate for the previously compared designs, the PV self-consumption is 59 % and 40 %, respectively. In both strategies, the ASHP capacity of the optimized designs ranges between 5.8 kW<sub>th</sub> and 8.3 kW<sub>th</sub>, while the thermal storage ranges between 256 l and 399 l. The ASHP and thermal storage capacities are similar to the ones presented in the deterministic design optimization of an ASHP with thermal storage in the United Kingdom under different electricity tariffs (8.5 kW<sub>th</sub> and 300 l) [298]. The fact that thermal storage is considered in every optimized design indicates that thermal storage is beneficial for improving the LCOX mean and LCOX robustness over the system lifetime. Similarly, Renaldi et al. [298] indicated that combining thermal energy storage with ASHP results in a smaller cost as opposed to not considering thermal storage. Nevertheless, this relatively small thermal energy storage capacity is only suitable for daily thermal energy storage, as the significant heat loss limits the application of seasonal storage in its current form.

**Table 4.4:** The characteristics of the three representative designs for the photovoltaic-battery-ASHP management strategies: an optimized mean design (lowest upper-bound LCOX mean), an intermediate design and a robust design (lowest upper-bound LCOX standard deviation).

	$N_{PV}$ kW <sub>p</sub>	$N_{bat}$ kWh	$N_{hp}$ kW <sub>th</sub>	$N_{thermal\ storage}$ l	$\mu_{LCOX}$ €/MWh	$\sigma_{LCOX}$ €/MWh
<b>without battery support</b>						
optimized mean	4.0		5.8	256	[396,807]	[57.6,88.8]
intermediate	5.8	3.3	8.3	399	[438,872]	[47.1,70.2]
robust	10.2	10.0	8.3	397	[473,980]	[34.5,65.1]
<b>with battery support</b>						
optimized mean	4.2		5.6	266	[397,807]	[54.6,88.8]
intermediate	5.5	7.5	6.9	325	[435,870]	[41.0,70.0]
robust	12.2	12.8	7.2	352	[465,980]	[27.7,58.7]

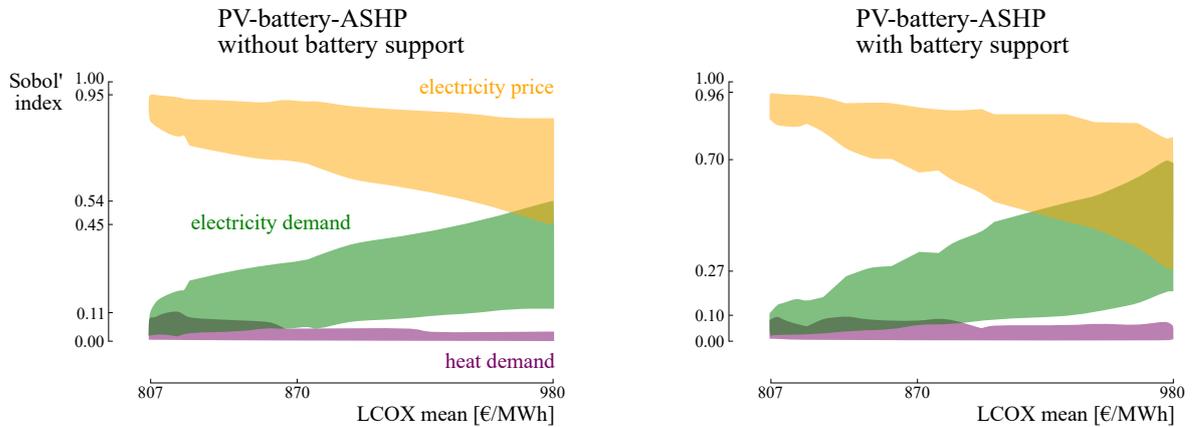
To understand the evolution of the LCOX standard deviation on the Pareto fronts for the two power management strategies, the imprecise Sobol' indices are quantified for the optimized designs (Figure 4.9). Clearly, the uncertainty on the future grid electricity price dominates the LCOX standard deviation. This conclusion can be made despite the overall epistemic uncertainty on the stochastic parameters, which induces the range on the imprecise Sobol' indices, as the imprecise Sobol' indices barely overlap. However, for the robust design of the strategy with battery support, the Sobol' index range for the electricity price and electricity demand clearly overlap. This indicates that for this specific design, it remains inconclusive



**Figure 4.8:** The PV self-consumption, annual grid electricity consumption and system capacities in function of the LCOX mean for the optimized designs for the two power management strategies of the PV-battery-ASHP configurations. In most optimized designs under the strategy with battery support, a significantly larger battery is considered, which results in a lower annual grid electricity consumption and a higher PV self-consumption than for the optimized designs under the strategy without battery support.

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

which parameter is the most dominant driver of the true-but-unknown LCOX standard deviation, due to the overall epistemic uncertainty on the stochastic parameters.



**Figure 4.9:** The evolution of the Sobol' indices for the stochastic parameters on the set of optimized designs indicate that the uncertainty on the future electricity price is the dominant uncertainty for the LCOX for most optimized designs. For the robust design, it remains inconclusive if the aleatory uncertainty on the electricity price or the electricity demand is the main driver, due to epistemic uncertainty.

The reducing trend in the imprecise Sobol' index of the electricity price implies a decreasing dependency on the grid electricity to comply with the energy demand. This reducing trend is larger under the battery support strategy, as the battery storage reduces the grid electricity consumption of the ASHP by covering part of the ASHP demand when the solar irradiance is insufficient (e.g., at night). The increasing trend on the importance of the aleatory uncertainty of the electricity demand corresponds to an increased photovoltaic array and battery capacity, for which its economic effectiveness depends on the match with the electricity demand.

Finally, the effect of the heat demand aleatory uncertainty on the LCOX is small. This is mainly due to the reduced weight of the heat demand by the Carnot coefficient in an exergy environment (Equation 4.31).

#### 4.3.5.2 Probability box characterization for optimized designs

In the previous section, the designs are configured to minimize the effect of aleatory uncertainty on the expected LCOX, while maintaining an optimized LCOX mean. To study the effect of the epistemic uncertainty on the LCOX over the system lifetime, the LCOX p-boxes

for a set of representative designs are analysed (i.e., optimized mean design, intermediate design and robust design, Table 4.4). The effect of the epistemic uncertainty on these three representing designs is comparable between both strategies (i.e., range on the statistical moments in Table 4.4 and width of the p-box in Figure 4.10). When comparing the three representative designs, the robust design is more affected by epistemic uncertainty. Nevertheless, given the overall epistemic uncertainty, the p-boxes of the three representing designs for both strategies overlap. Therefore, at this step, it remains inconclusive which design achieves the most favourable true-but-unknown CDF.

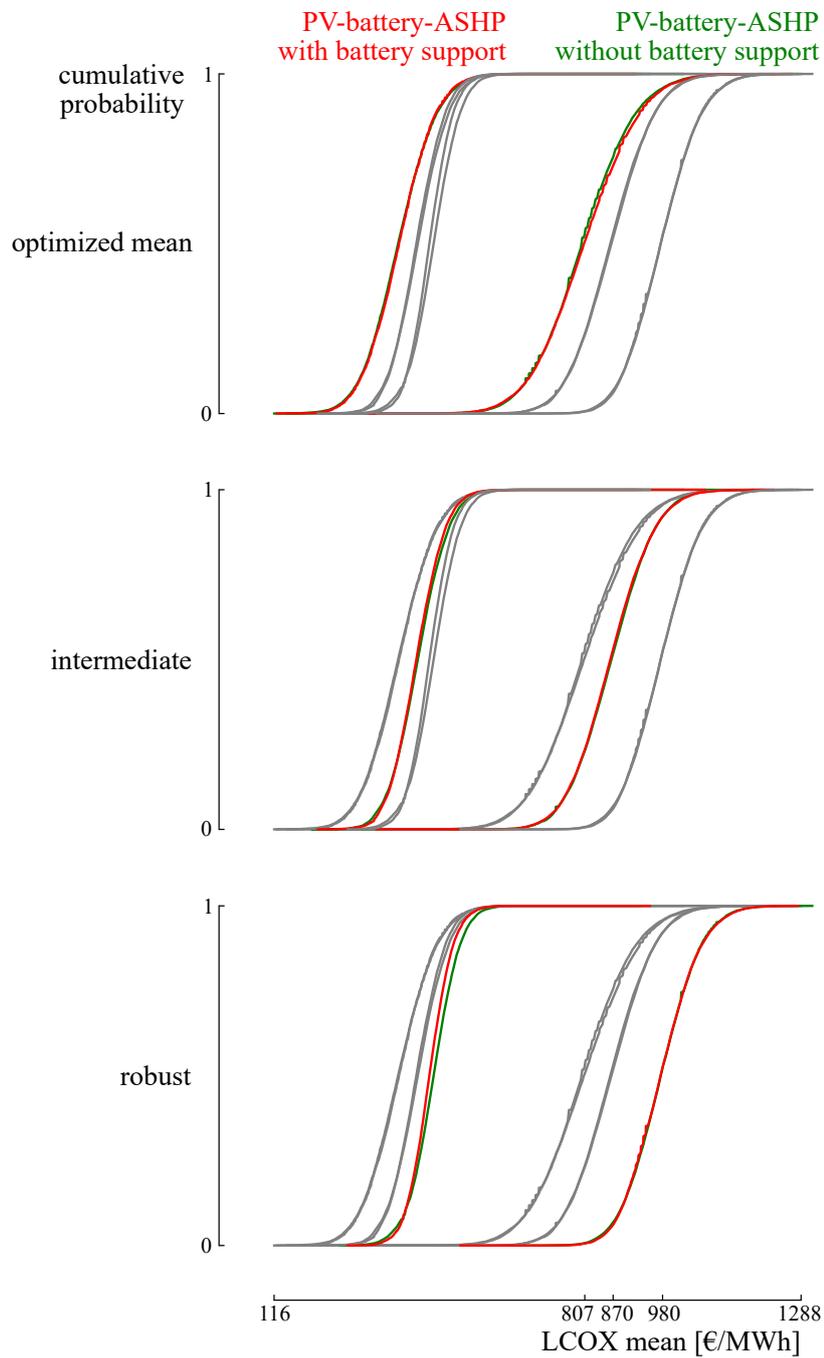
### 4.3.5.3 Global sensitivity analysis on the epistemic uncertainty

To determine the true-but-unknown CDF of the optimized designs, the epistemic uncertainty must be reduced. To characterize what contributes to the epistemic uncertainty on the LCOX of these designs, a global sensitivity analysis on the epistemic uncertainty of the p-boxes was performed. As the epistemic uncertainty affects both the mean and standard deviation (Table 4.4), two sets of Sobol' indices are generated for each p-box.

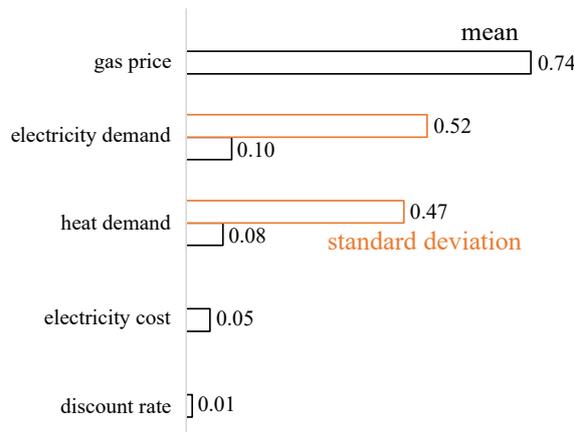
For the optimized LCOX mean design of the PV-battery-gas boiler configuration, the epistemic uncertainty on the LCOX mean (and thus the width of the p-box) is dominated by the epistemic uncertainty on the gas price (Figure 4.11). Hence, to determine the expected LCOX performance of this design, the main action is to determine the gas supplier and the corresponding contract. The uncertainty on the occupant behavior in electricity demand and heat demand contribute equally (i.e., a Sobol' index of 52 % and 47 %, respectively) to the epistemic uncertainty on the LCOX standard deviation. Therefore, adopting heat demand and electricity demand profiles for a specific household with information on the occupant behavior drastically improves the knowledge on the LCOX standard deviation and thus on the robust performance of the specific design.

For the optimized LCOX mean designs for both power management strategies of the PV-battery-ASHP configuration (Table 4.4), the epistemic uncertainty on the electricity price dominates the uncertainty on the mean LCOX (Figure 4.12, *opt. mean*). For the epistemic uncertainty on the standard deviation, the dominating contribution comes from the uncertainty on the occupant behavior regarding the heat demand. As both designs do not include a battery stack, there is no difference due to the power management strategy and

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage



**Figure 4.10:** The probability boxes for the optimized mean design, intermediate design and robust design for both PV-battery-heat pump configurations indicate that the difference in the effect of the epistemic uncertainty is negligible between configurations.



**Figure 4.11:** Top-5 total order Sobol' indices for the PV-battery-gas boiler design.

therefore the difference between the top-5 Sobol' indices is minimal.

For the intermediate designs, the epistemic uncertainties on the electricity price and electricity demand are the most significant contributors to the uncertainty on the LCOX mean (Figure 4.12, *int.*). Also, the epistemic uncertainty on the CAPEX for the heat pump gains significant importance. Hence, specifying the specific heat pump technology and its corresponding market condition becomes a significant action to reduce the epistemic uncertainty on the mean LCOX for these intermediate designs. For the uncertainty on the LCOX standard deviation, the epistemic uncertainty on the heat demand is again the main driver. Finally, for the robust designs, the epistemic uncertainty on the electricity demand dominates both the epistemic uncertainty on the LCOX mean and LCOX standard deviation (Figure 4.12, *robust*). Hence, gaining knowledge on the occupant behavior related to the electricity demand is the most effective measure to clarify both performance and robustness of these designs.

In conclusion, considering a specific electricity demand and heat demand profile for a specific household, and thus reducing the epistemic uncertainty of using generic profiles, are the main actions to determine the true-but-unknown LCOX standard deviation for these representative designs. Reducing the epistemic uncertainty related to the electricity demand is also the main action to clarify the true-but-unknown mean LCOX for the intermediate PV-battery-ASHP designs and the robust designs, while gaining information on the grid electricity contract and supplier is significant for both optimized LCOX mean

### 4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage

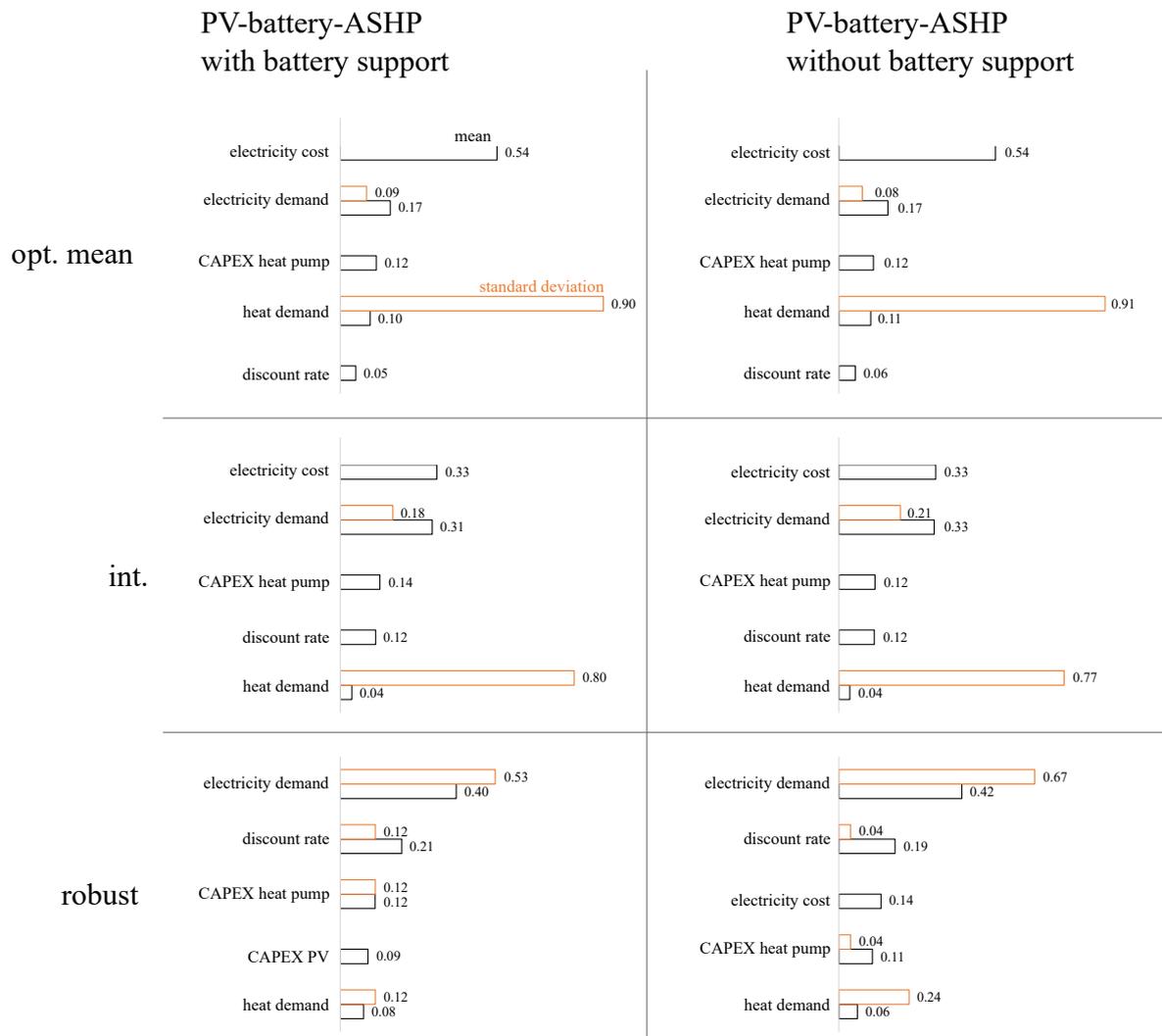


Figure 4.12: Top-5 total order Sobol' indices for the optimized PV-battery-ASHP designs.

designs. Finally, the epistemic uncertainty induced by considering generic ASHP technology is significant, while the effect of considering generic technology models for the other components is minor.

### 4.3.6 Conclusion

Sector coupling of heating and electricity improves the decarbonization of the heating sector. In this framework, Air-Source Heat Pumps (ASHP) provide an adequate approach to realise this coupling. For a photovoltaic-battery-heat pump configuration, a robust design optimization on the aleatory uncertainty of the Levelized Cost Of eXergy (LCOX) illustrates that the robustness towards LCOX aleatory uncertainty can be improved by 36 % when compared to a photovoltaic-battery-gas boiler configuration, at the expense of a 39 % increase in LCOX mean. The robustness is optimized when both excess PV electricity and stored electricity in the battery are considered to run the ASHP, as opposed to using the battery electricity only to cover the household electricity demand. Additionally, in these optimized photovoltaic-battery-heat pump designs, thermal energy storage is consequently considered by the optimizer for short-term energy storage (i.e., days), which indicates that thermal storage is beneficial for improving the LCOX mean and LCOX robustness over the system lifetime. In conclusion, PV-battery-ASHP configurations provide a cost-competitive, robust alternative to PV-battery-gas boiler configurations.

In between the optimized set of PV-battery-ASHP designs, it remains inconclusive which design results in optimized expected performance and which design is the most robust, due to the overall epistemic uncertainty present on the LCOX. To reduce the epistemic uncertainty effectively on the LCOX mean (and thus on the expected performance), gaining information on the grid electricity contract is the main action for the optimized LCOX mean designs. This action proves also significant for an intermediate design (i.e., a design which considers a trade-off between minimized LCOX mean and LCOX standard deviation), in addition to considering a specific electricity demand profile for a household. For the epistemic uncertainty on the LCOX standard deviation, the uncertainties on the occupant behavior regarding electricity demand and heat demand are the main drivers. Therefore, specific heat demand and electricity demand profiles enable characterizing the true-but-unknown robustness toward variability in the LCOX due to aleatory uncertainty. Finally, the epistemic uncertainty induced by considering generic technology models as opposed to a

### **4.3 Power-to-heat: photovoltaic-battery-heat pump with thermal storage**

---

specific type proves to be negligible in the epistemic uncertainty on the statistical moments of the LCOX, except for the ASHP.



# 5 Conclusions and perspectives

## Chapter overview

- Summary of the main results;
- Recommendations and guidance;
- Future perspectives.

## 5.1 Main conclusions

The main conclusion from this work is that **RDO provides an added value in the design of HRES**. By applying the RDO method to HRES, novel robust design alternatives are defined. These robust designs provide additional insights into the decision-making process. Generally, in HRES, the robust design alternatives are characterized by a higher penetration of renewable energy systems and the inclusion of energy storage. To illustrate, in a PV-battery-hydrogen power-to-power system, the economically preferred solution is a photovoltaic array (LCOE mean of 269 €/MWh), when the self-sufficiency ratio is irrelevant ( $\leq 30\%$ ). When a higher self-sufficiency ratio threshold is of interest, i.e., up to 59%, PV-battery designs and PV-battery-hydrogen designs provide the cost-competitive alternatives that are least sensitive to real-world uncertainty (LCOE standard deviation reduction up to 42%, when compared to the optimized PV array). In the context of heavy-duty mobility, a diesel-fueled bus fleet achieves an optimized LCOD mean. However, shifting towards a bus fleet that partly consists of hydrogen-fueled buses (54% of the fleet) reduces the LCOD standard deviation by 36%, the CI mean by 46% and CI standard deviation by 51%, at the expense

of an increase in LCOE mean by only 11 %. Finally, combining a PV-battery configuration with a gas boiler corresponds to the optimized LCOE mean. However, a PV-battery-heat pump with thermal storage system achieves an upper-bound LCOE standard deviation of 70.0 €/MWh, which reduces the standard deviation by 36 % as opposed to the PV-battery-gas boiler design with an upper-bound LCOE standard deviation of 109.8 €/MWh. In addition to these robust design alternatives, the global sensitivity analysis highlights the main drivers of the variance on the quantities of interest. Out of this analysis, specific guidelines can be deduced to further reduce this variance. To illustrate, in a directly-coupled PV-electrolyzer system, the main drivers of the LCOE uncertainty are the uncertainty on the CAPEX of the PV array, CAPEX of the electrolyzer and the discount rate. During system operation, the most significant contributing parameters to the variation are the operational expenditure and lifetime of the electrolyzer stack. Therefore, bulk manufacturing of the technology, promoting more demonstration projects and continuous high-quality maintenance are the most significant actions to improve the levelized cost of hydrogen robustness.

An additional conclusion is that **surrogate-assisted RDO is suitable to comply with HRES holistic design expectations in terms of computational efficiency**. When designing an HRES, a holistic approach is often suggested, considering technical, economic and environmental performance. To realize such a holistic optimization, the effect of each HRES component on each objective should be captured, resulting in a large stochastic dimension. Integrating sparse PCE approaches in surrogate-assisted RDO enables addressing this computational cost, resulting in computationally tractable optimization problems. Consequently, surrogate-assisted RDO can be applied to complex HRES in a holistic optimization context, without the need of scaling down the stochastic dimension through assumptions.

Finally, **surrogate-assisted RDO enables to consider more advanced uncertainty characterization techniques**. In current HRES stochastic evaluation, either interval uncertainty or probabilistic uncertainty is applied to characterize the parametric uncertainty. In these approaches, a single representation combines all types of uncertainty, without distinguishing between epistemic and aleatory uncertainty. Imprecise probability methods provide an elegant distinction between the epistemic and aleatory uncertainty on a parameter. Sparse augmented PCE propagates probability boxes in a computationally-efficient way, enabling the use of imprecise probabilities in an RDO context. In this approach, robust designs

can be characterized, with a robust performance towards the aleatory (irreducible) uncertainty, while the sensitivity analysis highlights the main actions to reduce the epistemic uncertainty.

## 5.2 Recommendation and guidance

Surrogate-assisted RDO provides alternative designs for complex HRES, in a holistic framework, at a reasonable computational cost. Some recommendations and guidance are provided when applying this method to other (HRES) models.

**Consider uncertainty in HRES design optimization** A large number of numerical models exist for HRES, with different levels of fidelity. These models are not computationally expensive, when compared to models in e.g., Computational Fluid Dynamics, that may take several days for a single evaluation. In the fields where such high-fidelity models are needed (e.g., aerospace engineering [68]), the importance of considering uncertainty on the parameters is well-established. Hence, the need for UQ in these fields encouraged the research on improving the computational efficiency in UQ and stochastic optimization. Moreover, multi-fidelity approaches were explored, where the main trends are captured on a low-fidelity model and the solution is refined on a high-fidelity model [299]. Thanks to these advancements, applying these methods results in computationally-efficient stochastic evaluations of HRES. In conclusion, with several well-developed and computationally-efficient UQ and RDO tools available, it is suggested to at least partially address, even in a simple format, uncertainty in HRES design.

**Simple uncertainty approaches are still better than a deterministic assumption** Defining the uncertainty on a parameter can be even more uncertain than an estimate on the true value of a parameter [300]. Especially in the context of HRES, where the knowledge on most of the parameters is limited to a handful of data points. However, characterizing the uncertainty with a precise distribution or interval is more valuable than a deterministic, best-guess value, as it provides an initial value on the variance of the quantity of interest. In addition, deriving the distribution or interval from a limited set of data points or expert judgement provides a rigorous estimate on the order of magnitude on the parameter un-

## Chapter 5. Conclusions and perspectives

---

certainty, which is more meaningful in RDO and global sensitivity analysis than adopting generic ranges (e.g., assuming 10 % variation on all parameters).

**Consult expert judgement when no data is available** When limited or no data is available, uncertainty ranges can initially be biased by a single data point. Therefore, expert judgement can be consulted to determine if the parameter will be a significant contributor to the uncertainty on the quantity of interest. If yes, expert judgement can initialize a conservative range on the parameter. When propagating this uncertainty through the system model, the Sobol' index can be evaluated to determine the importance of that parameter, based on the initial conservative range. If the Sobol' index is negligible, then the parameter can be safely considered as deterministic. If the Sobol' index is significant, it is useful to refine the initial range, avoiding biased results and avoiding robust design alternatives that mainly focus on reducing the dependency on this uncertain parameter in the quantity of interest.

**Trends are more important than actual design values** Linked to the second point, a simplified uncertainty characterization might not result in the actual stochastic behavior of the quantity of interest. Nevertheless, the trend in subsequent design changes on a Pareto front provides valuable information for the decision-making process.

**Use surrogate-assisted RDO to create a robust HRES model** Like any model, an HRES model is vulnerable to bugs. Especially when complex power management strategies need to be coded, an uncommon scenario in the strategy might not have been tested properly. The HRES model might work for nominal design conditions, which are often considered as inputs for the design variables during the development of the model. But, when odd designs are provided as input for the model, the model might fail to produce feasible results. As the UQ and optimization algorithm consider the model as a black box, the set of input samples is constructed without knowledge of the model during RDO. Hence, the design space and stochastic space is explored by the set of samples, including unusual designs (e.g., no PV array, but a large battery stack). Evaluating such a set of samples is an excellent debugging tool, as it indicates if the model provides feasible results over the entire stochastic design space. Moreover, due to the inherent nature of the optimizer, the optimizer will actively look for designs that result in non-feasible results. To illustrate, when the aim is to optimize

the self-sufficiency ratio, which, in the formulation in this thesis, cannot exceed 100 %, the optimizer will try to find designs that exceed this upper limit. If the model provides designs with a self-sufficiency ratio larger than 100 %, the model is wrongly defined for these designs.

**Extend towards existing energy system models** The HRES models are developed in Python. However, many energy system models exist which can be controlled externally, e.g., define the parameters by modifying an input *.txt* file, run the model via an executable file and read the results from an output *.txt* file. Examples include EnergyPLAN [122] and EnergyScopeTD [13]. The surrogate-assisted RDO can be connected to these models, to generate robust design alternatives in the specific energy system evaluated by these models (e.g., a Combined Heat and Power plant). This allows evaluating energy system models with different technical, spatial and temporal resolution, at different scales (e.g., the Belgian energy system [301]).

**Consider the trade-off between time and quality in your study** Energy system modelling is a time-intensive process. In RDO, an additional amount of time needs to be invested in the characterization of the uncertainties and the tuning of the RDO parameters. Hence, a significant amount of time needs to be invested to set up the case and evaluate the results, by people with a background in energy modelling and stochastic optimization. As both backgrounds are not always available in firms, and time is a precious good, a few recommendations are provided to perform RDO with reasonable quality and time invested. First, adopt existing component models, or generate simplified models (i.e., linearize the technical behavior of a system). Second, adopt ranges from existing literature. If the literature is not easily accessible, consult expert judgement within the firm to define the ranges. Moreover, use expert judgement to define which parameters are subject to uncertainty, i.e., the parameters most likely to contribute to the variance of the quantity of interest. Finally, perform the RDO with a low-order PCE, and refine the final optimized designs with a higher order. To make this work, a generalized database for HRES parameters and models are requested. The RHEIA framework, developed during this thesis, includes, in addition to the Python code to perform RDO, a set of models for typical HRES, and the uncertainty characterization for typical techno-economic and environmental parameters.

### 5.3 Future perspectives

The work consists of two main pillars: the surrogate-assisted RDO and the HRES. Therefore, both segments can be improved substantially. A non-exhaustive list of future perspectives is provided below.

**Characterize and propagate dependent uncertainties** In this thesis, independent uncertainties are adopted, which is a typical assumption in the majority of energy system studies [36]. However, parameter independence can result in biased conclusions (i.e., a variance on the quantity of interest that deviates from reality). Therefore, it is suggested to consider dependency between relevant uncertainties. To do so, the parameter correlations should be attributed. To illustrate, Carpaneto et al. [302] considered a correlation between the electrical load, thermal load and electricity price in a cogeneration system. Thereafter, either the dependent uncertainties need to be mapped into independent uncertainties using e.g., Nataf transformation, to construct the PCE, or the PCE algorithm needs to be modified to compute orthogonal polynomials for dependent uncertainties. To illustrate, Jakeman et al. [303] recently developed a method to compute orthonormal polynomials for dependent uncertainties using Gram-Schmidt orthogonalization.

**Consider dynamic models with uncertainties increasing over time** In energy system evaluation, parametric uncertainties exist that increase over time, e.g., the uncertainty on the energy demand for a dwelling. To address time-varying uncertainties, and thus to increase the fidelity towards reality, the complexity is increased in two ways: A method has to be established to characterize the variation of the uncertainties over time (e.g., Moret et al. [87] assumed a linear increase over time on the range of variation), and a time-dependent UQ method has to be applied (e.g., time-frozen PCE has been applied on a dynamic model of an ammonia synthesis process by Verleysen et al. [304]). Combined, the standard deviation on the quantity of interest can be quantified under time-varying uncertainties and the standard deviation can be used as an objective in a dynamic RDO procedure.

**Optimize the power management strategies** In this thesis, the design optimization was performed by quantifying the component capacities. For each design, a conventional power management strategy was performed, independent of the capacities considered in the respective design. An inner optimization can be performed, to optimize the power management strategy for each design (e.g., buy electricity from the grid when the electricity price is low, for a design with a significant battery size). This additional degree of freedom might further improve the expected performance and robustness in the quantity of interest, at the expense of an increase in computation cost (i.e., the inner optimization of the power management strategy will likely increase the computational time of a single evaluation of the model).

**Alternative optimization algorithms can be explored** NSGA-II has been applied in the presented cases, based on the proven effectiveness of the algorithm in surrogate-assisted RDO. Despite its proven effectiveness, alternative optimization algorithms can be explored to reach convergence faster (i.e. with less deterministic evaluations of the system model). To illustrate, NSGA-III is designed to improve the performance of problems with more than three objectives [305]. In addition, recent research suggests that multi-objective Cuckoo Search and multi-objective Firefly Algorithm achieve superior performance in the final Pareto front and in convergence speed [109]. Next to the convergence speed, alternative optimization algorithms can be implemented to handle discrete variables, as in more specific cases, the problem might convert into an optimization problem with discrete design variables (e.g., optimization of the number of solar panels for a specific household with known spatial constraints). Instead of rounding the design variables, which might lead to suboptimal designs, a discrete metaheuristic optimization algorithms can be implemented (e.g., Jumping Frog Particle Swarm Optimization [306]). In conclusion, the performance of alternative multi-objective optimization algorithms can be explored, while discrete optimization algorithms can be adopted for specific cases.

**Integrate Reliability-Based Robust Design Optimization (RBRDO)** In the current constraint handling method, a Gaussian distribution is assumed on the constraint parameter. In RBRDO, a probabilistic constraint is considered, which ensures that the probability of constraint violation is limited to the allowed probability of failure [35]. To quantify the probability of failure, several methods exist, such as the First-Order Reliability Method (FORM)

## Chapter 5. Conclusions and perspectives

---

and Second-Order Reliability Method (SORM) [307]. In the surrogate-assisted RDO adopted in this work, the training samples to construct a PCE for the quantities of interest can be used to construct a PCE for the constraint parameter. Through a large sampling ( $10^4$ ) of the PCE for the constraint parameter, the probability of failure can be easily quantified without additional model evaluations and thus a minimal increase in computational cost [308].

**Add socio-political objectives** To design an HRES, the performance indicators are generally classified into four categories: technical, economic, environmental and socio-political indicators. In a holistic design optimization, these four categories are considered in the objective function. Social objectives are scarcely considered in HRES optimization, but are of interest to guarantee a sustainable solution [25]. Whenever considered, the socio-political objectives are quantified using simplified functions. To illustrate, Sawle et al. [24] presented objective functions for the Human Development Index and job creation for an HRES using weights to define the importance of having surplus energy (e.g. a job creation factor of 0.14 jobs per GWh/year of surplus energy). Such weights are difficult to characterize, especially in a stochastic environment. To reach a similar fidelity on the social aspect than on e.g., the technical aspect of the model, which is generally well-developed due to the engineering background of the authors in most of the studies, it is suggested to include expertise from the socio-political field to characterise the uncertainties and objective function for the socio-political dimension.

# A Application of the methods on an illustrative example

## Overview

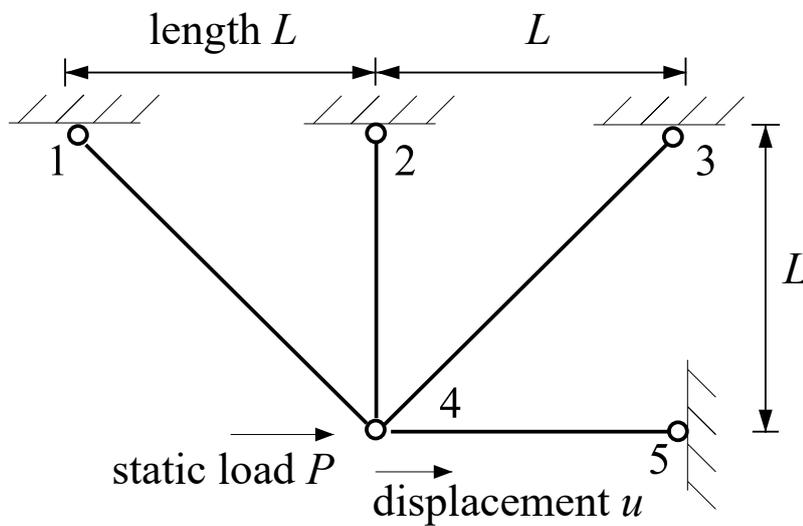
The application of the surrogate-assisted RDO framework is illustrated on an academic case. The appendix provides more information on:

- The polynomial order in surrogate-assisted RDO from Chapter 2;
- The number of training samples for the sparse PCE in the surrogate-assisted RDO from Chapter 3;
- The robust design characteristics for the four-bar truss;
- The imprecise probability method from Chapter 4.

To illustrate the application procedure of the proposed methods, and the advantages of a robust design, the methods have been applied on an illustrative example, adopted from Kang [38]. The non-linear model (Section A.1) is designed with a single design variable, which enables to clearly visualize the design space in the post-processing of the results. In the original case, only two parameters are considered uncertain. In this case, three additional parameters are considered uncertain (Section A.2), to complicate the convergence of the polynomial order (Section A.3) and to better illustrate the advantages of using the sparse PCE algorithm from Chapter 3 (Section A.4). The advantages of the robust design are illustrated in Section A.5. In Section A.6, the precise distributions are replaced by parametric probability-boxes, to illustrate the UQ method proposed in Chapter 4.

## A.1 Four-bar truss case

In the four-bar truss structure (Figure A.1), the displacement of the fourth node, induced by a static load  $P$ , is the quantity of interest. To control this displacement, the cross-sectional area of the first and third bar ( $A_1$ ), and the cross-sectional area of the second and fourth bar ( $A_2$ ) can be designed. Each bar is characterized by a Young modulus  $E$  and mass density  $\rho$ . The weight of the entire structure is fixed at 5, which implies that  $A_2 = \frac{5}{2\rho L} - \sqrt{2}A_1$ . Hence, the optimization problem is reduced to a single design variable  $A_1$ .



**Figure A.1:** By designing the cross-sectional areas of the bars in this four-bar truss, the displacement of the fourth node can be minimized. This figure is adapted from Kang [38].

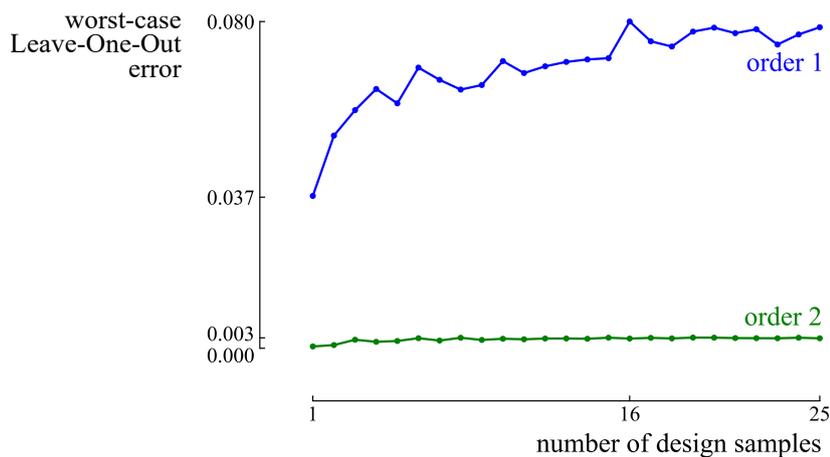
## A.2 Uncertainty characterization with precise probabilities

In this example, 5 stochastic parameters are considered, which are characterized by a Gaussian distribution. The first and third bar are defined by an uncertain Young modulus  $E_1 \sim \mathcal{N}(210, 21)$ . Similarly, the Young modulus for the second and fourth bar is subject to  $E_2 \sim \mathcal{N}(100, 5)$ . The fourth node is subject to a static load  $P \sim \mathcal{N}(1, 0.01)$ , the mass density of the materials equals  $\rho \sim \mathcal{N}(1, 0.01)$ , and the length  $L \sim \mathcal{N}(1, 0.001)$ .

### A.3 Determination of the polynomial order

To perform surrogate-assisted RDO (Subsection 2.2.4) on the four-bar truss example, the total degree of the multivariate polynomials needs to be defined (i.e., the order of the PCE). Based on the method proposed in Subsection 2.2.4.1, the aim is to find the polynomial order that leads to a sufficiently accurate PCE over the entire design space. To screen the design space, the LOO error is evaluated for a set of design samples. If the set of design samples captures the worst-case LOO over the entire design space, the size of the set of design samples can be considered sufficient. Hence, the size is gradually increased, and for each size, the worst-case LOO is stored (Figure A.2). Note that each set of design samples is generated independently from the previous one via LHS (i.e., the set with 3 design samples does not necessarily contain the 2 design samples generated for the set with 2 design samples). The highest worst-case LOO error (0.080) for a PCE with order 1 is found when the set contains 16 design samples. Despite that for each set, new design samples were generated via LHS, an increasing trend in the captured worst-case LOO with the number of design samples is visible. For a PCE with order 2, the highest worst-case LOO error found corresponds to 0.002. Following this result, a PCE with order 2 corresponds to a sufficiently accurate representation of the input-output relation over the entire design space.

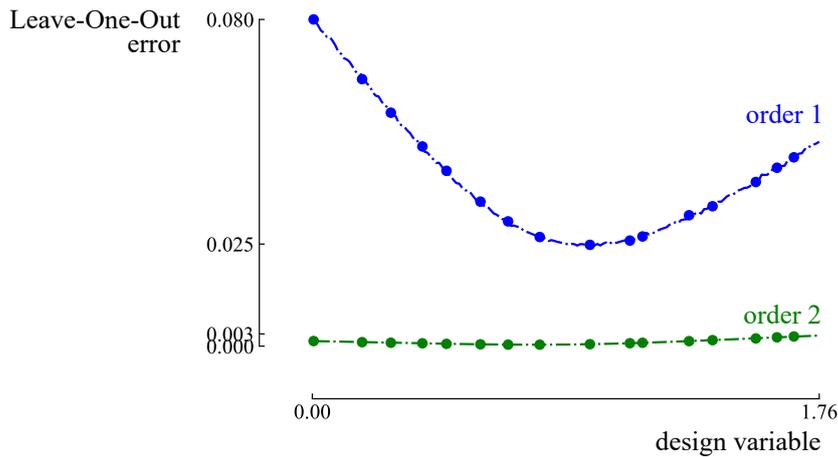
To evaluate if the worst-case LOO error found with the screening method, corresponds to the actual worst-case LOO error in the design space, 1000 random design samples are



**Figure A.2:** For each set of design samples, the worst-case Leave-One-Out (LOO) error is quantified for a Polynomial Chaos Expansion of order 1 and order 2. When considering a set of 16 design samples, the highest worst-case LOO error is retrieved.

## Appendix A. Application of the methods on an illustrative example

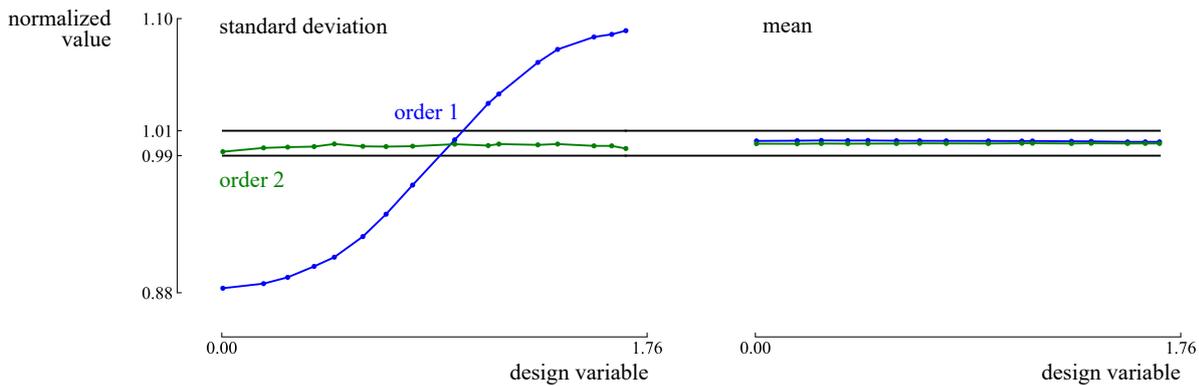
generated and the corresponding LOO errors are plotted for a polynomial order of 1 and 2 (dashed lines on Figure A.3). Clearly, the LOO error varies significantly over the design space for a PCE with order 1 (i.e., between 0.037 and 0.080). Indeed, the maximum LOO error (0.080 for order 1, 0.003 for order 2) was captured by the screening method. To illustrate, the LOO errors retrieved when considering a set of 16 samples are presented (dots on Figure A.3).



**Figure A.3:** The Leave-One-Out error (LOO) of a Polynomial Chaos Expansion (PCE) order 1 and 2 for 16 design samples generated through Latin Hypercube Sampling (dots). The dashed line corresponds to the LOO error from 1000 random design samples for each order. Considering a PCE order 1 leads to a significant variation in the LOO error over the design space.

Finally, the error on the statistical moments is evaluated over the design space. To do so, the displacement mean and standard deviation for the 16 design samples are retrieved from a PCE (order 1 and order 2) and from a MCS with  $10^5$  training samples (i.e., the reference value). For each design sample, the displacement mean from both PCEs is divided by the mean retrieved from MCS (Figure A.4). The same approach is applied for the standard deviation. The resulting normalized values illustrate that, for each design sample, the displacement mean from a PCE with order 1 and from a PCE with order 2 result in a difference below 1 % with the reference mean value. Hence, when only the displacement mean is of interest, a PCE with order 1 is sufficiently accurate. For the standard deviation retrieved from a PCE with order 1, the deviation goes up to 12 % at the lower bound of the design space (i.e., when the design variable  $A_1$  corresponds to 0). As expected, this design point corresponds to the PCE with the highest LOO error over the design space, as illustrated in Figure A.3. For a PCE with order 2, the standard deviation corresponds to an error below 1 % when compared to the reference value, and this over the entire set

of 16 design samples. Thus, by adopting a PCE with order 2, the surrogate-assisted RDO algorithm operates with accurate statistical moments over the design space. Note that the computational budget might not always be available to quantify reference values for the statistical moments through MCS. In that case, the accuracy of the PCE over the design space can be performed based on the LOO error.



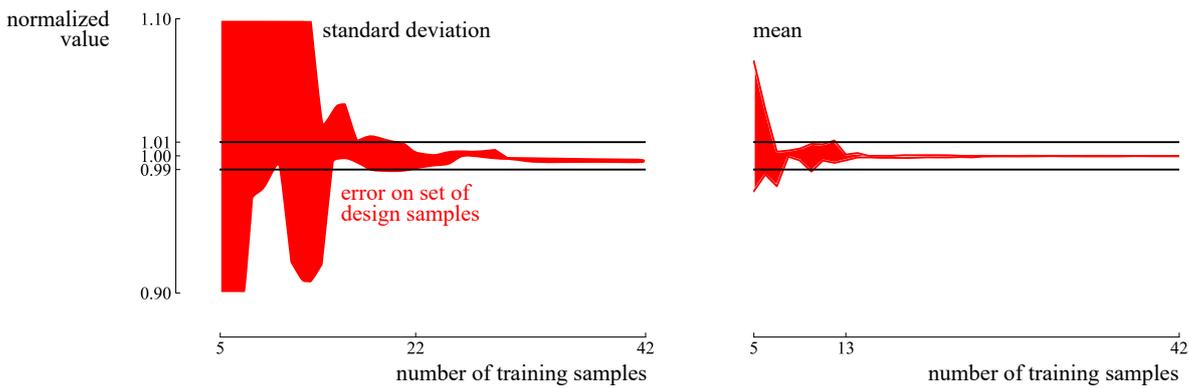
**Figure A.4:** The normalized values for the mean and standard deviation of the displacement over the design space indicate that with a Polynomial Chaos Expansion of order 2, the derived statistical moments deviate by less than 1 % from the reference values.

## A.4 Characterization of the sparse PCE

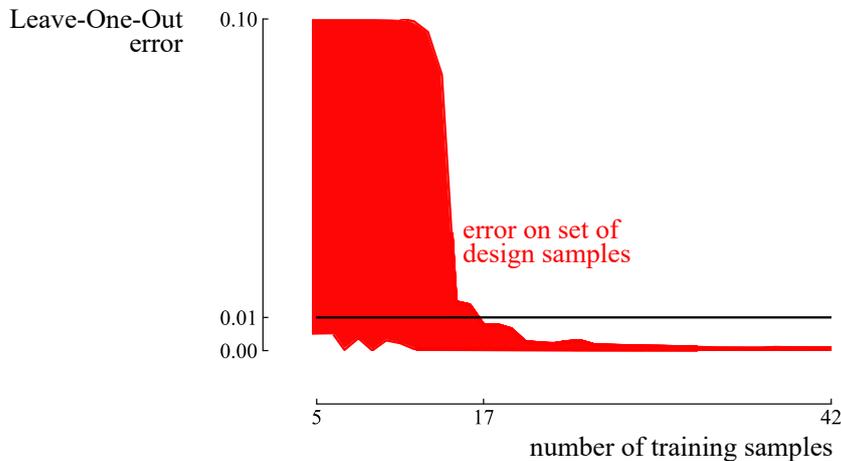
Based on Section A.3, a PCE with polynomial order 2 is sufficient to generate accurate statistical moments over the design space. Hence, 42 model evaluations are required to construct each PCE, according to the conventional truncation scheme (subsection 2.2.2.1). This number of training samples can be reduced by adopting the stepwise regression sparse PCE method (Section 3.2). In this method, the number of training samples needs to be defined a priori, without knowing if the size will suffice. For a gradually increasing size of training samples, the resulting mean and standard deviation is derived from the sparse PCE. These statistical moments are divided by the mean and standard deviation generated by MCS with  $10^5$  training samples. This approach is performed for the set of 16 design samples, generated via LHS (Section A.3). The normalized values for the mean illustrate that an accurate mean is found rather quickly: With 13 training samples, the means for the 16 design samples deviate below 1 % from the reference value (Figure A.5, where the red area is shaped by the lowest and highest value found for each set of training samples). For the standard deviation, the number of training samples required is higher (22). Nevertheless,

## Appendix A. Application of the methods on an illustrative example

to reach statistical moments with acceptable accuracy over the design space, only 52 % (22) samples are needed, when compared to the number of training samples needed when adopting the conventional truncation scheme (42). When a reference value can not be quantified by MCS due to computational limits, either the reference value can be replaced by the statistical moments derived from a conventional PCE, or the LOO error for the sparse PCE can be used (Figure A.6). In this case, the LOO error can be reduced below 0.01 with 17 training samples.



**Figure A.5:** With only 22 training samples, the error on the statistical moments remains below 1 % over the design space, when adopting the sparse Polynomial Chaos Expansion method.



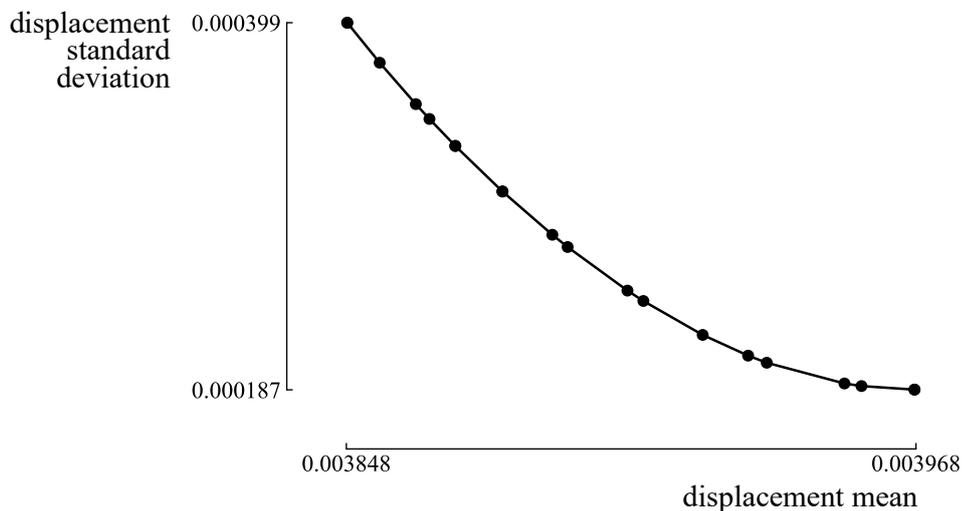
**Figure A.6:** The Leave-One-Out error of the sparse Polynomial Chaos Expansion algorithm over the design space lies below 0.01, when at least 17 training samples are generated.

### A.5 Robust design characteristics

The application of surrogate-assisted RDO results in a trade-off between minimizing the displacement mean and minimizing the displacement standard deviation (Figure A.7). The

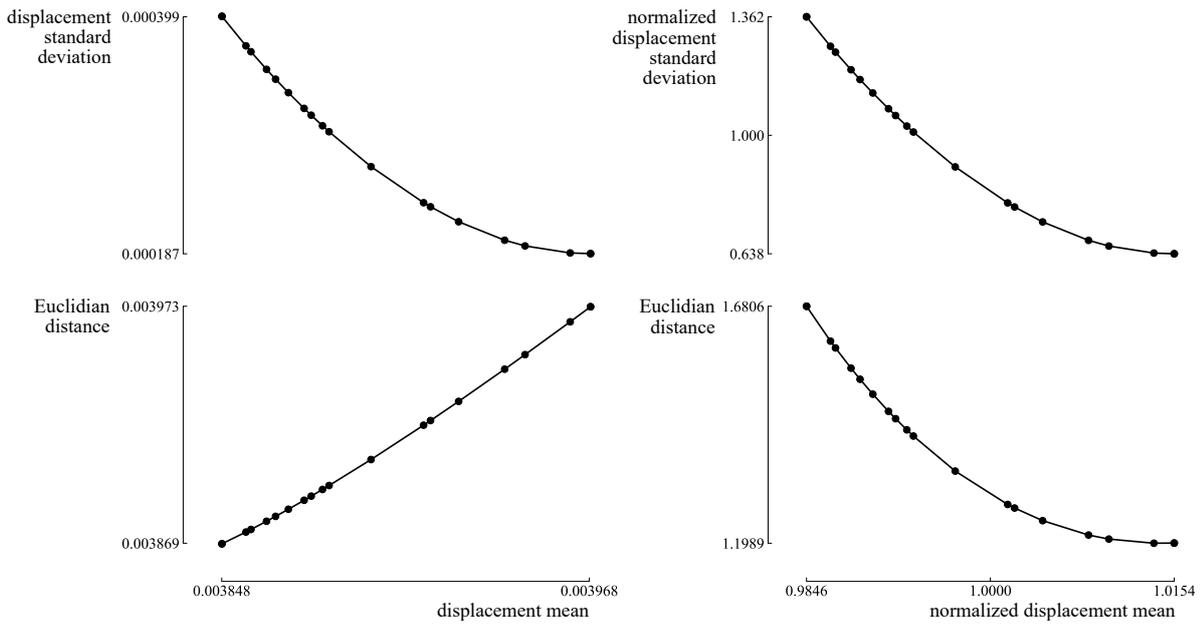
design with the optimized mean ( $A_1 = 1.767$ ,  $A_2 = 0$ ) achieves a displacement mean of  $3.85 \times 10^{-3}$  and a displacement standard deviation of  $3.99 \times 10^{-4}$ . Instead, the robust design ( $A_1 = 0.352$ ,  $A_2 = 2.001$ ) decreases the standard deviation by 53 %, down to  $1.87 \times 10^{-4}$ , at the expense of an increase in mean with 3 %, up to  $3.97 \times 10^{-3}$ . Hence, the robust design results in a significantly lower sensitivity towards the uncertainty, while the increase in mean displacement is tolerable. Thus, the robust design from RDO is a suitable alternative for the four-bar truss structure.

To decide on a design from the Pareto front, the knee-point method is applied on the absolute and normalized Pareto front (Figure A.8), with an utopia point defined at  $[0, 0]$ . For the typical Pareto front representation (left), the design with the optimized mean achieves the lowest Euclidian distance towards the utopia point. Instead, when the objectives are normalized (right), the robust design achieves the lowest distance towards the utopia point. Finally, the PDF (Figure A.9) and CDF (Figure A.10) for the optimized mean design and robust design are presented. The CDF illustrates that when the designer defines a constraint on the displacement of 0.00412, both design achieve the same probability (77 %) that in reality, the displacement will be lower or equal to 0.00412. When an upper limit is defined larger than 0.00412, the robust design ensures a higher probability to not violate the constraint, when compared to the optimized mean design.

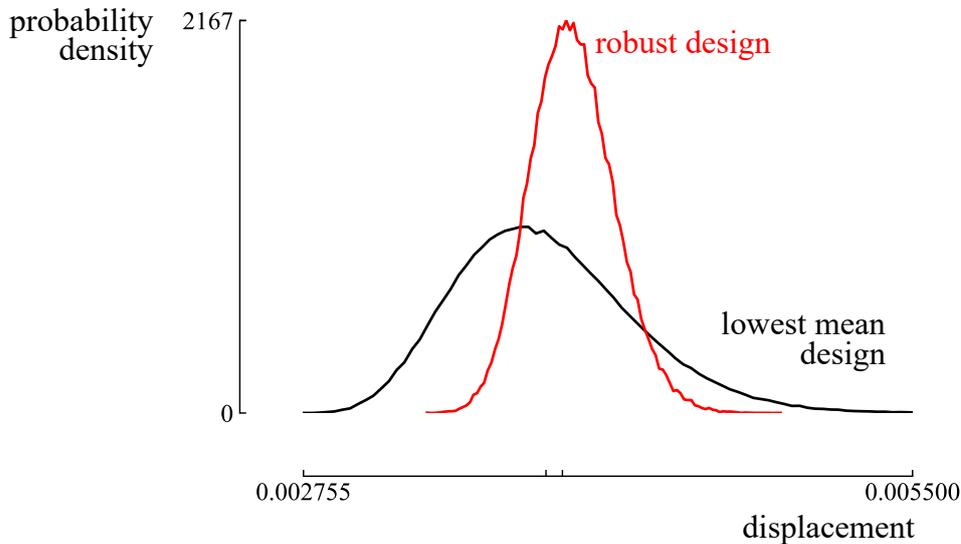


**Figure A.7:** A trade-off exists between minimizing the displacement mean and displacement standard deviation, where the displacement standard deviation can be significantly reduced (53 %), at the expense of a tolerable increase in displacement mean (3 %).

## Appendix A. Application of the methods on an illustrative example



**Figure A.8:** For a typical Pareto front representation, the design with the optimized mean achieves the lowest Euclidian distance towards the utopia point. Instead, when the objectives are normalized, the robust design achieves the lowest distance towards the utopia point. The utopia point is defined at  $[0, 0]$

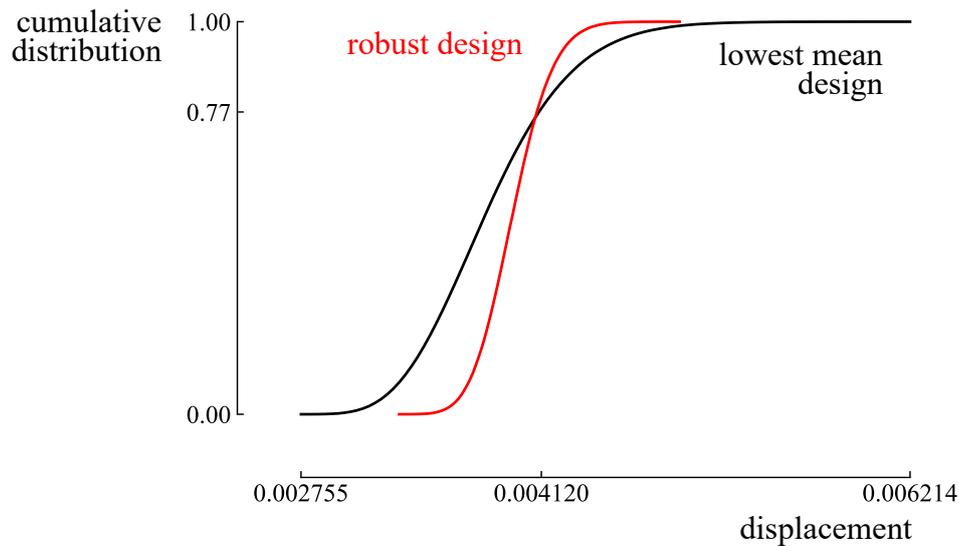


**Figure A.9:** The probability density functions of the design with optimized mean and the design with optimized standard deviation (i.e., robust design) illustrate that the displacement scatter is significantly smaller for the robust design.

## A.6 Uncertainty quantification with imprecise probabilities

In this example, the uncertainties described in Section A.2 are characterized by parametric p-boxes. Hence, the parameters are characterized by both epistemic and aleatory uncertainty.

## A.6 Uncertainty quantification with imprecise probabilities

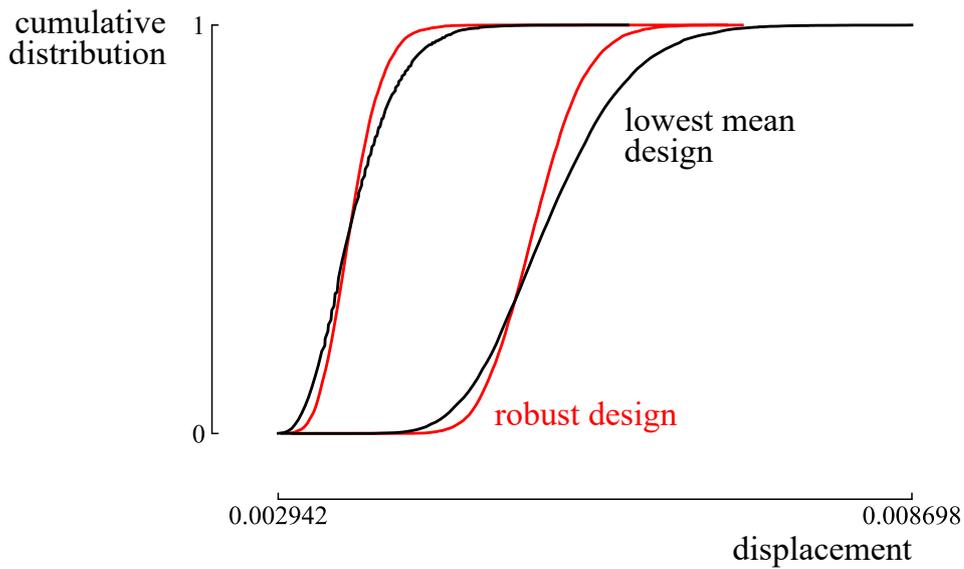


**Figure A.10:** The cumulative distribution functions for the robust design and the optimized mean design illustrate that the robust design ensures with a higher probability that the displacement will be below any value higher than 0.00412.

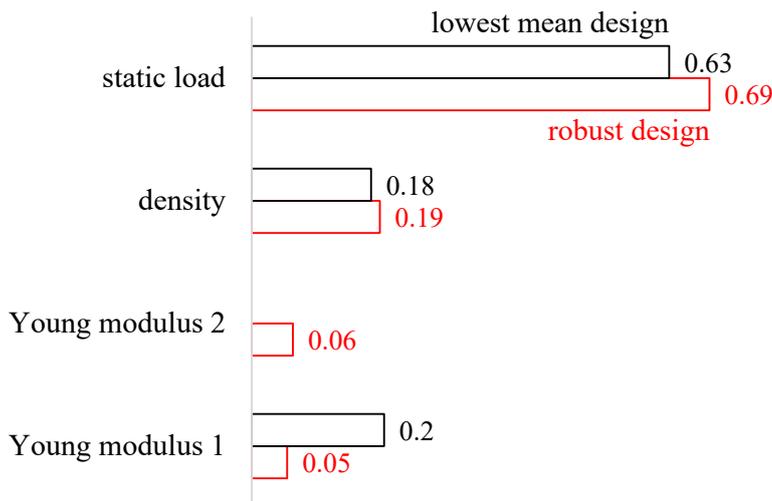
The aleatory uncertainty might relate to variability in the exact composition of the truss material (which affects the density and Young moduli), the manufacturing tolerances (which affects the actual lengths of the bars) and the variation in the static load. The epistemic uncertainty might relate to not knowing the material that will be used to construct the truss, and to inaccuracies in the method used to quantify the static load. Following the distributions considered in Section A.2, epistemic uncertainty is considered on the Young moduli  $\mu_{E_1} \in [200, 220]$  and  $\mu_{E_2} \in [100, 112]$ , static load  $\mu_P \in [1, 1.2]$  and density  $\mu_\rho \in [1, 1.1]$ .

The p-boxes are propagated through the system model, for the lowest mean design and for the robust design (found in Section A.5), using the method discussed in Section 4.2. This propagation results in the p-boxes for the considered designs (Figure A.11). Due to the epistemic uncertainty present, the p-boxes mainly overlap. To reduce the width of the p-boxes, the Sobol' indices on the range of the displacement mean, and thus on the width of the p-box, can be quantified (Figure A.12). These Sobol' indices illustrate that the width of the p-box is mainly dominated by the epistemic uncertainty on the static load (total-order Sobol' index of 0.63 for the lowest mean design, 0.69 for the robust design). Note that the epistemic uncertainty related to  $E_2$  is equal to zero for the lowest mean design, as  $A_2 = 0$  for that design. In conclusion, gaining more information on the static load is the main action to reduce the width of the p-box for both designs.

## Appendix A. Application of the methods on an illustrative example



**Figure A.11:** Due to the epistemic uncertainty, the parametric probability boxes on the displacement for the lowest mean design and the robust design overlap.



**Figure A.12:** The Sobol' indices illustrate that the epistemic uncertainty on the static load dominates the epistemic uncertainty on the displacement for both designs.

The imprecise Sobol' indices indicate the average importance of an input p-box (the average value) on the variance of the displacement, affected by epistemic uncertainty (the range on the Sobol' index). Despite the epistemic uncertainty, it is clear that the variance in the displacement of the lowest mean design is dominated by Young modulus 1, i.e.,  $S_{E_1}^T \in [0.968, 0.993]$  (Table A.1). For the robust design, both Young moduli dominate the variance in the displacement.

**Table A.1:** The imprecise Sobol' indices for the displacement of both designs.

parameter	imprecise Sobol' index	
	robust design	lowest mean design
Young modulus 1	[0.489, 0.608]	[0.968, 0.993]
Young modulus 2	[0.361, 0.474]	[0.000, 0.008]
density	[0.014, 0.026]	[0.002, 0.020]
static load	[0.011, 0.024]	[0.002, 0.018]
length	[0.000, 0.001]	[0.000, 0.001]



## B Physical behavior of system components

### Overview

The physical behavior is detailed on:

- a wind turbine array;
- a PV array;
- a PEM electrolyzer;
- a PEM fuel cell;
- an air-source heat pump.

This chapter provides additional information on the system component models. First, the construction of the hourly capacity factor for the wind turbines is explained, followed by the behavior of the current-voltage curves of the photovoltaic array under different solar irradiance and ambient temperature conditions. Thereafter, the current-voltage curves, power density and energy efficiency of the PEM electrolyzer and PEM fuel cell are elaborated. The chapter concludes with the efficiency profile of the scroll compressor and the coefficient of performance of the air-source heat pump, for a range of different ambient temperatures.

### B.1 Wind turbine array

To quantify the wind turbine array power output, the method proposed by Staffel et al. [229] is adopted. The method allows quantifying the power output of a wind turbine array at

## Appendix B. Physical behavior of system components

---

any location around the world using reanalysis models (i.e., output from global atmospheric simulations). Typical reanalysis datasets are Modern-Era Retrospective analysis for Research and Applications (MERRA) [309] and the European Centre for Medium-Range Weather Forecasts twentieth century reanalysis (ERA-20C) [310]. Starting from the grid points provided by reanalysis, the wind speeds for the wind farm are quantified through interpolation. The interpolation is performed using a non-parametric, locally weighted scatterplot smoother (LOESS). Following this spatial interpolation, a height extrapolation is required as well, as the wind speeds are initially provided at the height of reanalysis. Hence, a wind profile is adopted, allowing to quantify the wind speed at the wind turbine height. A logarithmic wind profile is adopted, which is considered accurate when wind speeds are within the atmospheric boundary layer and that the atmosphere has neutral stability (i.e., the air temperature drops with increasing altitude at a dry adiabatic lapse rate of  $-9.8^\circ\text{C}/\text{km}$ ). The wind profile is presented as follows:

$$w(h) = \left(\frac{u}{\kappa}\right) \log\left(\frac{h-d}{z}\right) = A \log(h-d) - A \log(z), \quad (\text{B.1})$$

where  $w$  is the wind speed at the desired height  $h$ ,  $u$  is the friction velocity,  $\kappa$  is the Von Karman constant (0.4),  $A = \frac{u}{\kappa}$  is the wind shear,  $d$  is the displacement height and  $z$  is the surface roughness. Based on wind speeds at three stationary heights, adopted from MERRA, the wind shear  $A$  and surface roughness  $z$  are estimated.

Using these wind speeds, the estimated capacity factor  $CF_{\text{sim}}$  can be quantified, which follows out of the aggregate power curve  $PC_A$  of the wind farm. The aggregate power curve is a smoothed transformation of a turbine power curve  $PC_T$ . To illustrate, the power curve for the Vestas V90 2 MW wind turbine has a cut-in speed of 4 m/s, cut-off wind speed of 25 m/s and rated wind speed of 13.5 m/s (Figure B.1). The smoothed transformation accounts for different wind speeds within a given hour, and for the spatial distribution of the wind turbines in the farm:

$$PC_A(x, \sigma) = \sum_{n=-4\sigma}^{4\sigma} PC_T(x-n) \tau(n, \sigma), \quad (\text{B.2})$$

which corresponds to the integral over the turbine power curve, evaluated by a wind speed

$x$  and multiplied by a Gaussian filter  $\Gamma$ :

$$\Gamma(n, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{n^2}{2\sigma}\right), \quad (\text{B.3})$$

where the width of the Gaussian filter  $\sigma$  is characterized such that it captures the increase in spatial and temporal variation with the increase in wind speed:

$$\sigma = 0.6 + 0.2w. \quad (\text{B.4})$$

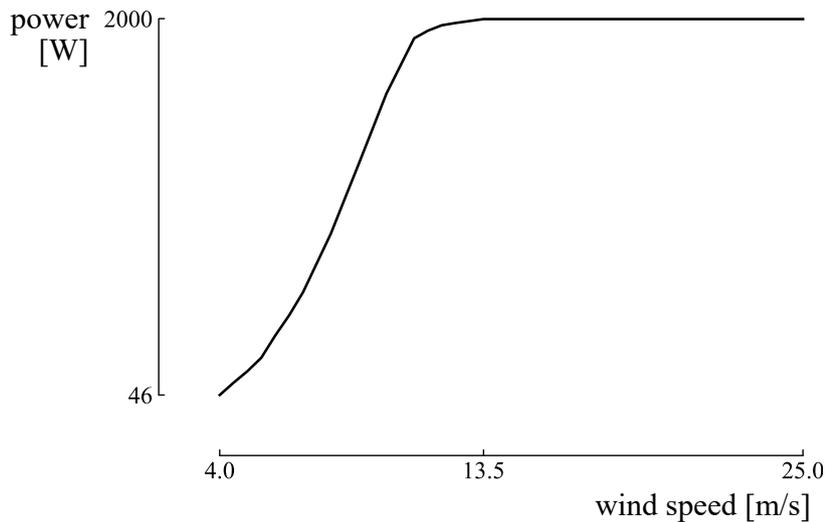
Finally, a bias correction is proposed, based on the observed capacity factors  $CF_{\text{obs}}$  and simulated capacity factors  $CF_{\text{sim}}$  for a country:

$$\epsilon_{CF} = \frac{CF_{\text{obs}}}{CF_{\text{sim}}}. \quad (\text{B.5})$$

This bias correction corrects the energy yield as follows:

$$P(w') = \epsilon_{CF} P(w), \quad (\text{B.6})$$

where  $w'$  corresponds to the corrected wind speed,  $P$  is the conversion function from wind speed to capacity factor,  $\overline{P(w')} = CF_{\text{obs}}$  and  $\overline{P(w)} = CF_{\text{sim}}$ . Additional information on the method is described by Staffell et al. [229].

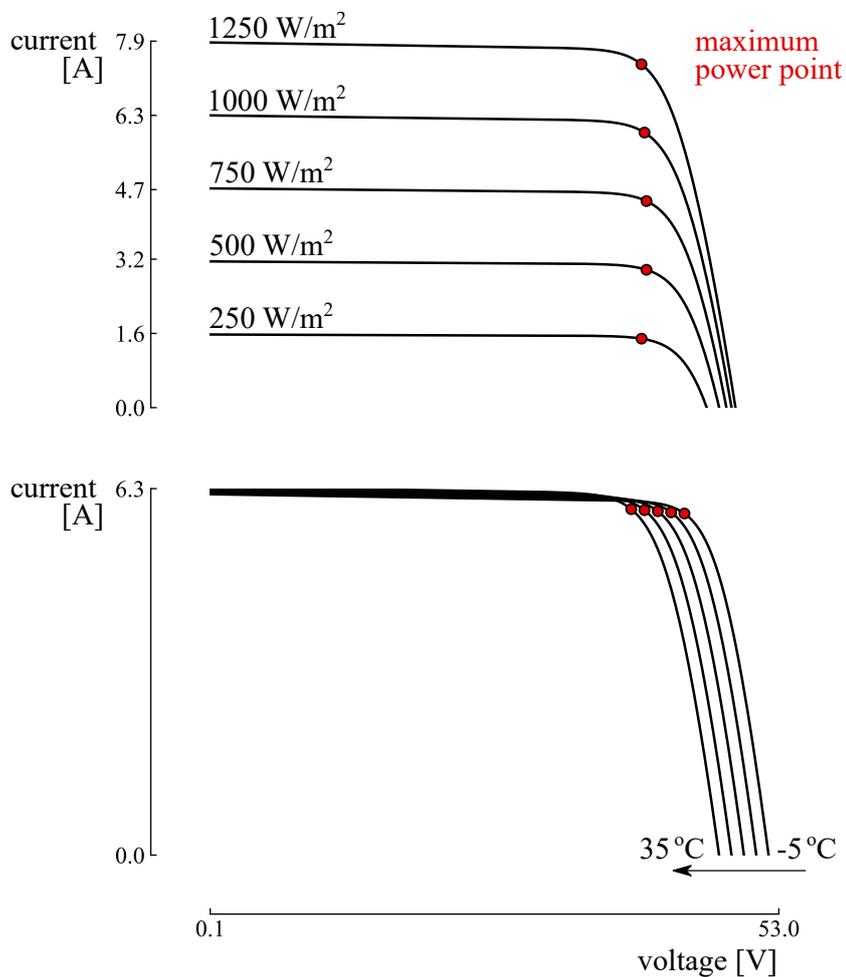


**Figure B.1:** The power curve of a Vestas V90 2000 wind turbine, with a cut-in, cut-off and rated wind speed of 4 m/s, 25 m/s and 13.5 m/s, respectively [311].

## B.2 Photovoltaic array

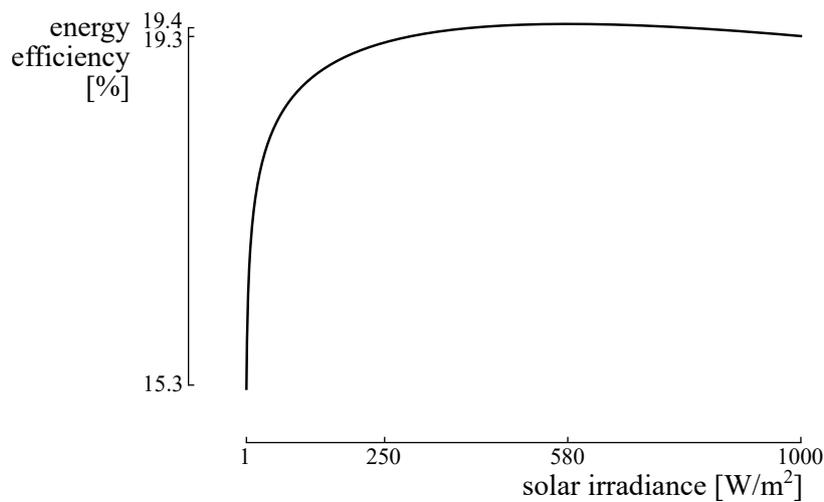
To evaluate the performance of a photovoltaic array, the characteristics are adopted from a typical monocrystalline silicon PV panel (Sunpower SPR X-19-240-BLK [191]):

- 240  $W_p$  nominal power;
- 5 % power tolerance;
- 42.8 V rated voltage;
- 5.61 A rated current;
- 50.6 V open-circuit voltage;
- 5.98 A short-circuit current.



**Figure B.2:** The current-voltage characteristics of the PV panel at different ambient temperatures (bottom figure) illustrate the reduced power output when the ambient temperature increases, while the power output increases with the solar irradiance (upper figure).

The corresponding current-voltage characteristic illustrates the power output, depending on the solar irradiance and ambient temperature (Figure B.2). The actual power output depends on the operating point. As the PV array is connected to a DC-DC converter with Maximum Power Point Tracking, the maximum power point is assumed as the operating point at each time step. At this maximum power point, the efficiency of the PV technology is approximately 19.3 % in a typical solar irradiance range of 250 W/m<sup>2</sup> - 1000 W/m<sup>2</sup>, with a maximum efficiency of 19.4 % near a solar irradiance of 580 W/m<sup>2</sup> (Figure B.3).

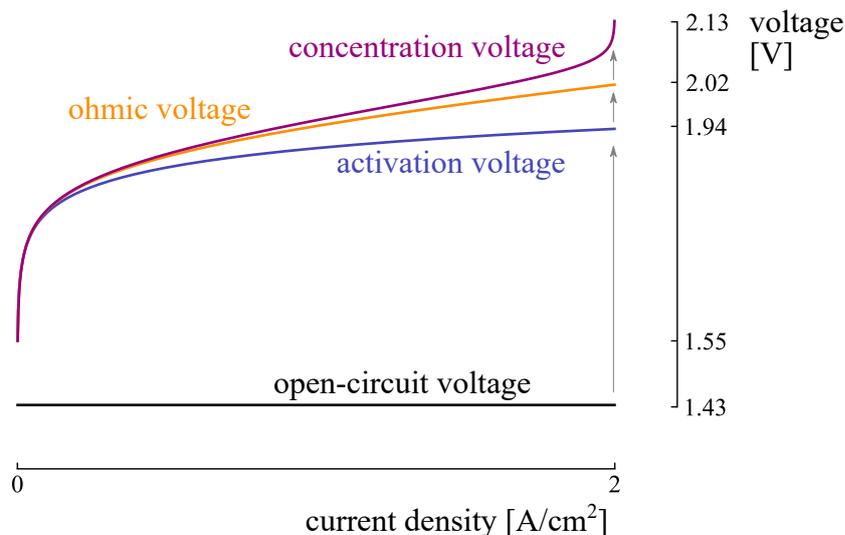


**Figure B.3:** The energy efficiency of the considered photovoltaic technology in function of the solar irradiance illustrates that the efficiency is near 19.3 % in a typical solar irradiance range of 250 W/m<sup>2</sup> - 1000 W/m<sup>2</sup>.

### B.3 PEM electrolyzer

Similar to the PV array, the power consumption of a PEM electrolyzer is characterized by a current-voltage characteristic. The characteristic is configured starting from the open-circuit voltage, and affected by the activation, ohmic and concentration voltage (Figure B.4). Clearly, the largest increase comes from the activation voltage, with a significant effect at low current densities. The ohmic voltage contributes moderately to the operating voltage, with a steady increase in contribution with respect to the current density. The contribution of the concentration voltage to the operating voltage is significant only at high current densities.

With the operating voltage increasing with the current density, the power density of the

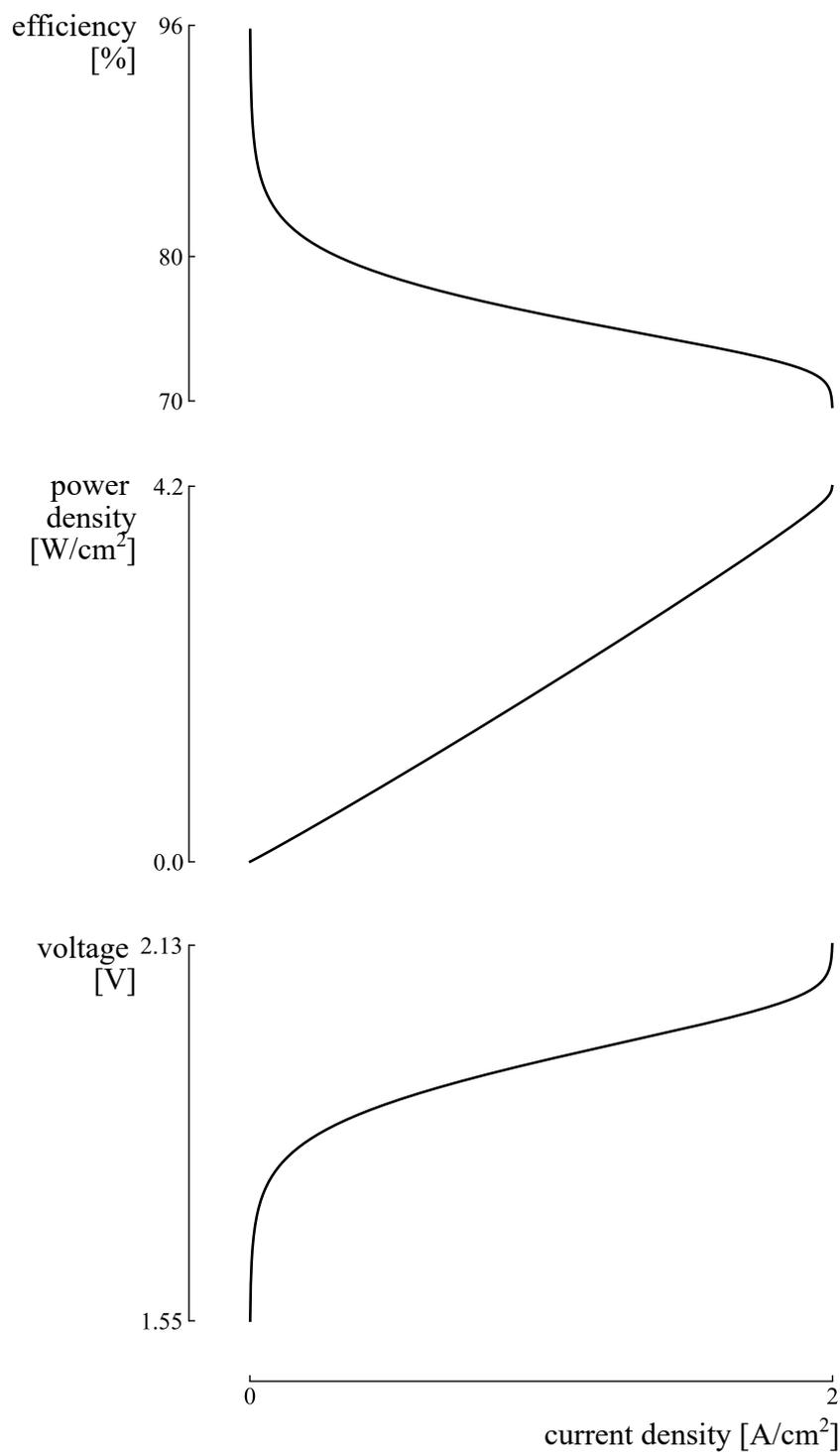


**Figure B.4:** The operating voltage of the electrolyzer is driven mainly by the activation voltage, with a moderate contribution from the ohmic voltage and a contribution from the concentration voltage only at high current densities.

PEM electrolyzer increases proportionally with the current density, up to  $4.2 \text{ W/cm}^2$  at the limiting current density of  $2 \text{ A/cm}^2$  (Figure B.5). The power applied to the PEM electrolyzer is used to quantify the energy efficiency, i.e., the energy content of the hydrogen produced, divided by the power applied. The energy efficiency is high (96 %) at low current density, but the hydrogen production is minimal. Instead, at the current density upper limit, the energy efficiency is reduced to 70 %. As the hydrogen flow rate is linearly proportional to the current, the shape of the efficiency curve with respect to the current density depends only on the operating voltage.

### B.4 PEM fuel cell

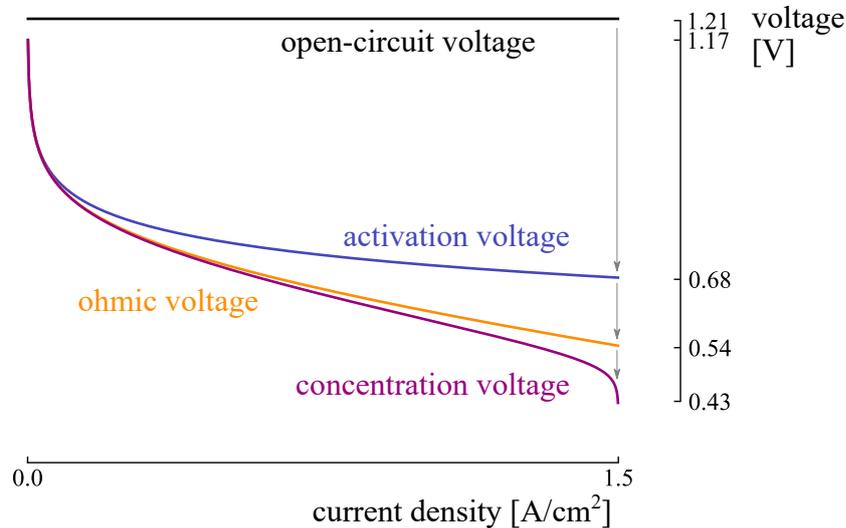
A PEM fuel cell operates in reverse when compared to PEM electrolyzer: while the electrolyzer uses power to split water into hydrogen and oxygen, the PEM fuel cell generates power by combining water and oxygen into water. Nevertheless, a PEM fuel cell has a similar structure to a PEM electrolyzer, i.e., an anode and cathode, separated by a solid electrolyte. Therefore, a PEM fuel cell is characterized by a current-voltage characteristic as well. Similarly, the operating voltage is characterized by the open-circuit voltage, affected by an activation voltage, ohmic voltage and concentration voltage (Figure B.6). In this case, the operating voltage is reduced due to these effects: the activation voltage significantly reduces



**Figure B.5:** The power density is proportional with the current density, as the power is characterized by the operating current and voltage. As the hydrogen flow rate is linearly proportional to the current, the shape of the energy efficiency curve with respect to the current density depends only on the operating voltage.

## Appendix B. Physical behavior of system components

the operating voltage at low current densities, the ohmic voltage reduces the operating voltage at intermediate current densities, while the concentration voltage mainly affects the operating voltage at high current densities.

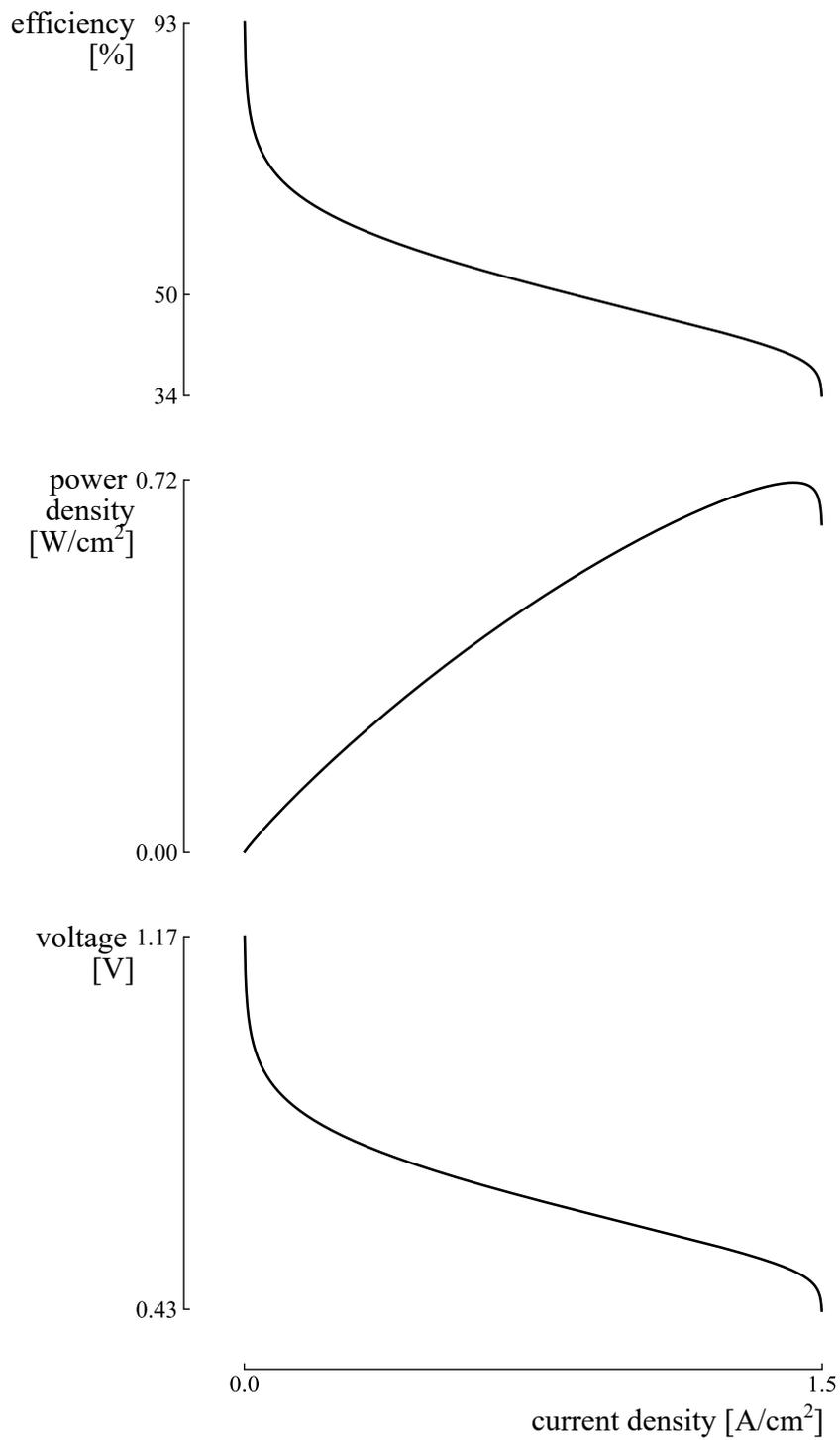


**Figure B.6:** The operating voltage of the fuel cell is driven mainly by the activation voltage, with a moderate contribution from the ohmic voltage and a contribution from the concentration voltage only at high current densities.

Similar to the PEM electrolyzer, the power density of the PEM fuel cell is proportional to the current density (Figure B.7). However, due to the significant operating voltage range (0.43 V - 1.17 V, as opposed to 1.55 V - 2.13 V for the PEM electrolyzer), the voltage curve affects the shape of the power density curve more for the PEM fuel cell than the voltage curve affects the shape of the power density curve for the PEM electrolyzer. The energy efficiency for a PEM fuel cell is the inverse of the energy efficiency for a PEM electrolyzer: the power produced, divided by the energy content of the hydrogen flow rate. The energy efficiency follows a similar shape to the energy efficiency for a PEM electrolyzer. However, the average efficiency is significantly lower, with an energy efficiency of 34 % at the limiting current density, as opposed to 70 % for the PEM fuel electrolyzer.

## B.5 Air Source Heat Pump

The Air Source Heat Pump extracts heat from the ambient air to heat up to water temperature. To extract heat from the ambient temperature, the R410a refrigerant is expanded, until the saturation temperature of the refrigerant reaches a lower state than the ambient air

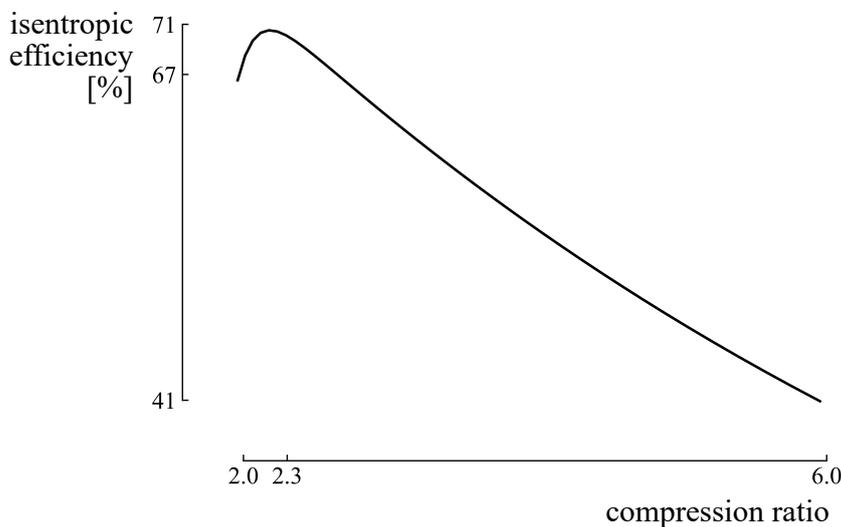


**Figure B.7:** The power density is proportional to the current density, as the power is characterized by the operating current and voltage. As the hydrogen flow rate is linearly proportional to the current, the shape of the energy efficiency curve with respect to the current density depends only on the operating voltage.

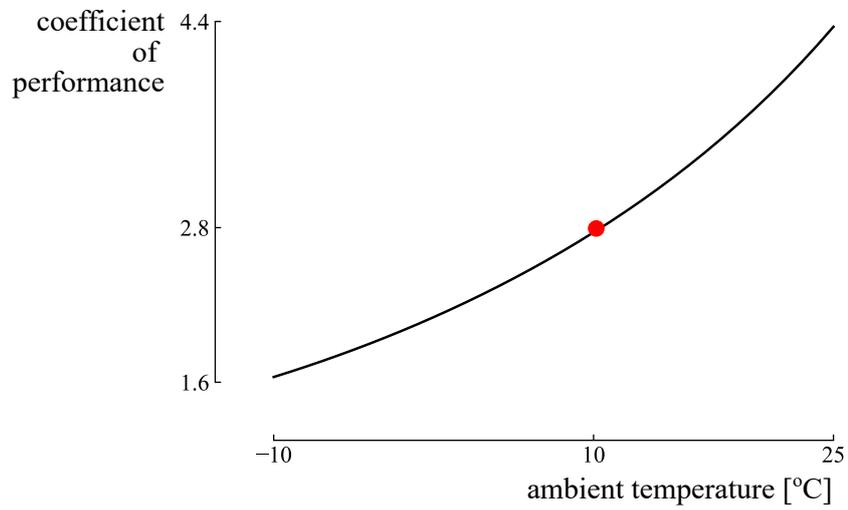
## Appendix B. Physical behavior of system components

---

temperature. After evaporation, the refrigerant is compressed up to a pressure that corresponds to a saturation temperature that reaches a value higher than the water temperature. Hence, a low ambient temperature requires a lower pressure at the evaporator, and thus a higher compression ratio to reach the pressure in the heat exchanger. However, a low compression ratio is preferred to reach the highest isentropic efficiency in the fixed-speed scroll compressor (Figure B.8). The peak efficiency of 71 % is achieved at a compression ratio of 2.3, and the efficiency drops down to 41 % at a compression ratio of 6. Following the decreased isentropic efficiency at high compression ratio, the COP is low at low ambient temperatures. To illustrate, the COP equals 1.6 at an ambient temperature of  $-10^{\circ}\text{C}$ , while the COP equals 2.8 at an ambient temperature of  $10^{\circ}\text{C}$  and 4.4 at an ambient temperature of  $25^{\circ}\text{C}$  (Figure B.9).



**Figure B.8:** The isentropic efficiency of the scroll compressor is the highest at a compression ratio of 2.3 (71 %), and reduces down to 40 % at a compression ratio of 6.



**Figure B.9:** The Coefficient Of Performance is proportional to the ambient air temperature.



# Bibliography

- [1] Statistics, IEA, “Key world energy statistics,” *Paris. International Energy Agency*, 2020.
- [2] E. Kabir, P. Kumar, S. Kumar, A. A. Adelodun, and K.-H. Kim, “Solar energy: Potential and future prospects,” *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 894–900, 2018.
- [3] British Petroleum, “Bp statistical review of world energy,” *BP Statistical Review, London, UK*, 2020.
- [4] International Energy Agency, “Renewables 2020: Analysis and forecast to 2025,” International Energy Agency, Tech. Rep., 2020.
- [5] J. Mertens, R. Belmans, and M. Webber, “Why the Carbon-Neutral Energy Transition Will Imply the Use of Lots of Carbon,” *C*, vol. 6, no. 2, p. 39, 2020.
- [6] X. Rixhon, G. Limpens, F. Contino, and H. Jeanmart, “Taxonomy of the fuels in a whole-energy system,” *Front. Energy Res. Sustain. Energy Syst. Policies*, 2021.
- [7] M. Deshmukh and S. Deshmukh, “Modeling of hybrid renewable energy systems,” *Renewable and sustainable energy reviews*, vol. 12, no. 1, pp. 235–249, 2008.
- [8] COGEN Europe, “IPCC confirms role of CHP in longterm CO2 reduction,” *Press release*, 2014.
- [9] D. Volkskrant, “Essent legt grote centrale stil,” *De volkskrant*, 2014.
- [10] D. Rekioua, “Hybrid renewable energy systems overview,” in *Hybrid Renewable Energy Systems*. Springer, 2020, pp. 1–37.
- [11] B. S. Palmintier, “Incorporating operational flexibility into electric generation planning: Impacts and methods for system design and policy analysis,” Ph.D. dissertation, Massachusetts Institute of Technology, 2013.
- [12] S. Collins, J. P. Deane, K. Poncelet, E. Panos, R. C. Pietzcker, E. Delarue, and B. P. Ó. Gallachóir, “Integrating short term variations of the power system into integrated energy system models: A methodological review,” *Renewable and Sustainable Energy Reviews*, vol. 76, pp. 839–856, 2017.
- [13] G. Limpens, “Optimisation of energy transition pathways: application to the case of Belgium,” PhD thesis, UCLouvain, Tech. Rep., 2021.
- [14] P. Arul, V. K. Ramachandaramurthy, and R. Rajkumar, “Control strategies for a hybrid renewable energy system: A review,” *Renewable and sustainable energy reviews*, vol. 42, pp. 597–608, 2015.
- [15] Y. He, M. Wang, and Z. Xu, “Coordinative low-voltage-ride-through control for the wind-photovoltaic hybrid generation system,” *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 2, pp. 1503–1514, 2019.
- [16] S. Surender Reddy and P. Bijwe, “Real time economic dispatch considering renewable energy resources,” *Renewable Energy*, vol. 83, pp. 1215–1226, 2015.

## Bibliography

---

- [17] H.-S. Lee, C. Tekin, M. Van Der Schaar, and J.-W. Lee, "Adaptive contextual learning for unit commitment in microgrids with renewable energy sources," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 688–702, 2018.
- [18] L. Yang, G. Li, Z. Zhang, X. Ma, and Y. Zhao, "Operations & maintenance optimization of wind turbines integrating wind and aging information," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 211–221, 2020.
- [19] A. S. Dagoumas and N. E. Koltsaklis, "Review of models for integrating renewable energy in the generation expansion planning," *Applied Energy*, vol. 242, pp. 1573–1587, 2019.
- [20] M. D. Al-falahi, S. D. Jayasinghe, and H. Enshaeh, "A review on recent size optimization methodologies for standalone solar and wind hybrid renewable energy system," *Energy Conversion and Management*, vol. 143, pp. 252–274, jul 2017.
- [21] K. Anoune, M. Bouya, A. Astito, and A. B. Abdellah, "Sizing methods and optimization techniques for PV-wind based hybrid renewable energy system: A review," *Renewable and Sustainable Energy Reviews*, vol. 93, pp. 652–673, 2018.
- [22] M. Iqbal, M. Azam, M. Naeem, A. S. Khwaja, and A. Anpalagan, "Optimization classification, algorithms and tools for renewable energy: A review," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 640–654, 2014.
- [23] T. Tezer, R. Yaman, and G. Yaman, "Evaluation of approaches used for optimization of stand-alone hybrid renewable energy systems," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 840–853, 2017.
- [24] Y. Sawle, S. Gupta, and A. K. Bohre, "Socio-techno-economic design of hybrid renewable energy system using optimization techniques," *Renewable energy*, vol. 119, pp. 459–472, 2018.
- [25] E. L. Eriksson and E. M. A. Gray, "Optimization and integration of hybrid renewable energy hydrogen fuel cell energy systems – A critical review," *Applied Energy*, vol. 202, pp. 348–364, sep 2017.
- [26] S. Sinha and S. Chandel, "Review of recent trends in optimization techniques for solar photovoltaic–wind based hybrid energy systems," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 755–769, 2015.
- [27] J. Lian, Y. Zhang, C. Ma, Y. Yang, and E. Chaima, "A review on recent sizing methodologies of hybrid renewable energy systems," *Energy Conversion and Management*, vol. 199, p. 112027, 2019.
- [28] L. Urbanucci, "Limits and potentials of Mixed Integer Linear Programming methods for optimization of polygeneration energy systems," *Energy Procedia*, vol. 148, pp. 1199–1205, 2018.
- [29] O. Erdinc and M. Uzunoglu, "Optimum design of hybrid renewable energy systems: Overview of different approaches," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 3, pp. 1412–1425, 2012.
- [30] M. Fadaee and M. A. M. Radzi, "Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3364–3369, 2012.
- [31] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [32] M. J. Mayer, A. Szilágyi, and G. Gróf, "Environmental and economic multi-objective optimization of a household level hybrid renewable energy system by genetic algorithm," *Applied Energy*, vol. 269, p. 115058, 2020.
- [33] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, vol. 4. IEEE, 1995, pp. 1942–1948.
- [34] A. Maleki, M. Ameri, and F. Keynia, "Scrutiny of multifarious particle swarm optimization for finding the optimal size of a PV/wind/battery hybrid system," *Renewable Energy*, vol. 80, pp. 552–563, 2015.

- [35] H. Yu, "Reliability-based design optimization of structures: methodologies and applications to vibration control," Ph.D. dissertation, Ecully, Ecole centrale de Lyon, 2011.
- [36] G. Mavromatidis, K. Orehounig, and J. Carmeliet, "A review of uncertainty characterisation approaches for the optimal design of distributed energy systems," *Renewable and Sustainable Energy Reviews*, vol. 88, pp. 258–277, 2018.
- [37] R. Nadimi and K. Tokimatsu, "Analyzing of renewable and non-renewable energy consumption via Bayesian inference," *Energy Procedia*, vol. 142, pp. 2773–2778, 2017.
- [38] Z. Kang, "Robust design optimization of structures under uncertainties," 2005.
- [39] J. A. Carta, P. Ramirez, and S. Velazquez, "A review of wind speed probability distributions used in wind energy analysis: Case studies in the canary islands," *Renewable and sustainable energy reviews*, vol. 13, no. 5, pp. 933–955, 2009.
- [40] A. N. Celik, "A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey," *Renewable energy*, vol. 29, no. 4, pp. 593–604, 2004.
- [41] Z. Zhou, J. Zhang, P. Liu, Z. Li, M. C. Georgiadis, and E. N. Pistikopoulos, "A two-stage stochastic programming model for the optimal design of distributed energy systems," *Applied Energy*, vol. 103, pp. 135–144, 2013.
- [42] F. Kaminsky, R. Kirchhoff, C. Syu, and J. Manwell, "A comparison of alternative approaches for the synthetic generation of a wind speed time series," 1991.
- [43] L. Hontoria, J. Aguilera, J. Riesco, and P. Zufiria, "Recurrent neural supervised models for generating solar radiation synthetic series," *Journal of Intelligent and Robotic Systems*, vol. 31, no. 1, pp. 201–221, 2001.
- [44] A. Malkawi and G. Augenbroe, *Advanced building simulation*. Routledge, 2004.
- [45] S. Mirkhani and Y. Saboohi, "Stochastic modeling of the energy supply system with uncertain fuel price—A case of emerging technologies for distributed power generation," *Applied energy*, vol. 93, pp. 668–674, 2012.
- [46] A. S. Siddiqui and C. Marnay, "Operation of distributed generation under stochastic prices," *INFORMS Annual Meeting*, 2005.
- [47] S. Burhenne, O. Tsvetkova, D. Jacob, G. P. Henze, and A. Wagner, "Uncertainty quantification for combined building performance and cost-benefit analyses," *Building and Environment*, vol. 62, pp. 143–154, 2013.
- [48] N. N. Taleb, *The black swan: The impact of the highly improbable*. Random house, 2007, vol. 2.
- [49] A. Maheri, "Multi-objective design optimisation of standalone hybrid wind-PV-diesel systems under uncertainties," *Renewable Energy*, vol. 66, pp. 650–661, 2014.
- [50] S. Fazlollahi, "Decomposition optimization strategy for the design and operation of district energy systems," EPFL, Tech. Rep., 2014.
- [51] F. Bustos, A. Toledo, J. Contreras, and A. Fuentes, "Sensitivity analysis of a photovoltaic solar plant in Chile," *Renewable Energy*, vol. 87, pp. 145–153, 2016.
- [52] B. Liu and Y. Wang, "Energy system optimization under uncertainties: A comprehensive review," *Towards Sustainable Chemical Processes*, pp. 149–170, 2020.
- [53] G. B. Dantzig, "Linear programming under uncertainty," *Management science*, vol. 1, no. 3-4, pp. 197–206, 1955.
- [54] A. Zakaria, F. B. Ismail, M. H. Lipu, and M. A. Hannan, "Uncertainty models for stochastic optimization in renewable energy applications," *Renewable Energy*, vol. 145, pp. 1543–1571, 2020.

## Bibliography

---

- [55] A. Soroudi and T. Amraee, "Decision making under uncertainty in energy systems: State of the art," *Renewable and Sustainable Energy Reviews*, vol. 28, pp. 376–384, 2013.
- [56] A. L. Soyster, "Convex programming with set-inclusive constraints and applications to inexact linear programming," *Operations research*, vol. 21, no. 5, pp. 1154–1157, 1973.
- [57] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton university press, 2009.
- [58] D. Bertsimas and M. Sim, "The price of robustness," *Operations research*, vol. 52, no. 1, pp. 35–53, 2004.
- [59] F. S. Gazijahani and J. Salehi, "Reliability constrained two-stage optimization of multiple renewable-based microgrids incorporating critical energy peak pricing demand response program using robust optimization approach," *Energy*, vol. 161, pp. 999–1015, 2018.
- [60] A. Parisio, C. Del Vecchio, and A. Vaccaro, "A robust optimization approach to energy hub management," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 98–104, 2012.
- [61] K. Akbari, M. M. Nasiri, F. Jolai, and S. F. Ghaderi, "Optimal investment and unit sizing of distributed energy systems under uncertainty: A robust optimization approach," *Energy and Buildings*, vol. 85, pp. 275–286, 2014.
- [62] E. Kuznetsova, C. Ruiz, Y.-F. Li, and E. Zio, "Analysis of robust optimization for decentralized microgrid energy management under uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 64, no. September 2015, pp. 815–832, 2015.
- [63] S. Moret, "Strategic energy planning under uncertainty," PhD thesis, EPFL, Tech. Rep., 2017.
- [64] S. H. Lee, W. Chen, and B. M. Kwak, "Robust design with arbitrary distributions using Gauss-type quadrature formula," *Structural and Multidisciplinary Optimization*, vol. 39, no. 3, pp. 227–243, 2009.
- [65] N. D. Lagaros, V. Plevris, and M. Papadrakakis, "Reliability based robust design optimization of steel structures," *International Journal for Simulation and Multidisciplinary Design Optimization*, vol. 1, no. 1, pp. 19–29, 2007.
- [66] V. Penadés-Plà, T. Garcia-Segura, and V. Yepes, "Robust design optimization for low-cost concrete box-girder bridge," *Mathematics*, vol. 8, no. 3, p. 398, 2020.
- [67] G. Lei, J. Zhu, Y. Guo, C. Liu, and B. Ma, "A review of design optimization methods for electrical machines," *Energies*, vol. 10, no. 12, p. 1962, 2017.
- [68] D. Lee, L. F. Gonzalez, J. Periaux, and K. Srinivas, "Efficient hybrid-game strategies coupled to evolutionary algorithms for robust multidisciplinary design optimization in aerospace engineering," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 2, pp. 133–150, 2011.
- [69] T. Ghisu, J. P. Jarrett, and G. T. Parks, "Robust design optimization of airfoils with respect to ice accretion," *Journal of Aircraft*, vol. 48, no. 1, pp. 287–304, 2011.
- [70] T. Ghisu, G. T. Parks, J. P. Jarrett, and P. J. Clarkson, "Robust design optimization of gas turbine compression systems," *Journal of Propulsion and power*, vol. 27, no. 2, pp. 282–295, 2011.
- [71] N. Bilel, N. Mohamed, A. Zouhaier, and R. Lotfi, "Multi-objective robust design optimization of a mechatronic system with uncertain parameters, using a polynomial chaos expansion method," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 231, no. 9, pp. 729–739, 2017.
- [72] J. Fang, Y. Gao, G. Sun, C. Xu, and Q. Li, "Multiobjective robust design optimization of fatigue life for a truck cab," *Reliability Engineering & System Safety*, vol. 135, pp. 1–8, 2015.
- [73] W. Tian, Y. Heo, P. De Wilde, Z. Li, D. Yan, C. S. Park, X. Feng, and G. Augenbroe, "A review of uncertainty analysis in building energy assessment," *Renewable and Sustainable Energy Reviews*, vol. 93, pp. 285–301, 2018.

- [74] J. Keirstead, M. Jennings, and A. Sivakumar, "A review of urban energy system models: Approaches, challenges and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3847–3866, 2012.
- [75] G. Limpens, H. Jeanmart, and F. Maréchal, "Belgian Energy Transition: What Are the Options?" *Energies*, vol. 13, no. 1, p. 261, 2020.
- [76] N. Lüthen, S. Marelli, and B. Sudret, "Sparse polynomial chaos expansions: Literature survey and benchmark," *SIAM/ASA Journal on Uncertainty Quantification*, vol. 9, no. 2, pp. 593–649, 2021.
- [77] D. Coppitters, "Robust design optimization of renewable Hydrogen and derived energy carrier systems (RHEIA)," Read the Docs, 2021, Available online: <https://rheia.readthedocs.io/en/latest/>. Accessed: 16th May 2021.
- [78] D. Coppitters, W. De Paepe, and F. Contino, "Surrogate-assisted robust design optimization and global sensitivity analysis of a directly coupled photovoltaic-electrolyzer system under techno-economic uncertainty," *Applied Energy*, vol. 248, pp. 310–320, 2019.
- [79] G. I. Schuëller and H. A. Jensen, "Computational methods in optimization considering uncertainties—an overview," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 1, pp. 2–13, 2008.
- [80] X. Du and W. Chen, "Towards a better understanding of modeling feasibility robustness in engineering design," *J. Mech. Des.*, vol. 122, no. 4, pp. 385–394, 2000.
- [81] M. Diez and D. Peri, "Robust optimization for ship conceptual design," *Ocean Engineering*, vol. 37, no. 11-12, pp. 966–977, 2010.
- [82] Y. Kanno, "On three concepts in robust design optimization: absolute robustness, relative robustness, and less variance," *Structural and Multidisciplinary Optimization*, pp. 1–22, 2020.
- [83] S. Giorgetti, D. Coppitters, F. Contino, W. D. Paepe, L. Bricteux, G. Aversano, and A. Parente, "Surrogate-Assisted Modeling and Robust Optimization of a Micro Gas Turbine Plant With Carbon Capture," *Journal of Engineering for Gas Turbines and Power*, vol. 142, no. 1, 2020.
- [84] K. Verleysen, D. Coppitters, A. Parente, W. De Paepe, and F. Contino, "How can power-to-ammonia be robust? Optimization of an ammonia synthesis plant powered by a wind turbine considering operational uncertainties," *Fuel*, vol. 266, p. 117049, 2020.
- [85] W. De Paepe, D. Coppitters, S. Abraham, P. Tsirikoglou, G. Ghorbaniasl, and F. Contino, "Robust Operational Optimization of a Typical micro Gas Turbine," *Energy Procedia*, vol. 158, pp. 5795–5803, feb 2019.
- [86] A. Der Kiureghian and O. Ditlevsen, "Aleatory or epistemic? Does it matter?" *Structural safety*, vol. 31, no. 2, pp. 105–112, 2009.
- [87] S. Moret, V. C. Gironès, M. Bierlaire, and F. Maréchal, "Characterization of input uncertainties in strategic energy planning models," *Applied energy*, vol. 202, pp. 597–617, 2017.
- [88] A. Buttler and H. Spliethoff, "Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 2440–2454, 2018.
- [89] Elia, "Electricity scenarios for Belgium towards 2050: Elia's quantified study on the energy transition in 2030 and 2040," Elia, Tech. Rep., 2017.
- [90] S. Wilcox and W. Marion, "Users manual for TMY3 data sets," National Renewable Energy Laboratory Golden, CO, Tech. Rep., 2008.
- [91] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo method*. John Wiley & Sons, 2016, vol. 10.

## Bibliography

---

- [92] P. Glasserman, *Monte Carlo methods in financial engineering*. Springer Science & Business Media, 2013, vol. 53.
- [93] S. Raychaudhuri, "Introduction to monte carlo simulation," in *2008 Winter simulation conference*. IEEE, 2008, pp. 91–100.
- [94] B. Sudret, "Polynomial chaos expansions and stochastic finite-element methods," *Risk and Reliability in Geotechnical Engineering*, vol. Chap. 6, no. 2003, pp. 265–300, 2014.
- [95] M. Stein, "Large sample properties of simulations using latin hypercube sampling," *Technometrics*, vol. 29, no. 2, pp. 143–151, 1987.
- [96] P. Bratley and B. L. Fox, "Algorithm 659: Implementing Sobol's quasirandom sequence generator," *ACM Transactions on Mathematical Software (TOMS)*, vol. 14, no. 1, pp. 88–100, 1988.
- [97] P. W. Glynn and D. L. Iglehart, "Importance sampling for stochastic simulations," *Management science*, vol. 35, no. 11, pp. 1367–1392, 1989.
- [98] B. Sudret, S. Marelli, and J. Wiart, "Surrogate models for uncertainty quantification: An overview," in *2017 11th European conference on antennas and propagation (EUCAP)*. IEEE, 2017, pp. 793–797.
- [99] P. Tsirikoglou, K. Kyprianidis, A. I. Kalfas, and F. Contino, "Optimization in probabilistic domains: an engineering approach," in *Nature-Inspired Computation and Swarm Intelligence*. Elsevier, 2020, pp. 391–414.
- [100] I. Kaymaz, "Application of kriging method to structural reliability problems," *Structural Safety*, vol. 27, no. 2, pp. 133–151, 2005.
- [101] Z. Deng, X. Hu, X. Lin, Y. Che, L. Xu, and W. Guo, "Data-driven state of charge estimation for lithium-ion battery packs based on Gaussian process regression," *Energy*, vol. 205, p. 118000, 2020.
- [102] B. Richard, C. Cremona, and L. Adelaide, "A response surface method based on support vector machines trained with an adaptive experimental design," *Structural Safety*, vol. 39, pp. 14–21, 2012.
- [103] H. Rabitz and Ö. F. Aliş, "General foundations of high-dimensional model representations," *Journal of Mathematical Chemistry*, vol. 25, no. 2, pp. 197–233, 1999.
- [104] D. Xiu and G. E. Karniadakis, "The Wiener–Askey polynomial chaos for stochastic differential equations," *SIAM journal on scientific computing*, vol. 24, no. 2, pp. 619–644, 2002.
- [105] R. Jin, X. Du, and W. Chen, "The use of metamodeling techniques for optimization under uncertainty," *Structural and Multidisciplinary Optimization*, vol. 25, no. 2, pp. 99–116, 2003.
- [106] G. G. Wang and S. Shan, "Review of metamodeling techniques in support of engineering design optimization," *Journal of Mechanical Design*, vol. 129, pp. 370–380, 2007.
- [107] T. Chatterjee, S. Chakraborty, and R. Chowdhury, "A Critical Review of Surrogate Assisted Robust Design Optimization," *Archives of Computational Methods in Engineering*, pp. 1–30, 2017.
- [108] T. Chatterjee, R. Chowdhury, and P. Ramu, "Decoupling uncertainty quantification from robust design optimization," *Structural and Multidisciplinary Optimization*, vol. 59, no. 6, pp. 1969–1990, 2019.
- [109] P. Tsirikoglou, S. Abraham, F. Contino, Ö. Bağcı, J. Vierendeels, and G. Ghorbaniasl, "Comparison of metaheuristics algorithms on robust design optimization of a plain-fin-tube heat exchanger," in *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2017, p. 3827.
- [110] K. Deb and H. Gupta, "Introducing robustness in multi-objective optimization," *Evolutionary Computation*, vol. 14, no. 4, pp. 463–494, 2006.
- [111] A. Gaspar-Cunha and J. A. Covas, "Robustness in multi-objective optimization using evolutionary algorithms," *Computational Optimization and Applications*, vol. 39, no. 1, pp. 75–96, 2008.

- [112] A. Saha and T. Ray, "Practical robust design optimization using evolutionary algorithms," *Journal of Mechanical Design*, vol. 133, no. 10, 2011.
- [113] U. Lee, S. Park, and I. Lee, "Robust design optimization (RDO) of thermoelectric generator system using non-dominated sorting genetic algorithm II (NSGA-II)," *Energy*, vol. 196, p. 117090, 2020.
- [114] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary computation*, vol. 2, no. 3, pp. 221–248, 1994.
- [115] X.-S. Yang, *Nature-inspired optimization algorithms*. Academic Press, 2020.
- [116] S. N. Ghoreishi, A. Clausen, and B. N. Joergensen, "Termination Criteria in Evolutionary Algorithms: A Survey." in *IJCCI*, 2017, pp. 373–384.
- [117] M. Dodson and G. T. Parks, "Robust aerodynamic design optimization using polynomial chaos," *Journal of Aircraft*, vol. 46, no. 2, pp. 635–646, 2009.
- [118] F. Xiong, B. Xue, Z. Yan, and S. Yang, "Polynomial chaos expansion based robust design optimization," in *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*. IEEE, 2011, pp. 868–873.
- [119] S. Chakraborty, T. Chatterjee, R. Chowdhury, and S. Adhikari, "A surrogate based multi-fidelity approach for robust design optimization," *Applied Mathematical Modelling*, vol. 47, pp. 726–744, 2017.
- [120] Y. He, M. Razi, C. Forestiere, L. Dal Negro, and R. M. Kirby, "Uncertainty quantification guided robust design for nanoparticles' morphology," *Computer Methods in Applied Mechanics and Engineering*, vol. 336, pp. 578–593, 2018.
- [121] P. Turati, N. Pedroni, and E. Zio, "Simulation-based exploration of high-dimensional system models for identifying unexpected events," *Reliability Engineering and System Safety*, vol. 165, pp. 317–330, 2017.
- [122] H. Lund, E. Münster, and L. H. Tambjerg, "EnergyPLAN," *Computer model for energy system analysis version*, vol. 6, 2015.
- [123] M. Aneke and M. Wang, "Energy storage technologies and real life applications—A state of the art review," *Applied Energy*, vol. 179, pp. 350–377, 2016.
- [124] A. Mohammadi and M. Mehrpooya, "A comprehensive review on coupling different types of electrolyzer to renewable energy sources," *Energy*, vol. 158, pp. 632–655, 2018.
- [125] A. Khalilnejad, A. Sundararajan, and A. I. Sarwat, "Optimal design of hybrid wind/photovoltaic electrolyzer for maximum hydrogen production using imperialist competitive algorithm," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 1, pp. 40–49, 2018.
- [126] A. Maroufmashat, F. Sayedin, and S. S. Khavas, "An imperialist competitive algorithm approach for multi-objective optimization of direct coupling photovoltaic-electrolyzer systems," *International Journal of Hydrogen Energy*, vol. 39, no. 33, pp. 18743–18757, 2014.
- [127] R. García-Valverde, N. Espinosa, and A. Urbina, "Optimized method for photovoltaic-water electrolyser direct coupling," *International Journal Of Hydrogen Energy*, vol. 36, no. 17, pp. 10574–10586, 2011.
- [128] B. A. Pinaud, J. D. Benck, L. C. Seitz, A. J. Forman, Z. Chen, T. G. Deutsch, B. D. James, K. N. Baum, G. N. Baum, S. Ardo, H. Wang, E. Miller, and T. F. Jaramillo, "Technical and economic feasibility of centralized facilities for solar hydrogen production via photocatalysis and photoelectrochemistry," *Energy and Environmental Science*, vol. 6, no. 7, pp. 1983–2002, 2013.
- [129] F. Sayedin, A. Maroufmashat, S. Sattari, A. Elkamel, and M. Fowler, "Optimization of Photovoltaic Electrolyzer Hybrid systems; Taking into account the effect of climate conditions," *Energy Conversion and Management*, vol. 118, pp. 438–449, 2016.
- [130] B. Zakeri and S. Syri, "Electrical energy storage systems: A comparative life cycle cost analysis," *Renewable and sustainable energy reviews*, vol. 42, pp. 569–596, 2015.

## Bibliography

---

- [131] A. Al-Karaghoul and L. L. Kazmerski, "Energy consumption and water production cost of conventional and renewable-energy-powered desalination processes," *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 343–356, 2013.
- [132] B. Olateju and A. Kumar, "A techno-economic assessment of hydrogen production from hydropower in Western Canada for the upgrading of bitumen from oil sands," *Energy*, vol. 115, pp. 604–614, 2016.
- [133] *MSX-60 and MSX-64 Photovoltaic Modules*, Solarex, 1997.
- [134] F. M. González-Longatt, "Model of Photovoltaic Module in Matlab," in *2DO Congreso Iberoamericano de estudiantes de ingeniería eléctrica, electrónica y computación (II CIBELEC 2005)*, 2005, pp. 1–5.
- [135] D. Bonkougou, Z. Koalaga, and D. Njomo, "Modelling and simulation of photovoltaic module considering single-diode equivalent circuit model in matlab," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 3, pp. 493–502, 2013.
- [136] Q. Kou, S. A. Klein, and W. A. Beckman, "A method for estimating the long-term performance of direct-coupled PV pumping systems," *Solar Energy*, vol. 64, no. 1-3, pp. 33–40, 1998.
- [137] R. García-Valverde, N. Espinosa, and A. Urbina, "Simple PEM water electrolyser model and experimental validation," *International Journal of Hydrogen Energy*, vol. 37, no. 2, pp. 1927–1938, 2012.
- [138] B. Paul, "Direct-coupling of the photovoltaic array and PEM electrolyser in solar-hydrogen systems for remote area power supply," Ph.D. dissertation, RMIT University, 2009.
- [139] Meteotest, "Handbook part II: Theory," *Global Meteorological Database Version 7. Software and Data for Engineers, Planers and Education*, 2017.
- [140] A. Tugirumubano, L. K. Kwac, M. S. Lee, J. H. Lee, J. J. Lee, S. A. Jeon, H. N. Jeong, and H. G. Kim, "Analysis of the parametric effect on the performance of a polymer electrolyte membrane electrolyzer," *International Journal of Mechanical Engineering and Robotics Research*, vol. 6, no. 1, pp. 1–5, 2017.
- [141] A. H. Fathima and K. Palanisamy, "Optimization in microgrids with hybrid energy systems - A review," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 431–446, 2015.
- [142] L. Reichenberg, F. Hedenus, M. Odenberger, and F. Johnsson, "The marginal system LCOE of variable renewables – Evaluating high penetration levels of wind and solar in Europe," *Energy*, vol. 152, pp. 914–924, 2018.
- [143] K. Branker, M. J. Pathak, and J. M. Pearce, "A review of solar photovoltaic levelized cost of electricity," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4470–4482, 2011.
- [144] K. Rajeshwar, R. D. McConnell, and S. Licht, *Solar Hydrogen Generation: Toward a Renewable Energy Future*. Springer, 2008.
- [145] C. A. Rodriguez, M. A. Modestino, D. Psaltis, and C. Moser, "Design and cost considerations for practical solar-hydrogen generators," *Energy Environ. Sci.*, vol. 7, no. 12, pp. 3828–3835, 2014.
- [146] F. Segura, V. Bartolucci, and J. M. Andújar, "Hardware/software data acquisition system for real time cell temperature monitoring in air-cooled polymer electrolyte fuel cells," *Sensors (Switzerland)*, vol. 17, no. 7, p. 1600, 2017.
- [147] F. Sayedin, A. Maroufmashat, R. Roshandel, and S. S. Khavas, "Optimal design and operation of a photovoltaic-electrolyser system using particle swarm optimisation," *International Journal of Sustainable Energy*, vol. 35, no. 6, pp. 566–582, 2014.
- [148] D. Coppitters, W. De Paepe, and F. Contino, "Robust design optimization and stochastic performance analysis of a grid-connected photovoltaic system with battery storage and hydrogen storage," *Energy*, vol. 213, p. 118798, 2020.

- [149] A. A. Bazmi and G. Zahedi, "Sustainable energy systems: Role of optimization modeling techniques in power generation and supply—A review," *Renewable and sustainable energy reviews*, vol. 15, no. 8, pp. 3480–3500, 2011.
- [150] D. C. Montgomery, *Design and analysis of experiments*. John Wiley & Sons, 2014.
- [151] G. Blatman, "Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis," Ph.D. dissertation, Clermont-Ferrand 2, 2009.
- [152] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [153] G. Blatman and B. Sudret, "Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach," *Comptes Rendus Mécanique*, vol. 336, no. 6, pp. 518–523, 2008.
- [154] —, "An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis," *Probabilistic Engineering Mechanics*, vol. 25, no. 2, pp. 183–197, 2010.
- [155] —, "Adaptive sparse polynomial chaos expansion based on least angle regression," *Journal of Computational Physics*, vol. 230, no. 6, pp. 2345–2367, 2011.
- [156] S. Abraham, M. Raisee, G. Ghorbaniasl, F. Contino, and C. Lacor, "A robust and efficient stepwise regression method for building sparse polynomial chaos expansions," *Journal of Computational Physics*, vol. 332, pp. 461–474, 2017.
- [157] Y. Zhang, P. E. Campana, A. Lundblad, and J. Yan, "Comparative study of hydrogen storage and battery storage in grid connected photovoltaic system: Storage sizing and rule-based operation," *Applied Energy*, vol. 201, pp. 397–411, sep 2017.
- [158] T. Taner, S. A. H. Naqvi, and M. Ozkaymak, "Techno-economic analysis of a more efficient hydrogen generation system prototype: a case study of PEM electrolyzer with Cr-C coated SS304 bipolar plates," *Fuel Cells*, vol. 19, pp. 19–26, 2019.
- [159] T. Taner, "Energy and exergy analyze of PEM fuel cell: a case study of modeling and simulations," *Energy*, vol. 143, pp. 284–294, 2018.
- [160] W. Saeed and G. Warkozek, "Modeling and analysis of renewable PEM fuel cell system," *Energy Procedia*, vol. 74, pp. 87–101, 2015.
- [161] R. Zahedi and M. M. Ardehali, "Power management for storage mechanisms including battery, supercapacitor, and hydrogen of autonomous hybrid green power system utilizing multiple optimally-designed fuzzy logic controllers," *Energy*, vol. 204, p. 117935, 2020.
- [162] Y. Han, H. Yang, Q. Li, W. Chen, F. Zare, and J. M. Guerrero, "Mode-triggered droop method for the decentralized energy management of an islanded hybrid PV/hydrogen/battery DC microgrid," *Energy*, vol. 199, p. 117441, 2020.
- [163] A. Khiareddine, C. B. Salah, D. Rekioua, and M. F. Mimouni, "Sizing methodology for hybrid photovoltaic/wind/hydrogen/battery integrated to energy management strategy for pumping system," *Energy*, vol. 153, pp. 743–762, 2018.
- [164] W. Zhang, A. Maleki, M. A. Rosen, and J. Liu, "Optimization with a simulated annealing algorithm of a hybrid system for renewable energy including battery and hydrogen storage," *Energy*, vol. 163, pp. 191–207, 2018.
- [165] D. Parra, G. S. Walker, and M. Gillott, "Modeling of PV generation, battery and hydrogen storage to investigate the benefits of energy storage for single dwelling," *Sustainable Cities and Society*, vol. 10, pp. 1–10, 2014.
- [166] M. A. Pellow, C. J. Emmott, C. J. Barnhart, and S. M. Benson, "Hydrogen or batteries for grid storage? A net energy analysis," *Energy & Environmental Science*, vol. 8, no. 7, pp. 1938–1952, 2015.

## Bibliography

---

- [167] M. Yue, H. Lambert, E. Pahon, R. Roche, S. Jemei, and D. Hissel, "Hydrogen energy systems: A critical review of technologies, applications, trends and challenges," *Renewable and Sustainable Energy Reviews*, vol. 146, p. 111180, 2021.
- [168] M. Castañeda, A. Cano, F. Jurado, H. Sánchez, and L. M. Fernández, "Sizing optimization, dynamic modeling and energy management strategies of a stand-alone PV/hydrogen/battery-based hybrid system," *International Journal of Hydrogen Energy*, vol. 38, no. 10, pp. 3830–3845, 2013.
- [169] W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, "pvlib python: a python package for modeling solar energy systems," *The Journal of Open Source Software*, vol. 3, p. 884, 2018.
- [170] T. Gurupira and A. Rix, "Pv simulation software comparisons: Pvsyst, nrel sam and pvlib," in *Conf.: SAUPEC*, 2017.
- [171] W. De Soto, S. A. Klein, and W. A. Beckman, "Improvement and validation of a model for photovoltaic array performance," *Solar Energy*, vol. 80, no. 1, pp. 78–88, 2006.
- [172] G. de Oliveira e Silva and P. Hendrick, "Lead–acid batteries coupled with photovoltaics for increased electricity self-sufficiency in households," *Applied Energy*, vol. 178, pp. 856–867, sep 2016.
- [173] S. Blaifi, S. Moulahoum, I. Colak, and W. Merrouche, "An enhanced dynamic model of battery using genetic algorithm suitable for photovoltaic applications," *Applied Energy*, vol. 169, pp. 888–898, may 2016.
- [174] B. Battke, T. S. Schmidt, D. Grosspietsch, and V. H. Hoffmann, "A review and probabilistic model of lifecycle costs of stationary batteries in multiple applications," *Renewable and Sustainable Energy Reviews*, vol. 25, pp. 240–250, sep 2013.
- [175] G. Díaz, J. Gómez-Aleixandre, J. Coto, and O. Conejero, "Maximum income resulting from energy arbitrage by battery systems subject to cycle aging and price uncertainty from a dynamic programming perspective," *Energy*, vol. 156, pp. 647–660, 2018.
- [176] C. Phurailatpam, B. S. Rajpurohit, and L. Wang, "Planning and optimization of autonomous DC microgrids for rural and urban applications in India," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 194–204, feb 2018.
- [177] I. Pawel, "The cost of storage - How to calculate the levelized cost of stored energy (LCOE) and applications to renewable energy generation," *Energy Procedia*, vol. 46, pp. 68–77, 2014.
- [178] P. Komor and J. Glassmire, "Electricity storage and renewables for island power: a guide for decision makers," International Renewable Energy Agency (IRENA), Tech. Rep., 2012.
- [179] Z. Abdin, C. J. Webb, and E. M. Gray, "Modelling and simulation of a proton exchange membrane (PEM) electrolyser cell," *International Journal of Hydrogen Energy*, vol. 40, no. 39, pp. 13 243–13 257, oct 2015.
- [180] A. Al-Sharafī, A. Z. Sahin, T. Ayar, and B. S. Yilbas, "Techno-economic analysis and optimization of solar and wind energy systems for power generation and hydrogen production in Saudi Arabia," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 33–49, 2017.
- [181] H. Görgün, "Dynamic modelling of a proton exchange membrane (PEM) electrolyzer," *International Journal of Hydrogen Energy*, vol. 31, no. 1, pp. 29–38, 2006.
- [182] K. Murugesan and V. Senniappan, "Investigation of water management dynamics on the performance of a Ballard-Mark-V proton exchange membrane fuel cell stack system," *Int. J. Electrochem. Sci*, vol. 8, pp. 7885–7904, 2013.
- [183] K. Ameer, A. Hadjaissa, M. Ait Cheikh, A. Cheknane, and N. Essounbouli, "Fuzzy energy management of hybrid renewable power system with the aim to extend component lifetime," *International Journal of Energy Research*, vol. 41, pp. 1867–1879, 2017.

- [184] M. Taghvaei, M. Radzi, S. Moosavain, H. Hizam, and M. H. Marhaban, "A current and future study on non-isolated DC–DC converters for photovoltaic applications," *Renewable and sustainable energy reviews*, vol. 17, pp. 216–227, 2013.
- [185] G. A. Rampinelli, A. Krenzinger, and F. Chenlo Romero, "Mathematical models for efficiency of inverters used in grid connected photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 578–587, 2014.
- [186] Open Energy Information, "Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States," Available online: <http://en.openei.org/datasets/dataset/commercial-and-residential-hourly-load-profiles-for-all-tmy3-locations-in-the-united-states>. Accessed: 16th June 2020.
- [187] M. Montero Carrero, I. R. Sánchez, W. D. Paepe, A. Parente, and F. Contino, "Is There a Future for Small-Scale Cogeneration in Europe ? Economic and Policy Analysis of the Internal Humid Air Turbine Cycles," *Energies*, vol. 12, no. 3, pp. 1–27, 2019.
- [188] VREG - Flemish Energy Regulator of the electricity and gas market, "Marktrapport 2019," Vlaamse Regulator van de Elektriciteits- en Gasmarkt, Tech. Rep., 2020, Available online: <https://www.vreg.be/sites/default/files/document/rapp-2020-10.pdf>. Accessed: 28th January 2021.
- [189] P. Balcombe, D. Rigby, and A. Azapagic, "Investigating the importance of motivations and barriers related to microgeneration uptake in the UK," *Applied Energy*, vol. 130, pp. 403–418, 2014.
- [190] N. Lukač, S. Seme, K. Dežan, B. Žalik, and G. Štumberger, "Economic and environmental assessment of rooftops regarding suitability for photovoltaic systems installation based on remote sensing data," *Energy*, vol. 107, pp. 854–865, jul 2016.
- [191] Sunpower, *X-Series residential solar panels: supplementary technical specifications*.
- [192] B. Guinot, B. Champel, F. Montignac, E. Lemaire, D. Vannucci, S. Sailler, and Y. Bultel, "Techno-economic study of a PV-hydrogen-battery hybrid system for off-grid power supply: Impact of performances' ageing on optimal system sizing and competitiveness," *International Journal of Hydrogen Energy*, vol. 40, no. 1, pp. 623–632, jan 2015.
- [193] W. G. Colella, B. D. James, J. M. Moton, G. Saur, and T. Ramsden, "Techno-economic Analysis of PEM Electrolysis for Hydrogen Production," Strategic Analysis inc. and National Renewable Energy Laboratory (NREL), Tech. Rep., 2014.
- [194] L. Placca, R. Kouta, J. F. Blachot, and W. Charon, "Effects of temperature uncertainty on the performance of a degrading PEM fuel cell model," *Journal of Power Sources*, vol. 194, no. 1, pp. 313–327, 2009.
- [195] T. Taner, "Energy and exergy analyze of PEM fuel cell: A case study of modeling and simulations," *Energy*, vol. 143, pp. 284–294, jan 2018.
- [196] S. Niaz, T. Manzoor, and A. H. Pandith, "Hydrogen storage: Materials, methods and perspectives," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 457–469, 2015.
- [197] Y. Budak and Y. Devrim, "Comparative study of PV/PEM fuel cell hybrid energy system based on methanol and water electrolysis," *Energy conversion and management*, vol. 179, pp. 46–57, 2019.
- [198] E. Ozden and I. Tari, "PEM fuel cell degradation effects on the performance of a stand-alone solar energy system," *International Journal of Hydrogen Energy*, vol. 42, pp. 13 217–13 225, 2017.
- [199] J. Kotowicz, D. Wecel, and M. Jurczyk, "Analysis of component operation in power-to-gas-to-power installations," *Applied energy*, vol. 216, pp. 45–59, 2018.
- [200] J. Wu, X. Z. Yuan, J. J. Martin, H. Wang, J. Zhang, J. Shen, S. Wu, and W. Merida, "A review of pem fuel cell durability: Degradation mechanisms and mitigation strategies," *Journal of Power Sources*, vol. 184, pp. 104–119, 2008.

## Bibliography

---

- [201] F. Abbasi and M. Abapour, "Reliability and cost optimization of multi-input bidirectional DC/DC converter implemented in renewable sources," *Majlesi Journal of Electrical Engineering*, vol. 12, pp. 57–65, 2018.
- [202] A. Allouhi, M. S. Buker, H. El-houari, A. Boharb, M. Benzakour Amine, T. Kousksou, and A. Jamil, "PV water pumping systems for domestic uses in remote areas: Sizing process, simulation and economic evaluation," *Renewable Energy*, vol. 132, pp. 798–812, mar 2019.
- [203] A. Kashefi Kaviani, G. H. Riahy, and S. M. Kouhsari, "Optimal design of a reliable hydrogen-based stand-alone wind/PV generating system, considering component outages," *Renewable Energy*, vol. 34, no. 11, pp. 2380–2390, 2009.
- [204] M. A. Ramli, A. Hiendro, and S. Twaha, "Economic analysis of PV/diesel hybrid system with flywheel energy storage," *Renewable Energy*, vol. 78, pp. 398–405, jun 2015.
- [205] S. M. Wilcox, "National solar radiation database 1991-2010 update: User's manual," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2012.
- [206] Vlaamse Regulator van de Elektriciteits- en Gasmarkt (Flemish Regulator of the Electricity and Gas Market), "Evolution of the household electricity bill," Available online: <https://infogram.com/evolutie-energiefactuur-gezin-1h7j4djlx9k94nr?live>. Accessed: 3th February 2020.
- [207] Triami Media BV, "Historic inflation Belgium - CPI inflation," 2019, Available online: [inflation.eu/inflation-rates/belgium/historic-inflation/cpi-inflation-belgium.aspx](http://inflation.eu/inflation-rates/belgium/historic-inflation/cpi-inflation-belgium.aspx). Accessed: 1st August 2020.
- [208] T. H. T. Nguyen, T. Nakayama, and M. Ishida, "Optimal capacity design of battery and hydrogen system for the DC grid with photovoltaic power generation based on the rapid estimation of grid dependency," *International Journal of Electrical Power & Energy Systems*, vol. 89, pp. 27–39, 2017.
- [209] A. Singh, P. Baredar, and B. Gupta, "Techno-economic feasibility analysis of hydrogen fuel cell and solar photovoltaic hybrid renewable energy system for academic research building," *Energy Conversion and Management*, vol. 145, pp. 398–414, 2017.
- [210] P. K. Rose and F. Neumann, "Hydrogen refueling station networks for heavy-duty vehicles in future power systems," *Transportation Research Part D: Transport and Environment*, vol. 83, p. 102358, 2020.
- [211] C. Wang, Z. Ye, Y. Yu, and W. Gong, "Estimation of bus emission models for different fuel types of buses under real conditions," *Science of the Total Environment*, vol. 640, pp. 965–972, 2018.
- [212] E. R. Grijalva and J. M. López Martínez, "Analysis of the Reduction of CO<sub>2</sub> Emissions in Urban Environments by Replacing Conventional City Buses by Electric Bus Fleets: Spain Case Study," *Energies*, vol. 12, no. 3, p. 525, 2019.
- [213] G. Zhang, J. Zhang, and T. Xie, "A solution to renewable hydrogen economy for fuel cell buses—a case study for zhangjiakou in north china," *International journal of hydrogen energy*, vol. 45, no. 29, pp. 14 603–14 613, 2020.
- [214] K. G. Logan, J. D. Nelson, and A. Hastings, "Electric and hydrogen buses: Shifting from conventionally fuelled cars in the UK," *Transportation Research Part D: Transport and Environment*, vol. 85, p. 102350, 2020.
- [215] G. Correa, P. Muñoz, T. Falaguerra, and C. Rodriguez, "Performance comparison of conventional, hybrid, hydrogen and electric urban buses using well to wheel analysis," *Energy*, vol. 141, pp. 537–549, 2017.
- [216] T. A. Gunawan, I. Williamson, D. Raine, and R. F. Monaghan, "Decarbonising city bus networks in ireland with renewable hydrogen," *International Journal of Hydrogen Energy*, 2020.
- [217] J. Burkhardt, A. Patyk, P. Tanguy, and C. Retzke, "Hydrogen mobility from wind energy—A life cycle assessment focusing on the fuel supply," *Applied energy*, vol. 181, pp. 54–64, 2016.
- [218] P. Spath and M. Mann, "Life Cycle Assessment of Renewable Hydrogen Production via Wind/Electrolysis: Milestone Completion Report," National Renewable Energy Lab., Golden, CO.(US), Tech. Rep., 2004.

- [219] co2 emissiefactoren, "Lijst emissiefactoren," 2020, Available online: <https://www.co2emissiefactoren.nl/lijest-emissiefactoren/>. Accessed: 4th February 2021.
- [220] M. Kayfeci, A. Keçebaş, and M. Bayat, "Hydrogen production," in *Solar Hydrogen Production*. Elsevier, 2019, pp. 45–83.
- [221] M. Minutillo, A. Perna, A. Forcina, S. Di Micco, and E. Jannelli, "Analyzing the levelized cost of hydrogen in refueling stations with on-site hydrogen production via water electrolysis in the Italian scenario," *International Journal of Hydrogen Energy*, 2020.
- [222] M. Rivarolo, O. Improta, L. Magistri, M. Panizza, and A. Barbucci, "Thermo-economic analysis of a hydrogen production system by sodium borohydride (NaBH<sub>4</sub>)," *International Journal of Hydrogen Energy*, vol. 43, no. 3, pp. 1606–1614, 2018.
- [223] M. Mohsin, A. Rasheed, and R. Saidur, "Economic viability and production capacity of wind generated renewable hydrogen," *International Journal of hydrogen energy*, vol. 43, no. 5, pp. 2621–2630, 2018.
- [224] T. Ayodele and J. Munda, "Potential and economic viability of green hydrogen production by water electrolysis using wind energy resources in South Africa," *International Journal of Hydrogen Energy*, vol. 44, no. 33, pp. 17 669–17 687, 2019.
- [225] Available online: <https://www.delijn.be/nl/overdelijn/organisatie/geschiedenis/>. Accessed: 16th February 2021.
- [226] B. Reuter, M. Faltenbacher, O. Schuller, and N. Whitehouse, "New Bus ReFuelling for European Hydrogen Bus Depots - Guidance Document on Large Scale Hydrogen Bus Refuelling," Fuel Cells and Hydrogen, Joint Undertaking, Tech. Rep., 2017.
- [227] B. E. carbon reduction in business, "Update on the digital meter and grid tariffing methodology in Flanders," Interreg 2 Seas Mers Zeeën - European Regional Development Fund, Tech. Rep., 2018.
- [228] "Renewables.ninja." [Online]. Available: <https://www.renewables.ninja/>
- [229] I. Staffell and S. Pfenninger, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy*, vol. 114, pp. 1224–1239, nov 2016.
- [230] N. Boccard, "Capacity factor of wind power realized values vs. estimates," *Energy Policy*, vol. 37, no. 7, pp. 2679–2688, 2009.
- [231] Monterey Gardiner, "Hydrogen and Fuel Cells Program Record 9013: Energy requirements for hydrogen gas compression and liquefaction as related to vehicle storage needs," Tech. Rep., 2019.
- [232] International Energy Agency, "Renewables 2019: Analysis and forecast to 2024," International Energy Agency, Tech. Rep., 2019.
- [233] T. F. Kristjansdottir, C. S. Good, M. R. Inman, R. D. Schlanbusch, and I. Andresen, "Embodied greenhouse gas emissions from pv systems in norwegian residential zero emission pilot buildings," *Solar Energy*, vol. 133, pp. 155–171, 2016.
- [234] R. H. Wiser and M. Bolinger, "2018 wind technologies market report," 2019.
- [235] E. A. Smoucha, K. Fitzpatrick, S. Buckingham, and O. G. Knox, "Life cycle analysis of the embodied carbon emissions from 14 wind turbines with rated powers between 50kW and 3.4 MW," *Journal of Fundamentals of Renewable Energy and Applications*, vol. 6, no. 4, p. 1000211, 2016.
- [236] X. Zhang, C. Bauer, C. L. Mutel, and K. Volkart, "Life Cycle Assessment of Power-to-Gas: Approaches, system variations and their environmental implications," *Applied Energy*, vol. 190, pp. 326–338, 2017.
- [237] A. Mayyas and M. Mann, "Manufacturing competitiveness analysis for hydrogen refueling stations," *International Journal of Hydrogen Energy*, vol. 44, no. 18, pp. 9121–9142, 2019.
- [238] International Stainless Steel Forum, "Stainless Steel and CO<sub>2</sub>: Facts and Scientific Observations," Tech. Rep., 2019.

## Bibliography

---

- [239] W. K. Biswas, V. Duong, P. Frey, and M. N. Islam, "A comparison of repaired, remanufactured and new compressors used in Western Australian small-and medium-sized enterprises in terms of global warming," *Journal of Remanufacturing*, vol. 3, no. 1, pp. 1–7, 2013.
- [240] Civitas, "Smart Choices for Cities—Clean Buses for Your City," *Police Note*, vol. 52, 2013.
- [241] Y. Ruf, M. Kaufmann, S. Lange, J. Pfister, F. Heieck, and A. Endres, "Fuel Cells and Hydrogen Applications for Regions and Cities," *Roland Berger: Brussels, Belgium*, 2017.
- [242] N. Pocard and C. Reid, "Fuel cell electric buses: An attractive value proposition for zero-emission buses in the united kingdom," *Balard: London, UK*, 2016.
- [243] Z. Liu, T. Li, Q. Jiang, and H. Zhang, "Life cycle assessment–based comparative evaluation of originally manufactured and remanufactured diesel engines," *Journal of Industrial Ecology*, vol. 18, no. 4, pp. 567–576, 2014.
- [244] R. Kawamoto, H. Mochizuki, Y. Moriguchi, T. Nakano, M. Motohashi, Y. Sakai, and A. Inaba, "Estimation of CO2 Emissions of internal combustion engine vehicle and battery electric vehicle using LCA," *Sustainability*, vol. 11, no. 9, p. 2690, 2019.
- [245] S. Dhanushkodi, N. Mahinpey, A. Srinivasan, and M. Wilson, "Life cycle analysis of fuel cell technology," *Journal of Environmental Informatics*, vol. 11, no. 1, pp. 36–44, 2008.
- [246] R. Fu, D. J. Feldman, and R. M. Margolis, "US solar photovoltaic system cost benchmark: Q1 2018," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2018.
- [247] H. C. Frey, N. M. Roupail, H. Zhai, T. L. Farias, and G. A. Gonçalves, "Comparing real-world fuel consumption for diesel-and hydrogen-fueled transit buses and implication for emissions," *Transportation Research Part D: Transport and Environment*, vol. 12, no. 4, pp. 281–291, 2007.
- [248] United Nations Framework Convention for Climate Change., "National emissions reported to the UNFCCC and to the EU Greenhouse Gas Monitoring Mechanism," 2020.
- [249] I. Duić, N. Štefanić, Z. Lulić, G. Krajačić, T. Pukšec, and T. Novosel, "EU28 fuel prices for 2015, 2030 and 2050. Deliverable 6.1: Future fuel price review," 2017.
- [250] T. Mayer, M. Semmel, M. A. G. Morales, K. M. Schmidt, A. Bauer, and J. Wind, "Techno-economic evaluation of hydrogen refueling stations with liquid or gaseous stored hydrogen," *International Journal of Hydrogen Energy*, vol. 44, no. 47, pp. 25 809–25 833, 2019.
- [251] R. P. Micena, O. R. Llerena-Pizarro, T. M. de Souza, and J. L. Silveira, "Solar-powered Hydrogen Refueling Stations: A techno-economic analysis," *International Journal of Hydrogen Energy*, vol. 45, no. 3, pp. 2308–2318, 2020.
- [252] I. Iordache, D. Schitea, and M. Iordache, "Hydrogen refueling station infrastructure roll-up, an indicative assessment of the commercial viability and profitability," *International Journal of Hydrogen Energy*, vol. 42, no. 8, pp. 4721–4732, 2017.
- [253] Ø. Ulleberg and R. Hancke, "Techno-economic calculations of small-scale hydrogen supply systems for zero emission transport in norway," *international journal of hydrogen energy*, vol. 45, no. 2, pp. 1201–1211, 2020.
- [254] W. H. A. Al-Zubaidi, P. Thongtanunam, H. K. Dam, C. Tantithamthavorn, and A. Ghose, "Workload-aware reviewer recommendation using a multi-objective search-based approach," in *Proceedings of the 16th ACM International Conference on Predictive Models and Data Analytics in Software Engineering*, 2020, pp. 21–30.
- [255] D. Coppitters, W. De Paepe, and F. Contino, "Robust design optimization of a photovoltaic-battery-heat pump system with thermal storage under aleatory and epistemic uncertainty," *Energy*, p. 120692, 2021.
- [256] T. Augustin, F. P. Coolen, G. De Cooman, and M. C. Troffaes, *Introduction to imprecise probabilities*. John Wiley & Sons, 2014.

- [257] I. Kozine and Y. V. Filimonov, "Imprecise reliabilities: experiences and advances," *Reliability Engineering & System Safety*, vol. 67, no. 1, pp. 75–83, 2000.
- [258] M. Beer, "Fuzzy Probability Theory." *Encyclopedia of complexity and systems science*, vol. 6, pp. 4047–4059, 2009.
- [259] S. Ferson, V. Kreinovich, L. Grinzburg, D. Myers, and K. Sentz, "Constructing probability boxes and Dempster-Shafer structures," Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), Tech. Rep., 2015.
- [260] X. Yang, Y. Liu, Y. Zhang, and Z. Yue, "Hybrid reliability analysis with both random and probability-box variables," *Acta Mechanica*, vol. 226, no. 5, pp. 1341–1357, 2015.
- [261] M. Beer, S. Ferson, and V. Kreinovich, "Imprecise probabilities in engineering analyses," *Mechanical systems and signal processing*, vol. 37, no. 1-2, pp. 4–29, 2013.
- [262] J. Zhang and M. D. Shields, "On the quantification and efficient propagation of imprecise probabilities resulting from small datasets," *Mechanical Systems and Signal Processing*, vol. 98, pp. 465–483, 2018.
- [263] M. S. Eldred and L. P. Swiler, "Efficient algorithms for mixed aleatory-epistemic uncertainty quantification with application to radiation-hardened electronics," *Part I: Algorithms and Benchmark Results, Sandia National Laboratories, SAND2009-5805, Albuquerque, NM, 2009.*
- [264] R. Schöbi and B. Sudret, "Global sensitivity analysis in the context of imprecise probabilities (p-boxes) using sparse polynomial chaos expansions," *Reliability Engineering & System Safety*, vol. 187, pp. 129–141, 2019.
- [265] R. Schöbi, "Surrogate models for uncertainty quantification in the context of imprecise probability modelling," Ph.D. dissertation, ETH Zurich, 2017.
- [266] T.-V. Hoang, B. Rosić, and H. G. Matthies, "Characterization and propagation of uncertainties associated with limited data using a hierarchical parametric probability box," *PAMM*, vol. 18, no. 1, p. e201800475, 2018.
- [267] K. Chowdhary and P. Dupuis, "Distinguishing and integrating aleatoric and epistemic variation in uncertainty quantification," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 47, no. 3, pp. 635–662, 2013.
- [268] X. Xie and R. Schenkendorf, "Robust process design in pharmaceutical manufacturing under batch-to-batch variation," *Processes*, vol. 7, no. 8, p. 509, 2019.
- [269] REN21, "Renewables 2020 global status report," REN21, Tech. Rep., 2020.
- [270] H. Lund, "Renewable heating strategies and their consequences for storage and grid infrastructures comparing a smart grid to a smart energy systems approach," *Energy*, vol. 151, pp. 94–102, 2018.
- [271] S. Murugan and B. Horák, "Tri and polygeneration systems-A review," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1032–1051, 2016.
- [272] G. Kyriakarakos, A. I. Dounis, S. Rozakis, K. G. Arvanitis, and G. Papadakis, "Polygeneration microgrids: a viable solution in remote areas for supplying power, potable water and hydrogen as transportation fuel," *Applied Energy*, vol. 88, no. 12, pp. 4517–4526, 2011.
- [273] M. A. V. Rad, R. Ghasempour, P. Rahdan, S. Mousavi, and M. Arastounia, "Techno-economic analysis of a hybrid power system based on the cost-effective hydrogen production method for rural electrification, a case study in Iran," *Energy*, vol. 190, p. 116421, 2020.
- [274] R. Luthander, J. Widén, D. Nilsson, and J. Palm, "Photovoltaic self-consumption in buildings: A review," *Applied energy*, vol. 142, pp. 80–94, 2015.
- [275] H. Lund, P. A. Østergaard, D. Connolly, and B. V. Mathiesen, "Smart energy and smart energy systems," *Energy*, vol. 137, pp. 556–565, 2017.

## Bibliography

---

- [276] H. Lund, P. A. Østergaard, M. Chang, S. Werner, S. Svendsen, P. Sorknæs, J. E. Thorsen, F. Hvelplund, B. O. G. Mortensen, B. V. Mathiesen *et al.*, “The status of 4th generation district heating: Research and results,” *Energy*, vol. 164, pp. 147–159, 2018.
- [277] H. Wolisz, T. Schütz, T. Blanke, M. Hagenkamp, M. Kohn, M. Wesseling, and D. Müller, “Cost optimal sizing of smart buildings’ energy system components considering changing end-consumer electricity markets,” *Energy*, vol. 137, pp. 715–728, 2017.
- [278] A.-L. Klingler, “The effect of electric vehicles and heat pumps on the market potential of PV+battery systems,” *Energy*, vol. 161, pp. 1064–1073, 2018.
- [279] G. Angenendt, S. Zurmühlen, J. Figgner, K.-P. Kairies, and D. U. Sauer, “Providing frequency control reserve with photovoltaic battery energy storage systems and power-to-heat coupling,” *Energy*, vol. 194, p. 116923, 2020.
- [280] G. Angenendt, S. Zurmühlen, F. Rücker, H. Axelsen, and D. U. Sauer, “Optimization and operation of integrated homes with photovoltaic battery energy storage systems and power-to-heat coupling,” *Energy Conversion and Management: X*, vol. 1, p. 100005, 2019.
- [281] B. M. Radhakrishnan and D. Srinivasan, “A multi-agent based distributed energy management scheme for smart grid applications,” *Energy*, vol. 103, pp. 192–204, 2016.
- [282] M. Barma, R. Saidur, S. Rahman, A. Allouhi, B. Akash, and S. M. Sait, “A review on boilers energy use, energy savings, and emissions reductions,” *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 970–983, 2017.
- [283] M. Liu, S. Wang, Y. Zhao, H. Tang, and J. Yan, “Heat-power decoupling technologies for coal-fired CHP plants: Operation flexibility and thermodynamic performance,” *Energy*, vol. 188, p. 116074, 2019.
- [284] D. Fischer and H. Madani, “On heat pumps in smart grids: A review,” *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 342–357, 2017.
- [285] C. P. Underwood, M. Royapoor, and B. Sturm, “Parametric modelling of domestic air-source heat pumps,” *Energy and Buildings*, vol. 139, 2017.
- [286] I. H. Bell, J. Wronski, S. Quoilin, and V. Lemort, “Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp,” *Industrial & engineering chemistry research*, vol. 53, no. 6, pp. 2498–2508, 2014.
- [287] R. Hendron and C. Engebrecht, “Building america house simulation protocols,” National Renewable Energy Laboratory, Tech. Rep., 2010.
- [288] R. Sangi, P. M. Martín, and D. Müller, “Thermoeconomic analysis of a building heating system,” *Energy*, vol. 111, pp. 351–363, 2016.
- [289] M. Di Somma, B. Yan, N. Bianco, G. Graditi, P. Luh, L. Mongibello, and V. Naso, “Multi-objective design optimization of distributed energy systems through cost and exergy assessments,” *Applied Energy*, vol. 204, pp. 1299–1316, 2017.
- [290] PricewaterhouseCoopers, “Vergelijking van de elektriciteitsen aardgasstarieven,” Vlaamse Regulator van de Elektriciteits- en Gasmarkt, Tech. Rep., 2019.
- [291] S. Gendebien, E. Georges, S. Bertagnolio, and V. Lemort, “Methodology to characterize a residential building stock using a bottom-up approach: a case study applied to Belgium,” *International Journal of Sustainable Energy Planning and Management*, vol. 4, pp. 71–88, 2014.
- [292] Climact, “Electrical energy savings scenarios for Belgium,” Climact, Tech. Rep., 2012.
- [293] D. S. Ryberg, J. Freeman, and N. Blair, “Quantifying interannual variability for photovoltaic systems in PVWatts,” National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2015.

- [294] M. Hofmeister and M. Guddat, "Techno-economic projections until 2050 for smaller heating and cooling technologies in the residential and tertiary sectors in the EU," *Joint Research Centre (European Commission)*, 2017.
- [295] K. K. Cao, A. N. Nitto, E. Sperber, and A. Thess, "Expanding the horizons of power-to-heat: Cost assessment for new space heating concepts with Wind Powered Thermal Energy Systems," *Energy*, vol. 164, pp. 925–936, 2018.
- [296] Vaillant, "allstor exclusiv," Available online: <https://www.vaillant.be/consumenten/producten/allstor-exclusiv-2048.html>. Accessed: 22th October 2020.
- [297] W. Liu, D. Klip, W. Zappa, S. Jelles, G. J. Kramer, and M. van den Broek, "The marginal-cost pricing for a competitive wholesale district heating market: A case study in the Netherlands," *Energy*, vol. 189, p. 116367, 2019.
- [298] R. Renaldi, A. Kiprakis, and D. Friedrich, "An optimisation framework for thermal energy storage integration in a residential heat pump heating system," *Applied energy*, vol. 186, pp. 520–529, 2017.
- [299] J. Beckers, D. Coppitters, W. De Paepe, F. Contino, J. Van Mierlo, and B. Verrelst, "Multi-Fidelity Design Optimisation of a Solenoid-Driven Linear Compressor," vol. 9, no. 2, 2020, p. 38.
- [300] D. Hattis, "Three candidate" laws" of uncertainty analysis," *Risk Analysis*, vol. 10, no. 11, 1990.
- [301] X. Rixhon, G. Limpens, D. Coppitters, H. Jeanmart, and F. Contino, "The Role of Electrofuels under Uncertainties for the Belgian Energy Transition," *Energies*, vol. 14, no. 13, p. 4027, 2021.
- [302] E. Carpaneto, G. Chicco, P. Mancarella, and A. Russo, "Cogeneration planning under uncertainty: Part I: Multiple time frame approach," *Applied Energy*, vol. 88, no. 4, pp. 1059–1067, 2011.
- [303] J. D. Jakeman, F. Franzelin, A. Narayan, M. Eldred, and D. Plfüger, "Polynomial chaos expansions for dependent random variables," *Computer Methods in Applied Mechanics and Engineering*, vol. 351, pp. 643–666, 2019.
- [304] K. Verleysen, A. Parente, and F. Contino, "How sensitive is a dynamic ammonia synthesis process? Global sensitivity analysis of a dynamic Haber-Bosch process (for flexible seasonal energy storage)," *Energy*, p. 121016, 2021.
- [305] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: Handling constraints and extending to an adaptive approach," *IEEE Transactions on evolutionary computation*, vol. 18, no. 4, pp. 602–622, 2013.
- [306] M. Gomez-Gonzalez, A. López, and F. Jurado, "Optimization of distributed generation systems using a new discrete PSO and OPF," *Electric power systems research*, vol. 84, no. 1, pp. 174–180, 2012.
- [307] Y.-G. Zhao and T. Ono, "A general procedure for first/second-order reliability method (FORM/SORM)," *Structural safety*, vol. 21, no. 2, pp. 95–112, 1999.
- [308] C. Hu and B. D. Youn, "Adaptive-sparse polynomial chaos expansion for reliability analysis and design of complex engineering systems," *Structural and Multidisciplinary Optimization*, vol. 43, no. 3, pp. 419–442, 2011.
- [309] M. M. Rienecker, M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim *et al.*, "MERRA: NASA's modern-era retrospective analysis for research and applications," *Journal of climate*, vol. 24, no. 14, pp. 3624–3648, 2011.
- [310] P. Poli, H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. Tan, C. Peubey, J.-N. Thépaut *et al.*, "ERA-20C: An atmospheric reanalysis of the twentieth century," *Journal of Climate*, vol. 29, no. 11, pp. 4083–4097, 2016.
- [311] M. Al-Addous, M. Saidan, M. Bdour, Z. Dalala, A. Albatayneh, and C. B. Class, "Key aspects and feasibility assessment of a proposed wind farm in Jordan," *International Journal of Low-Carbon Technologies*, vol. 15, no. 1, pp. 97–105, 2020.