

École polytechnique de Louvain

Could the PCE method be a better alternative to the Monte-Carlo method on the Power System Adequacy Assessment problem to make it more efficient computationally?

Power system adequacy assessment through PCE method

Author: **Alexandre TRIFIN**

Supervisors: **Francesco CONTINO, Emmanuel DE JAEGER**

Readers: **Diederik COPPITERS, François GLINEUR, Aurélia HER-
NANDEZ**

Academic year 2023–2024

Master [120] in Electrical Engineering

Abstract

With the relatively recent wide use of PCE in the energy sector, the PCE method is explored as a potential better alternative than the Monte-Carlo method on Power System Adequacy Assessment problems to make it more efficient computationally. Using Adequacy and PCE tools, a PCE process is created and the results are compared with the initial Monte-Carlo method. With a high Leave-One-Out (LOO) error, new methods are explored in the hope of improving the initial method. First, a transformation of the outputs LOE (Loss-of-Energy) and LOL (Loss-of-Load) into continuous output variables is performed, thanks to the extrapolation of monotonic curves for each profile (set of realizations of the uncertainty). Those extrapolations can be linear or quadratic. The LOO error is still high; profiles are divided in subprofiles (same realizations of uncertainties but different start times), the mean LOE and LOL of the subprofiles being the LOE and LOL of the profile. A certain combination of number of subprofiles, parameters of extrapolation and order of PCE permits to have decent results, according to the LOO error, to compare them with Monte-Carlo and to explore the analysis of a study case through PCE. Thanks to a decision tree, a hierarchy of units impacting the adequacy of the power system can be established and be served for deeper research.

Acknowledgement

First and foremost, I am profoundly grateful to Aurélia Hernandez and Diederik Coppitters for their unwavering support, insightful guidance, and unfailing availability. Their expertise and patience were instrumental in shaping this modest paper, and their encouragement kept me motivated throughout the process.

This master's thesis would be nothing without the framework given by Dr. Fransecso Contino and Dr. Emmanuel De Jaeger. I am also grateful for their constructive criticism and valuable suggestions. Their input has significantly enhanced the quality of this work.

Special thanks to Camille Vanhoenacker with whom I started this thesis and work until end February. This present thesis is still imbued with the formidable work you delivered when we worked together. I really hope that you are going well and that you are safe. You are a truly courageous hero.

In this acknowledgement, I would like to thank the UCLouvain, more specifically l'École Polytechnique de Louvain. My first year at Louvain-la-Neuve was in 2019, and, thinking back, I could not imagine how much I would have grown through those studies, by the challenging tasks (almost) always met, by the marvelous friends I encountered there, and by the incredible opportunities I had to discover the joy of being a student mentor and going to Benin and Japan. These 5 years will leave me with unforgettable memories of this life at Louvain-la-Neuve.

Finally, I would like to address my most thankful thoughts to all my friends who support me, in their own ways, through this final step of my Master's Degree. I am also gladly thankful to my family—my little sister, Lily-Rose; my father, Jacky; and my mother, Sandrine—for bringing me the will to never stop being curious and always supporting me in my projects. This last sentence of acknowledgement is reserved for my dearest and sweetest, Sébastien, who supported me in my lowest moments with his endless love and his solar joy, for which I can never be sufficiently grateful.

Contents

Acknowledgement	i
Contents	iii
List of Figures	v
List of Tables	vii
Nomenclature	viii
1 Introduction	1
2 Background	4
2.1 Introduction to Power System Adequacy Assessment	4
2.1.1 Preliminary concepts:	5
2.1.2 Deterministic method	6
2.1.3 Probabilistic method	7
2.1.4 Adequacy metrics	9
2.2 Overview of Monte-Carlo Method	11
2.2.1 Introduction	11
2.2.2 Convergence in Monte Carlo	12
2.2.3 Simulation Approaches	13
2.3 Overview of Polynomial Chaos Expansion (PCE) Method	15
2.3.1 Introduction	15
2.3.2 Definition	15
2.3.3 Metrics	16
2.3.4 Curse of Dimensionality	17
2.4 PCE in Power Adequacy Assessment	17
3 Methodology	19
3.1 Tools	19
3.1.1 Initial Power System Adequacy Assessment simulation . . .	19
3.1.2 PCE computation: Rheia package	23
3.2 Choice of Random Variables for PCE	24
3.2.1 Time-to-Failure (TTF) and Time-to-Repair (TTR)	24

3.2.2	Wind farm availability factor $\alpha_k(r)$	26
3.3	Connection of the Rheia package and the Power System Adequacy Assessment simulation	28
3.3.1	Inputs	28
3.3.2	Process Blocks	29
3.3.3	Outputs	31
4	First results	34
4.1	Study Case	34
4.2	Results with Monte-Carlo	39
4.3	Results with PCE	40
5	Continuous output variables, subprofiles and post-processing	46
5.1	Transformation into continuous output variables	46
5.1.1	Monotonic Curve	47
5.1.2	Extrapolation	48
5.1.3	Results	52
5.2	Subprofiles	52
5.2.1	Method	52
5.2.2	Results	56
5.3	Post-Processing	61
6	Discussion	70
7	Conclusion	76
	Bibliography	82
	Appendices	I
A	Calendar	II
B	Study Case IEEE 30-Bus System Data	IV
C	Source Code	XIV

List of Figures

3.1	Implemented Sequential Monte Carlo Process [18]	20
3.2	Wind turbine availability distribution	27
3.3	Maximum and available wind farm production over time for a certain wind profile (19) and a certain wind farm (Belwind)	28
3.4	PCE Loop	33
4.1	Study Case based on IEEE 30-bus Test Case (inspired by [26]) . . .	35
4.2	Buses Load	37
4.3	Kernel Distribution Estimate (KDE) of the LOL and LOE	40
4.4	PDFs for outputs (order 1), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	43
4.5	Sobol indices for outputs (order 1), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	43
4.6	PDFs for outputs (order 2), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	44
4.7	Sobol indices for outputs (order 2), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	44
4.8	PDFs for outputs (order 3), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	45
4.9	Sobol indices for outputs (order 3), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	45
5.3	Monotonic PCE Loop	51
5.4	Different subprofiles for identical profiles	53
5.5	4 different subprofiles for 1 profile	54
5.6	Subprofile PCE Loop	55
5.7	Evolution of LOE over $u_{ttf,gen1}$, 4 subprofiles	56
5.8	LOO error $\hat{\epsilon}_{LOO}$ depending on the number of subprofiles n_{sp} and the order p	58
5.9	PDFs for outputs (Order $p = 3$, number of subprofiles $n_{sp} = 4$), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	59

5.10	Sobol indices for outputs (Order $p = 3$, number of subprofiles $n_{sp} = 4$), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]	59
5.11	Kernel Distribution Estimate (KDE) of the LOL and LOE for Monte-Carlo (MC) and Polynomial Chaos Expansion (PCE)	61
5.12	LOE distribution divided in two parts: <i>System Working</i> (green) or <i>System Failing</i> (red)	62
5.13	Decision Tree with a maximum depth of 2, the first decision is based on the value of $u_{ttf,gen1}$ and the second ones are based on the value of $u_{ttf,gen2}$	63
5.14	Decision Tree Performance Metrics (Interpetability, Coverage, Cross-validation Coverage), the coverage and the cross-validation score does not increase significantly after 4 tree leaves	64
5.15	$(u_{TTF,gen1}, u_{TTR,gen1})$ -plane with System Working (green) and System Failing (red) (10572 samples)	65
5.16	(gen. 1 TTR, gen. 1 TTR)-plane with System Working (green) and System Failing (red) (10572 samples)	66
5.17	(gen 1 TTR, gen 1 TTR)-plane with System Working (green) and System Failing (red) (10572 samples), the points on the lines of 0.7, 0.8, 0.9, 0.95 show the points where the availability of the generator 1 is $\alpha_{gen} = 0.7$, $\alpha_{gen} = 0.8$, $\alpha_{gen} = 0.9$ and $\alpha_{gen} = 0.95$ respectively .	67
5.18	Decision Tree according to the availability of the 4 generators, "worst" and "best" stand for the variation of the ratio of Working Systems when the other uncertainties are taken into account (typically the wind farm availabilities α_{wf})	68

List of Tables

4.1	General values of the Study Case	36
4.2	Generators Main Parameters	38
4.3	Wind Farms Main Parameters	38
4.4	Monte-Carlo Main Results	39
4.5	Random Variables for PCE	41
4.6	Order 1 PCE Main Results	43
4.7	Order 2 PCE Main Results	44
4.8	Order 3 PCE Main Results	45
5.1	Order 3 PCE Linear Extrapolation Main Results	52
5.2	Order 3 PCE Quadratic Extrapolation Main Results	52
5.3	PCE Quadratic Extrapolation ($n_{extra} = 10$) with Subprofiles (1 uncertainty) Main Results	56
5.4	PCE Linear Extrapolation ($n_{extra} = 50$) with subprofiles (10 uncer- tainties) Main Results	57
5.5	Order $p = 3$, number of subprofiles $n_{sp} = 4$ PCE Main Results . . .	59
A.1	Calendar of the Master's thesis	II
B.1	General	IV
B.2	Buses	V
B.3	Branches	VI
B.4	Generators	VIII
B.5	Distributed Generators	IX
B.7	Distributed Renewable Power	IX
B.8	Hydrogen Technology	X
B.9	Hydrogen Demand	X
B.10	Electric Vehicles Part 1	XI
B.11	Electric Vehicles Part 2	XI
B.12	Load Shifting	XII
B.13	Voluntary Load Shedding	XII
B.14	Heat Pumps Part 1	XII

B.15 Heat Pump Part 2	XII
---------------------------------	-----

Nomenclature

$C(n, t)$	Hourly curtailed electrical power at node n [MWh/h]
$Ens(n, t)$	Hourly electrical energy-not-served at node n [MWh/h]
$F(l, t)$	Hourly flowing electricity through line l [MWh/h]
$P_{in}(n, t)$	Hourly power injection at node n [MWh/h]
$P_{gen}(g, t)$	Hourly production of conventional generator g [MWh/h]
$\Theta(n, t)$	Hourly phase angle of node n [MWh/h]
$b_{pu}(l)$	Susceptance of line l [pu]
$c_{gen}(g)$	Operational cost of conventional generator g [€/MWh]
$d_l(n)$	Share of classical electrical demand associated to node n [-]
$f_{max}(l)$	Maximum power that can transit through line l [MW]
$load(t)$	Hourly classical electrical demand [MWh/h] (excluding HPs, EVs, electrolyzers, and distributed production)
$load_{year}$	Yearly demand [TWh]
$p_{gen,min}(g)$	Minimum capacity of generator g [MW]
$p_{gen,max}(g)$	Maximum capacity of generator g [MW]
$res_{max}(r, t)$	Hourly power production of renewable power plant r without wind turbine outage [MWh/h]
s_{base}	Base power of the per-unit system [MVA]
n_{bus}	Number of bus nodes
G_n	Set of generators located at node n
$L_{to,n}$	Set of lines whose sign flow is defined to enter node n
$L_{from,n}$	Set of lines whose sign flow is defined to exit node n

R_n	Set of offshore wind farms located at node n
$\alpha_{gen}(g, t)$	Hourly availability of generator $g \in [0, 1]$
$\alpha_l(l, t)$	Hourly availability of transmission line $l \in [0, 1]$
$\alpha_k(r)$	Availability of offshore wind farm r on profile k
$p_{avail,max}(g, t)$	Hourly maximum available capacity of generator g [MWh/h]
$res(r, t)$	Hourly power production of offshore wind farm r [MWh/h]
$res_k(r, t)$	Hourly power production of offshore wind farm r on profile k [MWh/h]
u_{ttf}	Uniform random variable for TTF $\in [0, 1]$
u_{ttr}	Uniform random variable for TTR $\in [0, 1]$
u	Global uniform random variable $\in [0, 1]$
$f(t)$	Failure density distribution
$F(t)$	Failure distribution or Unreliability function
$R(t)$	Reliability function
$t_{failure}$	Time of failure
$h(t)$	hazard rate
λ	Poisson distribution mean
μ	Mean
σ	Standard Deviation
$\epsilon_{incr,rm}$	Rolling mean of ϵ_{incr}
ϵ_{incr}	Incremental epsilon
$\epsilon(n)$	Ratio of the variance on the mean of $LOEE(n)$
$\epsilon_{incr,threshold}$	$\epsilon_{incr,rm}$ limit fixed by the user
n_{limit}	Maximum number of Monte-Carlo years
n_{year}	Number of different outages scenarios available/Monte-Carlo years for a same power system
α	PCE Multi-indices
$\mathcal{A}$	Considered set of multi-indices for truncation scheme
u_α	PCE coefficients
Ψ_i	Orthogonal polynomial
$\mathcal{M}^{PCE}(\mathbf{X})$	PCE representation of the model
M	Stochastic dimension OR M of big M method
$\mathbf{X}$	Stochastic dimensions
$\mathcal{M}(\mathbf{X})$	Truncated PCE representation of the model
p	Polynomial order
u_0	First PCE coefficient

$S_i^{T,PC}$	Total-order Sobol' indices
$\hat{\epsilon}_{LOO}$	Leave-One-Out error [-]
x_0	Hour where the monotonic curve intersects the x-axis
x'_0	Hour where the extrapolated curve intersects the x-axis
n_{extra}	Number of monotonic curve last points for the extrapolation
k	Profile index $\in [1, n_{pro}]$
n_{pro}	Number of profiles for one process
k_s	Subprofile index $\in [1, n_{sp}]$
n_{sp}	Number of subprofiles for 1 profile
n_{simu}	Number of simulations/samples
pu	per unit
MC	Monte-Carlo
PCE	Polynomial Chaos Expansion
ENS	Energy-Not-Served or Energy-Not-Supplied [MWh/year]
$EENS$	Expected-Energy-Not-Supplied [MWh/year]
$LOLE$	Loss-Of-Load Expectation [hours/year]
$LOEE$	Loss-Of-Energy Expectation [MWh/year]
LOL	Loss-Of-Load [hours/year]
LOE	Loss-Of-Energy [MWh/year]
LOO	Leave-One-Out
LOP	Lack-Of-Power
LOI	Lack-Of-Interconnection
LO	Line-Overload
TTF	Time-To-Failure [hours]
TTR	Time-To-Repair [hours]
$MTTF$	Mean-Time-To-Failure [hours]
$MTTR$	Mean-Time-To-Repair [hours]
$VOLL$	Value-Of-Loss-Load [€/MWh]

Chapter 1

Introduction

Electricity is an important feature of development in our modern world, as it can bring energy over a long distance almost instantly and easily. Its vast applications have enabled societies to power houses, drive industries and fuel unprecedented technological advancements.

Nevertheless, the distribution of electricity faces different challenges, one of which is the choice of electricity mix to take account of economic realities, resource scarcity, and the quantities of released greenhouse gases in the atmosphere, which have a direct contribution to climate change. To minimize the quantities of greenhouse gases and their dependence on non-renewable sources, many countries have invested in renewable energy sources like solar energy, wind energy or hydroelectric energy. Those solutions, though imperfect, are realistic options to address the goals set by the country's policies concerning the environment and economic growth.

In a 2022 paper [3], the International Energy Agency (IEA) reported that between 2010 and 2020, the proportion of renewable energy in Belgium's total final energy increased from 6% to 12%. This increase was attributed to the growth of renewable electricity generation, primarily from photovoltaics (PV) and wind turbines, as well as an increase in the use of bioenergy. Also, according to this report, in 2021, Belgium had the sixth-highest offshore wind capacity in the world and is planning for a major expansion of offshore wind deployment. The Belgian target for 2030 is to reach 17.5% of gross final energy consumption for renewable energy. It does not totally match the EU-wide greenhouse gas emissions reduction target, which increased from 40% to 55%. Therefore, new targets for energy efficiency, renewable energy and emissions reduction will therefore must be set in Belgium. The target value for the 2030 fraction of renewable energy will, thus, almost certainly be superior to the value of 17.5% predicted.

An important aspect to take into consideration when dealing with renewable sources in a transmission network is the stochastic features added to the system. While the energy production from nuclear fission, coal or gas is highly predictable based on the quantity of materials used in the plants, the quantity of energy released from renewable energy sources is less predictable. This last quantity depends on the intensity of sunlight and the intensity of the wind, which are mostly stochastic values that we can approximate with models. Another important aspect to think of when dealing with uncertainties is the time for a generator, a line or any other element present in a transmission network to suffer an outage and, if it is out-of-service, the time to get repaired.

As those uncertainties will become more and more predominant with time and as the transmission network becomes more and more decentralized [28], the transmission network models must automatically become more complex to deal with the new complexities induced by the uncertainties. With the uncertainties, finding adequacy between energy produced and energy consumed becomes a relatively more complex task. As it can be seen from an article about the Integration of Hydrogen in Sequential Monte Carlo Power Systems Adequacy Assessment [18], a robust method that can be used to deal with those uncertainties is the classic Monte-Carlo method. The relatively recent use of the Polynomial Chaos Expansion (PCE) method to represent a random variable in terms of a polynomial function of other random variables in the Energy sector and the promising performances it could bring compared to the Monte-Carlo method [10][25][9][33] makes it the ideal candidate for wider use in transmission network simulations.

All this information together makes us converge toward the research question of this thesis:

Could the PCE method be a better alternative to the Monte-Carlo method on the Power System Adequacy Assessment problem to make it more efficient computationally?

The computation efficiency will be determined by the analysis of the trade-off between accuracy and time spent for the two different methods: Monte-Carlo and PCE. The purpose of this thesis is, then, to analyze the relevance of using the PCE method precisely instead of the Monte-Carlo method when dealing with power adequacy.

After this introduction, the thesis is divided into six main chapters: the background to observe the state-of-the-art, the work already executed to bring context

to the research question and the theoretical framework to provide mathematical and computational explanations of the adequacy and PCE models; the methodology to set the data and give information about how the analyses have been operated; the first results of the analyses; the new methodologies performed considering the first results and their new results; the discussion, giving a critical eye over the results from the chapter before; and the conclusion, which ends the thesis and provides a summary of the main results got. Those chapters are followed by the references and the appendices.

Chapter 2

Background

2.1 Introduction to Power System Adequacy Assessment

The most important goal of a power system is to efficiently deliver electrical energy to its consumers, which means to prioritize economic viability while ensuring a sufficient and satisfactory level of reliability, all while complying with environmental policies [13].

The concepts of economy and reliability are inversely related to each other. Not investing enough in the resources compromises reliability, leading to client discontent, whereas investing too much leads to improved reliability but also higher operating costs, which then increases tariff structure [24].

The Power System Reliability Assessment is a wide term that can include numerous concepts. Nevertheless, a division can be operated between two domains: the adequacy and the security [5], [1]. The adequacy assessment of a power system is, in simple terms, the analysis of the balancing of a power system, i.e., if the production can meet the demand in its most basic form. Other elements can be taken into account when dealing with adequacy assessment as irradiance for photovoltaic, the wind for the wind turbines, the maximum power flow of the lines. On the other hand, the security assessment is a matter of evaluating if the system can remain within its stability limits when an outage of a generator, line, etc. occurs [20].

Adequacy assessment evaluates the reliability of a generation system and helps guarantee its capacity to meet future load demands. Consequently, this type of assessment is considered a valuable planning tool for guiding investment-related decisions regarding the installation of new resources based on planning criteria.

Different developed models can be used to quantify the generation adequacy of the system by calculating indices such as LOLE (Loss-Of-Load Expectation) [hours/year] and LOEE (Loss-Of-Energy expectation) [MWh/year] based, respectively, on the number of hours or the quantity of energy for which the demand cannot be met [20]. The planning reserve margin of the system is determined based on the LOLE criterion to be satisfied. These assessments are needed for the short-term and long-term planning of the system to meet future load demand.

Determining the capacity value of variable renewable resources such as wind and solar is very important in the adequacy assessment framework because it quantifies the contribution of renewable energy resources to the reliability of the system [24]. According to a report of the European Commission about the Identification of Appropriate Generation and System Adequacy Standards for the Internal Electricity Market (2014), "*the System adequacy is the sum of generation and transmission adequacy*" [24].

The **Transmission adequacy** refers to the evaluation of its capability to effectively handle the flow generated by the power transfer from the generation source to the consumption centers [42].

Always according to this European Commission report, "*The first step in any adequacy evaluation is to produce the ‘states of the power system’ one at a time, verifying the ability of the system to supply the load power and energy*"[31], where "*states of the power system*" signifies a combination of loads, available generators and available transmission lines. [24]

2.1.1 Preliminary concepts:

The most common methodology for **Adequacy Assessment** consists in a calculation procedure (like Monte Carlo simulations), an analysis of the data from several sources like generator availability, wind regime, etc., and assumptions on the evolution of the system [24].

Two approaches can be used: **deterministic** or **probabilistic** [30] which lead to different metrics. Those metrics are defined to quantify the maturity of a system with respect to a target or standard, and can be used for comparison between different systems and on different time periods (because each metric has an associated standard value defined by the Transmission System Operator or by the National Regulatory Authority) [24].

2.1.2 Deterministic method

A deterministic method is a computational approach used to analyze or solve a problem where the outcome is entirely determined by the initial conditions and the parameters of the system. For power systems, a deterministic method involves using mathematical models and algorithms to quickly calculate certain aspects of the system's behavior, such as power flow or voltage levels.

Indeed, with this kind of methodology, the adequacy evaluation is based on the requirement for each outage event in the contingency set to ensure that the system's performance maintains the ability to meet demand through the provision of adequate frequency control reserves. These assessments are defined by selecting a discrete set of system configurations (such as network topology), a range of system operating conditions (like load distribution and generating unit dispatching), a list of possible outage events (like the unavailability of some generators or of some other system components such as lines or transformers), and performance evaluation criteria (such as values of voltage, frequency and, loading of the network elements inside their operating range).

These different operating scenarios are simulated using load flow and dynamic stability analysis tools to evaluate the system's performance under each scenario.

The deterministic method thus requires an exhaustive study based on careful thought and insight because of the large number of possible network configurations, the wide range of operating conditions, and the high number of conceivable outages. Therefore, the deterministic approach has evolved within the electrical power industry to decrease effort while still providing useful results.

This evolution brings several drawbacks:

- It only accounts for a limited number of specific outcomes based on predefined analysis criteria, disregarding numerous other potential scenarios.
- Each outcome is treated equally (with the same weight) without considering the likelihood of occurrence.
- It overlooks the interdependence between inputs, the varying impact of different inputs on outcomes, and other complexities, potentially leading to oversimplified models and decreased accuracy.
- It is too difficult to consider combinations of insufficient primary resources like wind or water with failures in system facilities, which, of course is a very likely event for systems relying on a lot of renewable energy resources.

Despite its limitations and potential inaccuracies, numerous organizations continue to utilize deterministic analysis, but they often complement it with a probabilistic modeling approach for better accuracy in evaluating particular cases and scenarios. Furthermore, while deterministic methods are favored for their quick execution, especially in real-time operational security assessment at the dispatching level, probabilistic methodologies are typically preferred for planning purposes [24].

2.1.3 Probabilistic method

The probabilistic method is an analytical approach used to account for uncertainty by considering the probability of different outcomes or events occurring [24].

In probabilistic methods applied to power systems, all possible system configurations are derived from factors representing the unavailability and/or failure probabilities associated with each element or component of the system. Those factors typically include the failure rates/probabilities of different system elements like generators, transmission lines, transformers or other critical components. By quantifying the probabilities of these individual failures, probabilistic methods can assess the overall reliability and performance of the power system under different operating conditions and scenarios.

Probabilistic criteria is a sort of generalization of the deterministic approach, since all possible constraining situations are examined and then, based on their results, risk indices are obtained. This probabilistic approach takes into consideration the random nature of loads, Renewable Energy Sources (RES) productions, and outages of generation/transmission equipment.

This approach solves some of the following problems caused by the deterministic approach:

- There is no discretion in the selection of scenarios and contingencies; the one limitation is accorded to the maximum number of examined situations.
- It is possible to assign a different weight to each case based on its probability of occurrence.
- It is possible to study a large interconnected system.

The primary challenge in probabilistic adequacy assessments lies in the high computational workload needed to analyze a large array of system states, each

with its own unique probability of occurrence. Nonetheless, advancements in computational power are effectively addressing this obstacle.

There are two main types of probabilistic methods [36]:

- Convolution of probability functions (basic concept: given the distribution probability of two variables $f_1(x_1)$ and $f_2(x_2)$, it finds the probability distribution of the sum/difference of these two variables $f_{12}(x_1 \pm x_2)$).
- The Monte Carlo model (which will be seen more in detail in subsection 2.2)

The main idea behind Monte Carlo is the following: imagine playing dice and randomly selecting outcomes to mimic possible scenarios in a complex system. Instead of fixed values, random variables capture uncertainties like component failures, renewable energy fluctuations, or weather-dependent loads.

Process:

- Random Sampling: Assign specific "unavailability ranges" to each element based on their failure probabilities. Generate random numbers for each element.
- Element Status: If a number falls within its unavailability range, the element is "disconnected" for that simulation. Otherwise, it functions normally.
- Repeat and analyze: Run numerous simulations (tens of thousands) with different random number sets, creating a diverse collection of system states.

Benefits:

- Flexibility: Can handle various uncertainties: component failures, renewable, weather-affected loads, etc.
- Detailed Insights: Reveals the entire range of possible outcomes, including rare events, unlike deterministic methods.

Drawback:

- Computational Cost: Requires many simulations for accurate results, increasing processing time.

2.1.4 Adequacy metrics

Depending on the adopted method to assess adequacy (deterministic or probabilistic), a particular Transmission System Operator will adopt one or more metrics in order to describe the adequacy of their system [24].

Certain metrics necessitate a probabilistic approach for calculation, while others can be derived using a deterministic approach. It is important to note that all probabilistic measures represent expectations—they are not deterministic values but rather the average value of a probability distribution, which serves as a modeled approximation of reality. These measures offer valuable insights into system adequacy by considering both the stochastic and deterministic aspects of the generation/transmission system and customer demand.

Adequacy can be seen as a balance between the system cost of LOE (Loss-Of-Energy) [MWh/year] (also written as *ENS* (Energy-Not-Served or Energy-Not-Supplied) in the literature) and the investments needed to reduce or avoid it. This is why an appropriate methodology for adequacy assessment is characterized by the ability to provide a metric that can be used for monetizing the system cost of LOE for at least one year.

LOE (Loss-Of-Energy) or ENS (Energy-Not-Served)

The LOE (Loss-Of-Energy) [MWh/year] or ENS (Energy-Not-Served or Energy-Not-Supplied) [MWh/year] are defined as the amount of energy not being served to meet the demand due to outages from the generators, lines, wind turbines, etc. and/or limited production from renewable energy sources. The implementation of LOE as a reliability measure is now increasing, being that power systems are actually energy systems [35]. Mathematically, the LOE for a whole year can be computed by the Eq. 2.1,

$$LOE = - \sum_{t=1}^{8760} \sum_{n=1}^{n_{bus}} (P_{in}(n, t) - d_l(n)load(t)) \quad (2.1)$$

where n_{bus} is the number of buses (nodes) of a power system, $P_{in}(n, t)$ is the hourly power injection (in MW) at a node n , $d_l(n)$ is the share of load on the node n [%] and $load(t)$ is the hourly demand (in MW). Note that 8760 is the number of hours in a whole year.

The LOE translates, thus, the idea of demand impossible to meet at a certain node of a power system.

Loss-of-Energy-Expectation (LOEE)

Also known as EENS (Expected-Energy-Not-Supplied), it represents the expected amount of electricity demand, expressed in MWh/year, that is anticipated being unmet/lost during a given year when the demand surpasses the available generation capacity [24].

This energy deficit can be categorized into several primary causes:

- **Lack-of-Power (LOP):** This occurs when there is insufficient generation capacity within the electrical system under consideration.
- **Lack-of-Interconnection (LOI):** This occurs in situations where generation capacity exists in one area, but network constraints prevent the supply from reaching the load in another area.
- **Line-Overload (LO):** This occurs when network elements become overloaded, making it impossible to supply the load.
- **Network splitting or isolated node:** This occurs when one or more network elements become unavailable, jeopardizing the electrical system's integrity and making it impossible to supply the load [24].

However, it is not possible to directly compare electricity systems of different sizes. This could be solved with LOEE normalization, which means dividing by the total energy requested. Mathematically, LOEE is expressed as the mean LOE of several outage scenarios on the same power system and is given by the Eq. 2.2.

$$LOEE = \frac{1}{n_{year}} \sum_{k=1}^{n_{year}} LOE_k \quad (2.2)$$

where n_{year} is the number of different scenarios simulated or Monte-Carlo year for a same power system.

LOL (Loss-Of-Load)

This metric (expressed in hours/year) corresponds to the actual occurrence/duration of a loss of load event (when the demand for electricity exceeds the available generation capacity), which results in customers experiencing interruptions or outages. It is therefore expressed in hours/year. It represents a specific event or occurrence of load shedding or blackout rather than a probabilistic measure of system reliability [6]. It can be represented mathematically by the Eq. 2.3.

$$LOL = \sum_{t=1}^{8760} a(t) \quad (2.3)$$

where $a(t)$ is defined by,

$$a(t) = \begin{cases} 1, & \text{if } \exists n \in n_{bus} \text{ s.t. } P_{in}(n, t) \leq d_l(n)load(t) \\ 0, & \text{otherwise} \end{cases}$$

LOLE (Loss-Of-Load-Expectancy)

This metric (expressed in hours/year) represents the yearly hours with LOEE occurrence. In other words, it is the expected frequency or likelihood of experiencing a Loss-Of-Load event over a specific period, typically a year.

LOLE is one of the most used metrics because of its simplicity. However, it does not provide any information about the magnitude of the problem (for example, if it is a blackout affecting the entire electricity system or minor load curtailments). Therefore, only homogeneous systems should be compared [24].

Mathematically, LOLE is expressed as the mean LOL of several outage scenarios on the same power system and is given by the Eq. 2.4.

$$LOLE = \frac{1}{n_{year}} \sum_{k=1}^{n_{year}} LOL_k \quad (2.4)$$

2.2 Overview of Monte-Carlo Method

2.2.1 Introduction

The Monte Carlo approach is based on a random sampling of scenarios, where each sampled scenario is analyzed afterwards.

For Adequacy Assessment, there are two categories of evaluation techniques: **analytical** and **simulation**.

Analytical techniques utilize analytical models to represent the system and evaluate the indices from these models using mathematical solutions.

Monte Carlo simulation methods use simulations of the actual process and random behavior of the system in order to estimate the indices. The problem is thus treated as a series of experiments.

Typically, analytical techniques are more efficient when complex operating conditions are not taken into account or when the failure probabilities of components

are minimal, indicating high system reliability. However, when there are complex operating conditions involved or if the number of severe events is increasing, Monte Carlo methods are preferable.

Here are the main advantages of Monte Carlo methods [24]:

- They can include system effects or processes that might require approximation in analytical methods.
- The necessary number of samples for achieving a desired level of accuracy remains unaffected by the system's size, making those simulation methods more suitable for evaluating large-scale systems.
- They have the capability to model probability distributions associated with component failure and restoration activities.
- They can not only compute reliability indices in the form of expected values but also determine the distributions of these indices.
- They can simulate non-electrical system factors like reservoir operating conditions and weather effects.

2.2.2 Convergence in Monte Carlo

In power system reliability evaluation, different reliability indices have different convergence speeds. It has been found that the coefficient of variation of the LOEE index has the lowest rate of convergence. For that reason, this coefficient of variation is often used to provide appropriate accuracy [6].

In this thesis, the Monte-Carlo simulation ends when the condition,

$$\epsilon_{incr,rm} \leq \epsilon_{incr,treshold}$$

is respected, while $n_{year} \leq n_{limit}$ (n_{year} being the number of Monte-Carlo scenarios (or Monte-Carlo years) and n_{limit} being the maximum number of Monte-Carlo years (here $n_{limit} = 2000$)), is respected, where $\epsilon_{incr,treshold}$ is the maximum error tolerance for Monte-Carlo (for example $\epsilon_{incr,treshold} = 0.001$).

- $\epsilon_{incr,rm}$ is the rolling mean of ϵ_{incr} i.e., the mean of the 30 last ϵ_{incr} .
- ϵ_{incr} is defined by: $\epsilon_{incr} = \frac{|\epsilon(n) - \epsilon(n-1)|}{\epsilon(n)}$

- $\epsilon(n)$ is defined by $\epsilon(n) = \frac{\sigma(LOEE(n))}{\mu(LOEE(n))}$ with $\mu(LOEE(n))$, the mean of the *LOE* after 1, 2, ..., n scenarios and $\sigma(LOEE(n))$, the standard deviation of the *LOE* after 1, 2, ..., n scenarios.
- $LOEE(n)$ is the Loss-Of-Energy-Expectation after n scenarios (mean of n *LOE*).

The relative difference for ϵ_{incr} comes from the methodology used by Elia (The Belgian transmission system operator for high-voltage electricity (30 000–380 000 volts), located in Brussels, Belgium [14]). The rolling mean $\epsilon_{incre,rm}$ comes from Aurélia Hernandez, from UCLouvain [41]. The purpose is to smooth the irregular peaks, which can give too precocious convergence.

2.2.3 Simulation Approaches

Monte Carlo simulations are used as a tool for sampling the system states. There exist various simulation methods, which can be defined as **non-sequential methods** or **sequential methods**. The non-sequential method samples a set of system states randomly, whereas the sequential method obtains a sequence of system states where each system state depends on the previous system state [6].

There are three primary approaches within Monte Carlo simulation for adequacy assessment, each offering a different perspective on system behavior:

- **State Sampling Approach** (non-sequential method): the system is modeled as a collection of discrete states, each representing a particular configuration or condition of the system.
 - Random sampling is performed to select different states of the system based on their probabilities of occurrence.
 - The simulation then evaluates system performance by analyzing the behavior of the system as it transitions between different states over time.
 - This approach is particularly useful for analyzing system reliability in terms of the probability of being in a certain state or configuration at any given time.
- **State Duration Sampling Approach** (sequential method): extends the State Sampling Approach by considering not only the states themselves but also the periods of time for which the system remains in each state.
 - Random sampling is used to select both the states and the periods of time associated with each state based on their respective probabilities.

- TTF and TTR are directly integrated into the simulation by considering not only the states themselves but also the periods of time associated with each state.
 - Random sampling is used to select both the states and the periods of time, taking into account the probabilities of state transitions as well as the distributions of TTF and TTR.
 - This approach provides a more detailed understanding of system behavior by incorporating information about the length of time spent in different states.
 - It is particularly useful for analyzing reliability metrics such as outage periods of time and frequency of state transitions.
- **System State Transition Sampling Approach** (sequential method): the focus is on modeling the transitions between different system states rather than the states themselves.
 - Random sampling is used to simulate the sequence of state transitions that the system undergoes over time.
 - This approach is valuable for analyzing the dynamics of system behavior, including the frequency and timing of state transitions.
 - It provides insights into how the system evolves over time and how different factors influence the likelihood of transitioning between states.

The State Duration Method

In this thesis, the State Duration Sampling Approach is the one explored. The state duration method is a sequential method that creates chronological state histories for each of the system's components individually. A state history of a component is created by drawing random variables from the TTF distribution and the TTR distribution in sequence whenever a failure or a repair of the component happens. A complete system state history is obtained by combining the individual state histories together. It is common to assume that the TTFs and TTRs are exponentially distributed, although other distributions are equally applicable. As an initial starting point, all components are assumed to be available. If the simulation is performed for a short time span, the assumption may lead to an overestimation of the system's reliability. However, as the reliability indices, which are calculated from a sequential simulation, often require a large quantity of samples to converge, the effect is often negligible [6].

2.3 Overview of Polynomial Chaos Expansion (PCE) Method

2.3.1 Introduction

One of the first appearances of the PCE method comes from a paper of 1991 [32] which make use of a "stochastic finite element method" (SFEM) to solve boundary value problems with uncertain coefficients. A random output is represented as a polynomial series from the input variables, thanks to the main element of this method, which is the "polynomial chaos expansion" (PCE). Following this paper, applications of that method can be found in various fields like geotechnics [15]; [40]; [16]; [21]; [37]; [39]; [4].

Since the early 2010s, engineering problems make wide use of PCE to analyze uncertainty quantification in numerous fields like electromagnetism, fluid dynamics and heat transfer problems. The PCE has improved over the recent years to permit the emergence of non-intrusive approaches [38].

2.3.2 Definition

The explanations below come mainly from the Rheia documentation [8].

The Polynomial Chaos Expansion (PCE) representation of the model under consideration consists of a series of orthogonal polynomials Ψ_i with corresponding coefficients

$$\mathcal{M}^{PCE}(\mathbf{X}) = \sum_{\alpha \in \mathcal{A}} u_{\alpha} \Psi_{\alpha} \approx \mathcal{M}(\mathbf{X}) \quad (2.5)$$

where α are the multi-indices and $\mathcal{A}$ is the considered set of multi-indices, for which the size is defined by a truncation scheme (this size is theoretically infinite but finite in practice). The polynomial family that is orthogonal w.r.t the assigned probability distributions are known for classic distributions, e.g., a random variable input following a normal distribution, its resulting family is that of Hermite polynomials. A truncation scheme is adopted, which limits the polynomial order up to a chosen degree. This constrains the number of multi-indices in the set to:

$$|\mathcal{A}^{M,p}| = \frac{(p+M)!}{p! M!}, \quad (2.6)$$

where corresponds to the polynomial order and $M = |\mathbf{X}|$ is the stochastic dimension, i.e., number of uncertainties (parameters and variables). Consequently,

$|\mathcal{A}^{M,p}|$ coefficients are present in a full PCE. To quantify these coefficients, least-squares minimization is applied based on actual system model evaluations. To ensure a well-posed Least-Square Minimization, $2|\mathcal{A}^{M,p}|$ model evaluations are required. When the coefficients are quantified, the mean (μ) and standard deviation (σ) of the model output of interest are analytically derived:

$$\mu = u_0 \quad (2.7)$$

$$\sigma^2 = \sum_{i=1}^P u_i^2 \quad (2.8)$$

Beside to the mean and the variance, profitable information on the outputs of the contribution of each stochastic input can be extracted. The most used method is by Sobol' indices. Through post-processing of the coefficients, following the initial model evaluations, the Sobol' indices allow a global sensitivity analysis of the stochastic inputs and the PCE provides an analytical solution. The total-order Sobol' indices ($S_i^{T,PC}$) quantify the total impact of a stochastic input, even considering all mutual interactions.

$$S_i^{T,PC} = \sum_{\alpha \in A_i^T} u_\alpha^2 / \sum_{i=1}^P u_i^2 \quad (2.9)$$

$$A_i^T = \{\alpha \in A | \alpha_i > 0\}$$

For every coefficient that is characterized by considering input parameter i , among others, at an order > 0 is added to the total Sobol' index S_i [11].

2.3.3 Metrics

No method enables to know in advance the optimal polynomial degree for the PCE. This would often be performed empirically by testing with different truncation schemes before the number of samples becomes unmanageable and if there is a certain convergence for the quantities of interest (typically, with $p = 2, 3, 4$) [38].

A trade-off between relatively accurate error estimation and manageable computations can be obtained thanks to the *leave-one-out (LOO) cross-validation*. Different sets of points are used to create a PCE expansion and compare it with the the original computational model to compute the error. From the full experimental design $\mathcal{X}$, LOO cross-validation removes one point apart, let us say $\mathbf{x}^{(i)}$ and builds a PC expansion written by $\mathcal{M}^{PC \setminus i}(\cdot)$ with the $n - 1$ other points, i.e. from the ED $\mathcal{X} \setminus \mathbf{x}^{(i)} \stackrel{\text{def}}{=} \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i-1)}, \mathbf{x}^{(i+1)}, \dots, \mathbf{x}^{(n)}\}$. The predicted residual error at that point reads [38]:

$$\Delta_i \stackrel{\text{def}}{=} \mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{PC \setminus i}(\mathbf{x}^{(i)}) \quad (2.10)$$

2.3.4 Curse of Dimensionality

Engineering problems have commonly between $M = 10$ and 50 stochastic inputs for their computational models. Considering this, a large number of unknown coefficients (which is also the size of the truncation set $P = \text{card} \mathcal{A}^{M,p}$) must be found, even with a low-order PCE. When choosing $p = 3$, $M = 10$ and $M = 50$ stochastic inputs lead to 286 and 23 436 terms to find respectively [38].

2.4 PCE in Power Adequacy Assessment

In the literature, attempts to use PCE for Power Systems affected by various uncertainties, considering the increasing importance of renewable energy sources in the electrical mix, are rather recent.

The authors of the article *Solving Optimal Power Flow with Non-Gaussian Uncertainties via Polynomial Chaos Expansion* (2017) [27] modeled uncertainties for optimal power flow problems. By modeling uncertainties via continuous random variables of finite variance, they proposed a certain formulation of the uncertain OPF problem. In this article, the uncertainties can follow distributions different from the Gaussian, be multivariate and/or correlated. The PCE has been used to translate this infinite-dimensional random variable problem as a finite-dimensional random variable problem. An analysis of the sensitivity is also provided. Finally, a comparison with Monte-Carlo is set.

The system is composed of N buses and M lines under DC power flow conditions, and the net power injection $p \in \mathbb{R}^N$ is decomposed as follows:

$$p = d + Cu \quad (2.11)$$

where d represents fixed power generation ($d_i > 0$) to or demand ($d_i < 0$) from the network. The vector $u \in \mathbb{R}^{N_g}$ represents the N_g degrees of freedom in the later Optimal Power Flow (OPF) problem. $C \in \mathbb{N}_0^{N \times N_g}$ maps the free power injections u to their corresponding bus index. Uncertainties are assumed to enter via the fixed power d which may include uncertainties of the loads and uncertainty due to wind power or solar power.

This work formulates the OPF problem subject to non-Gaussian multivariate uncertainties entirely in terms of random variables such that the DC power

flow equations are always satisfied. Using PCE, the problem is reformulated as a SOCP with PCE coefficients as decision variables that are solved in a single numerical run. The optimal solution admits different interpretations: it yields the coefficients of power generation in terms of a polynomial basis of standard random variables; it yields a feedback policy in terms of the power fluctuation, computed only based on the stochasticity of the uncertainty. Immediately upon knowing the realization of the fluctuation, the corresponding power injection is known. The suggested approach is applied to a modified IEEE 300-bus test system and compared to the Monte Carlo simulation, where the DC-OPF is solved for each sample. Future work will study the relation between the proposed approach and automatic generation control, as well as extensions to the nonlinear AC-OPF problem.

The authors of the article *Probabilistic Power Flow Analysis Based on Arbitrary Polynomial Chaos Expansion for Networks with Uncertain Renewable Sources* (2018) [22] present an application of the arbitrary polynomial chaos (aPC) expansion technique to the Probabilistic Power Flow problem. This method is rather interesting because the basis polynomial sets can be constructed from measured data instead of following a certain distribution. A technique called "collocation technique" is used to reduce the computational time and other techniques are discussed.

From the numerical experiment results, it can be seen that the PPF analysis using the aPC basis is more suitable than the gPC-based approach when an inappropriate basis is employed. For a system with a large number of random variables, all PCs-based methods are found to suffer from the curse of dimension. It can be noticed that many aPC coefficients are almost zero. This implies that many bases may be unnecessary. This observation can be used to improve the computation. Enhancement of the method to handle such cases and still have a good balance between computation efficiency and the accuracy of results is under investigation.

As for the previous article, the main objective is to find every bus voltage phasor of the power system network that satisfies the power balance and other related constraints, but not to address power system adequacy assessment with LOLE and LOEE. The objective of this thesis will be to show the usage of PCE with systems, depending on time, where a conventional generator can possibly produce as many as their maximum power capacity or zero in case of an outage. Offshore wind production does not follow a statistical distribution but historical data.

Chapter 3

Methodology

3.1 Tools

In this section, the two main tools for the PCE computation of outputs from Power System Adequacy Assessment simulations will be briefly described.

3.1.1 Initial Power System Adequacy Assessment simulation

The Initial Power System Adequacy Assessment simulation and its Monte Carlo process have been written by Aurélia Hernandez (UCLouvain) as part of her thesis, under the direction of François Vallée (UMons) and Emmanuel De Jaeger (UCLouvain) [18]. Sequential Monte Carlo (MC) simulations are used to assess the adequacy of power systems. The main reason was that they could be applied to more complex systems while maintaining reasonable computational time. Moreover, sequentiality was chosen in order to consider seasonality. The MC process is depicted in Fig. 3.1. For each MC year, the hourly availability of the generating units is defined, as well as the hourly renewable energy sources production in time series. These time series will be used to perform a MP-DC-OPF (Multi-Period Direct-Current Optimal-Power-Flow), the results of which will be used to extract the adequacy indicators. Let us note that a *simulation* stands for the performance of one MP-DC-OPF for one MC year, and that a *process* stands for the whole procedure of running simulations until we observe a convergence of the LOEE values.

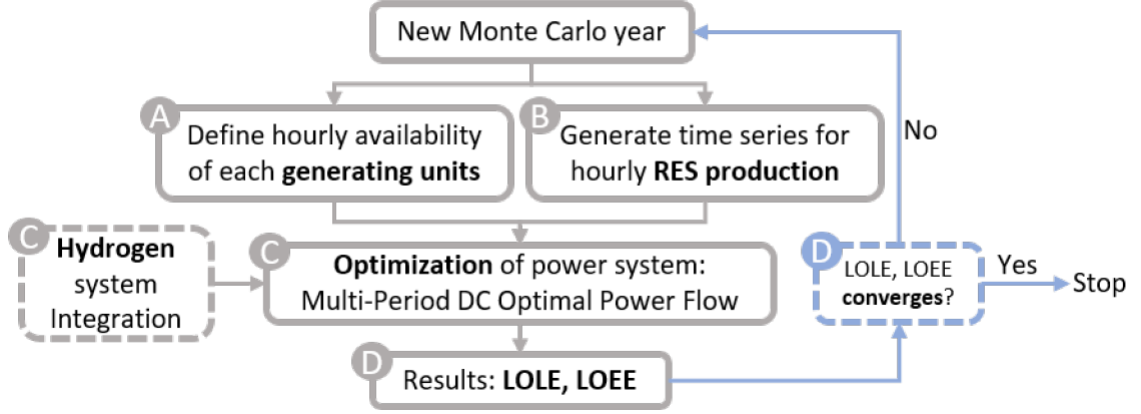


Figure 3.1: Implemented Sequential Monte Carlo Process [18]

Firstly, let us note that the programming language used is Julia. Julia offers the same easy-to-read and easy-to-understand syntax that is offered by Python. However, the mere fact that Julia is compiled means that it will automatically have an advantage over Python in terms of execution speed [45] which is necessary when dealing with long-time (typically $t = 8760$) simulations.

In an isolated simulation (without being integrated into a MC process), the first step is typically to extract the *static parameters* from an external Excel file. These static parameters correspond to all the data that does not vary from one year to another, e.g., the maximal active power a traditional power plant can produce, the maximal power flow a branch can support, or the load share of a certain bus.

The second step will be to generate uncertain parameters, which will be the hourly availability of each generating units and the time series for hourly renewable energy source production. Those uncertain parameters are based on probability density functions or historical data for which we select, randomly, a certain profile. Those uncertainties are the reasons why a Monte Carlo process is executed because, for two simulations, the outputs are unlikely to be similar, and we thus want to find the associated probability density function of the LOE.

Without entering too much into details and letting the interested reader read the Aurélia Hernandez's thesis for more information (not yet published), the main constraints and the objective function for the MP-DC-OPF will be presented:

Objective function

The costs related to the operation of conventional generating units P_{gen} are minimized and detailed in 3.1. $c_{gen}(g)$ describes the fixed cost [€], the linear marginal cost [€/MWh] and/or the quadratic marginal cost [€/MWh²] of a conventional generator g .

$$Min \sum_{t=1}^{8760} \left(\sum_{g=1}^{n_g} c_{gen}(g) P_{gen}(g, t) \right) \quad (3.1)$$

The optimization is a linear optimization and is operated thanks to the proprietary software *Gurobi* [17] which is used as a Julia Package.

Main constraints

Every constraint of the simulation will not be presented, as it is not the main goal of this master's thesis but the most important. In the case of a relatively simple power system test case:

Electrical balance

At each hour t of the year and at each node n of the system, electrical energy has to comply with the conservation of energy. The power produced by the conventional generators P_{gen} , the offshore wind farms res , the power flowing in and out F has to satisfy the electrical load l . If unbalances occur, production is curtailed C (excess) or energy is not served Ens (lack).

$$\begin{aligned} \sum_{g \in G_n} P_{gen}(g, t) + \sum_{r \in R_n} res(r, t) + \sum_{l \in L_{to, n}} F(l, t) - \sum_{l \in L_{from, n}} F(l, t) \\ + Ens(n, t) - C(n, t) = d_l(n)load(t), \forall n, \forall t \end{aligned} \quad (3.2)$$

$d_l(n)load(t)$ consists of the pure electrical demand. It represents the load seen at the high voltage transmission interface.

Conventional generators

Production of conventional generators P_{conv} are limited by their minimum/maximum power capacity $p_{conv, min/max}$, and by their availability $\alpha_{gen} \in [0; 1]$ at each time step. The unavailabilities represent unplanned (forced) outages.

$$\begin{aligned} \alpha_{gen}(g, t) p_{gen, min}(g) &\leq P_{gen}(g, t) \\ &\leq \alpha_{gen}(g, t) p_{gen, max}(g), \forall g, \forall t \end{aligned} \quad (3.3)$$

Power load flow

Electricity can flow through electrical lines in order to flow from generating units to consuming units. However, the flows are constrained by the physics of the load flow. In this case, the approximation of a DC load flow is considered, i.e., the power flowing through a line $F(k, t)$ is directly linked to the susceptance of the line $b_{pu}(l)$, and to the phase angle θ difference between the two nodes $n_{from}(l)$ and $n_{to}(l)$ bounding line l . The big M method [23] is used to represent two cases: (i) when the line is functional ($\alpha_l(i, t) = 1$), the physics of DC load flow is applied to the line, and the original equation of the DC load flow $F(l, t) = s_{base} b_{pu}(l)(\Theta(n_{from}(l), t) - \Theta(n_{to}(l), t))$ is obtained from 3.4, and (ii) when the line is non-functional ($\alpha_l(i, t) = 0$), the phase angles of the nodes bounding the line are independent and do not have to be equal, which allows lines in parallel to operate normally. Also, because only the relative phase angle values matter and their absolute values, a phase angle node n_{ref} is set as the reference node, and its value is set to zero.

$$\begin{aligned} -(1 - \alpha_l(l, t))M &\leq F(l, t) - s_{base} b_{pu}(l)(\Theta(n_{from}(l), t) - \Theta(n_{to}(l), t)) \\ &\leq (1 - \alpha_l(l, t))M, \forall l, \forall t \end{aligned} \quad (3.4)$$

$$\Theta(n_{ref}, t) = 0, \forall t \quad (3.5)$$

These flows are limited by the maximum capacity of the lines f_{max} and their availability α_l in 3.6. The phase angle difference between two adjacent nodes bounding a line is also limited by $\Delta\Theta_{max}$ in 3.7.

$$\begin{aligned} -\alpha_l(l, t)f_{max}(l) &\leq F(l, t) \\ &\leq \alpha_l(l, t)f_{max}(l), \forall l, \forall t \end{aligned} \quad (3.6)$$

$$\begin{aligned} -\Delta\Theta_{max} &\leq \Theta(n_{from}(l), t) - \Theta(n_{to}(l), t) \\ &\leq \Delta\Theta_{max}, \forall l, \forall t \end{aligned} \quad (3.7)$$

Energy-not-served

The energy-not-served, i.e., the energy lacking to serve the load, is bounded at each node by the local consumption (classical electrical load). The coefficient $d_l(n)$ describes the share of load depending on the node n .

$$Ens(n, t) \leq d_l(n)load(t), \forall n, \forall t \quad (3.8)$$

Other constraints

As previously said, the other constraints will not be detailed. The Electrical Imports, the Curtailment, the Electrolyzers, the Hydrogen-to-Power units, the Load Shifting, the Voluntary load shedding, the Flexible demands: mobility, heat and hydrogen, the PHS storage management, the Battery storage management, the Hydrogen balance constraint, the Hydrogen Pipelines, the Hydrogen Imports and the Hydrogen storage management are not taken into account in this thesis.

3.1.2 PCE computation: Rheia package

The Robust design optimization of renewable Hydrogen and dErIved energy cArrier systems (RHEIA) [11] framework provides multi-objective optimization (deterministic and stochastic) and uncertainty quantification algorithms. These algorithms can be applied to hydrogen-based energy systems, which are included in RHEIA. In addition, RHEIA allows the use of other models connected to the algorithms as well.

RHEIA, in its entirety, is a Python package. The uncertainty quantification procedure provides the mean and standard deviation of the quantity of interest, the Sobol' indices by constructing a PCE, and the capacity to provide the PDFs and CDFs of this quantity of interest. The mean value and the corresponding uncertainty (random variable inputs) of the model parameters are characterized in two files, `design_space.csv` and `stochastic_space.csv`, respectively.

A model parameter is a parameter that cannot be controlled by the designer in the context of an optimization (e.g., the availability of a wind farm). This parameter can be deterministic or uncertain.

In the `design_space.csv` file, the model parameters that need to be studied in the model are defined. When performing uncertainty quantification, the `design_space.csv` file consists only of model parameters. In this file, the mean values of the model parameters are given. The uncertainty on model parameters can be allocated through the file `stochastic_space.csv`. This file is required when performing uncertainty quantification, where several parameters are subjected to uncertainty. For every model parameter defined in `design_space.csv`, an uncertainty can be defined. Those uncertainties can follow uniform or Gaussian distributions.

To run the uncertainty quantification, some other important parameters must be set, like the polynomial order corresponding to the maximum polynomial degree in the PCE or the objective of interest to provide the mean and the standard

deviation of this objective of interest.

More information is available in the documentation of Rheia [11].

3.2 Choice of Random Variables for PCE

In the original Monte-Carlo process of the Power System Adequacy Assessment, various parameters were fixed randomly, either following a certain random distribution (time-to-failure, time-to-repair of generators, lines, wind turbines,...) or following historical data (offshore wind production at a time t). These *uncertainties* are the reason we use the Monte-Carlo process to find a convergence, in terms of LOE, in a certain power system. A certain set of fixed uncertainties for a Monte-Carlo year is called, by extension, a *scenario*.

In order to realize a PCE process, all these uncertainties will, ideally, become Random Variables for PCE to be set as design variables and model parameters as explained in the previous section 3.1.2. A certain set of realized uncertainties for one PCE year is called a *profile*.

3.2.1 Time-to-Failure (TTF) and Time-to-Repair (TTR)

Before discussing the Time-to-Failure (TTF) [hours] and the Time-to-Repair (TTR) [hours], some definitions must be given.

The failure density distribution $f(t)$ is the rate at which a component (traditional generator, line, wind turbines,...) fails, i.e. the probability density function (PDF) of the component's failure is: $f(t) = \frac{dF(t)}{dt}$.

The failure distribution or unreliability function $F(t)$ is the cumulative function (CDF) of the component's failure: $F(t) = \int_{-\infty}^t f(t')dt'$. It represents the probability that a component fails before time t , i.e., $F(t) = P[t_{failure} \leq t]$.

The reliability function $R(t)$ represents the probability that a component does not fail before time t , i.e. $R(t) = P[t_{failure} > t] = 1 - F(t)$.

The hazard rate $h(t)$ is the rate at which component fails at time t given that it is functional until t in 3.9.

$$\begin{aligned}
h(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P[t < t_{\text{failure}} \leq t + \Delta t | t_{\text{failure}} > t] \\
&= \lim_{\Delta t \rightarrow 0} \frac{P[t < t_{\text{failure}} \leq t + \Delta t]}{P[t_{\text{failure}} > t]} \\
&= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{F(t + \Delta t) - F(t)}{R(t)} \\
&= \frac{dF(t)}{dt} \frac{1}{R(t)} \\
&= \frac{f(t)}{R(t)}
\end{aligned} \tag{3.9}$$

Usually, the rate at which generators are represented by the negative exponential distribution. In probability theory and statistics, the exponential distribution is the **probability distribution of the time between events in a Poisson point process**, i.e., a process in which events occur continuously and independently at a constant average rate. It has the key property of being memoryless [46].

The principal characteristic of this distribution is that **its hazard rate is constant**, i.e., $h(t) = cst = \lambda$, which gives in 3.10:

$$\begin{aligned}
h(t) &= \lambda \\
\leftrightarrow \frac{f(t)}{R(t)} &= \lambda \\
\leftrightarrow \frac{dF(t)}{dt} \frac{1}{R(t)} &= \lambda \\
\leftrightarrow \frac{dF(t)}{dt} \frac{1}{1 - F(t)} &= \lambda \\
\leftrightarrow \frac{dF(t)}{dt} &= \lambda - \lambda F(t) \\
\leftrightarrow \frac{dF(t)}{dt} + \lambda F(t) &= \lambda \\
\leftrightarrow F(t) &= 1 - e^{-\lambda t}
\end{aligned} \tag{3.10}$$

giving $f(t) = \frac{dF(t)}{dt} = \lambda e^{-\lambda t}$ if $t \geq 0$ and 0 if $t < 0$,
and $R(t) = 1 - F(t) = e^{-\lambda t}$

By computing the inverse function of $F(t)$, which is $F^{-1}(u)$. We find a function which give a sample from $f(t)$ with u being a random sample from the uniform

distribution.

By setting $\lambda = \frac{1}{MTTF}$ or $\lambda = \frac{1}{MTTR}$, depending on if we want to compute the TTF or the TTR, we find in 3.11 and 3.12:

$$TTF = -MTTF \log(1 - u_{ttf}) \quad (3.11)$$

$$TTR = -MTTR \log(1 - u_{ttr}) \quad (3.12)$$

with MTTF and MTTR, respectively, standing for Mean-Time-To-Failure [hours] and Mean-Time-To-Repair [hours] and u_{ttf} and u_{ttr} being uniform random variables following $\mathcal{U}(0, 1)$.

One can be tempted to provide uncertainties on the MTTF and the MTTR for the simplicity of computation (the MTTF and the MTTR are directly accessible from the dataset that describes the power system). This can be discussed even more if we set defined values of u_{ttf} and u_{ttr} for the whole process, letting MTTF and MTTR be the sole uncertainties of TTF and TTR. To stay as close as possible to the original model, we will work with u_{ttf} and u_{ttr} as sources of uncertainties for the TTF and TTR.

Finally, let us note that, in the Monte-Carlo process, the TTR and TTF were always changing for a whole year while, in the PCE process, those stay constant during the whole year for a certain generator e.g., let us take a generator A during a Monte-Carlo year. It can start working with a TTF of 3000 hours, then be down and have a TTR of 45 hours, be working again with a TTF of 2600 hours, be down and have a TTR of 75 hours, etc. During the PCE process, as we must know the exact number of variables, we must put a constant u_{ttf} and u_{ttr} for a certain generator for a certain year; thus, if the generator starts working with a TTF of 2500 hours, it can have a TTR of 35 hours when being down but will have again a TTF of 2500 hours, and again a TTR of 35 hours after, etc.

3.2.2 Wind farm availability factor $\alpha_k(r)$

In the original model, each wind turbines of each offshore wind farm has its own TTF and TTR. If we take the 11 Belgian wind farms as references (in project and already built), a total of 633 different wind turbines must be considered, giving a total of 1266 random variables for the PCE, only from the offshore wind power contribution.

A reasonable method to deal with such a high number of variables is to approximate the number of wind turbines available for one wind farm such that

$$res_k(r, t) = \alpha_k(r) * res_{max}(r, t) \quad (3.13)$$

where, in 3.13, $res_{max}(r, t)$ is the wind power production (depending on a certain wind profile) for a certain hour t for a certain wind farm r considering that all wind turbines are always working, $\alpha_k(r)$ is the wind farm availability factor for a certain wind farm r on a certain profile k , $res_k(r, t)$ is, finally, the multiplication of those two factors, meaning that the wind power production can now be different for two different profiles.

The next question for this section is to find a way to describe the new uncertainty $\alpha_k(r)$ thanks to a random distribution. A method to do so is to observe the ratio of functional wind turbines to the total number of wind turbines through a Monte-Carlo year and create a histogram based on the ratio observed for each hour of the year. With this histogram, a random distribution can be fit to accurately describe the general availability of the wind farms throughout the year. The normal random distribution managed to be a more accurate description of the wind farm availability, as we can see on the Figure 3.2. The mean and the standard deviation (Std) found depend on the MTTF and the MTTR given in the source data.

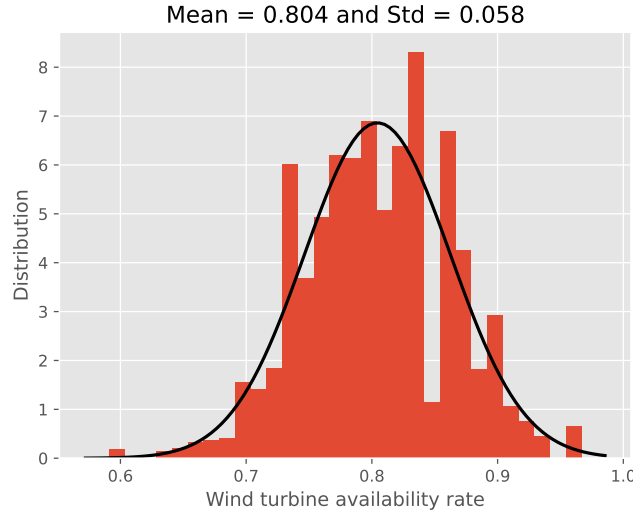


Figure 3.2: Wind turbine availability distribution

In the next graph on the Figure 3.3, we analyze the power produced by a certain wind farm, called Belwind, and for a certain wind profile, which is here the wind profile 19. The red curve describes the maximal wind farm power for a certain hour, i.e., when all the wind turbines composing a wind farm are functional. The blue curve describes the wind farm power when the availability rate $\alpha_k(r)$ of the wind farm is 0.804, i.e., 80.4% of all the wind turbines are working at a certain hour. As this availability rate $\alpha_k(r)$ follows a normal distribution, a prediction

interval of 95% can be given as the sky blue area on the graph. With a standard deviation of 0.058 and a mean of 0.804, we observe that the range of values, in this prediction interval, for the availability rate $\alpha_k(r)$ is between 0.688 and 0.92.

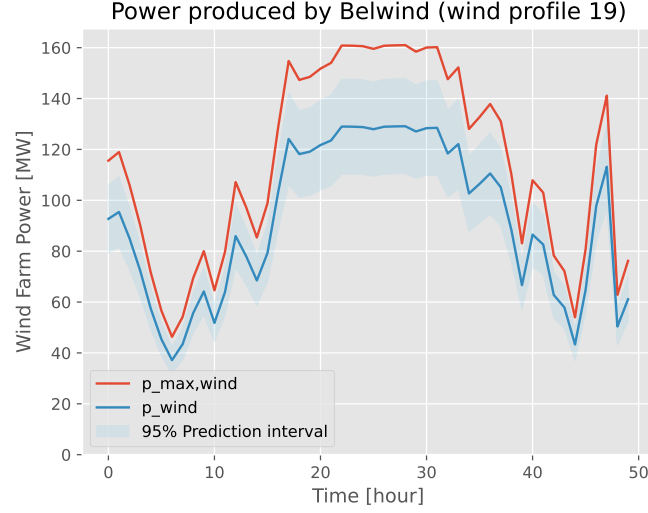


Figure 3.3: Maximum and available wind farm production over time for a certain wind profile (19) and a certain wind farm (Belwind)

3.3 Connection of the Rheia package and the Power System Adequacy Assessment simulation

In this section, the connection between the Rheia package and the Power System Adequacy Assessment simulation will be presented. Figure 3.4 will be frequently mentioned to express the relation between every block of the PCE process.

3.3.1 Inputs

Study Case Data

The *Study Case Data* is an excel file (or xlsx file) describing a certain power system with its traditional generators, its lines, its renewable energy sources and other important elements to take into account when modeling the power grid e.g the maximum power of the generators, the MTTF and MTTR of the lines, etc.

(see "Study Case Data" in Figure 3.4)

PCE Input Distributions

The PCE Input Distributions are two files, the `design_space.csv` and the `stochastic_space.csv` which are described more precisely in the subsection 3.1.2. As we saw in the section 3.2 about the choice of random variables for PCE, it is possible to link the uncertainties from the Monte-Carlo with uniform distributions (for the TTF and TTR) and with normal distributions (for the wind farm availability) (see "PCE Input Distributions" in Figure 3.4).

Wind Profiles

The "Wind Profiles" input is composed of 100 different wind profiles for each Belgian wind farm already built or in project (Belwind, C-Power, Mermaid, Nobelwind, Norther, Northwester2, Northwind, P1-2800-15, P1-3500-15, P23-2800-15, P23-2800-15, Rentel and Seastar). A wind profile describes the historical wind power production of each wind turbine of a certain wind farm, each hour, for a year, e.g., as the Belwind wind farm is composed of 56 wind turbines, a certain wind profile of Belwind (let us say the number 43) gives a matrix of $56 * 8760 = 490\,560$ values (8760 being the number of hours in a year). Considering the high number of values, these are stored in *npz* files thanks to their storage efficiency, speed efficiency and convenience compared to other file formats [29]. During the Monte-Carlo process, each simulation chooses, randomly, a wind profile index that will be the same for each wind farm. On the other hand, the same wind profile will be kept for the entire process with the PCE to avoid distorting the results obtained by the PCE (see "Wind Profiles" in Figure 3.4).

3.3.2 Process Blocks

Thanks to Julia offering faster computations than Python and easy-to-understand syntax [45], and Python offering a wide range of different packages (including Rheia), the main approach to connect the Rheia package and the Power System Adequacy Assessment simulation is to use both programming languages, the whole structure being programmed in Python and controlling Julia with the *julia* package. As it can be seen in the Figure 3.4, the data are exchanged thanks to *xlsx* files between Julia and Python.

PCE Initialization

The first step before entering the Julia environment is the initialization of the uncertainties for the number of desired profiles. Based on the design variable characteristics, a set of design samples is created through Latin Hypercube Sampling with Rheia e.g. if we have the uniform random variable u_{ttf} following the distribution $\mathcal{U}(0, 1)$, we can have $u_{ttf} = 0.518684$ for a profile and $u_{ttf} = 0.182747$ for the next profile. The number of profiles is defined with the expansion number $|\mathcal{A}^{M,p}|$ defined by:

$$|\mathcal{A}^{M,p}| = \frac{(p+M)!}{p! M!}$$

with p corresponding to the polynomial order and M is the stochastic dimension, i.e., the number of uncertainties. The total number of profiles is twice the expansion number to train the PCE; then, $n_{pro} = 2|\mathcal{A}^{M,p}|$ with n_{pro} , the number of different profiles. All the realized uncertainty values for each profile are stored in an Excel file called `PCE_data.xlsx` to be reused in the "Stochastic data generator" block (see "PCE Initialization" in Figure 3.4)

Read data, extract and store

This block is the first block being programmed in Julia. The purpose of this block is simply to read, extract and store the data from the input data discussed previously (Study Case Data and PCE data). It is operated only once in the process to provide the parameters of the study case for the other blocks in the Julia environment (see "Read data, extract and store" in Figure 3.4)

Model Initialization Jump

The purpose of this block is to initialize a first *static model*, based on the parameters stored in the block "Read data, extract and store". All the variables, constraints and the objective function are created based on the parameters (called here *static data* in the PCE process) not subject to uncertainties from the "Read data, extract and store" block. This block is computed once in the process to gain time not recreate the same variables, constraints and objective for the next profile (see "Model Initialization Jump" in Figure 3.4).

Stochastic data generator

The purpose of this block is to provide uncertainties for the deterministic model. Thanks to the realized uncertainty values for each profile and other useful *static*

data like the MTTF and the MTTR of every generator, line, etc., we can update the constraints for the next block. At first, the profile index k is set to 1 and every time this block is reached, new values of TTF, TTR and wind farm availability will be generated based on the corresponding profile k in `PCE_data.xlsx`. The loop ends when the profile index k exceeds the number of different profiles n_{pro} (see "Stochastic data generator" in Figure 3.4).

Update constraints + Optimization

In this block, based on the static model and the stochastic data we developed previously, it is possible to update the constraints of the deterministic model to include uncertainties in the model. The objective function is finally optimized, and it returns values of LOE, LOL and cost. The profile index k is incremented. The constraints based on the uncertainties will change for every profile, giving each time different values of LOE, LOL and cost (see "Update constraints + Optimization" in Figure 3.4).

Store Outputs

Every time the LOE, LOL and the cost from a different profile are found, they are stored in a Excel file called `Results_for_PCE.xlsx` which will be used in the next block (see "Store Outputs" in Figure 3.4).

PCE Computation

In this final block, the Julia environment and all the LOE, LOL and the cost of each profile are read, extracted and stored in Python thanks to the Excel file `Results_for_PCE.xlsx`. The PCE computation consists of finding the full PCE basis, determining the A matrix, getting the coefficient through OLS (Ordinary Least Squares), acquiring the statistics (mean and standard deviation), computing the Sobol Indices and the LOO (Leave-One-Out) error (see "PCE Computation" in Figure 3.4).

3.3.3 Outputs

Getting the PCE Statistics, the Sobol Indices and the LOO error allowed us to read crucial information about the accuracy of the PCE process. The PCE Statistics give information similar to the Monte-Carlo statistics, giving an approximate mean and standard deviation from the process. The Sobol Indices describe the contribution of each stochastic parameter to the variance of the model outputs and provide

valuable information on the model behavior under uncertainties [11]. The LOO (Leave-One-Out) error describes the error in prediction when a sample is removed from a trained model and used as a test sample. For a system to be viable, a LOO error less than 5% is a regular-used rule of thumbs ($< 30\%$ is considered acceptable in this thesis) (see "PCE Statistics, Sobol Parameters, LOO error" in Figure 3.4).

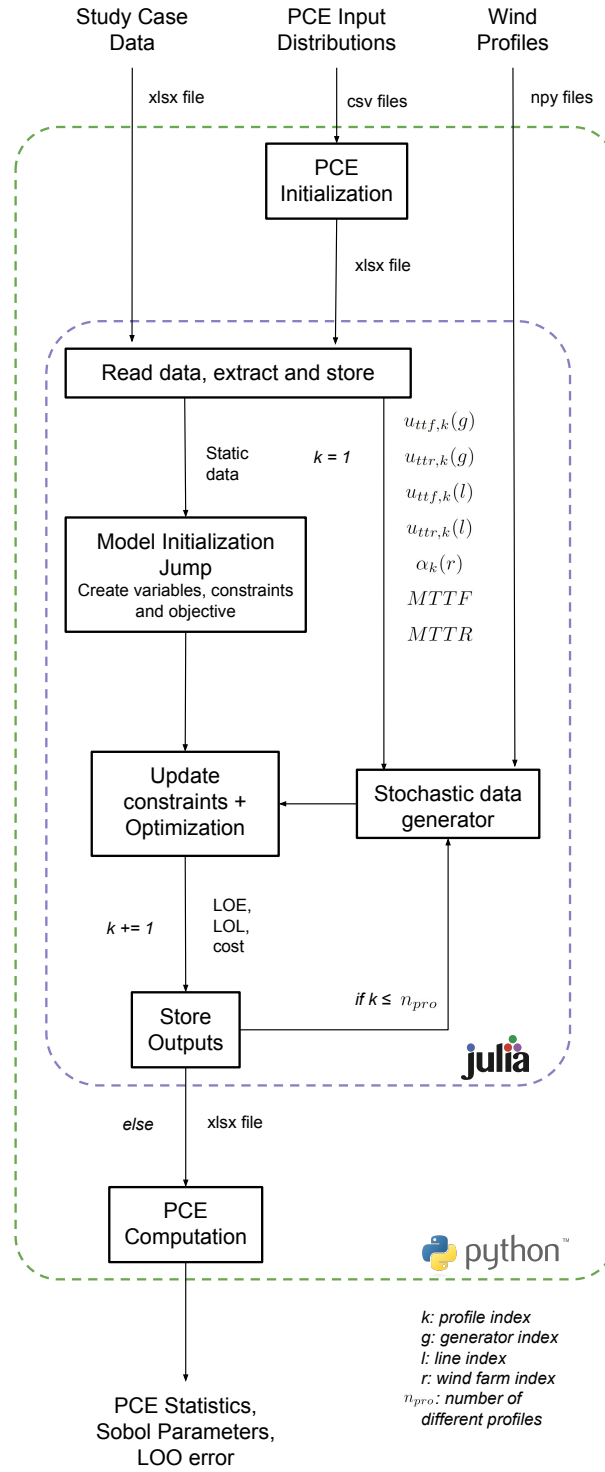


Figure 3.4: PCE Loop

Chapter 4

First results

4.1 Study Case

The study case used in this thesis is based on the IEEE 30-Bus System, whose data can be found on the Grainger College of Engineering (Information Trust Institute - University of Illinois) website [19]. According to this source, The IEEE 30-bus test case represents a simple approximation of the American Electric Power System as it was in December 1961. It is worth mentioning that the 30-bus test case does not have line limits initially. The system presents 2 active power generators, 4 synchronous condensers and 30 nodes and has been chosen instead of other more complex test cases (i.e., more nodes and units involved) to permit relatively fast computations for the Monte-Carlo and the PCE processes.

The main changes made on the IEEE 30-bus Test Case, to provide the study case used in this thesis, are the addition of offshore wind farms (based on the already built or in-project Belgian wind farms) on nodes 1, 2 and 13 (see Figure 4.1), the change of position of the two active power generators (from buses 1 and 2 to, respectively, buses 9 and 12) and their subdivision into 2 independent generators at the same node.

All the raw data is given in Appendix B, but general information about the study case will be provided here. It is worth saying that the only electricity generation sources are the conventional generators and the offshore wind turbines (thus, no importations, no solar power, no onshore wind power, no hydraulic power). In addition, there is no energy storage system (no pumped-storage hydroelectricity, no hydrogen technology, no battery) and other complexities in the system (no electrical vehicles, no load shifting, no load shedding, no heat pumps).

The Figure 4.1 gives a visual representation of the Study Case. The figure is largely inspired by [26], but a new copy has been made to offer better image resolution and to show the modifications from the initial test case.

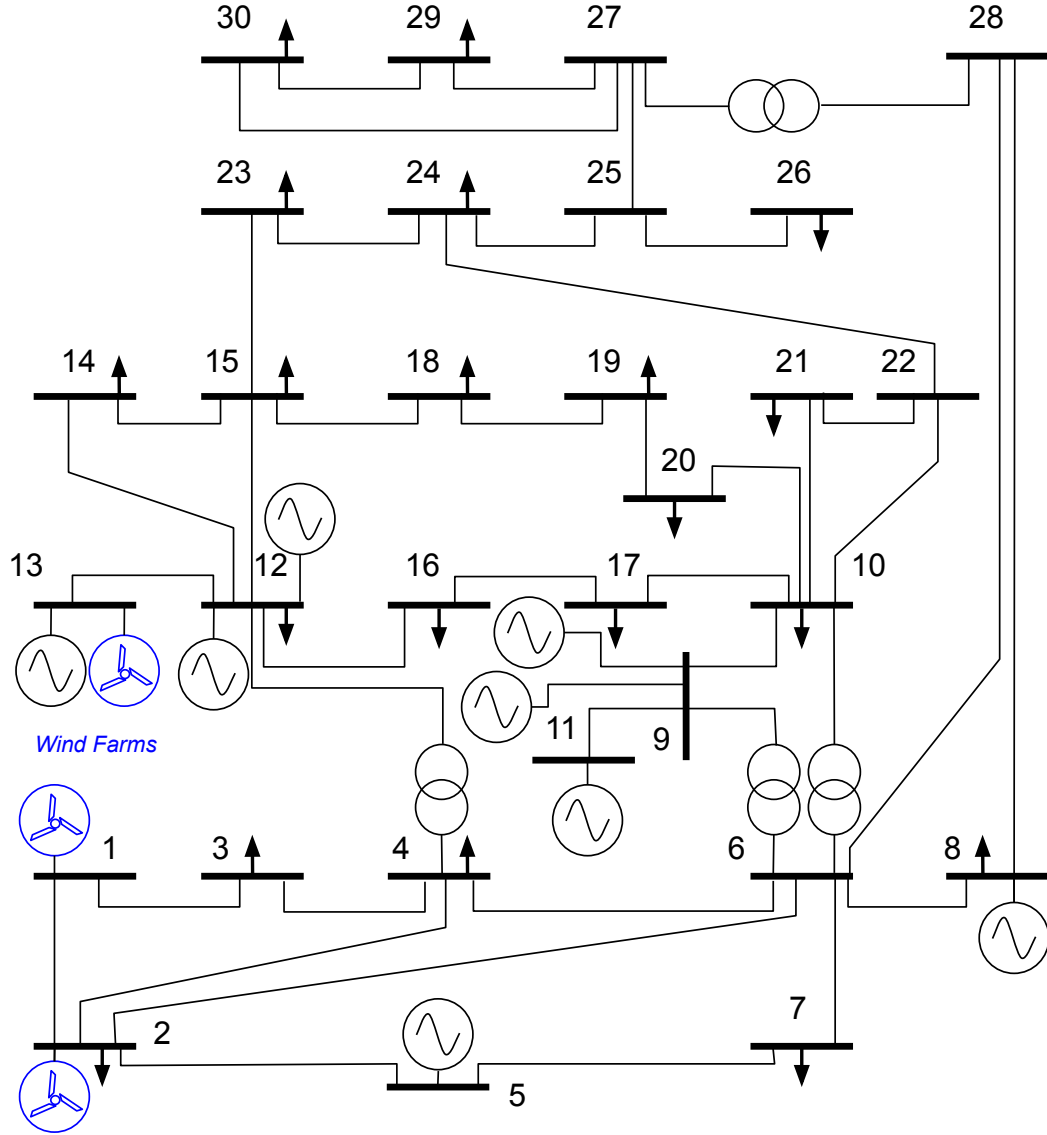


Figure 4.1: Study Case based on IEEE 30-bus Test Case (inspired by [26])

The general information about the Study Case is given in the following Table 4.1.

Table 4.1: General values of the Study Case

Parameter	Value
Value-Of-Loss-Load $VOLL$ [€/MWh]	10 000
Yearly Demand $load_{year}$ [TWh]	1.48955
Base Power for per-unit s_{base} [MVA]	100
Reference Bus	1

- **Value-Of-Loss-Load $VOLL$** (expressed in [€/MWh]): it describes how many euros by MWh must be paid when the demand is not satisfied at a certain node.
- **Yearly demand $load_{year}$** (expressed in [TWh]): the sum of the demand of each node for a whole year.
- **Base Power for per-unit s_{base}** (expressed in [MVA]): Base power of the per-unit system
- **Reference Bus:** Reference node for which the voltage angle $\Theta = 0^\circ$

The demand per hour is variable over time and follows the following arbitrary rule:

$$load(t) = \frac{load_{year}}{8760} (1 + 0.4\sin(0.1t) + 0.15\sin(0.6t))$$

giving a maximum possible load value of $1.55 \frac{load_{year}}{8760}$ and a minimum possible load value of $0.45 \frac{load_{year}}{8760}$.

Buses

The Buses (also called "Nodes" in the model context) are the points connected to the lines, the generators, the wind farms. Most of the buses have a non-zero energy demand (described as a load share $d_l(n)$ (in %) of the entire load for one hour $load(t)$ (in MWh)) which is represented in the Figure 4.2. Each bus also has a certain type (PQ Bus, PV Bus or Slack Bus), a certain active power injection p_d (in MW), a certain reactive power injection q_d (in MVAR), a certain conductance g_s (in pu), a certain susceptance b_s (in pu) and a certain base KV described in the Appendix B. Each node is in the same area; the voltage magnitudes v_m (in pu), the voltage angles v_a (in degrees), the maximum voltage v_{max} (in pu) and the minimum voltage v_{min} (in pu) are the same for each bus, in per unit (respectively, $v_m = 1pu$, $v_a = 0^\circ$, $v_{max} = 1.06pu$ and $v_{min} = 0.94pu$)

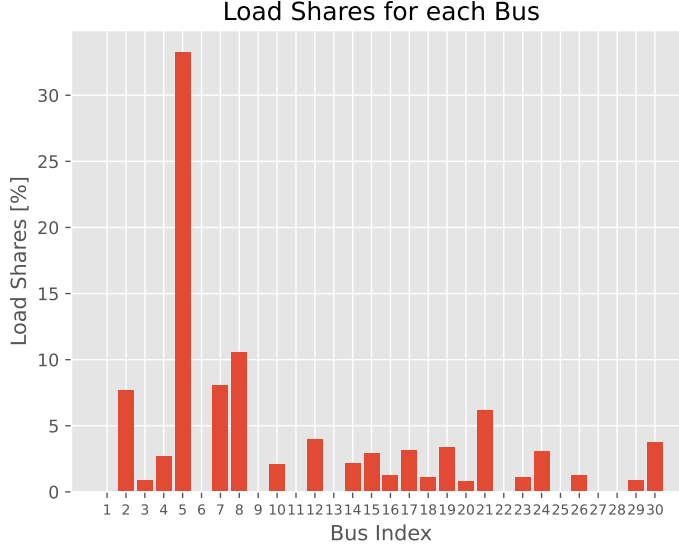


Figure 4.2: Buses Load

Branches

The Branches (also "Lines") are described by the two nodes they are connecting. The lines connecting a generator of a wind farm to a node, and the lines connecting a node to the demand source are considering as ideal and, thus, not taken into account. A line includes several parameters used in the Power Adequacy Assessment simulation as the resistance r (in MVA), the reactance x (in MVA), the ratio (if there are transformers on the line) (in %), the power capacity of the line f_{max} (in MW). Those parameters are described precisely for each line in the Appendix B. Each line has a maximum change of phase angle clockwise ang_{min} (in degrees) and counterclockwise ang_{max} (in degrees) of $ang_{min} = -30^\circ$ and $ang_{max} = 30^\circ$. The MTTF (in hours) and the MTTR (in hours) are also always identical, with respective values of 20 000 hours and 24 hours.

Generators

The generators can be divided into two categories, the conventional generating units (producing non-zero active power, i.e. $p_{max} > 0$) connected to buses 9 and 12 and the synchronous condensers (producing zero active power, i.e. $p_{max} = 0$) connected to buses 5, 8, 11 and 13. Each generator has its active power range limit between p_{min} (in MW) and p_{max} (in MW), its quadratic marginal cost c_2 (in €/MWh²), its linear marginal cost c_1 (in €/MWh) and its fixed cost c_0 (in

€) described in Table 4.2. Each generator has also a reactive power range limit between q_{min} (in MVAR) and q_{max} (in MVAR), a generator voltage v_g (in pu) (always $v_g = 1pu$), and fixed MTTF (in hours) and MTTR (in hours) (respectively 2000 hours and 50 hours). The other parameters, less important, are given in Appendix B.

Table 4.2: Generators Main Parameters

Bus	p_{max} [MW]	p_{min} [MW]	c_2 [€/MWh ²]	c_1 [€/MWh]	c_0 [€]
9	180.67	0	0	18.42	0
9	180.67	0	0	18.42	0
12	61.33	0	0	52.18	0
12	61.33	0	0	52.18	0
5	0	0	0	0	0
8	0	0	0	0	0
11	0	0	0	0	0
13	0	0	0	0	0

Offshore Wind Farms

As explained previously, the Wind Farms were not initially present in the IEEE 30-Bus test case initially, they were placed on buses 1, 2, and 13, as can be observed in Table 4.3. Each wind farm has a maximum active power p_{max} , corresponding to the situation where the wind is optimal and where every wind turbine of the wind farm is working. Each wind farm has a number of wind turbines n_{wt} , and each wind turbine has the same MTTF (in hours) and MTTR (in hours) of 1060 hours and 290 hours, respectively.

Table 4.3: Wind Farms Main Parameters

Name	Bus	p_{max} [MW]	n_{wt}
Belwind	2	16.099	56
C-Power	2	33.158	54
Mermaid	2	22.143	28
Nobelwind	2	15.534	50
Norther	2	37.338	44
Northwester2	2	20.570	23
Northwind	2	22.369	72
Rentel	2	29.062	42
Seastar	2	23.724	30
P1-3500-15	13	72.8	49

P23-3500-15	1	278.2	185
-------------	---	-------	-----

4.2 Results with Monte-Carlo

To apply the Monte-Carlo method to our study case, we consider the u_{ttf} and the u_{ttr} of the generators, lines and wind turbines as evaluated parameters from a uniform random distribution $\mathcal{U}(0, 1)$. A new u_{ttf} or u_{ttr} is generated if, respectively, the TTR or the TTF of a generator, line or wind turbine is completed. The wind profiles are also chosen randomly at each new Monte-Carlo year (from 100 different wind profiles).

The Monte-Carlo simulation ends when the condition of the Eq. 4.1 is satisfied (see Section 2.2).

$$\epsilon_{incr,rm} \leq \epsilon_{incr,threshold} \quad (4.1)$$

while $n_{year} \leq n_{limit}$ (n_{year} being the number of Monte-Carlo years and n_{limit} being the maximum number of Monte-Carlo years (here $n_{limit} = 2000$)), is respected, where $\epsilon_{incr,threshold}$ is the maximum error tolerance for Monte-Carlo (in our case $\epsilon_{incr,threshold} = 0.001$).

With all those conditions, we have the following results in Table 4.4, which will have a value of Gold Standard throughout the thesis. The empirical probability distribution functions of the LOL and the LOE are represented in Figures 4.3a and 4.3b respectively. The LOLE and LOEE values are given by the red dotted lines and are $LOLE = 12.87$ hours/year and $LOEE = 105.56$ hours/year respectively.

Table 4.4: Monte-Carlo Main Results

Name	Values
Number of Monte-Carlo years n_{year}	198
LOLE [hours/year]	12.89
LOEE [MWh/year]	105.56
Mean cost [€/year]	$1.037 \cdot 10^7$
LOL standard deviation [hours/year]	0.297
LOE standard deviation [MWh/year]	24.54
Cost standard deviation [€/year]	$2.949 \cdot 10^6$

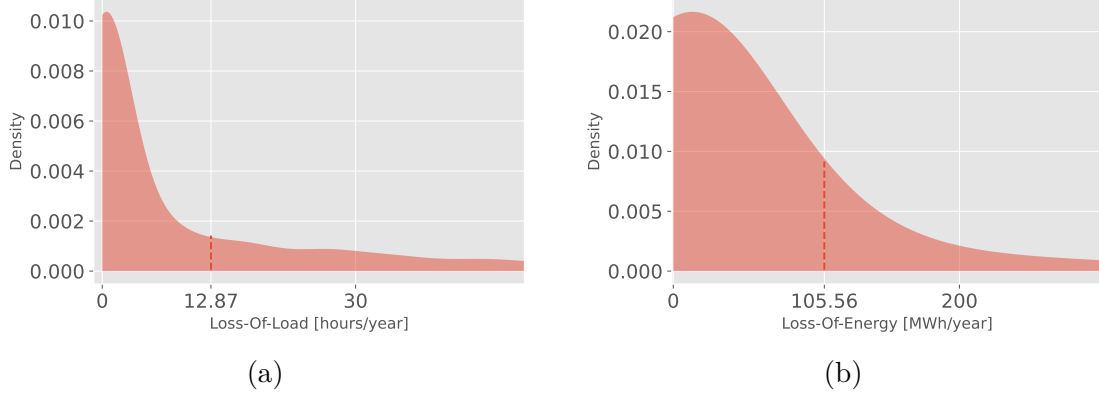


Figure 4.3: Kernel Distribution Estimate (KDE) of the LOL and LOE

4.3 Results with PCE

The PCE method is applied to the Study Case. A first assumption is to consider the TTF and TTR of a generator, line, etc. as constant for the entire simulated year. The reason is because the number of uncertainties must stay the same for all the PCE process. In the Monte-Carlo process, new u_{ttf} and u_{ttr} are generated at each repair/failure of a generator, line, etc., giving a variable number of u_{ttf} and u_{ttr} at each simulation. In the PCE process, the u_{ttf} and the u_{ttr} are given in advance of the year simulation and remain identical for the whole year, so only 2 variables (u_{ttf} and u_{ttr}) are necessary to describe the whole availability profile of a generator, line, etc.

A second assumption is to consider only 10 uncertainties. Those uncertainties are Random Variables for PCE and are given in Table 4.5. The main reason to consider only 10 uncertainties and no more is because of the curse of dimensionality (see subsection 2.3.4) leading to a large number of unknown coefficients, even using a low-order PC expansion. The first eight uniform random variables have been chosen because they determine the TTF and the TTR of the 4 conventional generating units from the Study Case (each generator with $p_{max} > 0$, none of the synchronous condensers). It is easy to observe that they have the highest impact on the outputs (LOL, LOE and cost), as, in case of an outage, a generator like generator 1 can reduce the maximum total power of the generators from 484MW to 303.33MW (37.3% reduction). The availability of the two more important wind farms (P1-3500-15 and P23-3500-15), in terms of maximum power p_{max} , has also been taken into account to consider uncertainties from the wind farms. The other wind farms have a constant availability rate of 0.804 (see subsection 3.2.2) to keep a reasonable number of random variables for PCE. All the lines and all the

other elements normally subject to uncertainties in the Monte-Carlo process are considered as ideal ($MTTF \rightarrow \infty$ and $MTTR = 0$).

All the assumptions are summarized:

- Instead of setting uncertainties on the TTF and TTR of each wind turbine in each wind farm, a global availability α_{wt} is used to describe the percentage of working wind turbines for one wind farm.
- As opposed to Monte-Carlo, the TTR and the TTF of each generator are fixed as constants for a whole simulated year. This helps reduce the uncertainty on the number of different TTR and TTF for a whole simulated year, as new TTR and new TTF were generated after each failure/repair in the Monte-Carlo process.
- Instead of working with 100 different wind profiles in the Monte-Carlo process, 1 wind profile is considered for the entire PCE process to compare two different samples more efficiently.
- 10 uncertainties are considered: the TTF and TTR of the 4 generators with $p_{max} > 0$, based on the random uniform variables u_{ttf} , u_{ttr} and the availability α_{wt} of the two wind farm having the highest p_{max} , based on a normal distribution.
- The lines and synchronous condensers are considered as ideal ($MTTF \rightarrow \infty$ and $MTTR = 0$) considering their mean availability at 99.88% and 97.56% ($MTTF/(MTTF + MTTR)$). The availability of the other wind farms are considered constant on the average, thus always 80.4% of availability (see subsection 3.2.2).

The main goal of those assumptions was to limit the number of uncertainties, which can rapidly require an unmanageable number of samples because of the curse of dimensionality (see subsection 2.3.4). The 10 uncertainties have been chosen in order to match the results obtained from the Monte-Carlo process as accurately as possible.

Table 4.5: Random Variables for PCE

Name	Variables	Distribution
Uniform random variable for gen. 1 TTF	$u_{ttf,gen1}$	$\mathcal{U}(0, 1)$
Uniform random variable for gen. 1 TTR	$u_{ttr,gen1}$	$\mathcal{U}(0, 1)$

Uniform random variable for gen. 2 TTF	$u_{ttf,gen2}$	$\mathcal{U}(0, 1)$
Uniform random variable for gen. 2 TTR	$u_{ttf,gen2}$	$\mathcal{U}(0, 1)$
Uniform random variable for gen. 3 TTF	$u_{ttf,gen3}$	$\mathcal{U}(0, 1)$
Uniform random variable for gen. 3 TTR	$u_{ttf,gen3}$	$\mathcal{U}(0, 1)$
Uniform random variable for gen. 4 TTF	$u_{ttf,gen4}$	$\mathcal{U}(0, 1)$
Uniform random variable for gen. 4 TTR	$u_{ttf,gen4}$	$\mathcal{U}(0, 1)$
Normal random variable for P1-3500-15 avail.	$\alpha_{wt,P1}$	$\mathcal{N}(0.804, 0.058)$
Normal random variable for P23-3500-15 avail.	$\alpha_{wt,P23}$	$\mathcal{N}(0.804, 0.058)$

With this choice of random variables for PCE, the following results are observed for the first 3 orders in the following Tables 4.6, 4.7 and 4.8 and Figures 4.4, 4.5, 4.6, 4.7, 4.8 and 4.9.

As an order 4 PCE will require the use of 2002 different profiles (more than n_{limit} , the maximum number of Monte-Carlo years for the Monte-Carlo methods), this order will not be discussed, nor will the orders above.

A Leave-One-Out error considered as acceptable is highly problem-dependent and depend on the number of considered points [38]. As reference for the Adequacy Assessment problems, an acceptable LOO error is set arbitrarily at 5% (tolerable until 30% in this thesis). In each of the three cases (and for each of the outputs), the Leave-One-Out error is too high compared to the acceptable LOO error we set at 5%. Those results clearly show that the PCE method cannot be used without fundamental changes to the method that will be investigated in the next section.

The first source of high error is because the LOL and LOE with the Monte-Carlo process cannot be negative, thing that can be possible with the PCE process, considering the proximity with the negative limit. A solution could be to define a "potential LOL" (negative LOL) and "potential LOE" (negative LOE) instead of having a zero LOL and LOE. The solution developed in section 5.1 is to develop the monotonic curve from a simulation and extrapolate the curve to find those potential LOL and potential LOE.

The second source of high error is the time-dependency of the problem. A generator, line, etc. can be, randomly, set as failing or working at the beginning of a simulated year. The solution to deal with this problem, developed in section 5.2, is to work with multiple "subprofiles" for a profile. Those subprofiles share the same evaluated random variables values but have different startup time between them.

Those solutions are explored in section 5.

Table 4.6: Order 1 PCE Main Results

Name	Values
Number of profiles/full basis n_{pro}	22
LOL mean [hours/year]	2.42
LOL standard deviation [hours/year]	6.02
LOE mean [MWh/year]	56.52
LOE standard deviation [MWh/year]	101.77
Cost mean [€/year]	$4.252 \cdot 10^6$
Cost standard deviation [€/year]	$1.074 \cdot 10^6$
Leave-One-Out error (LOL) [%]	124.53
Leave-One-Out error (LOE) [%]	102.75
Leave-One-Out error (Cost) [%]	97.66

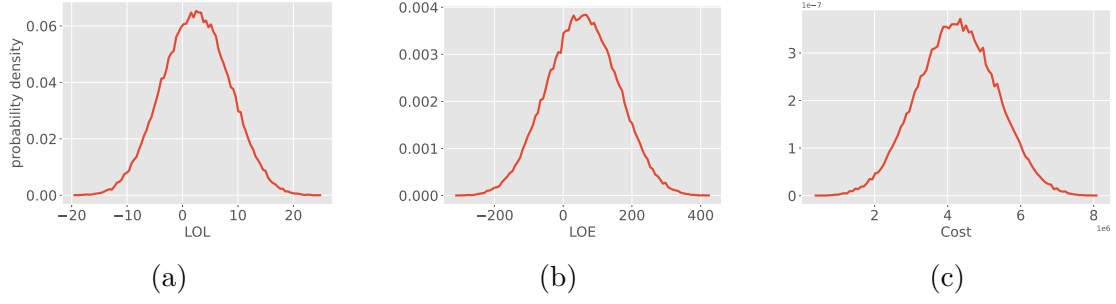


Figure 4.4: PDFs for outputs (order 1), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

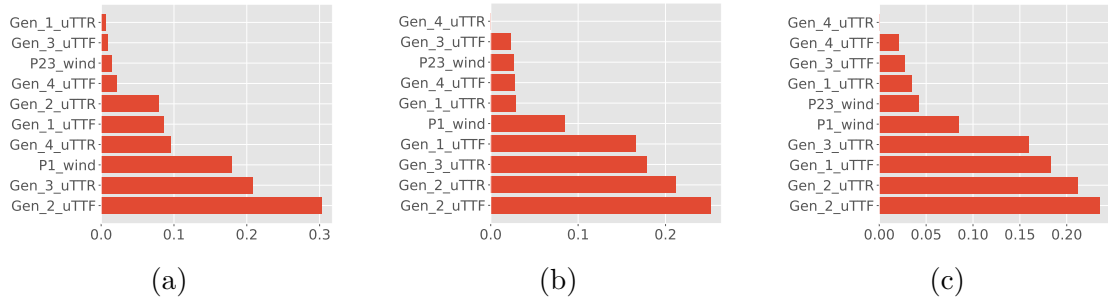


Figure 4.5: Sobol indices for outputs (order 1), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

Table 4.7: Order 2 PCE Main Results

Name	Values
Number of profiles/full basis n_{pro}	132
LOL mean [hours/year]	1.605
LOL standard deviation [hours/year]	16.99
LOE mean [MWh/year]	20.56
LOE standard deviation [MWh/year]	392.38
Cost mean [€/year]	$4.246 \cdot 10^6$
Cost standard deviation [€/year]	$2.368 \cdot 10^6$
Leave-One-Out error (LOL) [%]	137.12
Leave-One-Out error (LOE) [%]	171.43
Leave-One-Out error (Cost) [%]	116.2

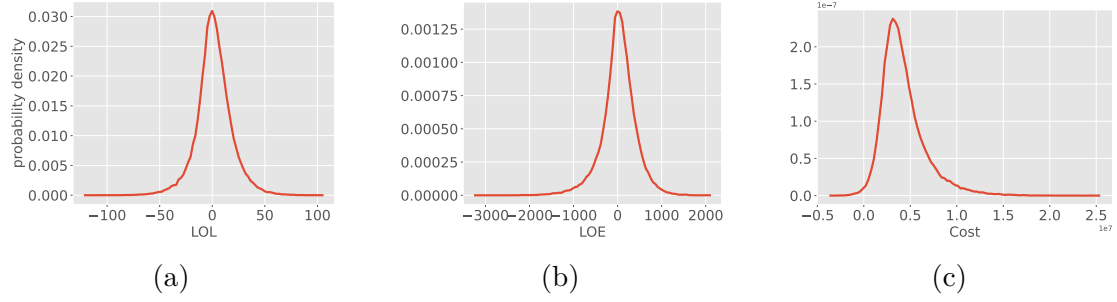


Figure 4.6: PDFs for outputs (order 2), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

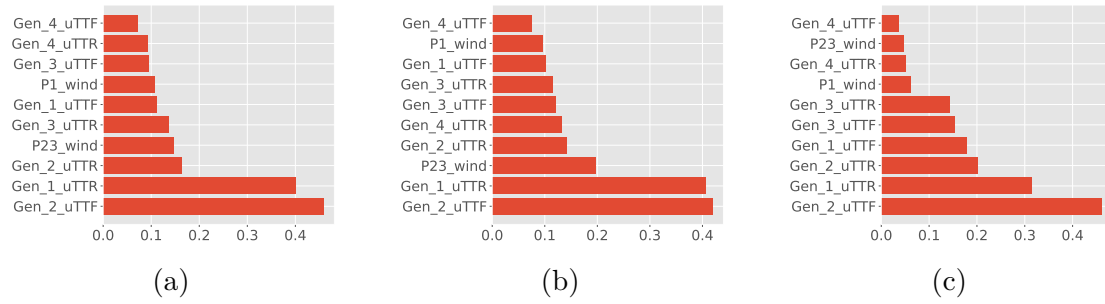


Figure 4.7: Sobol indices for outputs (order 2), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

Table 4.8: Order 3 PCE Main Results

Name	Values
Number of profiles/full basis n_{pro}	572
LOL mean [hours/year]	3.003
LOL standard deviation [hours/year]	13.04
LOE mean [MWh/year]	75.53
LOE standard deviation [MWh/year]	449.9
Cost mean [€/year]	$4.42 \cdot 10^6$
Cost standard deviation [€/year]	$4.588 \cdot 10^6$
Leave-One-Out error (LOL) [%]	94.29
Leave-One-Out error (LOE) [%]	128.52
Leave-One-Out error (Cost) [%]	126.1

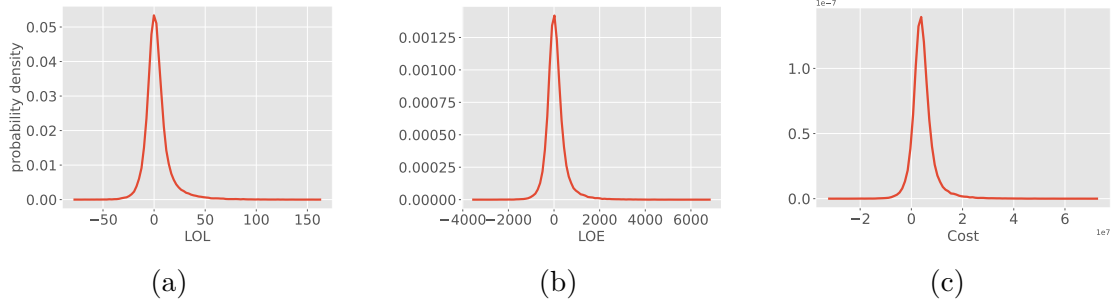


Figure 4.8: PDFs for outputs (order 3), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

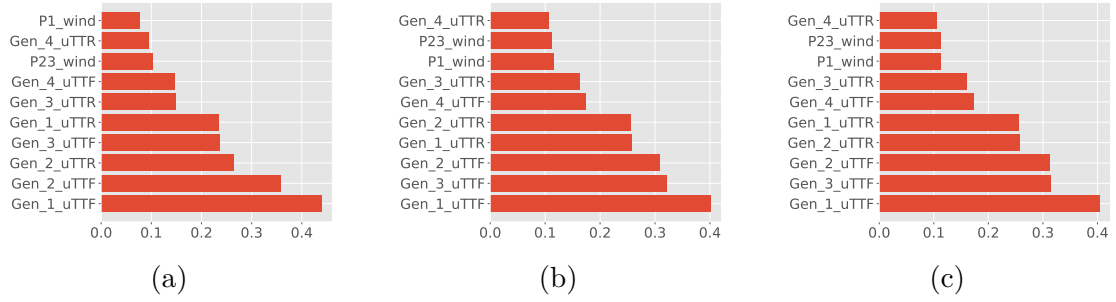


Figure 4.9: Sobol indices for outputs (order 3), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

Chapter 5

Continuous output variables, subprofiles and post-processing

In this chapter, some areas of improvement to make the PCE method less erroneous with a reasonable number of profiles will be discussed. The main developed areas of improvement and their limitations will be developed in the next sections. Finally, an analysis of the best result, according to the LOO error, is provided in the last section.

Before diving directly into the new methods and their results, it is worth nothing that it is possible to divide the possible modifications into 3 groups:

- *The change of the inputs* e.g., changing the values of uncertainty feeding the simulations.
- *The change of the outputs* e.g., changing the extracted values of LOE, LOL, and/or cost.
- *The change of the PCE method* e.g., using another PCE method than the standard method, e.g., Multi-element generalized polynomial chaos method that could divide the output space into different subspaces to adapt more to discontinuities or a Sparse PCE to speed-up the PCE computation, relying on only a few model evaluations.

The first method developed in Section 5.1 is based on a change of the outputs.

5.1 Transformation into continuous output variables

An important problem concerning the outputs LOE and LOL is the fact that those values cannot be negative ($LOL \geq 0$ and $LOE \geq 0$) and that we are performing on

the negative limit, making the PCE method difficult to use as it does consider the existence of negative LOL and LOE . One could think of using the cost instead of the LOE or LOL as an objective of interest for PCE, but, as the three values are linked, if the LOE or LOL are strictly positive ($LOE > 0$ and $LOL > 0$), the cost will be orders of magnitude higher than the case where the LOE and LOL are zero because of the high value of $VOLL$ (Value-Of-Loss-Load) [€/MWh].

A solution to address this problem of negative LOE and LOL could be to make them possible outputs for PCE computation, then, create *negative LOE* and *negative LOL* . A first solution could be to work with *monotonic curves*.

5.1.1 Monotonic Curve

A monotonic curve is a curve of maximum available power at a certain hour t subtracted by the demand at this same hour t . The result values are sorted to make a power versus time graph, describing a descending monotonic curve. Mathematically, it can be expressed as:

$$monotonic(t) = sorted(p_{avail,max}(t) - load(t)) \quad (5.1)$$

More specifically, in our study case, it is expressed as:

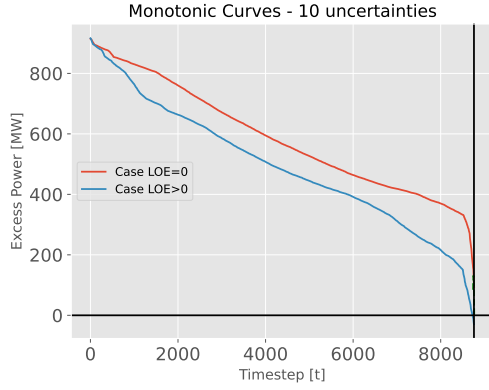
$$monotonic(t) = sorted\left(\sum_{g=1}^{n_g} p_{avail,max}(g, t) + \sum_{r=1}^{n_{wt}} res(r, t) - load(t)\right) \quad (5.2)$$

It is worth noting that the expression of the monotonic curve could be more complex considering more complex study cases, e.g., study cases with batteries or other systems of storage.

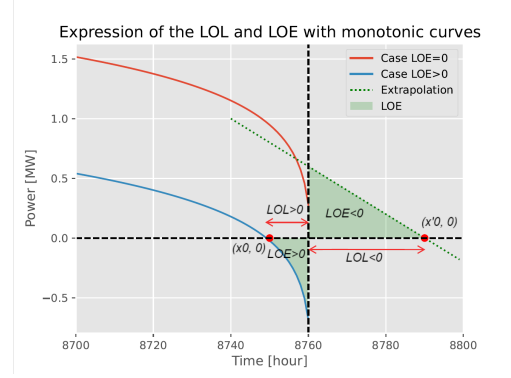
Once plotted (see Figure 5.1a), two different cases can be observed; the case where $LOE = 0$ and the case where $LOE \geq 0$. For the case where $LOE \geq 0$, the LOE and the LOL can be found. The LOE is the negative integral of the curve for the interval between the intersection with the x-axis and the last hour ($t = 8760$) (see Eq. 5.3) and the LOL is the difference between the last hour ($t = 8760$) and the hour where the curve intersects the x-axis (see Eq. 5.4 and 5.1b).

$$LOE = - \int_{x_0}^{8760} monotonic(t) dt \quad (5.3)$$

$$LOL = 8760 - x_0 \quad (5.4)$$



(a) Monotonic Curves



(b) LOE and LOL with Monotonic Curves (zoom)

Keeping the same idea, when the $LOE = 0$, the associated monotonic curve is extrapolated until this extrapolated curve intersects the x-axis. The negative LOE is the negative integral of the extrapolation for the interval between the last hour ($t = 8760$) and the intersection with the x-axis (see Eq. 5.5) and the negative LOL is the difference between the hour when the extrapolation intersects the x-axis and the last hour ($t = 8760$) (see Eq. 5.6). A graphical explanation is given in Figure 5.1b.

$$LOE = - \int_{8760}^{x'_0} \text{monotonic}_{\text{extrapolation}}(t) dt \quad (5.5)$$

$$LOL = -(x'_0 - 8760) \quad (5.6)$$

5.1.2 Extrapolation

The next question concerns how to perform this extrapolation. The more natural way consists of choosing a common function, e.g., a linear function or quadratic function, applying a regression on a certain number of last points n_{extra} and, then, extrapolate. As there is no general rule, the function and the number of last points giving the lower LOO error determine the final choice.

One can think of using tangent extrapolation. The problem with this method is that, even with two points next to each other on the curve, depending on which point we apply the tangent, it can give relatively highly different LOEs. The reason is that, in discrete functions with a constant interval in time, the derivative is given by the difference in power, which can be randomly important or not between two neighbor points.

Linear Extrapolation

The linear extrapolation is based on a linear regression on the n_{extra} last points of the monotonic curve. This number n_{extra} ($n_{extra} \geq 2$) is chosen arbitrarily according to the general shape of the monotonic curves. Taking a value too low might make the extrapolation too sensitive to point-to-point fluctuations, i.e., see the tangent extrapolation problem. On the other hand, taking a value too high might not translate efficiently the behavior of the curves in the last few hours. A trade-off can be found according to the LOO error found for different n_{extra} .

Based on the n_{extra} last points of the monotonic curve, we obtain the following linear regression:

$$y = at + b \quad (5.7)$$

The LOL is computed as the difference between the intersection of this line and the x-axis:

$$LOL = -(x'_0 - 8760) = -(\frac{-b}{a} - 8760) \quad (5.8)$$

The LOE is computed as the negative integral below the extrapolated part of the monotonic curve:

$$\begin{aligned} LOE &= - \int_{8760}^{x'_0} \text{monotonic}_{extrapolation}(t) dt \\ &= - \int_{8760}^{-\frac{b}{a}} (at + b) dt \\ &= \frac{b^2}{2a} + 8760(\frac{8760a}{2} + b) \end{aligned} \quad (5.9)$$

Quadratic Extrapolation

The quadratic extrapolation is based on a quadratic regression on the n_{extra} last points of the monotonic curve. This number n_{extra} ($n_{extra} \geq 3$) is chosen arbitrarily according to the general shape of the monotonic curves. As for the linear regression, taking a value too low might make the extrapolation too sensitive to point-to-point fluctuations, i.e., see the tangent extrapolation problem. On the other hand, taking a value too high might not translate the behavior of the curves efficiently, in the last few hours. A trade-off can be found according to the LOO error found for different n_{extra} .

Based on the n_{extra} last points of the monotonic curve, we obtain the following quadratic regression:

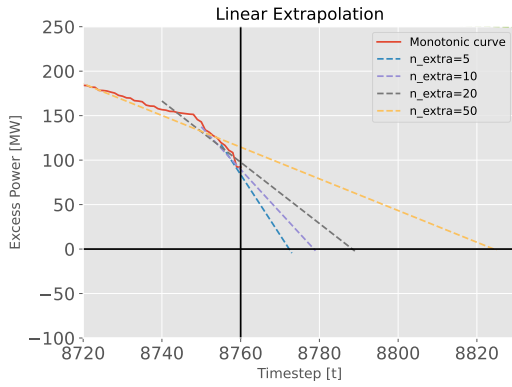
$$y = at^2 + bt + c \quad (5.10)$$

with the constraint that a must be negative to keep the extrapolated part decreasing monotonically. The LOL is computed as the difference between the intersection of this line and the x-axis:

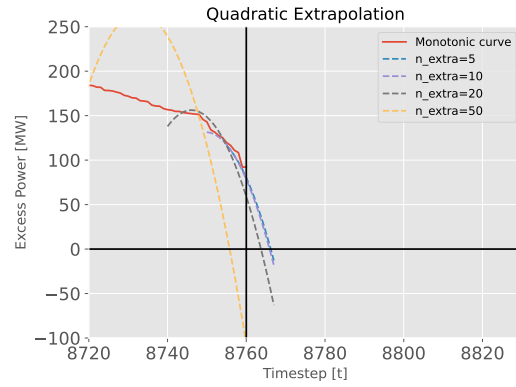
$$LOL = -(x'_0 - 8760) = -\left(\frac{-b - \sqrt{b^2 - 4ac}}{2a} - 8760\right) \quad (5.11)$$

The LOE is computed as the negative integral below the extrapolated part of the monotonic curve (with $x'_0 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$):

$$\begin{aligned} LOE &= - \int_{8760}^{x'_0} \text{monotonic}_{\text{extrapolation}}(t) dt \\ &= - \int_{8760}^{x'_0} (at^2 + bt + c) dt \\ &= \frac{a}{3}(x'^3_0 - 8760^3) + \frac{b}{2}(x'^2_0 - 8760^2) + c(x'_0 - 8760) \end{aligned} \quad (5.12)$$



(a) Linear Extrapolation



(b) Quadratic Extrapolation

Integration in the PCE process

Figure 5.3 shows the integration of those extrapolations in the general scheme of Figure 3.4. The "Negative LOE and LOL generator" block is located in the Python environment. It takes as input a text file with all the monotonic curve values (power over time sorted) for a certain profile. Each profile generates an additional text file. The block also takes, as input, the Excel (xlsx) file containing the outputs. This block changes the values LOE and LOL when their initial value is zero to change them toward a negative value, thanks to the associated monotonic curve. The output of the block is a modified Excel file with positive and negative LOE and LOL, and those values are transferred to the "PCE Computation" block.

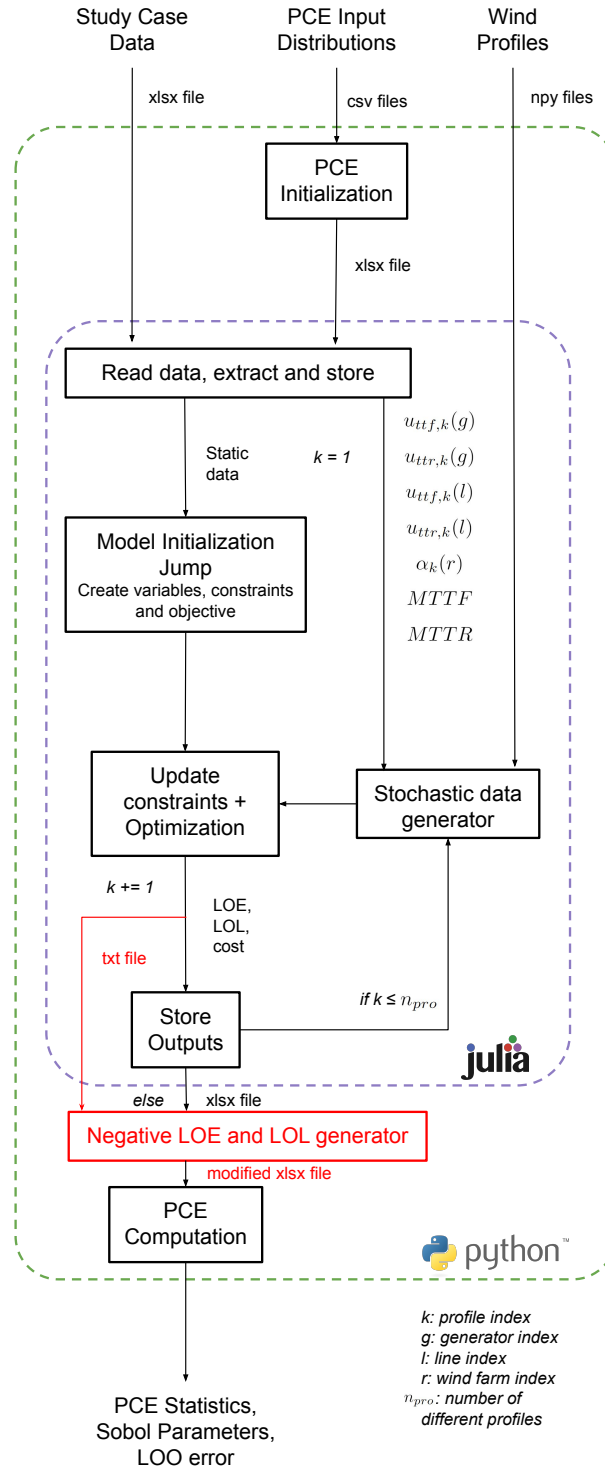


Figure 5.3: Monotonic PCE Loop

5.1.3 Results

This subsection gives the result of the presented methodology with the two different extrapolations (linear in Table 5.1 and quadratic in Table 5.2), for various numbers of last points n_{extra} (order 3 PCE process with the 10 random variables for PCE given in section 4.3). Let us note that to avoid inconvenience from almost-flat monotonic curves for the last hours, upper bounds of $a = -3$ and $a = -0.5$ are fixed for, respectively, the linear and the quadratic extrapolation.

Table 5.1: Order 3 PCE Linear Extrapolation Main Results

Number of last points n_{extra}	LOO error (LOE) [%]
5	87.71
10	154.99
20	130.38
50	88.77

Table 5.2: Order 3 PCE Quadratic Extrapolation Main Results

Number of last points n_{extra}	LOO error (LOE) [%]
5	130.43
10	120.22
20	117.7
50	127.65

It is observed that, regardless of n_{extra} and if the extrapolation is linear or quadratic, the LOO error remains high ($> 85\%$) for each test. This method alone is, thus, not sustainable for computing PCE. It is mainly because, even if the LOE the LOL can now be negative value to make the outputs continuous, other uncertainties like the start up time must be considered. Two generators having the same TTF and TTR for the whole year might produce LOE or not depending on when the failure occurs, e.g., if it is during a high peak of wind power production or during a low peak of wind power production.

5.2 Subprofiles

5.2.1 Method

One problem with the last methodology is its dependency on the moment the generators have an outage. Indeed, for the same profile, depending on when we start the time, e.g., in the first hour of January or in the 45th hour of September,

the LOE, LOL and cost outputs can have different values. This problematic is explained graphically in Figures 5.4a and 5.4b. If a generator is down during a peak of wind power production, this outage has no repercussions on the LOE. On the other hand, if this generator (with the same TTF and TTR) is down during a drop in wind power production, this outage no longer guarantees that demand will be met, and this generates LOE.

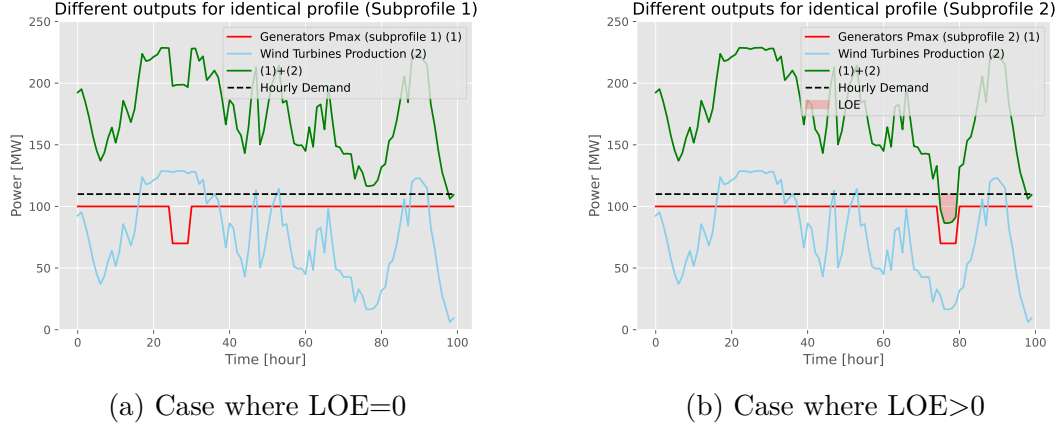


Figure 5.4: Different subprofiles for identical profiles

A solution to address this problem can be to consider, for a certain profile, a certain number n_{sp} of *subprofiles*. Between two *subprofiles* of the same profile, the TTF, the TTR and the availability rates of generators or wind farms would be identical but the sole difference is the moment the startup is set. This startup is different for each generator, and the whole considered time is always 8760 hours. Having the n_{sp} subprofiles for each profile, we can apply the extrapolation method for zero LOE and LOL and take the average LOE, LOL and cost of those subprofiles to create the profile output. Different generator profiles are plotted in Figure 5.5 to show the difference between different subprofiles for one profile.

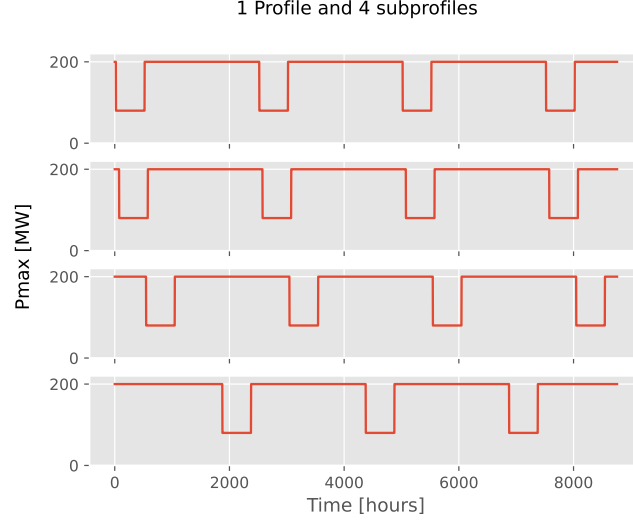


Figure 5.5: 4 different subprofiles for 1 profile

Figure 5.6, based on the Figure 3.4, shows the added blocks to make the whole new process. The "Random startup hours" create an availability profile for each generator over 5 years and pick a startup hour randomly between the first hour of the 1st year and the last of the 4th year to be able to perform on a whole year. This block is operated after the "Stochastic data generator" block and after the "Store Outputs" if the number of subprofiles k_s for a profile is not equal to the subprofile fixed for each profile n_{sp} .

It is worth noting that the monotonic curve data of each subprofile is stored in a text file for the block "Negative LOE and LOL generator". After this block, the outputs mean is extracted from all the subprofiles according to the Eq. 5.13 (also correct for LOL and cost).

$$LOE_k = \frac{1}{n_{sp}} \sum_{k_s=1}^{n_{sp}} LOE_k(k_s) \quad (5.13)$$

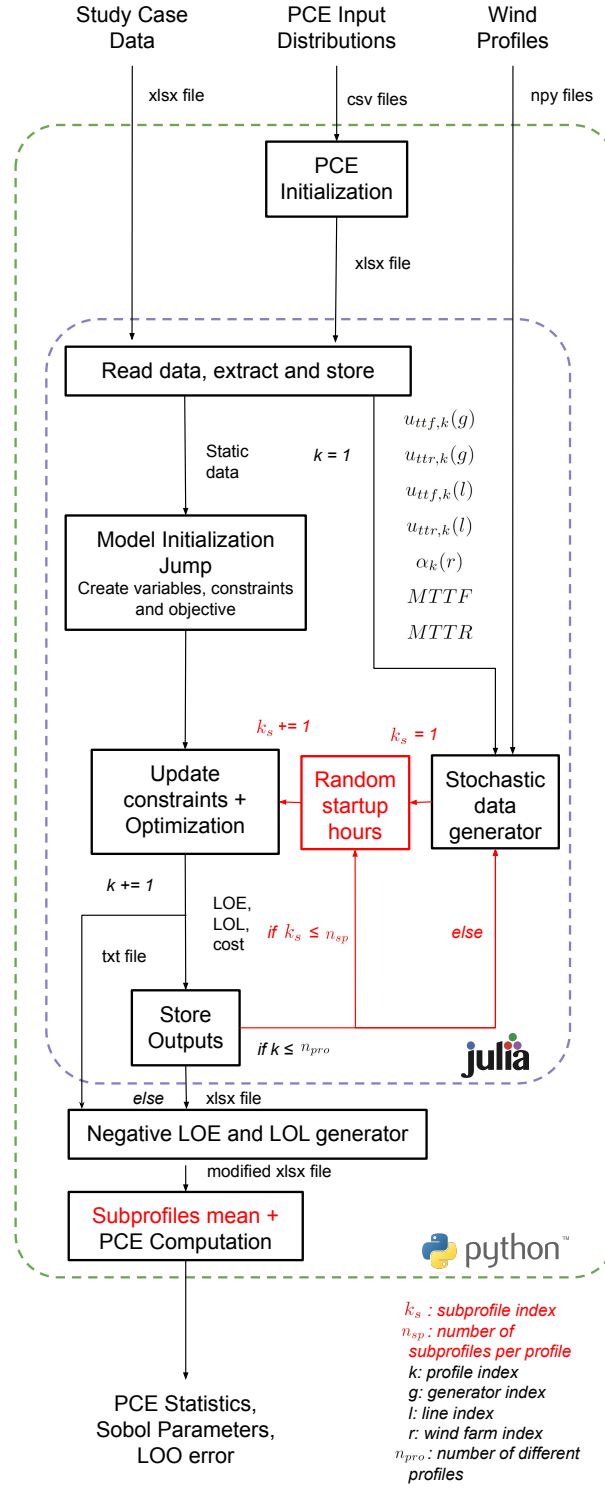


Figure 5.6: Subprofile PCE Loop

5.2.2 Results

In this subsection, the results for various orders and numbers of subprofiles are given for one uncertainty ($u_{ttf,gen1}$) in Table 5.3 and then, for 10 uncertainties in Table 5.4 (same uncertainties as in section 4.3). The primary reason to show the results for one uncertainty is to show the range of possible LOE for the same profile and to show the evolution of the LOE according to the value of $u_{ttf,gen1}$. Let us note that the yearly demand has been increased by a factor 1.67 for the case with 1 uncertainty to pressurize the system and have positive and negative LOE. The number of simulations is given by $n_{simu} = n_{pro} * n_{sp}$. The best extrapolation (linear or quadratic and n_{extra}) w.r.t. the LOO error has been chosen for the case with 1 uncertainty and the case with 10 uncertainties.

Table 5.3: PCE Quadratic Extrapolation ($n_{extra} = 10$) with Subprofiles (1 uncertainty) Main Results

PCE order p	n_{sp}	n_{simu}	LOO error (LOE) [%]
2	2	12	94.51
2	3	18	78.49
2	4	24	48.48
3	2	16	8.96
3	3	24	13.52
3	4	32	11.49

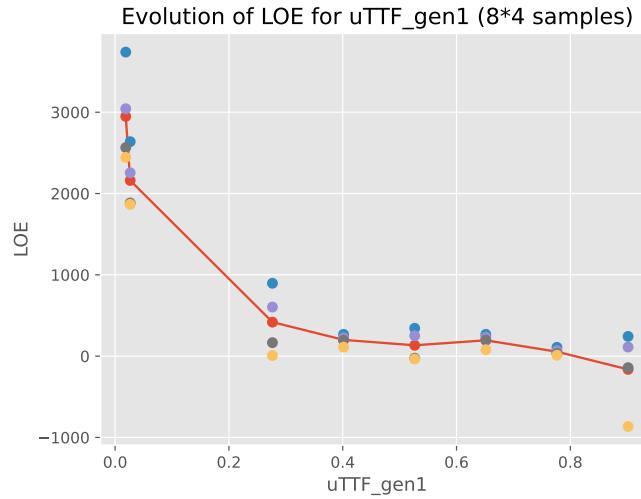


Figure 5.7: Evolution of LOE over $u_{ttf,gen1}$, 4 subprofiles

A remarkably low LOO error is observed in the case where only one uncertainty is involved and in order 3. The Figure 5.7 shows the evolution of LOE for the value of this uncertainty. The curve is decreasing monotonically as the value of $u_{tff,gen1}$ increases. The result is expected because a value of $u_{tff,gen1}$ close to 1 means high value of TTF, thus, more time before encountering an outage. The values of the subprofiles are represented on the Figure by the color points (from the highest LOE to the lowest: blue, purple, grey and yellow). It shows the range of possible values of LOE depending of $u_{tff,gen1}$.

Table 5.4: PCE Linear Extrapolation ($n_{extra} = 50$) with subprofiles (10 uncertainties) Main Results

PCE order p	n_{sp}	n_{simu}	LOO error (LOE) [%]
2	2	264	55.57
2	3	396	32.12
2	4	528	35.08
2	5	660	32.07
3	2	1144	63.74
3	3	1716	39.14
3	4	2288	28.54
3	5	2860	29.77

With 10 uncertainties, a relative decrease of the LOO error $\hat{\epsilon}_{LOO}$ is observed when more and more subprofiles are involved (see Figure 5.8). For the PCE at order 2, the LOO error is 55.57% with 2 subprofiles and is, at best, 32.07% with 5 subprofiles. For the PCE at order 3, the LOO error is 63.74% with 2 subprofiles and is, at best, 28.54% with 4 subprofiles. Due to the high number of simulations compared to the Monte-Carlo process, no order above, nor higher number of subprofiles is considered. In the Figure 5.8, the LOO error does not seem to decrease significantly with more than 3 subprofiles. Not enough data is available to determine if the LOO error will converge toward a certain value with the number of subprofiles n_{sp} and the order p and, if so, at what pace.

The main PCE results of the best case w.r.t. the LOO error, among the cases discussed in Table 5.4 are given in table 5.5 (case $p = 3$ and $n_{sp} = 4$). The PDFs of the LOL, LOE and cost are given in Figures 5.9a, 5.9b and 5.9c respectively. The Sobol' indices of the LOL, LOE and cost are given in Figures 5.10a, 5.10b and 5.10c respectively.

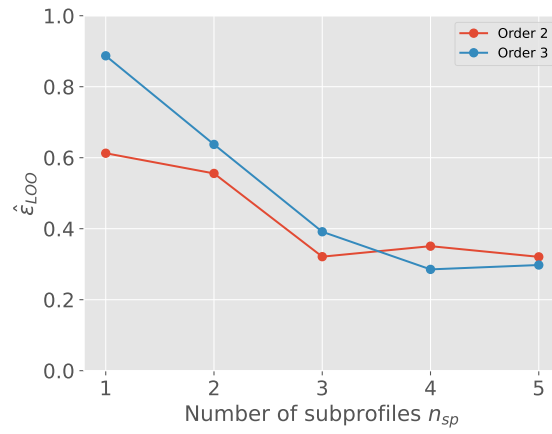


Figure 5.8: LOO error $\hat{\epsilon}_{LOO}$ depending on the number of subprofiles n_{sp} and the order p

Table 5.5: Order $p = 3$, number of subprofiles $n_{sp} = 4$ PCE Main Results

Name	Values
Number of profiles/full basis n_{pro}	572
Number of simulations n_{simu}	2288
Mean LOL [hours/year]	-24.75
LOL standard deviation [hours/year]	24.98
Mean LOE [MWh/year]	-1634.43
LOE standard deviation [MWh/year]	1788.28
Mean cost [€/year]	$4.453 \cdot 10^6$
Cost standard deviation [€/year]	$4.272 \cdot 10^6$
Leave-One-Out error (LOL) [%]	36.76
Leave-One-Out error (LOE) [%]	28.54
Leave-One-Out error (Cost) [%]	117.61

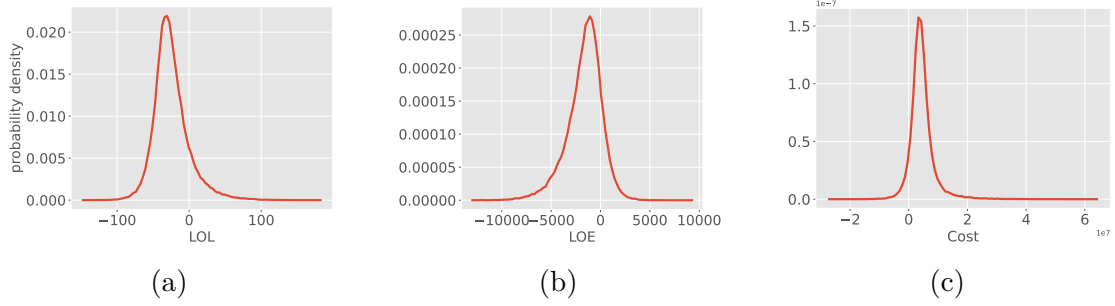


Figure 5.9: PDFs for outputs (Order $p = 3$, number of subprofiles $n_{sp} = 4$), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

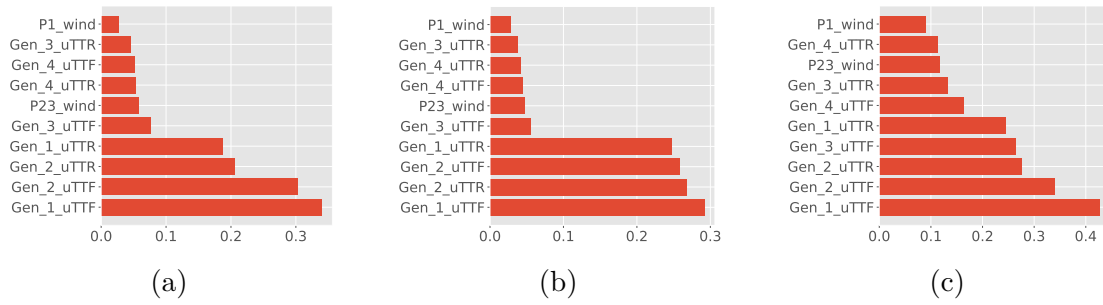


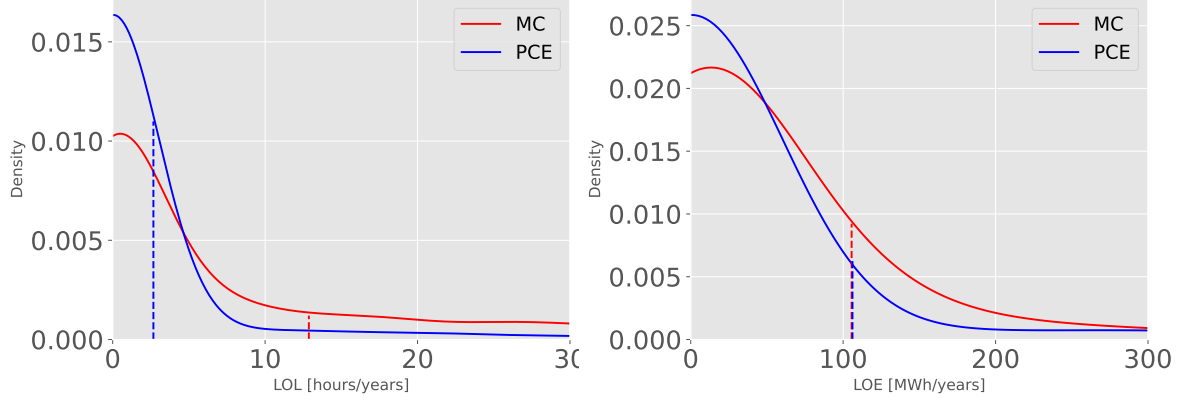
Figure 5.10: Sobol indices for outputs (Order $p = 3$, number of subprofiles $n_{sp} = 4$), (a) LOL [hours/year], (b) LOE [MWh/year], (c) Cost [€/year]

Even considering the relative high LOO error on the outputs with the linear extrapolation and the subprofiles, the LOO error of the LOE (28.54%) and the LOO error of the LOL (36.76%) are considered decent in this thesis, and an analysis of the adequacy through the PCE is given in section 5.3. Solutions and future work in perspective to reduce the LOO error are given in chapter 6. Considering the high LOO error of the cost (117.61%) w.r.t. the LOO error of the other outputs, this output is not taken into account for the analysis.

Firstly, as explained in Section 2.3, the total-order Sobol' indices quantify the total impact of a stochastic input parameter, including all mutual interactions. Among the different inputs considered, 4 different inputs seem to have the most impact on the LOL and the LOE, considering their Sobol' indices compared to those of the 6 other inputs. The $u_{ttr,gen1}$, $u_{ttr,gen2}$, $u_{ttf,gen2}$ and $u_{ttf,gen1}$ have the highest Sobol' indices for LOL (> 0.15 , see 5.10a) and for LOE (> 0.2 , see 5.10b). The meaning is that the TTF and the TTR of generator 1 and generator 2 are relatively highly related to the final LOL and LOE. This result is rather logical because they are the generators with the highest p_{max} and an outage of one of those generators will have the highest chance to not be able to balance the production and the demand, compared to a generator with a lower p_{max} at the same time.

Secondly, the given PDFs on the Figures 5.9a (LOL) and 5.9b (LOE) can be compared to the PDFs obtained thanks to the original Monte-Carlo process, represented in Figures 4.3a (LOL) and 4.3b (LOE). As no negative LOL and LOE values are generated in the Monte-Carlo process, all the generated negative LOE and LOE from the PCE are nullified and new means are computed. The Kernel Distribution Estimate (KDE) is used to represent the distributions, as found in [18]. The blue dotted lines give the mean LOL/LOE from the PCE distribution while the red dotted lines give the mean LOL/LOE from the MC distribution. Those values are, $LOLE_{PCE} = 2.67$ hours/year, $LOEE_{PCE} = 106.14$ MWh/year, $LOLE_{MC} = 12.87$ hours/year and $LOEE_{MC} = 105.56$ MWh/year.

While $LOEE_{MC}$ and $LOEE_{PCE}$ are relatively close, $LOLE_{MC}$ and $LOLE_{PCE}$ are relatively different from each other. Generally speaking, the distributions from the PCE are more concentrated on values relatively close to zero, while the distributions from Monte-Carlo are more widespread for large LOL/LOE. This can be partly explained by the fact that every uncertainty is not represented in the PCE process, e.g., the u_{ttf} and u_{ttr} of the lines, which gives more chance to not have any LOE and LOL on a whole year. Another part of the explanation can be the relatively high value of the LOO error, discussed previously.



(a) Distribution of the LOL with MC and PCE, the dotted lines provide the mean values, $LOLE_{MC} = 12.87$ hours/year and $LOLE_{PCE} = 2.67$ hours/year

(b) Distribution of the LOE with MC and PCE, the dotted lines provide the mean values, $LOEE_{MC} = 105.56$ MWh/year and $LOEE_{PCE} = 106.14$ MWh/year

Figure 5.11: Kernel Distribution Estimate (KDE) of the LOL and LOE for Monte-Carlo (MC) and Polynomial Chaos Expansion (PCE)

5.3 Post-Processing

When dealing with Polynomial Chaos Expansion (PCE), if a sufficient number of samples, denoted as n_{pro} , are available, the PCE can be trained. Similar to functions, the trained PCE can return outputs for LOL, LOE, and cost, based on the given inputs. With multiple samples, it is possible to plot distributions, as illustrated in Figures 5.9a, 5.9b, and 5.9c. In this section, since we aim to identify which uncertainties contribute most to the prevention of LOE/LOL, we focus on analyzing LOE as the output, in relation to its inputs.

In Figure 5.12, the distribution is divided into two parts. The part in green is called *System Working* and represent the cases where the LOE is negative or zero. The part in red is called *System Failing* and represent the cases where the LOE is positive. It is worth noting that, in the Study Case considered, a system has a success rate of 83.4%, i.e., be a *System Working*.

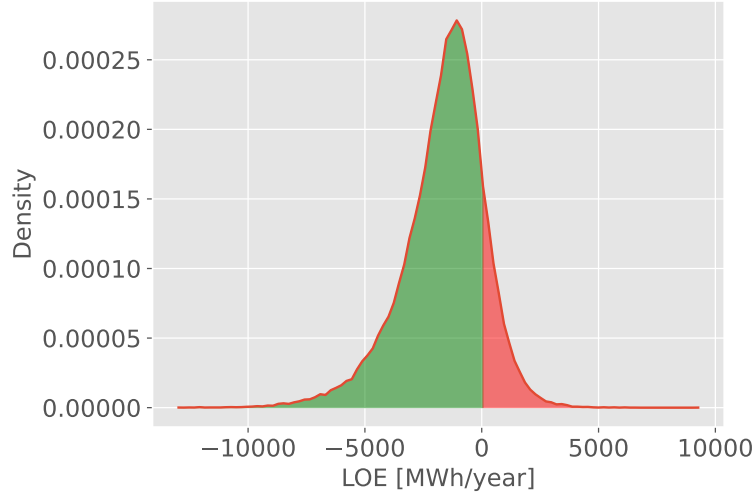


Figure 5.12: LOE distribution divided in two parts: *System Working* (green) or *System Failing* (red)

Considering the output with its inputs, it can be possible to build a *Decision Tree* as seen in Figure 5.13. Decision Trees (DTs) are a non-parametric supervised learning technique utilized for both classification and regression tasks. The objective is to develop a model that predicts the value of a target variable by deriving straightforward decision rules from the data features. This approach can be viewed as a piecewise constant approximation of the target variable [34]. From the 572 already existing samples in the training set, 10 000 new samples have been generated to increase representativeness.

In the Figure 5.13, the decision nodes are the boxes with a decision in the first line, the second line "samples" is the number of samples on which the decision is applied, the third line "value" is the ratio of samples that provide Systems Working or Systems Failing and the fourth line is the "class" which can be "sucess" or "fail" depending on whether the majority of the samples provide Systems Working or Systems Failing. The leaves are the boxes with three lines at the end of the tree. A first decision is made to divide the 10 572 samples, based on the value of $u_{ttf,gen1}$. If this value is above 0.18 ($TTF \approx 397$ hours for the generator 1, see Eq. 3.11), the system has approximately 89.7% chance of being a *System Working*, while the system has approximately 55.7% chance of being a *System Working*. A second decision is made, based on the value of $u_{ttf,gen2}$. For the case where $u_{ttf,gen1} \leq 0.18$, if $u_{ttf,gen2} \leq 0.24$ ($TTF \leq 549$ hours for generator 2), the system has 23.6% to be a *System Working* and 65.7% if $u_{ttf,gen2} \geq 0.24$. For the case where $u_{ttf,gen1} \geq 0.18$, if $u_{ttf,gen2} \leq 0.17$ ($TTF \leq 373$ hours for generator 2), the system has 65.5% to be a *System Working* and 94.6% if $u_{ttf,gen2} \geq 0.17$. This Decision Tree shows that

by dividing the space of the inputs by three decisions, scenarios giving a high percentage of Systems Working or a low percentage of Systems Failing can be found. $u_{tff,gen1}$ and $u_{tff,gen2}$ are important variables when it comes to making those decisions. Exactly as for the explanations of the Sobol' indices from Figure 5.10b, the meaning is that the TTF of the generator 1 and generator 2 are relatively highly related to the final LOL and LOE. This result is rather logical because they are the generators with the highest p_{max} and an outage of one of those generators will have a greater impact on the LOE than a generator with lower p_{max} at the same time.

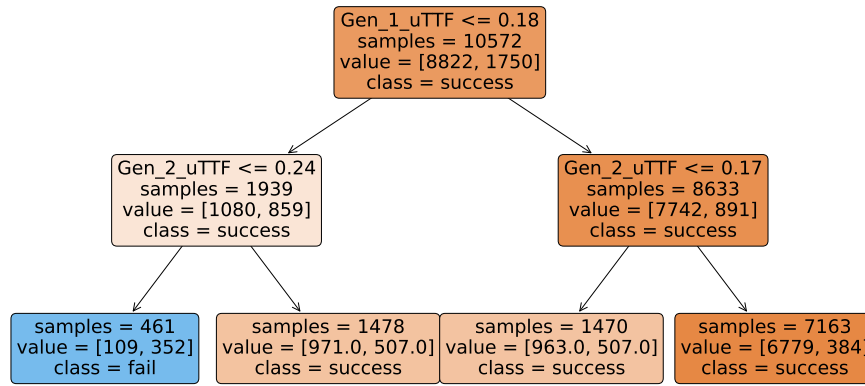


Figure 5.13: Decision Tree with a maximum depth of 2, the first decision is based on the value of $u_{tff,gen1}$ and the second ones are based on the value of $u_{tff,gen2}$

The next step should be to represent the Decision Tree Performance Metrics by 3 different performance metrics, applied to each tree leaf in Figure 5.14.

- *Interpretability:* This refers to how "easily" the decision tree model can be understood. In this context, interpretability is inversely related to the number of unique features (uncertainties) used by the tree. If a tree uses fewer features to make its decisions, it is considered more interpretable.
- *Coverage:* In this context, this is the accuracy of the decision tree when tested on the entire dataset. It indicates how well the model is performing in classifying the data points correctly.

- *Cross-Validation Coverage:* This is the average accuracy of the decision tree across multiple folds of cross-validation. Cross-validation helps in assessing the model's performance and robustness by training and testing it on different subsets of the data. In Figure 5.14, the coverage of a 5-fold cross-validation analysis is shown.

Figure 5.14 shows the score of the metrics described below in the Decision Tree from Figure 5.13. It is observed that the coverage and the cross-validation coverage scores do not increase significantly after 4 tree leaves, while the interpretability tends to decrease with the number of tree leaves. Fully grown trees tend to overfit data (produce non-significant splits), which can lower the prediction performance and make the algorithm heavier [7]. Having more than 4 tree leaves is, thus, not necessary, as having more makes the algorithm heavier and does not increase the cross-validation coverage and coverage significantly. 3 tree leaves, as displayed in Figure 5.13, also provide relatively similar cross-validation coverage and coverage score to 4 tree leaves.

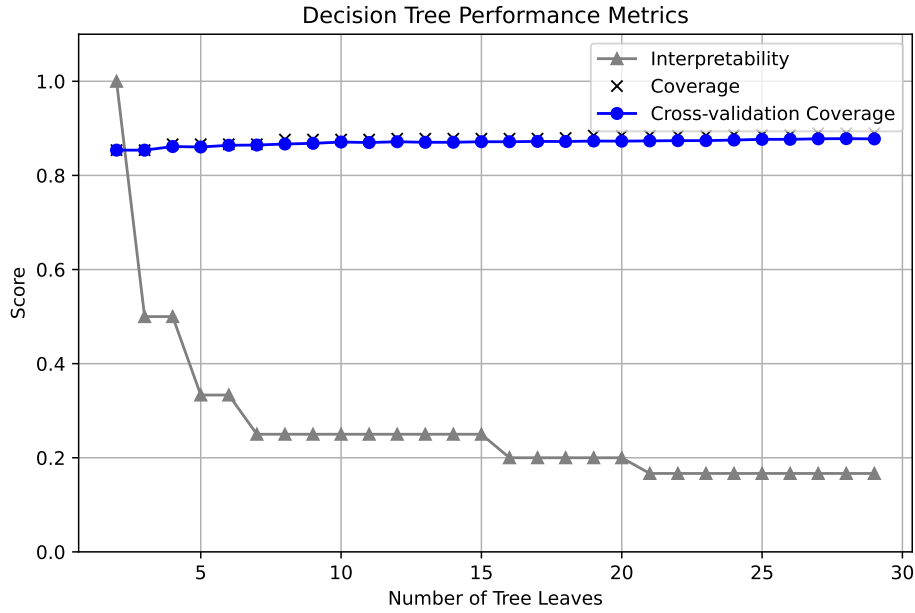


Figure 5.14: Decision Tree Performance Metrics (Interpetability, Coverage, Cross-validation Coverage), the coverage and the cross-validation score does not increase significantly after 4 tree leaves

The Figures 5.15 and 5.16 show the impact of the TTR and TTF of the generator 1 on the LOE, a red point representing a System Failing and a green

point representing a System Working in the $(u_{TTF,gen1}, u_{TTR,gen1})$ -plane for Figure 5.15 and in the (gen. 1 TTF, gen. 1 TTR)-plan for the Figure 5.16 (see Eq. 3.11 and 3.12 to understand how to pass from $u_{TTF,gen1}/u_{TTR,gen1}$ to gen. 1 TTF/TTR). Globally, a remarkable concentration of Systems Failing compared to the global success rate can be observed when $u_{TTF,gen1}$ is a low value (typically $u_{TTF,gen1} < 0.2$ or $TTF < 446$ hours) and $u_{TTR,gen1}$ is a high value (typically $u_{TTR,gen1} > 0.6$ or $TTR > 46$ hours). This result is logical, as the TTR and the TTF of the generators 1 and 2 are highly related to the output LOE. A high TTR and a low TTF imply a low availability of the generator on the grid, thus increasing the chance of not being able to provide electricity when needed.

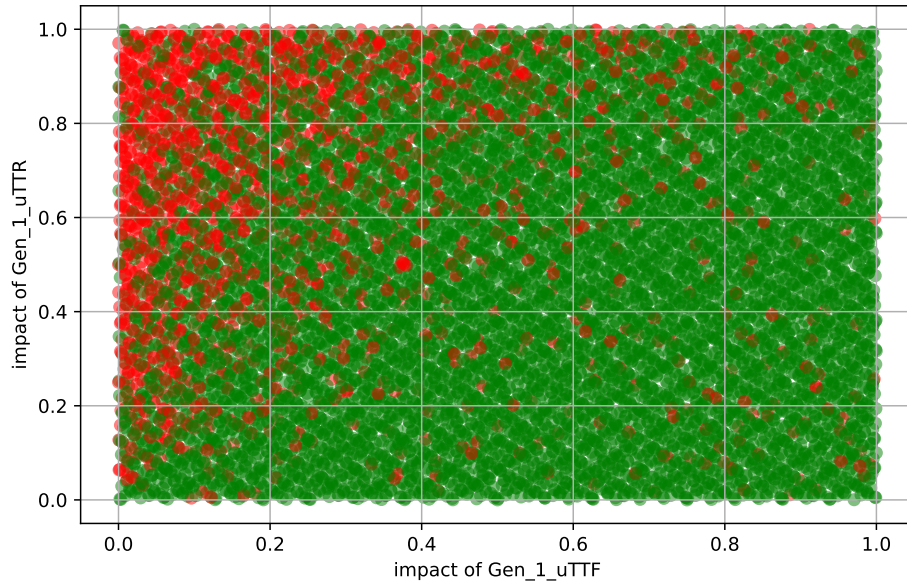


Figure 5.15: $(u_{TTF,gen1}, u_{TTR,gen1})$ -plane with System Working (green) and System Failing (red) (10572 samples)

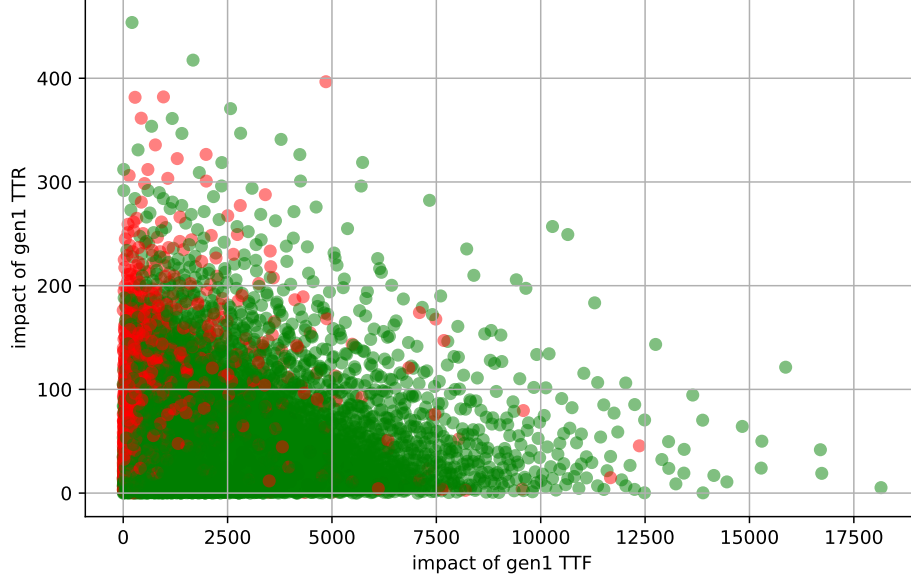


Figure 5.16: (gen. 1 TTR, gen. 1 TTR)-plane with System Working (green) and System Failing (red) (10572 samples)

When dealing with generators, it can be relevant to link their TTF and TTR under the availability α_{gen} of the generator, giving the mean time they are working on the total time (see Eq. 5.14).

$$\alpha_{gen} = \frac{TTF}{TTF + TTR} \quad (5.14)$$

By representing TTR as a function of TTF ($TTR = TTF^{\frac{1-\alpha_{gen}}{\alpha_{gen}}}$), it is possible to represent the lines where the generator 1 has an availability of $\alpha_{gen} = 0.7$, $\alpha_{gen} = 0.8$, $\alpha_{gen} = 0.9$ and $\alpha_{gen} = 0.95$ in Figure 5.17. To analyze this graph, we can imagine that a line (let us say 0.9), divides the space of points into two subspaces. The subspace "above" the line is the subspace where the availability of generator 1 is below 90%, i.e., $TTR > \frac{1}{9}TTF$. On the other hand, the subspace "below" the line is the subspace where the availability of generator 1 is above 90%, i.e., $TTR < \frac{1}{9}TTF$. Quite logically with the analysis already done, the availability of the generator 1 seems to have a significant impact on the LOE as a higher concentration of Failing Systems is present in the subspace where the availability of generator 1 < 90% than in the subspace where the availability of generator 1 > 90%.

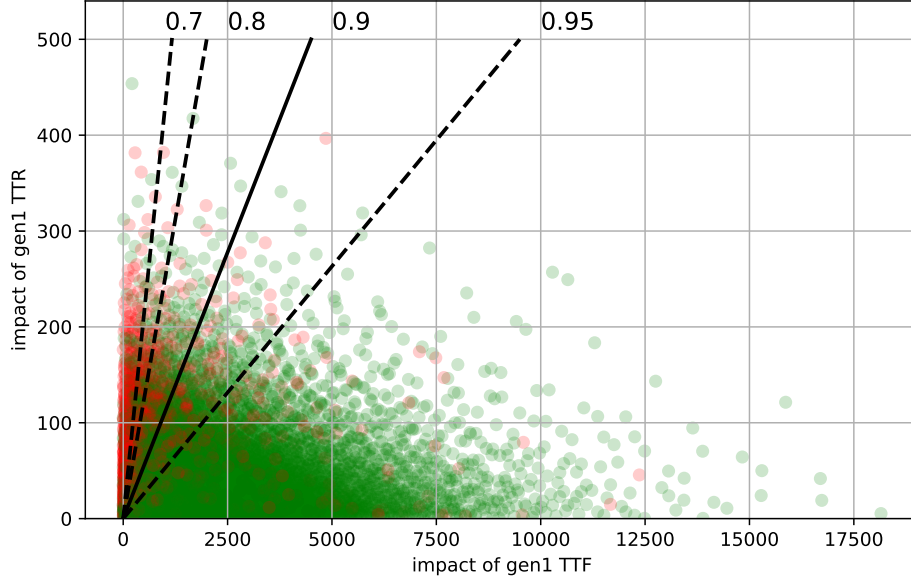


Figure 5.17: (gen 1 TTR, gen 1 TTR)-plane with System Working (green) and System Failing (red) (10572 samples), the points on the lines of 0.7, 0.8, 0.9, 0.95 show the points where the availability of the generator 1 is $\alpha_{gen} = 0.7$, $\alpha_{gen} = 0.8$, $\alpha_{gen} = 0.9$ and $\alpha_{gen} = 0.95$ respectively

The data presented in Figure 5.17 can only show the division between high availability ($> 90\%$) and low availability ($< 90\%$) for one generator. This choice of division at 90% of availability for each generator is arbitrary and is lower than the mean availability of each generator ($MTTF/(MTTF + MTTR) = 97.56\%$). Considering this division between high availability ($> 90\%$) and low availability ($< 90\%$) as a decision, it is possible to display a decision tree based on the availability of multiple generators as presented in Figure 5.18. For each level, a decision is made based on whether the availability of a certain generator is greater or less than 90%. The percentage of Working Systems for all the Systems of a leafnode is written in red and the percentage of samples from the 10 572 samples is given. The labels "worst" and "best" correspond to the percentage of Working Systems with the worst/best decisions, according to the percentage of Working Systems, with the remaining uncertainties ($\alpha_{wt,P1}$ and $\alpha_{wt,P23}$ for the level 5 (5th level of division, starting from the top) and $\alpha_{wt,P1}$, $\alpha_{wt,P23}$, the availability of the generator 3 and the availability of the generator 4 for the two leafnodes at level 3). The decision on the availability of a wind farm α_{wt} is based on whether the availability is greater than or less than the average wind farm availability 0.804.

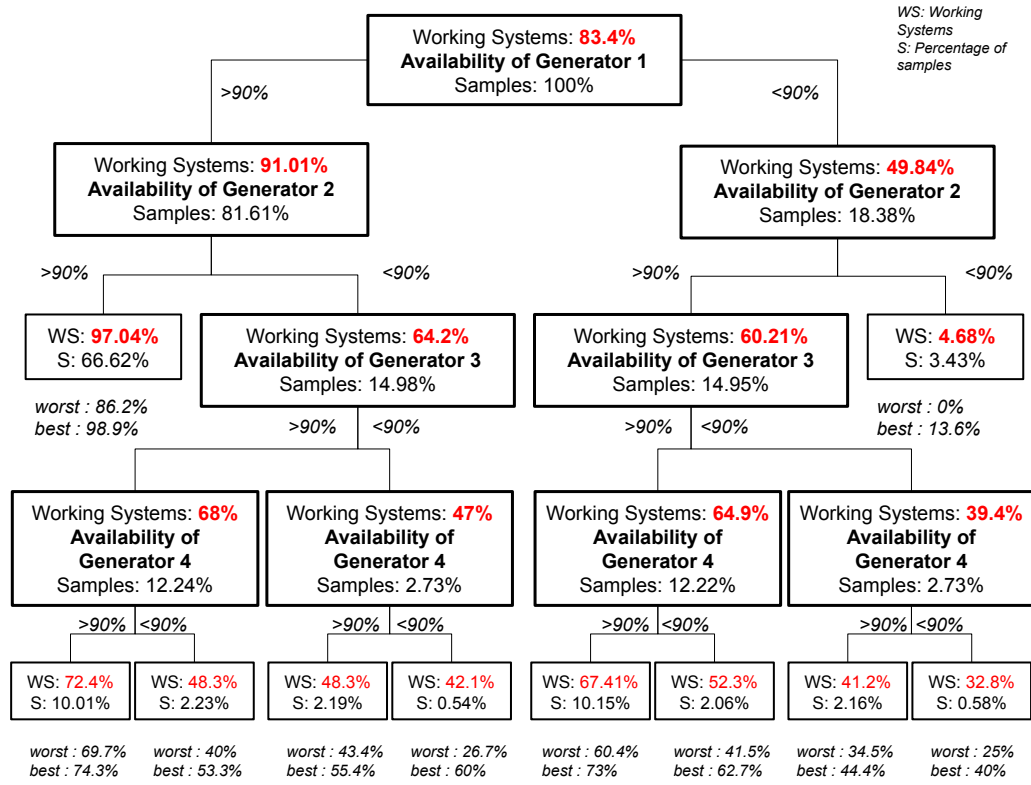


Figure 5.18: Decision Tree according to the availability of the 4 generators, "worst" and "best" stand for the variation of the ratio of Working Systems when the other uncertainties are taken into account (typically the wind farm availabilities α_{wf})

Figure 5.18 shows that the main parameters to control when dealing with adequacy in the Study Case are the availability of generator 1 and the availability of generator 2. If their respective availabilities are greater than 90%, there is 97.04% chance to have a Working Systems out of the possible Systems for those two decisions. Even with availabilities of generators 3 and 4 less than 90% and availabilities of the wind farms less than 80.4%, the percentage of Working Systems is 86.2% (worst scenario). On the other hand, if the availabilities of generators 1 and 2 are less than 90%, there is only 4.68% to obtain a Working System. The results are less certain when a high availability for generator 1/2 is combined with a low availability for generator 2/1, with Working Systems percentage between 60% and 65%. Those scenarios can be divided according to decisions based on the availability of generators 3 and 4. Depending on the scenarios, a shade of Working Systems percentages can be observed from 32.8% when the generator 2 has a high availability while generators 1, 3 and 4 have a low availability to 72.4% when the generators 1, 3 and 4 have a high availability while the generator 2 has a low availability.

Without surprise, the availability of generators 1 and 2 has a greater impact on the ratio of Working Systems to all Systems. It is mainly because those generators have the highest p_{max} w.r.t. the p_{max} of the other generators. When an outage of one of those generators occurs, there is a higher chance of not meet the demand than with other generators. This chance is increased as availability is low because there are more hours for which the demand cannot be met. The impact of generators 3 and 4 is relatively lighter compared to this of the generators 1 and 2. As explained, it is due to the difference in p_{max} , the generators 3 and 4 having a lower p_{max} than that of the generators 1 and 2. The impact of the availability of wind farms is relatively light compared to the other uncertainties. When the demand is not met at a certain hour, it is frequently because one or multiple outages of generators occur during a drop of wind power production. As the wind power production is at a drop, e.g., there is almost no wind, the availability of the wind farm does not offer a high variation of power compared to the variation of power when a generator is working or failing.

It is worth saying that the Working Systems percentages are obtained from a finite number of samples; thus, problems of under-representation can occur when dealing with a relatively small number of samples.

Chapter 6

Discussion

Observing the results of the previous section, the difficulty of finding an adequate PCE for the proposed Power System Adequacy Assessment simulation is present. No article from literature or book found on the casual communication channels deals with Power System Adequacy Assessment and PCE at the same time. The usual use of PCE in the context of Power Grid context is when dealing with the modification of the power injection at each node with stochastic demand and renewable energy sources [27][22] and not the Loss-of-Load (or Loss-of-Energy) of a Power System over a certain time.

The PCE is first applied to the Study Case with the following assumptions summarized:

- Instead of setting uncertainties on the TTF and TTR of each wind turbine in each wind farm, a global availability α_{wt} is used to describe the percentage of working wind turbines for one wind farm.
- As opposed to Monte-Carlo, the TTR and the TTF of each generator are fixed as constants for a whole simulated year. This helps reduce the uncertainty on the number of different TTR and TTF for a whole simulated year, as new TTR and new TTF were generated after each failure/repair in the Monte-Carlo process.
- Instead of working with 100 different wind profiles in the Monte-Carlo process, 1 wind profile is considered for the entire PCE process to compare two different samples more efficiently.
- 10 uncertainties are considered: the TTF and TTR of the 4 generators with $p_{max} > 0$, based on the random uniform variables u_{ttf} , u_{ttr} and the availability α_{wt} of the two wind farms having the highest p_{max} , based on a normal distribution.

- The lines and synchronous condensers are considered as ideal ($MTTF \rightarrow \infty$ and $MTTR = 0$) considering their mean availability at 99.88% and 97.56% ($MTTF/(MTTF + MTTR)$). The availability of the other wind farms are considered constant on the average, thus always 80.4% of availability (see subsection 3.2.2).

The main goal of those assumptions was to limit the number of uncertainties, which can rapidly require an unmanageable number of samples because of the curse of dimensionality (see subsection 2.3.4). The 10 uncertainties have been chosen in order to match the results obtained from the Monte-Carlo process as accurately as possible.

Applying the PCE naively to the Study Case with the presented assumptions leads to a high LOO error ($> 90\%$ minimum) for order lower than $p = 4$. This high error shows the impossibility of the PCE to represent the distribution of LOL, LOE and cost accurately in that way. The given PDFs and Sobol' indices have, thus, no real meaning according to the Study Case. Two main reasons can explain this problem. First, in the Monte-Carlo process, the LOE and the LOL cannot be negative while the PCE does consider the existence of negative LOE and LOL by continuity and because of their relative proximity to the negative limit for LOE and LOL. Secondly, for a generator having a certain TTF and a certain TTR, the time the outages occur can have a large impact on the LOE.

Considering this, some new methods have been investigated. Firstly, one to transform the LOE and LOL into continuous output variables thanks to the monotonic curves. This monotonic curve is extrapolated (with a linear or a quadratic function) if $LOE = 0$, to find a "potential LOL" and a "potential LOE" i.e., negative LOL and LOE. This approach alone did not significantly increase the accuracy of the PCE method, keeping a LOO error in the same range as what was computed naively. Secondly, in addition to work with extrapolations, it has been decided to work with subprofiles in order to represent the possible scenarios more accurately, depending on the time where the outages of the generators can occur, and take the mean output out of those subprofiles. Using a linear extrapolation of the monotonic curves, from a regression over the last $n_{extra} = 50$ points, and using $n_{sp} = 4$ subprofiles in order $p = 3$, reduces the LOO error at 28.54% for the LOE (36.76% for the LOL).

It must also be said that, in the present definition of the monotonic curve, the maximal power flow of the lines $F(l, t)$ is not considered, leading to possible mismatch between the monotonic curves and the true "potential LOL" and "potential LOE". Other elements not taken into account in the Study Case (batteries, storage, international interconnections, etc.) could also add to this mismatch for future

works.

Let us note that the alternative of creating a mapping between the initial LOE distribution and a possible distribution working with the PCE has been proposed. The major problem with this technique seems to be that if we have a profile giving $LOE = 0$, the information on the "vulnerability" of having a negative LOE is lost, i.e., two profiles giving $LOE = 0$ can have a high difference in "potential LOE".

While the results for only one uncertainty $u_{ttf, gen1}$ were relatively interesting, with a huge reduction of the LOO error for the LOE with multiple subprofiles in order 3 ($\hat{\epsilon}_{LOO} = 8.96\%$ if $n_{sp} = 2$ and $p = 3$ with quadratic extrapolation from quadratic regression over $n_{extra} = 10$ last points), the results were less encouraging for 10 uncertainties with LOO error $> 25\%$. A hypothesis for this reason could be the high complexity of the system with the uncertainties interacting with each other, giving a high diversity of possible subprofiles. Indeed, in the Study Case, there is not only one generator anymore but 4 generators that can possibly have outages at the same time, possibly giving high LOE. They can also never be down at the same time, and when wind production can compensate for the potential loss, give zero LOE. Then, a wide range of LOE can be obtained at the end for one same profile. Considering all the different scenarios, with 10 uncertainties, 5 subprofiles are not nearly enough to capture a representative value of LOL and LOE for a profile. Considering the high number of simulations (2860) for an order 3 PCE process with 5 subprofiles for each profile compared to Monte-Carlo (198), this methodology must be adapted if more subprofiles are added, even more if more uncertainties are considered in the Study Case or other Study Cases that are more complex.

An important point to discuss is also the choice of the parameters for the presented methodology to reduce the LOO error. The extrapolation of the monotonic curve by a linear or quadratic function has been studied but other studies can be conducted to verify the sustainability of extrapolation by other common functions like logarithmic, square root, etc., which could approximate the potential LOL and LOE better, possibly following more accurately the behavior of the monotonic curves. Depending on the function, we choose the extrapolate with, the number of last points n_{extra} to perform the regression over must be considered. A too low n_{extra} could lead to overfitting (more sensitive to points individually) where a too high n_{extra} could lead to underfitting (more sensitive to the general shape of the whole monotonic curve). The empirical value of $n_{extra} = 50$ (and compared with $n_{extra} = 5$, $n_{extra} = 10$ and $n_{extra} = 20$) for the analysis in Section 5.3 has been chosen but a systematic algorithm could be imagined to find the

optimal number of last points considered for the regression, giving the minimal LOO error. As explained in section 2.3, it is not possible to know in advance how to choose the maximal polynomial degree. The crude approach would consist of testing several truncation schemes of increasing degree (e.g., $p = 2, 3, 4$) [38], keeping the number of samples manageable. The number of subprofiles n_{sp} is also a parameter to control. It has been explained that $n_{sp} = 5$ subprofiles can somewhat not be sufficient when dealing with a large number of uncertainties but can keep the number of samples required to perform the PCE manageable. Based on the Figure 5.8 in Section 5.2, showing the evolution of the LOO error over the number of subprofiles n_{sp} for orders $p = 2$ and $p = 3$, the question of the LOO error convergence has merit to being asked. Future studies could show if increasing the number of subprofiles could really significantly reduce the LOO error or keep the LOO error at a certain range of values. The LOO error in itself could also be questioned as a relevant metric to deal with Power System Adequacy Assessment.

The number of samples needed to perform the PCE with a decent LOO error, compared to the number of samples needed to converge toward the LOEE value with Monte-Carlo, makes an interesting parallel with the research question, which is:

Could the PCE method be a better alternative to the Monte-Carlo method on the Power System Adequacy Assessment problem to make it more efficient computationally?

Comparing the number of simulations with the PCE process giving the lowest LOO error of 28.54% for LOE ($p = 3$, $n_{sp} = 4$, linear extrapolation with $n_{extra} = 50$), 2288 simulations are required compared with 198 with the Monte-Carlo method. The complexity of the two algorithms is $O(n_{simu})$ and $O(n_{year})$ respectively and it takes approximately 87 sec to perform a simulation or a Monte-Carlo year with an AMD Ryzen 5 5600U with Radeon Graphics 2.30 GHz. Even if the PCE enables the analytical derivation of distribution parameters and sensitivity indices, such as Sobol' indices, from the surrogate model's coefficients [25], which is a valuable advantage compared to Monte-Carlo, it does require more than 10 times more simulations in a simple Study Case and considering only 10 uncertainties. A solution to speed up the process and make the PCE process less greedy in simulations can be explored in the change of the PCE method, making it more suitable for the Adequacy Assessment problem.

The change in the PCE method has been briefly studied through different existing methods. The *Sparse PCE* [25] can approximate computer models with many input parameters, relying on only a few model evaluations, it seems interesting

to make the computations faster as fewer evaluations are needed. It could be a great future work to reduce the number of simulations required by the PCE process and compete with the Monte-Carlo process. The *Adaptive sampling* is a sequential optimal sampling that is capable of adding new samples based on the most recent basis polynomials and also re-utilizing the old set of samples, it also seems rather interesting because the function can adapt the distribution of the points and focus the resources on where the complexity is the highest. Another idea was to make use of the *Multi-Elements PCE*, whose purpose is to decompose the space of random inputs when the relative error in variance becomes greater than a threshold value. In each subdomain or random element, we then employ generalized polynomial chaos. Several methods exist in the literature:

- Adaptive multi-element generalized polynomial chaos method [44].
- Multi-element polynomial chaos expansion based on automatic discontinuity detection for nonlinear systems [12].
- Multi-element non-intrusive Polynomial Chaos method using agglomerative clustering based on the derivatives to study irregular and discontinuous Quantities of Interest [43].
- Anisotropic multi-element polynomial chaos expansion for high-dimensional non-linear structural problems [2].

Those methods have not been deeply analyzed in our model and can be possible insights for future works.

As previously said, one of the greatest advantages of the PCE compared to Monte-Carlo is to enable analytical derivation of distribution parameters and sensitivity indices, such as Sobol' indices, from the surrogate model's coefficients [25]. Another advantage of Monte-Carlo is its advantage to generate new samples with more simulations than the number required to train the model and perform analysis, as shown in the subsection 5.3. Even considering the high LOO error, a complete analysis with a Decision Tree is presented in Figure 5.18. It shows the generators having the most impact on the power system, for which, a low availability could cause a high probability of LOE. In this trivial example, we could guess that the generators having the highest p_{max} could give the highest rate of Systems Failing if their availabilities were lower than 90%. Nevertheless, this analysis could be performed on a more complex Study Case like the Belgian Grid. Analysis performed with the PCE on real Grid could offer a better understanding of the impacts of uncertainties, which can help make more robust decisions concerning power grid management, infrastructure planning and energy policies [47].

Finally, it is important to note that uncertainties could occur at many levels in a power system. The initial Power System Adequacy Assessment simulation only considers as uncertainties the possible outages of wind turbines, generators and lines and the production of renewable energy. Other uncertainties can be set e.g., by the load share, the value-of-the-loss-load (VOLL), etc.

Chapter 7

Conclusion

Through this thesis, we explore the context of the Power System Adequacy Assessment with different methods, the deterministic and the probabilistic ones. The adequacy metrics are also given. An overview of the Monte-Carlo Method and the Polynomial Chaos Expansion (PCE) is given, along with the use of the PCE method in a Power System context. The first methodology is based on a Power System Adequacy Assessment simulation [18], working originally with the Monte-Carlo method. The PCE method replaces the Monte-Carlo method thanks to the Python package Rheia [11]. The chosen Study Case is a modified IEEE 30-Bus System, including 11 wind farms and other various modifications. The Monte-Carlo method is used on that Study Case to find the LOLE and the LOEE, and it serves as the gold standard for the whole thesis. The PCE method is also applied, removing the uncertainties for Monte-Carlo and fixing 10 uncertainties on main units (8 on generators and 2 on wind farms).

As the results were not optimal, it has been decided to explore diverse possibilities to reduce the LOO error. By extrapolating on the monotonic curve of each profile to obtain negative LOL and LOE when LOL and LOE are normally zero, new distributions were found but with a new non-optimal LOO error. By using subprofiles and monotonic curve extrapolation, relatively optimal results (according to the LOO error) are found for 10 uncertainties because of the high degree of interactions between uncertainties, needing more subprofiles to capture relevant results but being computationally orders of magnitude less efficient than the Monte-Carlo method. An analysis of the best case, according to the LOO error is provided, showing the large impact of the availability of the more important generators, in terms of p_{max} , on the adequacy of the Study Case.

In that field of knowledge, areas of future research can be operated to make efficient use of the PCE method.

Firstly, improvements of the current method can be performed in particular on the steps to reduce LOO error. The monotonic curves could include, one way or another, the limitations of the maximum power flow through the lines and other elements from more complex power systems (batteries, storage, etc.). Other common functions to extrapolate the monotonic curves can be considered, as the logarithmic and/or the root-squared functions. The number of last points n_{extra} considered for the regression could be automatically set between two bounds to minimize the LOO error. A further analysis can also be operated to explore if the LOO error is effectively decreasing with higher order p and number of subprofiles n_{sp} than presented in this thesis.

Other research can be conducted in the PCE method. Considering the high number of simulations, compared to Monte-Carlo, required to obtain a decent LOO error, exploring the Sparse PCE method. It is often observed a posteriori a PCE process that the non zero coefficients in the expansion form a sparse subset. A suitable sparse basis could be used instead of computing useless terms in the expansion which are eventually negligible [38]. The Sparse PCE method could reduce drastically the number of simulations required to perform the PCE process and, thus, compete with the Monte-Carlo in terms of performance.

When the PCE is trained, the PCE method offers the advantages of analyzing the sensibility of the outputs to certain values of inputs and of generating multiple inputs-outputs to study the tendencies of a model. A further interesting work could be to perform PCE on a more complex Study Case like the Belgian grid following the methodology described in this thesis. This could offer a better understanding of the impacts of uncertainties, which can help make more robust decisions concerning power grid management, infrastructure planning and energy policies.

Bibliography

- [1] R.N. Allan. “Power-system reliability in perspective”. English. In: *Electronics and Power* 30 (3 Mar. 1984), 231–236(5). ISSN: 0013-5127. URL: <https://digital-library.theiet.org/content/journals/10.1049/ep.1984.0118>.
- [2] A.A. Basmaji et al. “Anisotropic multi-element polynomial chaos expansion for high-dimensional non-linear structural problems”. In: *Probabilistic Engineering Mechanics* 70 (2022), p. 103366. ISSN: 0266-8920. DOI: <https://doi.org/10.1016/j.probengmech.2022.103366>. URL: <https://www.sciencedirect.com/science/article/pii/S0266892022000996>.
- [3] *Belgium 2022 Energy Policy Review*. International Energy Agency (IEA). France, 2022.
- [4] Marc Berveiller, Bruno Sudret, and Maurice Lemaire. “Stochastic finite element, A non intrusive approach by regression”. In: *European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique* 15 (Mar. 2006). DOI: 10.3166/remn.15.81–92.
- [5] R. Billinton and R. N. Allan. *Reliability Evaluation of Power Systems*. New York: Plenum Press, 1996.
- [6] Roy Billinton and Wenyuan Li. *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*. Springer Science+Business Media, 1994.
- [7] Yves Brostaux. *Analyse Multivariée 1: exploration et structuration de données (ULiège)*. [Online; accessed 30-May-2024]. 2024.
- [8] Diederik Coppitters. “Robust design optimization of hybrid renewable energy systems”. PhD thesis. Mons: University of Mons and Vrije Universiteit Brussel, Nov. 2021.
- [9] Diederik Coppitters, Ward De Paepe, and Francesco Contino. “Robust design optimization and stochastic performance analysis of a grid-connected photovoltaic system with battery storage and hydrogen storage”. In: *Energy* 213.118798 (2020). <https://doi.org/10.1016/j.energy.2020.118798>.

- [10] Diederik Coppitters et al. “Energy, Exergy, Economic and Environmental (4E) analysis of integrated direct air capture and CO₂ methanation under uncertainty”. In: *Fuel* 344.127969 (2023). <https://doi.org/10.1016/j.fuel.2023.127969>.
- [11] Diederik Coppitters et al. “RHEIA: Robust design optimization of renewable Hydrogen and dErIved energy cArrier systems”. In: *Journal of Open Source Software* 7.75 (2022), p. 4370. DOI: 10.21105/joss.04370. URL: <https://doi.org/10.21105/joss.04370>.
- [12] Juliette Dréau, Benoit Magnain, and Alain Batailly. “Multi-element polynomial chaos expansion based on automatic discontinuity detection for nonlinear systems”. In: *Journal of Sound and Vibration* 567 (2023), p. 117920. ISSN: 0022-460X. DOI: <https://doi.org/10.1016/j.jsv.2023.117920>. URL: <https://www.sciencedirect.com/science/article/pii/S0022460X23003693>.
- [13] Nadine El Dabaghi and Jasmina Vucetic. “7 - Electrical power systems”. In: *Pressurized Heavy Water Reactors*. Ed. by Jovica R. Riznic. Vol. 7. JSME Series in Thermal and Nuclear Power Generation. Elsevier, 2022, pp. 275–286. ISBN: 978-0-12-822054-2. DOI: <https://doi.org/10.1016/B978-0-12-822054-2.00006-4>. URL: <https://www.sciencedirect.com/science/article/pii/B9780128220542000064>.
- [14] Elia. *Elia - About the company*. <https://www.elia.be/en/company>. [Online; accessed 27-May-2024]. 2024.
- [15] Roger G. Ghanem. “Stochastic Finite Elements with Multiple Random Non-Gaussian Properties”. In: *Journal of Engineering Mechanics-asce* 125 (1999), pp. 26–40. URL: <https://api.semanticscholar.org/CorpusID:17986560>.
- [16] Dan M. Ghiocel and Roger G. Ghanem. “Stochastic Finite-Element Analysis of Seismic Soil–Structure Interaction”. In: *Journal of Engineering Mechanics* 128.1 (2002), pp. 66–77. DOI: 10.1061/(ASCE)0733-9399(2002)128:1(66). eprint: <https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%290733-9399%282002%29128%3A1%2866%29>. URL: <https://ascelibrary.org/doi/abs/10.1061/%28ASCE%290733-9399%282002%29128%3A1%2866%29>.
- [17] *Gurobi Optimization*. <https://www.gurobi.com/>. accessed on 20/12/21.
- [18] Aurélia Hernandez, François Vallée, and Emmanuel De Jaeger. *Integration of Hydrogen in Sequential Monte Carlo Power Systems Adequacy Assessment*. 2023.
- [19] *IEEE 30-Bus System*. <https://icseg.iti.illinois.edu/ieee-30-bus-system/>. accessed on 17/05/24.

- [20] Øystein Stake Laengen. “Application of Monte Carlo Simulation to Power System Adequacy Assessment”. PhD thesis. Norwegian University of Science and Technology, June 2018.
- [21] Rémi Lafargue and D. Clouteau. “An iterative solver for stochastic problems of solid-structure interaction”. In: *Revue Européenne des Éléments Finis* (Jan. 2004).
- [22] Jirasak Laowanitwattana and Sermsak Uatrangjit. “Probabilistic power flow analysis based on arbitrary polynomial chaos expansion for networks with uncertain renewable sources”. In: *IEEEJ Transactions on Electrical and Electronic Engineering* 13.12 (2018), pp. 1754–1759. DOI: <https://doi.org/10.1002/tee.22737>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/tee.22737>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/tee.22737>.
- [23] Cococcioni M. and Fiaschi L. *The Big-M method with the numerical infinite*. 15. M. Optim Lett, 2021. DOI: <https://doi.org/10.1007/s11590-020-01644-6>.
- [24] AF Mercados, E-Bridge, and REF-E. *Identification of Appropriate Generation and System Adequacy Standards for the Internal Electricity Market - Final Report*. Tech. rep. B-1049 Brussel: European Commission, Mar. 2016.
- [25] Roeland De Meulenaere et al. “Uncertainty Quantification for Thermodynamic Simulations with High-Dimensional Input Spaces Using Sparse Polynomial Chaos Expansion: Retrofit of a Large Thermal Power Plant”. In: *applied sciences* 13.10751 (2023). <https://doi.org/10.3390/app131910751>.
- [26] Farshad Mohammadi et al. “Allocation of Centralized Energy Storage System and Its effect on Daily Grid Energy Generation Cos”. In: *ResearchGate* (Sept. 2016).
- [27] Tillmann Mühlpfordt et al. “Solving optimal power flow with non-Gaussian uncertainties via polynomial chaos expansion”. In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. 2017, pp. 4490–4496. DOI: 10.1109/CDC.2017.8264321.
- [28] Talha Bin Nadeem et al. “Distributed energy systems, A review of classification, technologies, applications, and policies”. In: *Energy Strategy Reviews* 48 (2023), p. 101096.
- [29] *NPY format*. <https://numpy.org/devdocs/reference/generated/numpy.lib.format.html>. accessed on 15/05/24.
- [30] *Power System Reliability Analysis (TB 26)*. CIGRE WG 03 of SC 38 (Power system analysis and techniques). CIGRE. 1987.

- [31] *POWER SYSTEM RELIABILITY ANALYSIS (Volume 2) Composite Power System Reliability Evaluation (TB 70)*. Task Force 38.03.10. CIGRE. 1992.
- [32] Ghanem R.G. and Spanos P.D. *Stochastic Finite Elements, a Spectral Approach*. New York: Springer-Verlag, 1991. DOI: <http://dx.doi.org/10.1007/978-1-4612-3094-6>.
- [33] Xavier Rixhon et al. “The Role of Electrofuels under Uncertainties for the Belgian Energy Transition”. In: *Energies* 14.4027 (2021). <https://doi.org/10.3390/en14134027>.
- [34] scikit-learn. *scikit-learn: 1.10. Decision Trees*. <https://scikit-learn.org/stable/modules/tree.html>. [Online; accessed 30-May-2024]. 2024.
- [35] Abdullah M. Al-Shaalan. “Reliability evaluation in generation expansion planning based on the expected energy not served”. In: *Journal of King Saud University - Engineering Sciences* 24.1 (Jan. 2012), pp. 11–18.
- [36] Gord Stephen. *Probabilistic Resource Adequacy Suite (PRAS) v0.6 Model Documentation*. Tech. rep. National Renewable Energy Laboratory, May 2021.
- [37] Bruno Sudret. “Global sensitivity analysis using polynomial chaos expansion”. In: *Reliability Engineering and System Safety* 93 (July 2008), pp. 964–979. DOI: 10.1016/j.ress.2007.04.002.
- [38] Bruno Sudret. “Polynomial chaos expansions and stochastic finite element methods”. In: CRC Press, Nov. 2015. Chap. 6, pp. 265–300.
- [39] Bruno Sudret, Marc Berveiller, and Maurice Lemaire. “A stochastic finite element procedure for moment and reliability analysis”. In: *European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique* 15 (May 2012). DOI: 10.3166/remn.15.825-866.
- [40] Bruno Sudret and Armen Der Kiureghian. “Stochastic Finite Element Methods and Reliability A State-of-the-Art Report”. In: (Jan. 2000).
- [41] UCLouvain. *Aurélia Hernandez*. <https://uclouvain.be/en/directories/aurelia.hernandez>. [Online; accessed 27-May-2024]. 2024.
- [42] UCTE. *System Adequacy Methodology*. 2009.
- [43] Nicolas Vauchel, Éric Garnier, and Thomas Gomez. “A multi-element non-intrusive Polynomial Chaos method using agglomerative clustering based on the derivatives to study irregular and discontinuous Quantities of Interest”. In: *Journal of Computational Physics* 473 (2023), p. 111763. ISSN: 0021-9991. DOI: <https://doi.org/10.1016/j.jcp.2022.111763>. URL: <https://www.sciencedirect.com/science/article/pii/S0021999122008269>.

- [44] Xiaoliang Wan and George Em Karniadakis. “An adaptive multi-element generalized polynomial chaos method for stochastic differential equations”. In: *Journal of Computational Physics* 209.2 (2005), pp. 617–642. ISSN: 0021-9991. DOI: <https://doi.org/10.1016/j.jcp.2005.03.023>. URL: <https://www.sciencedirect.com/science/article/pii/S0021999105001919>.
- [45] Mason White. “The Need for Speed: Julia Vs. Python”. In: *ResearchGate* (Jan. 2022).
- [46] Wikipedia. *Poisson distribution* — *Wikipedia, The Free Encyclopedia*. <http://en.wikipedia.org/w/index.php?title=Poisson%20distribution&oldid=1223426380>. [Online; accessed 25-May-2024]. 2024.
- [47] Yijun Xu et al. “Propagating Uncertainty in Power System Dynamic Simulations Using Polynomial Chaos”. In: *IEEE Transactions on Power Systems* 34 (Aug. 2018), pp. 338–348. DOI: 10.1109/TPWRS.2018.2865548.

Appendices

Appendix A

Calendar

On the following Table A.1, the dates of meetings and subjects of weekly work are listed. People involved are referenced following the list:

- AT: Alexandre Trifin (Writer)
- CV: Camille Vanhoenacker (ex-Writer)
- AH: Aurélia Hernandez (Reader)
- DC: Diederik Coppitters (Reader)
- FC: Francesco Contino (Supervisor)
- EDJ: Emmanuel De Jaeger (Supervisor)

Table A.1: Calendar of the Master's thesis

Dates	Object	People
11/10	First Meeting online	AT, CV, AH, DC, EDJ
11/10 - 08/11	Literature research	AT, CV
08/11	Tools introduction (Julia, Gurobi, Rheia)	AT, CV, AH, DC
08/11 - 22/11	Tools setup and software handling	AT, CV
22/11	Adequacy: theoretical background (1)	AT, CV, AH
22/11 - 27/11	Question of research	AT, CV
27/11	Adequacy: theoretical background (2)	AT, CV, AH
28/11	PCE: theoretical background	AT, CV, AH, DC
28/11 - 23/01	Test model research + Adequacy and PCE	AT, CV
23/01	Presentation of the first semester work	AT, CV, AH, DC, FC, EDJ
23/01 - 31/01	Test connection adequacy + PCE	AT, CV

31/01	Meeting about the tests	AT, CV, AH, DC
31/01 - 14/02	Poster preparation for the poster session	AT, CV
14/02	Poster session	CV, AH, DC, FC
14/02 - 19/02	Background writing	AT, CV
19/02	First face-to-face meeting	AT, CV, AH, DC, FC, EDJ
19/02 - 27/02	PCE loop+random var.+PCE+MC	AT, CV
27/02	Meeting about the week's work	AT, CV, AH, DC
27/02 - 06/03	MC+PCE loop+random var.+results	AT
06/03	Meeting about the week's work	AT, AH, DC, FC, EDJ
06/03 - 12/03	Wind turbine availability + ME-PCE	AT
12/03	Meeting about the week's work	AT, AH, DC
12/03 - 20/03	Multi-Element PCE Research	AT
20/03	Meeting about the week's work	AT, AH, DC, FC
20/03 - 29/03	Monotonic curves problems (1)	AT
29/03	Meeting about the week's work	AT, AH, DC
29/03 - 10/04	Monotonic curves problems (2)	AT
10/04	Meeting about Monotonic curves problem	AT, AH
11/04	Meeting about Monotonic curves problem	AT, DC
11/04 - 17/04	Mono. (tan., lin., quad.)+LOE Subspace	AT
17/04	Meeting about the week's work	AT, AH, FC, EDJ
17/04 - 25/04	Focus on one parameter	AT
25/04	Meeting about the week's work	AT, AH
25/04 - 29/04	Focus on one parameter	AT
29/04	Meeting about the week's work	AT, AH, FC
29/04 - 13/05	Subprofiles	AT
13/05	Meeting about the week's work	AT, DC
13/05 - present	Writing of the thesis	AT
22/05	Little meeting	AT, AH
27/05	Review of the Master's thesis	AT, AH, DC

Appendix B

Study Case IEEE 30-Bus System Data

General

Table B.1: General

Parameter	Value
$VOLL$ [€/MWh]	10000
Yearly Demand [TWh]	1.48955
Max Power Imports [MW]	0
S_B [MVA]	100
Ref Bus	1

- $VOLL$: Value of Loss-Load [€/MWh]
- S_B : Base power of the per-unit system [MVA]
- Ref Bus: Reference node (voltage angle $\Theta = 0$)
- Max Power Imports [MW]: Simultaneous import limit of electricity [MW]
- Yearly Demand [TWh]: Yearly electrical demand [MWh/year]

Buses

Table B.2: Buses

bus_i	type	p_d	q_d	g_s	b_s	area	v_m	v_a	baseKV	zone	v_{max}	v_{min}	$d_l\%$
1	3	0	0	0	0	1	1	0	132	1	1.06	0.94	0
2	2	21.7	12.7	0	0	1	1	0	132	1	1.06	0.94	7.657021877
3	1	2.4	1.2	0	0	1	1	0	132	1	1.06	0.94	0.846859562
4	1	7.6	1.6	0	0	1	1	0	132	1	1.06	0.94	2.681721948
5	2	94.2	19	0	0	1	1	0	132	1	1.06	0.94	33.23923783
6	1	0	0	0	0	1	1	0	132	1	1.06	0.94	0
7	1	22.8	10.9	0	0	1	1	0	132	1	1.06	0.94	8.045165843
8	2	30	30	0	0	1	1	0	132	1	1.06	0.94	10.58574453
9	1	0	0	0	0	1	1	0	1	1	1.06	0.94	0
10	1	5.8	2	0	19	1	1	0	33	1	1.06	0.94	2.046577276
11	2	0	0	0	0	1	1	0	11	1	1.06	0.94	0
12	1	11.2	7.5	0	0	1	1	0	33	1	1.06	0.94	3.952011291
13	2	0	0	0	0	1	1	0	11	1	1.06	0.94	0
14	1	6.2	1.6	0	0	1	1	0	33	1	1.06	0.94	2.187720536
15	1	8.2	2.5	0	0	1	1	0	33	1	1.06	0.94	2.893436838
16	1	3.5	1.8	0	0	1	1	0	33	1	1.06	0.94	1.235003529
17	1	9	5.8	0	0	1	1	0	33	1	1.06	0.94	3.175723359
18	1	3.2	0.9	0	0	1	1	0	33	1	1.06	0.94	1.129146083
19	1	9.5	3.4	0	0	1	1	0	33	1	1.06	0.94	3.352152435
20	1	2.2	0.7	0	0	1	1	0	33	1	1.06	0.94	0.776287932
21	1	17.5	11.2	0	0	1	1	0	33	1	1.06	0.94	6.175017643
22	1	0	0	0	0	1	1	0	33	1	1.06	0.94	0
23	1	3.2	1.6	0	0	1	1	0	33	1	1.06	0.94	1.129146083
24	1	8.7	6.7	0	4.3	1	1	0	33	1	1.06	0.94	3.069865914
25	1	0	0	0	0	1	1	0	33	1	1.06	0.94	0
26	1	3.5	2.3	0	0	1	1	0	33	1	1.06	0.94	1.235003529
27	1	0	0	0	0	1	1	0	33	1	1.06	0.94	0
28	1	0	0	0	0	1	1	0	132	1	1.06	0.94	0
29	1	2.4	0.9	0	0	1	1	0	33	1	1.06	0.94	0.846859562
30	1	10.6	1.9	0	0	1	1	0	33	1	1.06	0.94	3.740296401

- bus_i : index of the bus
- type: type of the bus (1: PQ bus, 2: PV bus, 3: Slack bus)
- p_d : active power injection of the bus [MW]
- q_d : reactive power injection of the bus [MVAR]
- g_s : conductance [pu]

- b_s : susceptance [pu]
- area: area of working
- v_m : voltage magnitude [pu]
- v_a : voltage angles [pu]
- baseKV: base KV
- zone: zone of working
- v_{max} : maximum voltage [pu]
- v_{min} : minimum voltage [pu]
- $d_l\%$: Load share at each bus [%]

Branches

Table B.3: Branches

bus_{from}	bus_{to}	r	x	b	rateA	ratio	ang_{min}	ang_{max}	$MTTF$	$MTTR$
1	2	0.0192	0.0575	0.0528	138	0	-30	30	20000	24
1	3	0.0452	0.1652	0.0408	152	0	-30	30	20000	24
2	4	0.0570	0.1737	0.0368	139	0	-30	30	20000	24
3	4	0.0132	0.0379	0.0084	135	0	-30	30	20000	24
2	5	0.0472	0.1983	0.0418	144	0	-30	30	20000	24
2	6	0.0581	0.1763	0.0374	139	0	-30	30	20000	24
4	6	0.0119	0.0414	0.0090	148	0	-30	30	20000	24
5	7	0.0460	0.1160	0.0204	127	0	-30	30	20000	24
6	7	0.0267	0.0820	0.0170	140	0	-30	30	20000	24
6	8	0.0120	0.0420	0.0090	148	0	-30	30	20000	24
6	9	0.0000	0.2080	0.0000	142	0.978	-30	30	20000	24
6	10	0.0000	0.5560	0.0000	53	0.969	-30	30	20000	24
9	11	0.0000	0.2080	0.0000	142	1.000	-30	30	20000	24
9	10	0.0000	0.1100	0.0000	267	1.000	-30	30	20000	24
4	12	0.0000	0.2560	0.0000	115	0.932	-30	30	20000	24
12	13	0.0000	0.1400	0.0000	210	1.000	-30	30	20000	24
12	14	0.1231	0.2559	0.0000	29	0	-30	30	20000	24
12	15	0.0662	0.1304	0.0000	29	0	-30	30	20000	24
12	16	0.0945	0.1987	0.0000	30	0	-30	30	20000	24
14	15	0.2210	0.1997	0.0000	20	0	-30	30	20000	24
16	17	0.0524	0.1923	0.0000	38	0	-30	30	20000	24
15	18	0.1073	0.2185	0.0000	29	0	-30	30	20000	24
18	19	0.0639	0.1292	0.0000	29	0	-30	30	20000	24

19	20	0.0340	0.0680	0.0000	29	0	-30	30	20000	24
10	20	0.0936	0.2090	0.0000	30	0	-30	30	20000	24
10	21	0.0348	0.0749	0.0000	30	0	-30	30	20000	24
10	22	0.0727	0.1499	0.0000	29	0	-30	30	20000	24
21	22	0.0116	0.0236	0.0000	29	0	-30	30	20000	24
15	23	0.1000	0.2020	0.0000	29	0	-30	30	20000	24
22	24	0.1150	0.1790	0.0000	26	0	-30	30	20000	24
23	24	0.1320	0.2700	0.0000	29	0	-30	30	20000	24
24	25	0.1885	0.3292	0.0000	27	0	-30	30	20000	24
25	26	0.2544	0.3800	0.0000	25	0	-30	30	20000	24
25	27	0.1093	0.2087	0.0000	28	0	-30	30	20000	24
28	27	0.0000	0.3960	0.0000	75	0.968	-30	30	20000	24
27	29	0.2198	0.4153	0.0000	28	0	-30	30	20000	24
27	30	0.3202	0.6027	0.0000	28	0	-30	30	20000	24
29	30	0.2399	0.4533	0.0000	28	0	-30	30	20000	24
8	28	0.0636	0.2000	0.0428	140	0	-30	30	20000	24
6	28	0.0169	0.0599	0.0130	149	0	-30	30	20000	24

- bus_{from} : bus on one side which the line will be connected with
- bus_{to} : bus on the other side which the line will be connected with
- r : Resistance [pu] with base sb=100 [MVA]
- x : Reactance [pu] with base sb=100 [MVA]
- b : Susceptance [pu] with base sb=100 [MVA]
- rateA: Power capacity of the line [MW]
- ratio: ratio of the transformers [%]
- ang_{max} : Maximum voltage angle of bus [°]
- ang_{min} : Minimum voltage angle of bus [°]
- $MTTF$: Mean-Time-To-Failure of each line [hours]
- $MTTR$: Mean-Time-To-Repair of each line [hours]

Generators

Table B.4: Generators

bus	p_g	q_g	q_{max}	q_{min}	v_g	mBase	status	p_{max}	p_{min}	$c(n-1)$	...	$c0$	$MTTF$	$MTTR$
9	135.5	5	10	0	1	100	1	180.67	0	0	18.42	0	2000	50
9	135.5	5	10	0	1	100	1	180.67	0	0	18.42	0	2000	50
12	46	3	46	-40	1	100	1	61.33	0	0	52.18	0	2000	50
12	46	3	46	-40	1	100	1	61.33	0	0	52.18	0	2000	50
5	0	0	40	-40	1	100	1	0	0	0	0	0	2000	50
8	0	15	40	-10	1	100	1	0	0	0	0	0	2000	50
11	0	9	24	-6	1	100	1	0	0	0	0	0	2000	50
13	0	9	24	-6	1	100	1	0	0	0	0	0	2000	50

- bus: bus connected to the generator
- p_g : active power of the generator [MW]
- q_g : reactive power of the generator [MVAR]
- q_{max} : max reactive power of the generator [MVAR]
- q_{min} : min reactive power of the generator [MVAR]
- v_g : voltage of the generator [pu]
- mBase : Base power of the per-unit system [MVA]
- status : 1 if in service, 0 if not
- p_{max} : Max capacity of each generator [MW]
- p_{min} : Min capacity of each generator [MW]
- $c(n-1)$: Quadratic marginal cost of each generator [€/MWh²]
- ...: Linear marginal cost of each generator [€/MWh]
- $c0$: Fixed cost of each generator [€]
- $MTTF$: Mean-Time-To-Failure of each generator [hours]
- $MTTR$: Mean-Time-To-Repair of each generator [hours]

Distributed Generators

Table B.5: Distributed Generators

Type	Power_MW	LF
Gas CHP	0	0
Biomass	0	0
Waste	0	0

- Power_MW : Max capacity of each distributed generator
- LF: load factor

Renewable Energy Sources

Name	Node	Pmax_MW	Nb_wt	MTTF	MTTR
Belwind	2	160.9901642	56	1060	290
C-Power	2	331.5832506	54	1060	290
Mermaid	2	221.4320856	28	1060	290
Nobelwind	2	155.3413866	50	1060	290
Norther	2	373.3842055	44	1060	290
Northwester2	2	205.7096543	23	1060	290
Northwind	2	223.6915966	72	1060	290
Rentel	2	290.6296123	42	1060	290
Seastar	2	237.2486631	30	1060	290
P1_3500_15	13	728	49	1060	290
P23_3500_15	1	2782	185	1060	290

- Name: Name of offshore wind farms
- Nb_wt: Number of wind turbines in each offshore wind farm
- Pmax_MW : Max capacity of each wind farm [MW]
- MTTF: Mean-Time-To-Failure of wind turbines in the concerned wind farm [hours]
- MTTR: Mean-Time-To-Repair of wind turbines in the concerned wind farm [hours]

Distributed Renewable Power

Table B.7: Distributed Renewable Power

Psolar_MW	Ponshore_MW	Phydro_MW
0	0	0

- Psolar_MW: PV capacity [MW]
- Ponshore_MW: Wind onshore capacity [MW]
- Phydro_MW: Hydro capacity [MW]

Imports

Empty

Pumped Hydro Storage (PHS)

Empty

Hydrogen Technology

Table B.8: Hydrogen Technology

Elec. η [-]	H2toP η [-]	Compression η [-]	€ Elec. [€/MWh]	€ H2 Impo. [€/MWh]
0.7	0.43	0.09	0	150

- Elec. η [-]: Efficiency of electrolyzers [-]
- H2toP η [-]: Efficiency of hydrogen-to-power units [-]
- Compression η [-]: Electrical consumption for H2 compression to 350 bars [MWh_el/MWh_H2]
- € Elec. [€/MWh]: Operational cost of electrolyzers (not including the fuel cost and compression needs) [€/Mwh_el]
- € H2 Impo. [€/MWh]: Cost of H2 imports [€/Mwh_h2]

Batteries

Empty

Hydrogen Pipelines

Empty

Hydrogen Imports

Empty

Hydrogen Demand

Table B.9: Hydrogen Demand

Annual H2 demand [kt/year]	H2Flex share [-]	H2Fixed share [-]
0	0.5	0.5

- Annual H2 demand [kt/year]: Yearly H2 demand [kt_H2/year]
- H2Flex share [-]: Share of flexible H2 demand [-]
- H2Fixed share [-]: Share of fixed H2 demand [-]

Electrolyzer

Empty

Hydrogen to Power

Empty

Hydrogen Storage

Empty

Electric Vehicles

Table B.10: Electric Vehicles Part 1

V0 share [-]	V1H share [-]	V1M share [-]	V2H share [-]	Number of EVs [-]
1	0	0	0	0

Table B.11: Electric Vehicles Part 2

Charger Power [kW]	Yearly Distance [km/year/car]	Car Efficiency [kWh/km]
7	15000	0.18

- Number of Evs [-]: Number of EVs in 2030 [cars]
- V0 share [-]: Share of V0 vehicles [-]
- V1H share [-]: Share of V1H vehicles [-]
- V2H share [-]: Share of V2H vehicles [-]
- V1M share [-]: Share of V1M vehicles [-]
- Charger Power [kW]: Power of EV charger [MW]
- Yearly Distance [km/year/car]: [km/year/car]
- Car Efficiency [kWh/km]: [kwh/km]

Load Shifting

Table B.12: Load Shifting

Max shift pow. [MW]	Max daily shift ener. [MWh/day]	Shift cost [€/MWh]
0	0	50

- Max shift pow. [MW]: Max shiftable power [MW_{el}]
- Max daily shift ener. [MWh/day]: Maximum daily shiftable energy [MWh_{el}/day]
- Shift cost [€/MWh]: Cost to shift energy [€/Mwh]

Voluntary Load Shedding

Table B.13: Voluntary Load Shedding

Max shed power [MW]	Max daily shed energy [MWh/day]	Shed cost [€/MWh]
0	0	300

- Max shed power [MW]: Max voluntary shedable power [MW]
- Max daily shed energy [MWh/day]: Max voluntary shed energy per day [MWh/day]
- Shed cost [€/MWh]: Cost to shed energy [€/MWh]

Heat Pumps

Table B.14: Heat Pumps Part 1

HP0 share [-]	HP1H share [-]	HP1M share [-]	Number of HPs [-]
1	0	0	0

Table B.15: Heat Pump Part 2

Yearly Demand [TWh/year]	COP [-]	HP Thermal Power [kW _{th}]
0	3	10

- HP0 share [-]: Share of HP0 heat-pumps [-]
- HP1H share [-]: Share of HP1H heat-pumps [-]

- HP1M share [-]: Share of HP1M heat-pumps [-]
- Yearly Demand [TWh/year]: Yearly electrical demand from heat-pumps [MWh/year]
- Number of Hps [-]: Number of HPs in 2030 [HPs]
- COP [-]: COP of heat-pumps [W_{th}/W_{el}]
- HP Thermal Power [kW_{th}]: Elec power of installed HP [MW_{el}]

Appendix C

Source Code

Please access the present link:
<https://github.com/atrifin/MPDCOPF-PCE>

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl