Adaptive Multi-Agent Systems for Wind Power Forecasting

Tanguy Esteoule

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Tanguy ESTEOULE

Le 3 décembre 2019

Prévision de production de parcs éoliens par systèmes multi-agents auto-adaptatifs

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Informatique et Télécommunications**

Unité de recherche :
IRIT : Institut de Recherche en Informatique de Toulouse

Thèse dirigée par
Marie-Pierre GLEIZES et Carole BERNON

Jury

M. Pierre Pinson, Rapporteur
M. Yves Demazeau, Rapporteur
M. Gauthier Picard, Examinateur
Mme Marie-Pierre Gleizes, Directrice de thèse
Mme Carole Bernon, Co-directrice de thèse
M. Julien Guépet, Co-directeur de thèse

THESIS

presented at

Université Toulouse III - Paul Sabatier

F.S.I. FACULTÉ DES SCIENCES ET INGÉNIERIE

to obtain the title of

DOCTEUR DE L'UNIVERSITÉ DE TOULOUSE

delivered by

UNIVERSITÉ TOULOUSE III - PAUL SABATIER

Mention INTELLIGENCE ARTIFICIELLE

by

TANGUY ESTEOULE

Doctoral school: Mathématiques, Informatique, Télécommunication de Toulouse
Laboratory: Institut de Recherche en Informatique de Toulouse
Team: Systèmes Multi-Agents Coopératifs

Adaptive Multi-Agent Systems for Wind Power Forecasting

JURY

Pierre PINSON	*Professor, Technical University of Denmark*	(Reviewer)
Yves DEMAZEAU	*Director of Research, Université Grenoble Alpes*	(Reviewer)
Gauthier PICARD	*Professor, Mines de Saint-Étienne*	(Examiner)
Marie-Pierre GLEIZES	*Professor, Université Toulouse III*	(Supervisor)
Carole BERNON	*Associate Professor, Université Toulouse III*	(Co-Supervisor)
Julien GUÉPET	*Chief Technology Officer, *SWIFT*	(Co-Supervisor)

Tanguy Esteoule

Prévision de production de parcs éoliens par Systèmes Multi-Agents Auto-Adaptatifs

Directrice de thèse : Marie-Pierre Gleizes, Professeur, Université de Toulouse
Co-Directeurs : Carole Bernon, Maître de conférences, Université de Toulouse
Julien Guépet, Directeur de la technologie, *SWIFT

Résumé

Le recours aux énergies renouvelables, notamment l'éolien, est une des solutions communément retenues pour limiter l'aggravation du changement climatique en cours. La variabilité et l'intermittence de ces sources d'énergie constituent la principale contrainte à gérer pour assurer l'intégration des énergies renouvelables sur le réseau électrique. Ce problème peut être en partie résolu par l'amélioration des prévisions de production à court et moyen termes.

La théorie des AMAS (Adaptive Multi-Agent Systems) propose de résoudre des problèmes complexes par auto-organisation pour lesquels aucune solution algorithmique n'est connue. Le comportement local et coopératif des agents permet au système de s'adapter à un environnement dynamique pour maintenir le système dans un état de fonctionnement adéquat. Dans cette thèse, cette approche est appliquée à la prévision de production de parcs éoliens. Plus précisément, nous étudions l'intégration de données à plus fine échelle (les parcs éoliens pour une région ou les éoliennes pour un parc) dans le modèle de prévision.

Nous proposons donc une méthode prenant en compte des données locales dans la prévision globale et plus précisément les interdépendances entre la production des éoliennes et des parcs. L'étude a mené à la conception de deux systèmes multi-agents auto-adaptatifs : AMAWind-Turbine prévoyant la production d'un parc en utilisant les données des éoliennes, et AMAWind-Farm prévoyant la production d'une région en utilisant les données des parcs. Ces systèmes ont été testés en conditions réelles sur cinq parcs éoliens actuellement en cours d'exploitation. Les expérimentations effectuées ont validé le bon fonctionnement des systèmes et ont permis d'observer une baisse d'erreur de prévision.

Tanguy Esteoule

ADAPTIVE MULTI-AGENT SYSTEMS FOR WIND POWER FORECASTING

Supervisor: Marie-Pierre Gleizes, Professor, Université de Toulouse
Co-Supervisors: Carole Bernon, Associate Professor, Université de Toulouse
Julien Guépet, Chief Technology Officer, *SWIFT

Abstract

The use of renewable energies, particularly wind power, is one of the solutions commonly admitted to limit the worsening of ongoing climate change. The variability and intermittency of these energy sources are the main constraints to be managed to ensure the integration of renewable energies into the electricity grid. This problem can be partly solved by improving production forecasts in the short and medium term.

The theory of AMAS (Adaptive Multi-Agent Systems) proposes to solve complex problems by self-organization for which no algorithmic solution is known. The local and cooperative behavior of the agents enables the system to adapt to a dynamic environment for maintaining the system in an adequate operating state. In this thesis, this approach is applied to the forecasting of wind farm production. More specifically, we are studying the integration of finer scale data (wind farms for a region or wind turbines for a farm) into the forecast model.

Therefore, we propose a method that takes into account local data in the global forecast and more precisely the interdependencies between wind turbine and wind farm productions. The study led to the design of two adaptive multi-agent systems: AMAWind-Turbine forecasting the production of a wind farm using wind turbine data, and AMAWind-Farm forecasting the production of a region using wind farm data. These systems have been tested in real conditions on five wind farms currently operating. The experiments carried out validated the proper functioning of the systems and showed a decrease in forecasting error.

Remerciements

This is the end, beautiful friend... ♫

Le doctorat est souvent dépeint comme un exercice très personnel. Il est vrai que jamais auparavant je ne m'étais autant remis en question et n'avais autant exprimé de doutes sur mon travail. Malgré cette lutte intérieure, il ne faut pas oublier les autres personnes derrière cette thèse qui m'ont guidé, corrigé, conseillé, soutenu et celles qui m'ont permis de rendre ces moments agréables. Pour faire le lien avec mon sujet de recherche, on pourrait dire que toutes les interactions coopératives que j'ai eu ont fait émerger ce travail conséquent, quelque chose de bien plus grand que ce que je pensais pouvoir effectuer à mon simple niveau. Voici donc quelques mots doux pour toutes ces personnes. Ne rougissez pas, tout est mérité.

Tout d'abord, je tiens à remercier Pierre Pinson, Yves Demazeau et Gauthier Picard d'avoir accepté d'évaluer ma thèse. Merci pour vos remarques et vos apports.

Marie-Pierre, merci de m'avoir suivi tout au long de cette thèse. Malgré toutes tes responsabilités tu as toujours su trouver du temps pour me conseiller. Merci également de prendre tant soin de cette équipe. Carole, merci pour ta rigueur et ta disponibilité. Cette thèse serait bien moins claire sans tes remarques pertinentes. Un grand merci plus particulièrement pour la fin de thèse où tu as relu avec grand soin ce manuscrit ! Un grand merci également à Pierre pour ta constante bonne humeur et pour ta volonté de transmettre ta passion. Tes sifflements me manquent déjà ! Alexandre, merci pour ton aide et tes conseils en début de thèse. Merci plus globalement à l'équipe SMAC qui m'a montré une autre facette de l'informatique, bien plus philosophique que je le croyais en sortant des études. C'est impossible de nommer tout le monde, j'en oublierais forcément. Mais merci à tous les anciens et actuels stagiaires/doctorants/post-doc/permanents pour tous ces bons moments et cette convivialité permanente. Les parties endiablées d'Avalon me manquent déjà !

Julien, merci pour ta rigueur et ton sens de la pédagogie, qui m'ont apporté aussi bien sur le plan scientifique qu'humain. Morgane, merci de m'avoir laissé ma chance et bravo pour tout ce que tu as entrepris avec tant d'intégrité et d'humilité. Merci à meteo*swift que j'ai eu la chance de voir grandir. Malgré les nombreux obstacles vous avez toujours su rebondir et l'histoire ne fait que commencer. Vivement la suite, pour que je vous montre mon vrai niveau au baby !

Merci aux PPV, mes amis grenoblois, dont certains me supportent depuis très très longtemps. Une belle amitié qui perdure malgré la distance et le temps ! J'espère fêter le Nouvel An encore de nombreuses fois avec vous.. ou pas !

Merci à tous mes amis du CPP et de l'Ensimag, toutes ces années d'études sont passées en un rien de temps avec vous. Merci également à toute la bande des thésards du BDS qui m'ont

sans doute motivé sans le savoir à suivre cette voie. Bonne chance à ceux qui soutiennent après moi et qui risquent à tout moment de recevoir une question gênante d'un docteur fraichement diplômé...

Merci également à mes anciens colocs à qui je dédicace, entre deux tranches de pâté croute, un petit "Couuurs Merrrry".

Merci à tous les membres de ma grande et belle famille. Merci à toi maman pour m'avoir transmis ta créativité et toi papa pour ta curiosité. Merci à vous deux de m'avoir partagé votre passion pour la musique qui a toujours rythmé ma vie. J'en profite pour faire un coucou à Victoria et au petit Noah qui vient tout juste de naître !

Merci également aux personnes que je n'ai pas citées et qui ne se retrouvent pas dans ces remerciements. Si vous lisez ces mots c'est que vous contribuez un peu à cette thèse en la lisant.

Et bien évidemment, le pompon sur la Garonne, merci Audrey pour égayer ma vie depuis tant d'années. Je savais dès le départ que la thèse ne serait pas une partie de plaisir mais grâce à ton soutien permanent j'ai pu arriver au bout sans encombres. Une nouvelle page se tourne, je ne sais pas quelle sera la suite mais j'ai hâte de passer ces moments avec toi !

Contents

General Introduction **1**

Contribution of the thesis 1

Manuscript organization 2

I Context and State of the Art **5**

1 Energy Overview **7**

1.1 History of energy 8

1.2 Overview of the electricity sector 9

1.2.1 Principles of electricity production 9

1.2.2 Different primary sources 10

1.2.2.1 Non-renewable resources 10

1.2.2.2 Renewable resources 11

1.2.3 Main electricity key figures 11

1.2.3.1 Current electricity mix is mainly based on fossil fuels 11

1.2.3.2 An energy future focused on renewable energy 12

1.3 Renewable energies 14

1.3.1 A growing sector 15

1.3.2 Main issues of concern 16

1.3.2.1 Intermittency 16

1.3.2.2 Variability 17

1.3.2.3 A still limited energy storage 17

1.4 Electricity markets 18

1.4.1 Market players 19

1.4.2 Functioning of electricity markets 19

1.4.3 Specificity of renewable energies on the market 21

1.4.4 Grid balancing 21

2 Wind Power Forecasting **23**

2.1 Wind energy 24

2.1.1 The wind as energy producer 24

2.1.1.1 Wind definition 24

2.1.1.2 Statistical distribution 26

2.1.1.3 Physical phenomena 27

2.1.2 Wind turbines 28

2.1.2.1 General principles 29

2.1.2.2 Wind turbine types 29

2.1.2.3 Theoretical studies 31

2.1.2.4 Power curve 32

2.1.2.5 The production in practice 33

2.1.2.6 Curtailments 34

2.1.2.7 Effects to be considered 35

2.2 Wind power forecasting 37

2.2.1 Different uses of forecasts 37

2.2.2 Forecast formats 38

2.2.3 Offshore specificity 39

2.3 Physical models 39

2.3.1 Numerical Weather Prediction (NWP) 40

2.3.1.1 The weather: a chaotic system 40

2.3.1.2 Description of the NWP models 40

2.3.1.3 The particularities of the wind forecasting 41

2.3.2 Downscaling method 42

2.4 Statistical models 44

2.4.1 Time series models 44

2.4.2 Machine Learning methods 45

2.4.2.1 Linear methods 46

2.4.2.2 Support Vector Machine (SVM) 46

2.4.2.3 Artificial Neural Networks 46

2.4.2.4 Ensemble methods 47

2.5 The spatial forecasting methods 48

2.5.1 Short scale dependencies 48

2.5.2 Large scale dependencies 49
2.6 Analysis of the current models 50
2.6.1 Analysis criteria 50
2.6.2 Evaluation of the criteria 51
2.6.3 Synthesis 51

II Contributions **55**

3 Integration of Interdependencies in Forecasting through Cooperation **57**
3.1 Interdependencies in wind power forecasting 58
3.1.1 Short scale 58
3.1.2 Large scale 59
3.2 Multi-Agent Systems 61
3.2.1 Agent 61
3.2.2 Environment 63
3.2.3 Properties of MAS 63
3.2.4 MAS for energy and forecasting 64
3.3 AMAS: Adaptive Multi-Agent Systems 64
3.3.1 Introduction 65
3.3.1.1 Complex systems 65
3.3.1.2 Emergence 65
3.3.2 Interaction and cooperation 66
3.3.3 Functional adequacy 67
3.3.4 Adaptation 68
3.3.5 Applications of AMAS 69
3.3.6 Relevance of the AMAS approach to the problem 70

4 AMAWind-Turbine: a MAS for Wind Power Forecasting at Turbine-Level **73**
4.1 Objectives 74
4.2 Environment 74
4.3 Entities 75
4.3.1 Grid Point entity 77
4.3.2 Grid Point Hour entity 77
4.3.3 Wind Turbine 78
4.4 A cooperative agent: Wind Turbine Hour agent 78
4.4.1 Description 78

4.4.2 Local goal 79
4.4.3 Behavior 79
4.4.4 Non Cooperative Situations 80
4.4.4.1 Uselessness 81
4.4.4.2 Incompetence 81
4.5 Criticality implementation 82
4.5.1 Percentiles computation 83
4.5.2 Generic criticality function 84
4.5.3 Subcriticalities computation 85
4.5.4 Final criticality 86
4.6 Technical choices 87
4.6.1 Software architecture 87
4.6.2 Agent initialization 88
4.6.2.1 Choice of initial forecasts 89
4.6.2.2 Increment choice 89
4.6.3 Agent scheduling and stop condition 89
4.7 Conclusion 90

5 AMAWind-Farm: a MAS for Wind Power Forecasting at Farm-Level 91
5.1 Objectives 91
5.2 Entities 92
5.2.1 Grid Point and Grid Point Hour entities 93
5.2.2 Farm 93
5.3 A cooperative agent: Farm Hour agent 94
5.3.1 Description 94
5.3.2 Local goal 95
5.3.3 Behavior 95
5.4 Implementation differences 95
5.4.1 A criticality adapted to heterogeneous data 96
5.4.2 A variable increment 97
5.5 Conclusion 97

III Experimental Evaluation 99

6 Methodology of the Experiments 101
6.1 General description of the experiments 101

6.2 Data description 102
6.2.1 Production data 102
6.2.2 Weather forecasts 103
6.3 Data preprocessing 105
6.3.1 Filtering of errors 105
6.3.2 Filtering of maintenance or failure cases 105
6.4 Evaluation criteria 106
6.4.1 Criteria related to the system validation 106
6.4.2 Criteria related to wind power forecasting 107
6.5 Model validation by cross-validation 108

7 AMAWind-Turbine Experiments 111
7.1 Description 111
7.2 System validation 113
7.2.1 Impact of the initial forecast 113
7.2.2 Computation times 114
7.2.3 Criticality evolution 115
7.2.4 Relationship between criticality and forecast error 116
7.2.5 Number of cycles and convergence 117
7.3 Evaluation of forecast accuracy 119
7.3.1 Reference forecasting methods 119
7.3.1.1 Choice of a machine learning algorithm 119
7.3.1.2 Method integrating interdependencies 121
7.3.2 Forecast errors 122
7.3.3 Detailed results 123
7.4 General synthesis 124

8 AMAWind-Farm Experiments 127
8.1 Description 127
8.2 System validation 129
8.2.1 Wind farm selection 129
8.2.2 Impact of the normalization 129
8.2.3 Impact of the initial forecast 131
8.2.4 Computation times 131
8.2.5 Criticality evolution 132
8.2.6 Relationship between criticality and forecast error 132

8.2.7 Number of cycles and convergence 133
8.3 Evaluation of forecast accuracy 133
8.3.1 Reference forecasting methods 134
8.3.2 Forecast errors 134
8.3.3 Detailed results 135
8.4 General synthesis 136

Conclusion and Perspectives **139**

General Conclusion **139**

Contribution **141**

Perspectives **142**

Bibliography **145**

List of Figures **159**

List of Tables **163**

Glossary **165**

General Introduction

IN 2018, 75% of the world's electricity production came from energy sources considered as non-renewable (coal, nuclear, natural gas and oil). A global effort is underway to mitigate climate change through an energy transition to a sustainable model. To achieve the objective of not exceeding a 2°C increase by 2100, the United Nations Conference on Climate Change COP21 has set a goal of 27% renewable energy in the overall energy supply in the European Union by 2030 compared to 16% in 2018.

Wind power will play a key role in the energy transition because the source of energy is unlimited and the exploitation of this resource does not emit greenhouse gases during electricity production. In 2018, wind energy covered more than 4% of the global electricity production. Currently, the annual growth rate of wind power activity is about 12%, it corresponds to an additional installed capacity of 50.1 GW in 2018.

However, due to the wind variability and the still high cost of electricity storage, we cannot depend solely on wind energy. In order to obtain an efficient energy mix, a precise estimate of electricity production and consumption is required to regulate the electricity grid. Electricity markets are thus organized as electricity pools, gathering production and consumption offers in order to dynamically find the quantities and prices for electricity generation and consumption maximizing social welfare. Wind power producers propose energy offers based on forecasts. The market clearing is designed to match production offers and consumption bids through an auction process. Since power producers are financially responsible for any deviation from these contracts, improving wind power forecasting accuracy enables to reduce the penalties they incur.

Contribution of the Thesis

In this thesis, we explore the problem of integrating local interdependencies between wind turbine production into wind farm forecasts. The dependencies between the productions of close turbines (e.g., the wake effect that occurs when a wind turbine disrupts the wind behind it and can cause a decrease in production of nearby wind turbines) represent additional information rarely integrated into the wind power forecasting models. By applying the Adaptive Multi-Agent System approach, we designed AMAWind-Turbine (*Adaptive Multi-Agent system for Wind power forecasting at Turbine-level*), an AMAS designed to forecast the production of a wind farm by taking into account the short-scale interdependencies between wind turbines composing this farm.

The interdependency problem can also be applied on a larger scale by integrating interdependencies between wind farms into regional forecasts (i.e a set of wind farms).

In this case, it is in particular the existing correlation between the productions of wind farms located in similar wind zones that makes it possible to improve the forecast at global scale. We have developed AMAWind-Farm (*Adaptive Multi-Agent system for Wind power forecasting at Farm-level*) which is an AMAS designed to forecast the production of a region by taking into account the large-scale interdependencies between wind farms in this region.

We evaluate our approach through experiments on both systems. This thesis being the result of a partnership with **SWIFT*, a company specialized in wind power forecasting, we are able to test our systems on real data in an operational context and compare our results with several methods currently used. From the analysis of the results, we highlight the advantages and limitations of our approach and point out the perspectives offered by this work.

Manuscript Organization

This manuscript is composed of three main parts.

The first part presents the context and the field of application of this thesis, wind power forecasting, and the motivations behind this work. This part is divided into two chapters:

- ▷ Chapter 1 introduces the *current context of the energy sector* and focuses on renewable energies by detailing their limitations. It also describes the functioning of the electricity markets. This chapter helps to understand the need for production forecasting.
- ▷ Chapter 2 presents a *state of the art of wind power forecasting*. The chapter first introduces generalities on wind energy and forecasting. It then describes the main approaches currently used by companies for forecasting. Finally this chapter specifies the problem that is retained in this thesis, which is the consideration of spatial interdependencies in the forecasts, and compares the approaches presented before with respect to criteria related to the problem.

Part 2 provides the theoretical and technical bases and the formal description of the two Adaptive Multi-Agent Systems (AMAS) designed in this thesis: AMAWind-Turbine and AMAWind-Farm. This part is detailed in the following chapters:

- ▷ Chapter 3 first describes the problem of interdependency in production forecasting by decomposing it on a short and large scales. After an introduction to Multi-Agent Systems (MAS), the chapter then proposes to study a particular type of MAS, *Adaptive Multi-Agent Systems* (AMAS), as a potential approach to take into account the interdependencies at both scales.
- ▷ Chapter 4 presents *AMAWind-Turbine* (Adaptive Multi-Agent system for Wind power forecasting at Turbine-level): an AMAS designed to forecast the production of a wind farm by taking into account the short-scale interdependencies between wind turbines. Its design and the behaviors of its agents are explained.
- ▷ Chapter 5 presents *AMAWind-Farm* (Adaptive Multi-Agent system for Wind power forecasting at Farm-level): an AMAS designed to forecast the production of a region (i.e.

several wind farms) by taking into account the large-scale interdependencies between wind farms. Since the design of AMAWind-Farm has a lot in common with the design of AMAWind-Turbine, this shorter chapter focuses on the differences between the two systems.

The last part, part 3, presents the experimental aspect of this thesis through three chapters:

- Chapter 6 describes the *common experimental conditions for the evaluation* of the two models. It presents the data used, the preprocessing applied to data, the evaluation criteria and the cross-evaluation method.
- Chapter 7 presents the *experiments carried out on AMAWind-Turbine*. First of all, the system is studied according to several validation criteria to test its proper functioning. Then a comparison of the results against reference algorithms is performed on real production data. Finally, a general synthesis points out the properties and limitations of the system.
- Chapter 8 details the *experiments done on AMAWind-Farm*. It follows the same pattern as the previous chapter by adding specific criteria to this system. The results are also compared with state-of-the-art methods and the results are finally discussed.

Finally, I conclude this manuscript with a synthesis of the work carried out in the context of this thesis and an enumeration of some interesting work perspectives.

Context and State of the Art

1 Energy Overview

This chapter aims to introduce the context necessary to understand the need for production forecasting. Firstly, it describes the historical, current and future context of the energy sector and then focuses on renewable energies. Finally, it presents the electricity markets.

1.1 History of energy 8
1.2 Overview of the electricity sector 9
1.2.1 Principles of electricity production 9
1.2.2 Different primary sources 10
1.2.2.1 Non-renewable resources 10
1.2.2.2 Renewable resources 11
1.2.3 Main electricity key figures 11
1.2.3.1 Current electricity mix is mainly based on fossil fuels 11
1.2.3.2 An energy future focused on renewable energy 12
1.3 Renewable energies 14
1.3.1 A growing sector 15
1.3.2 Main issues of concern 16
1.3.2.1 Intermittency 16
1.3.2.2 Variability 17
1.3.2.3 A still limited energy storage 17
1.4 Electricity markets 18
1.4.1 Market players 19
1.4.2 Functioning of electricity markets 19
1.4.3 Specificity of renewable energies on the market 21
1.4.4 Grid balancing 21

1.1 History of energy

Homo erectus was the first being to use fire. Fire, this natural source of energy, domesticated by "caveman" is a precious legacy of this period of pre-history known as Acheulean. The use of fire by mankind certainly dates back almost a million years with the oldest traces of known fires found in a cave in South Africa [1]. Between the discovery of fire and the time when pre-historic men managed to "make fire", they recovered the fire dying from natural fires, transported it and maintained it.

The use and control of fire by mankind is a major turning point in the history of human evolution. The benefits of using this first energy are immense and energy has become an essential resource for human survival and development. Since then and for hundreds of thousands of years, wood, used as fuel for fire, has been the only source of energy. In ancient times, Man began to use the power of water as an energy source. Hydropower was used in China at least 2000 years ago; the waterwheel was invented in ancient Greece and Rome, and in the year 13 B.C., the Roman engineer and writer Marcus Vitruvius Pollio described a grain mill driven by a waterwheel and a cogwheel gear [2].

The first uses of wind as an energy source are older, and mainly used in the maritime sector for the transport of goods in sailing ships along the Nile as early as 5000 BC. The use of wind as an energy source through the first windmills dates back to the 7$^{\text{th}}$ century BC, during which time the King of Babylon designed an irrigation system for the Mesopotamian plain using wind energy [3].

Until the beginning of the industrial era, wood, wind and water were the only energy sources commonly used by humans. As early as the Middle Ages, the first uses of coal were reported. Marco Polo informed in his travel diaries that in China, people used coal for heating and cooking food. But the widespread use of coal as an energy source came much later in the 18$^{\text{th}}$ century. In 1769, Watt filed a patent for his steam engine, which transforms the steam produced by heating water from coal combustion into mechanical energy. It was only at the end of this century that coal became the main source of energy, exceeding the use of wood [4]. This period corresponds to the *first industrial revolution* that transformed society through the mechanization of work. The rise of coal, combined with the steam engine, transformed society, placing industry at the heart of the economic structure of society and gradually replacing agriculture.

Almost a century later, the emergence of new sources of energy as oil or gas (known for a long time but not yet widely used) is leading to a new and profound transformation of society. The development of electricity during the 19$^{\text{th}}$ century led, at the end of the century, to a profound transformation of the energy used to produce electricity. Electrification and the growing demand for oil and gas following Daimler's invention of the internal combustion engine in 1886 led to the *second industrial revolution*. This revolution gradually sees coal being accompanied and replaced in some uses by the use of oil and gas as a primary energy source. These fuel changes were self-evident due to the competitive advantages in terms of energy density and price as well as the uses made of them. It was only in 1965 that the combined use of oil and gas surpassed that of coal.

Today, we are witnessing a new transformation of the energy sector with the beginning of a shift from a production system based on fossil fuels, from the first and second industrial revolutions, to a system that tends to be increasingly carbon-free. This is what Rifkin calls the *third industrial revolution* [5]. According to him, an industrial revolution is caused by new means of transport (electric vehicle), new means of communication (Internet) and new sources of energy. In this revolution, the new energy sources are renewable energies based on wind, sun and water, which have already been used for several centuries. However, they are now mainly used to produce electricity, which will be the source of energy at the heart of this revolution.

1.2 Overview of the electricity sector

Electricity is the physical phenomenon arising from the behavior of electrons and protons that is caused by the attraction of particles with opposite charges and the repulsion of particles with the same charge. Electricity has become a pillar of our current way of life. A large interconnected infrastructure allows it to be transported over long distances.

This section describe the means of generating electricity as well as the energy sources mainly used and then presents current figures and future trends.

It should be noted that in this section and more generally in this thesis, we will use the terms energy *production* and *consumption*. Strictly speaking, these terms have no physical meaning. In reality, we cannot "consume" or "produce" energy, by the principle of energy conservation we only transform energy [6]. In practice, "producing electricity" means transforming an energy source into electricity and injecting it into an electricity grid. Conversely, "consuming electricity" means the withdrawing of this electricity from the grid and often the transforming into another form (thermal, mechanical, etc.).

1.2.1 Principles of electricity production

The majority of the electrical energy produced in the world is generated by alternators, an electromechanical machine that converts mechanical energy into electrical energy in the form of alternating current. The alternator functioning is based on Faraday's law, which states that an electromotive force (voltage) occurs in an electrical circuit when it is stationary in a variable magnetic field or when the circuit is movable in a variable or permanent magnetic field. The alternator consists of a stator which is a stationary set of wire coil windings, inside which an electromagnet called rotor revolves. The rotation of the electromagnet inside the stator coils generates alternating current inside these coils [7].

Generally, a movement is generated from a primary source, and this mechanical energy is in turn transformed into electrical energy by the alternator. Mechanical energy is usually generated by turbines, driven by:

- Pressurized water vapour, produced by heating water with a thermal energy source (combustion of oil, coal, natural gas, nuclear fission of uranium 235 or concentration of solar energy).
- Water in hydroelectric power plants.

- ▷ Wind with wind turbines (grouped into wind farms).

Electricity can also be produced using solar panels in the case of photovoltaic plants. The photons transmitted by the sun excite the electrons of the materials when they reach the panels and this generates electricity.

1.2.2 Different primary sources

We have seen that electricity is generated from different primary sources. The different sources can be divided into two categories: *non-renewable* and *renewable* resources. Renewable energies are energy sources whose natural renewal is fast enough that they can be considered inexhaustible on the scale of human time. They come from cyclical or constant natural phenomena induced by the astronomical objects: the Sun essentially for the heat and light it generates, but also the attraction of the Moon (tides) and the heat generated by the Earth (geothermal energy). Their renewable nature depends on the speed at which the source is consumed, on the one hand, and on the speed at which it is renewed, on the other.

1.2.2.1 Non-renewable resources

Most of the electricity produced uses fossil fuels [8]. These are carbon-rich fuels, mainly hydrocarbons, produced by the methanization of living beings that have been dead and buried in the ground for several million years [9]. Usually, a thermal power plant is used to produce electricity. The ignition of a fuel causes water to expand. The heat source (nuclear fission, coal, incineration...) heats (directly or indirectly) water, which changes from liquid to vapour state. The steam thus produced is admitted into a turbine, which drives an alternator. At the turbine outlet, the steam is condensed in a condenser fed by a cold source (seawater, fresh river water, etc.) and returns to a liquid state. The resulting condensate is then returned to the water supply system for a new vaporization cycle [10].

The main fuels used to produce electricity are:

Coal — Coal is a fossil rock mined in coal mines as a fuel and formed from the partial degradation of plant organic matter.

Oil — Oil is a liquid rock of natural origin, a mineral oil composed of a multitude of organic compounds, mainly hydrocarbons, trapped in particular geological formations.

Natural gas — Natural gas, or fossil gas, is a gaseous mixture of hydrocarbons naturally present in some porous rocks.

Nuclear fuel — Nuclear fuel is a product that contains fissile materials. Nuclear fission is the phenomenon by which a heavy atomic nucleus (i.e. forming a large number of nucleons such as uranium, plutonium, etc.) is split into two or more lighter nuclides. This nuclear reaction is accompanied by the emission of neutrons and a very high energy release per fissioned atom. The main fuel is Uranium 235 but current research aims at using Uranium 238, which corresponds to 99.3% of the Uranium present on Earth, with breeder reactor or to use Thorium as a fuel.

1.2.2.2 Renewable resources

Renewable resources originate from cyclical or constant natural phenomena whose natural renewal is fast enough that they can be considered inexhaustible on the scale of human time. The most used sources are the following:

Hydropower — The kinetic energy of the water flow, natural or generated by the difference in level, is transformed into mechanical energy by a hydraulic turbine, then into electrical energy by an alternator.

Wind — A wind turbine is a device that transforms the kinetic energy of the wind into mechanical energy, which is then most often transformed into electrical energy. We will discuss this in more detail in section 2.1.

Solar photovoltaic — Photovoltaic energy is an electrical energy produced from solar radiation through photovoltaic panels or solar power plants. The photovoltaic cell is the basic electronic component of the system. It uses the photoelectric effect to convert electromagnetic waves (radiation) emitted by the Sun into electricity.

Biomass — Biomass is plant or animal material used for energy production, heat production, or in various industrial processes as raw material for a range of products. It can be purposely grown energy crops (e.g., miscanthus, switchgrass), wood or forest residues, waste from food crops (wheat straw, bagasse), horticulture (yard waste), food processing (corn cobs), animal farming (manure, rich in nitrogen and phosphorus), or human waste from sewage plants.

1.2.3 Main electricity key figures

Electricity mix is the distribution of the different primary energy sources that make up the total electricity production.

In this section we will present and comment on the current electricity mix key figures. We will then look at future trends and several scenarios in terms of electricity production.

1.2.3.1 Current electricity mix is mainly based on fossil fuels

Electricity in 2018 accounts for *19%* of the world final consumption of energy [11]. The distribution of electricity generation is shown in figure 1.1. It is important to note that more than *74%* of the electricity mix comes from non-renewable sources including *38%* of coal. Hydroelectricity has been a major source of renewable electricity for many years, accounting for *19%* of the energy mix. Non-renewable energies are also dominant for primary energies consumed in 2017 with a rate of *89.6%*. Oil represents *34%* of the energy consumed, although it can be used for electricity it is mainly used in transport and industry [12].

The high proportion of fossil energies in the energy mix poses two main problems:

- Ecological, the level of greenhouse gases released has led and continues to lead to global climate change

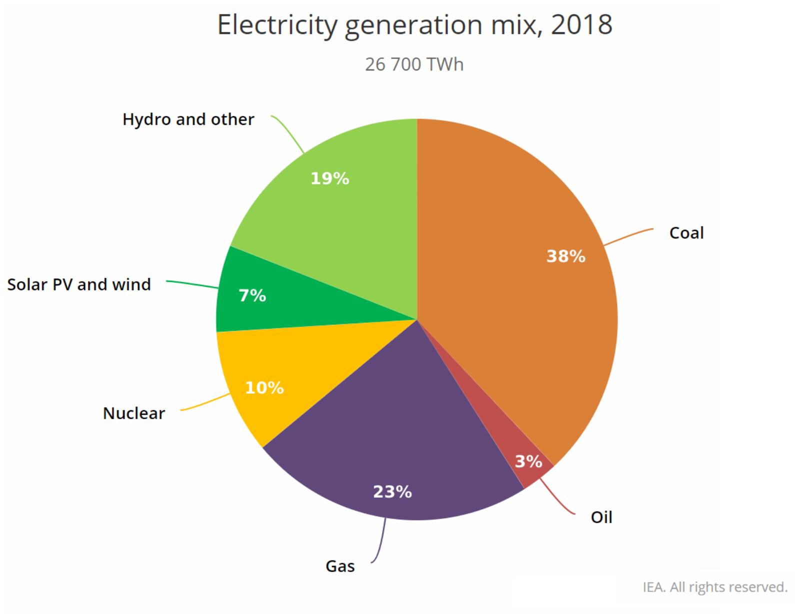

Figure 1.1 — World electricity generation mix in 2018 [8].

- Economic, limited resources in a world of continuous growth is mathematically impossible.

1.2.3.2 An energy future focused on renewable energy

The future is uncertain with regard to energy. Many factors enter the equation: political/economic situation, collective awareness, short-term consequences of climate change, new production technologies, etc. The limitation of fossil energy use seems inevitable, but the transition will be more or less abrupt. On the one hand, the estimated reserves stocks are quite high (see Table 1.1) and do not encourage rapid change. On the other hand, the extreme increase in emissions of CO_2 (See figure 1.2) and other greenhouse gas since the industrial era and its consequences on climate, alarm most of the scientific community [13].

Table 1.1 — Estimated number of years of production at this rate by source [12], [14].

Source	**Coal**	**Oil**	**Natural gas**	**Nuclear fuel (U235)**
Years of production	132	50	51	88

The International Energy Agency (IEA) in its World Energy Outlook 2018 [11] provides a way of exploring different possible futures, the levers that could bring them about, and the interactions that arise across a complex energy system. The organization proposes to compare three possible scenarios:

- *Current Policies Scenario*: a scenario in which there is no change in policies from today. This can lead to increasing strains on almost all aspects of energy security and a major additional rise in energy-related CO_2 emissions.

- *New Policies Scenario*: a scenario in which the policies follow the targets announced by

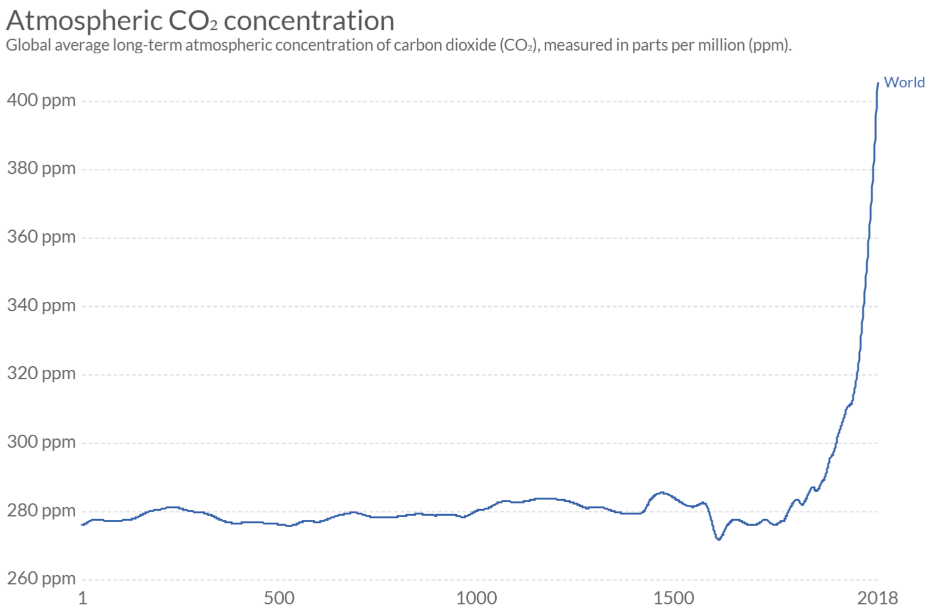

Figure 1.2 — Atmospheric CO_2 concentration estimated from gas microbubbles trapped in ice cores [15].

governments. Although the picture brightens, there is no reduction in global energy-related CO_2 emissions.

- *Sustainable Development Scenario*: a scenario in which accelerated clean energy transitions put the world on track to meet goals related to climate change, universal access and clean air.

Figure 1.3 shows the comparison between New Policies Scenario (a) and Sustainable Development Scenario (b) for electricity generation by sector. It should be noted that the Sustainable Development Scenario essentially aims at a significant decrease in coal, a significant increase in solar and wind energy and a decrease in total production (from about 40000 TWh to 36000 TWh in 2040).

There are an infinite number of other possible scenarios. In France, we can cite the *negawatt* scenario, hoping to achieve 100% renewable and carbon neutral energy by 2050 based on three elements [16]:

- Gradually making a transition to renewable energy
- Reducing losses by improving the energy efficiency of the conversion
- Changing behavior to consume less energy. For example, for a piece of equipment, it is possible to reduce its use, better size it or share it with other users.

Figure 1.4 shows the projections of the *negawatt* scenario for the primary energies consumed. The scenario is characterized by the sharp decrease in consumption while current policies aim for endless growth. This scenario seems difficult to achieve without a drastic change in the current economic model.

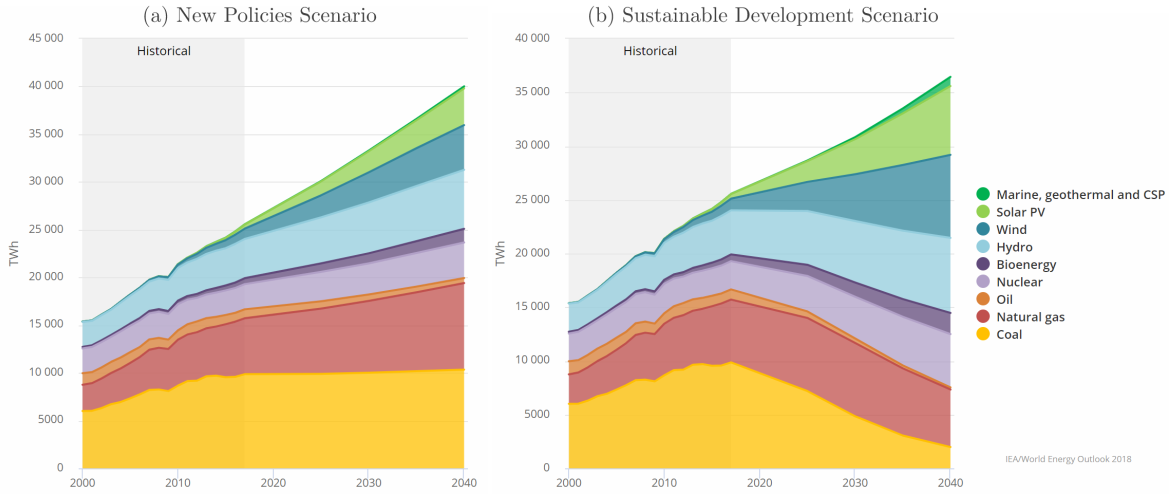

Figure 1.3 — IEA World electricity generation scenarios.

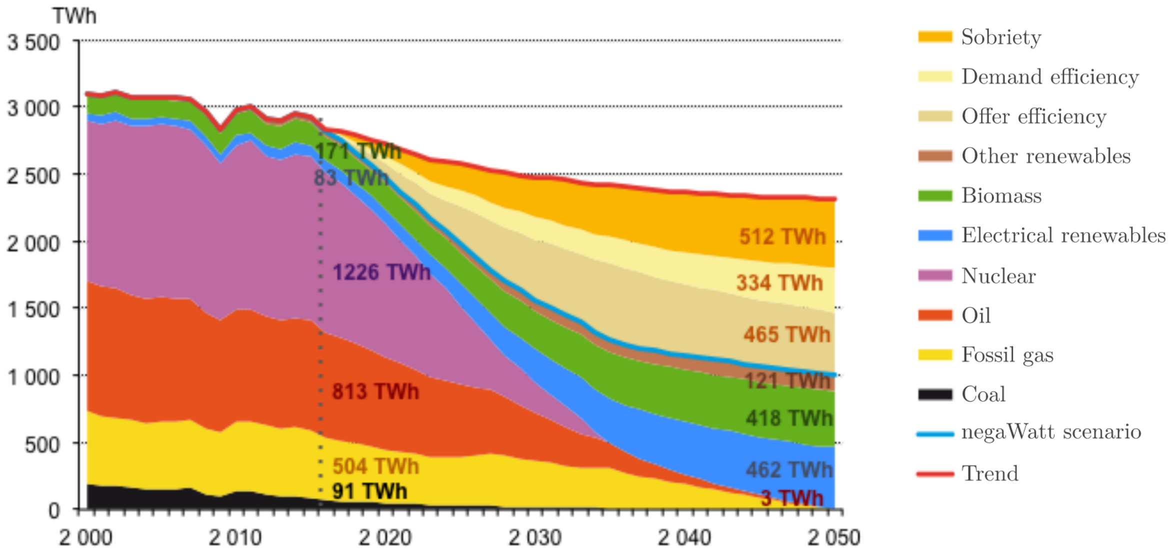

Figure 1.4 — French primary energy balance of the negaWatt scenario [16].

We have seen that renewable energies will be at the heart of the energy future. The next section will present these energies and the various problems they raise in more detail.

1.3 Renewable energies

The global energy market is undergoing a major transformation with the massive arrival of renewable energies. In this section we will see the current state of the sector and the related problems.

1.3.1 A growing sector

As seen earlier, the transition from fossil fuel use to renewable energy is one of the ways to reduce the impact of climate change and to have sustainable energy.

For electricity production, the main renewable resources are hydropower, wind turbines, biomass and solar panels. Table 1.2 shows global electricity production by sector in 2000, 2018 and the change between 2017-2018. In 2000, hydroelectricity was almost the only renewable source of electricity, today it is still the vast majority with 16%. However, its development is limited (-1% since 2000) because many sites are already being exploited and the impact on the environment is considerable (dams modify the entire landscape and watercourse, retain alluvium that enriches the soil, etc.).

The significant increase in solar panels and wind power (31.2% and 12.2% respectively) is due to the ease of implementation (compared to a dam) and unlimited resources. Besides, continuous improvements in solar and wind technologies keep decreasing costs and improving competitivity.

Table 1.2 — World electricity production by renewable source [12].

	2000	**2018**	**Variation 2017-2018**
Hydro	17%	16%	+3.1%
Wind	0%	5%	+12.2%
Biomass	1%	3%	+7.4%
Solar PV	0%	2%	+31.2%

The renewable energy rate in total electricity generation by country is presented in figure 1.5 (dark red if 0% and dark green if 100%). Overall, we observe a majority of "red" countries that have electricity production oriented towards non-renewable energies. Although several countries with high rates of renewable energy are observed, the figures are to be put into perspective. For example, some African countries are green because they currently have very low per capita consumption (for example, in 2016, a French person consumed on average 51 times more electricity than a Nigerian person [17]). Other countries such as Canada benefit from high hydroelectricity and a lower demand due to low population density.

Global investments in renewable energy by sector are shown in figure 1.6. Apart from large-scale hydropower, which is not in the graph, the majority of investments are currently in wind and solar power. Investments in renewable energy have increased significantly since 2004. The decreases observed are partly due to lower equipment prices, particularly in photovoltaics. As a result, installations are becoming cheaper and investments are decreasing. The second explanation is the irregularity of investment worldwide. For example, in 2015, China invested $119.1 billion in renewable energy and then this figure fell by 26% the following year. The Chinese government is now focusing on grid investment and electricity market reform to ensure that installed renewable energies can generate their full potential [19].

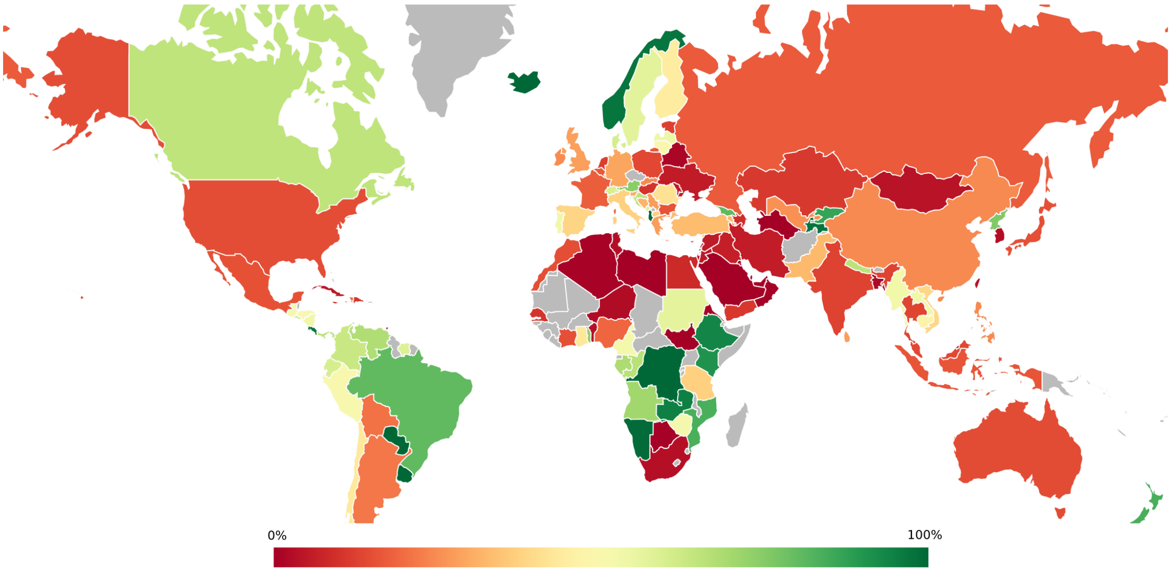

Figure 1.5 — Percentage of renewable energy in electricity generation by country in 2018 [18].

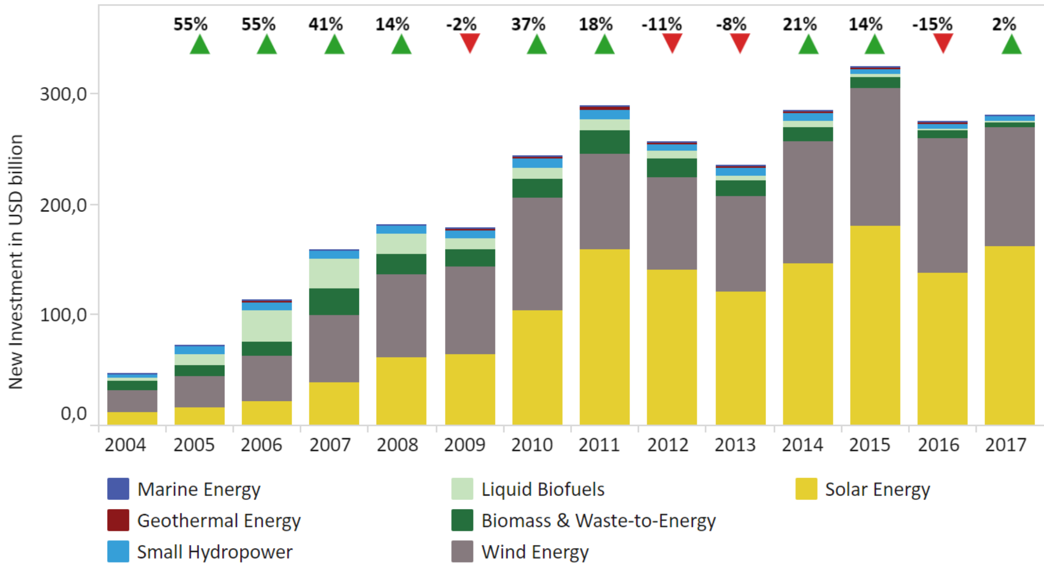

Figure 1.6 — Global trends in renewable energy investment [20].

1.3.2 Main issues of concern

Renewable energies are on the rise and are reaching ever more attractive prices. However, they have some issues that prevent a massive transition.

1.3.2.1 Intermittency

Renewable energies are based on natural phenomena that are not permanent and often irregular, this is called *intermittency*. This is particularly the case for wind power, photovoltaics and run-of-the-river hydroelectricity because they depend on weather conditions and the day/night cycle. Dams are not considered intermittent because they are controllable. For the wind, we can observe entire areas without wind for several days (for example, the Siberian anticyclone can extend over all of Europe for several days).

Moreover, production is not always correlated to demand. For example, solar energy produces nothing at night and less in winter when there is a higher consumption due to lighting and heating.

1.3.2.2 Variability

Variability is the fact that production can vary greatly over short periods of time. Wind energy is variable over short time scales with gusting wind producing peaks and troughs in power output that can cause voltage problems, because of the unevenness of the power being put onto the grid. To compensate for production deficits, additional production is used. However, not all means of production can satisfy this variability, for example, nuclear reactors are unable to start so quickly.

Therefore, in the case of renewable energies depending on weather conditions, it is almost impossible to produce energy in a continuous and controlled way, at least with the current technologies. A way to solve this problem is to store the excess energy and consume it when necessary, but electricity storage technology is still expensive. Another way is to rely on a large scale network, allowing to mitigate the variability.

1.3.2.3 A still limited energy storage

In France in 2018, because of their intermittent nature, a wind turbine produced on average 21.1% of its rated power while for a solar panel this rate is only 14% [21]. Moreover, these means of production do not allow production to be adapted to demand. In most cases, the need to instantly meet electricity demand requires coupling a wind farm with rapidly scalable electricity sources such as fossil fuel (coal or gas-fired power plants). An alternative to backup power plants, at least to compensate for short-term variations in wind generation, is to store energy in surplus periods, which is returned during low periods.

Electricity is difficult to store in sufficient quantities and at affordable costs to meet our energy needs. Direct solutions require "resistance-free" conductors called superconductors in which the electricity stored theoretically circulates without loss. These materials, which are currently available at very low temperatures of a few degrees Kelvin, are reserved for specific applications and small quantities. Indirect solutions provide only partial, expensive and often local solutions.

Here are the main current solutions or technology under development [22]:

Pumped-Storage Power Plant — The energy is stored as potential energy between two basins of different altitudes, the efficiency is estimated at between 75% and 80% (the energy efficiency of a cycle is the ratio between the amount of energy recovered and the amount of energy initially sought to be stored). This solution is already used but such an installation is physically difficult to install because it requires the construction of two dams (e.g., France has only seven stations, with a combined capacity of 7 GW [23]).

Battery Energy Storage Systems (BESS) — In batteries, storage is carried out in electrochemical form. Different battery technologies coexist depending on the oxido-reducing pairs they use. Currently, battery storage technology efficiency is typically around 80% to more than 90% for newer lithium ion devices. Moreover, it allows several charge and discharge cycles per day with an almost instantaneous response time. A major negative point is that batteries use relatively rare minerals such as lithium in their manufacture. The large-scale development of these batteries is likely to generate significant geopolitical and environmental pressure for the extraction of these minerals. A rise in the prices of these metals and, at the same time, in the price of batteries is to be feared.

Compressed air energy storage (CAES) — In this type of storage, air is compressed in a tank using an electric compressor. Currently the efficiency reaches 55% but this technology is not yet mature enough.

Power-to-gas — Power-to-gas technology allows storage through the transformation of water into hydrogen and oxygen by electrolysis of water. Hydrogen can then be injected into natural gas networks or converted back into electricity via a fuel cell. The conversion to electricity currently achieves low overall efficiency rates (between 30 and 40%) while the conversion to methane (which can be added directly to the existing gas network) currently achieves between 54 and 65% efficiency.

Flywheels — With inertial flywheels, the energy is in the form of kinetic energy through a rotary mechanical movement of a mass around a fixed axis. The mass is coupled via an electric motor that accelerates the rotating mass storing energy. Conversely, when the mass drives the motor, it delivers a torque that transforms the motor into an electric generator. It is an efficient storage medium for short periods of time (a few minutes maximum) with an efficiency of 85%.

Superconducting Magnetic Energy Storage (SMES) — SMES allows energy to be stored in the form of a magnetic field created by the circulation of a direct current in a cooled superconducting ring under its critical temperature. This technology is not yet mature enough for large-scale deployment. It allows quick adjustments by restoring electricity with excellent efficiency (95 %) for short periods of time (< 5 mn).

These are the main technologies currently in use or under development. But there are many other research tracks using existing technologies such as a train going up a hill [24] or a crane stacking blocks of used concrete [25].

Electricity storage will be a major issue in the energy transition. It is a highly studied field with many promising leads. At the moment, no technology is a perfect solution that would be mature enough to be used on a large scale, cost-effective enough and renewable enough for a long-term vision. The fate of highly variable energies is linked to that of electrical storage, as it allows them to be better integrated into the grid.

1.4 Electricity markets

This section introduces the electricity markets using the French market as an example and then describes the different players and the overall functioning of the market. Then we will focus on the specificities of renewable energies on the market and explain the mechanisms to balance the electricity grid.

1.4.1 Market players

Historically, because of the significant investments required, the energy sector was mainly controlled by a large territorial monopoly. In the European Union, the sector began to be liberalized in the 2000s following European directives with several main players:

- *Producer*, the one who produces energy.
- *Transmission System Operator (TSO)*, it is the operator of the high-voltage and very high-voltage electricity grid (respectively 63kV or 90kV and 225kV or 400kV), used for inter-regional and international electricity transmission. In one country, there may be only one TSO (e.g., RTE in France) or several (e.g., Germany is divided into four zones each managed by a different operator).
- *Distribution System Operator (DSO)*, it is the operator of the low-voltage (< 63kV) electricity grid to which most end customers are physically connected, at low or medium voltage (e.g., 95% by ENEDIS in France).
- *Supplier*, the one who markets electrical energy to his customers (individuals or companies) without necessarily producing or distributing it. (e.g., 80% of the residential sector by EDF in France in 2018, the others being considered as "alternative" suppliers [26]).

Depending on the country or the region, only one actor may be in charge of all or part of these tasks. In the European Union, countries no longer have the right to have a national monopoly occupying all roles. In France, this was the case for EDF in the past which is still the leading producer and supplier of electricity and which holds a majority stake in RTE and ENEDIS.

1.4.2 Functioning of electricity markets

Since the late 1990s, the European Union has been gradually organizing the liberalization of the internal electricity market. This has led to the creation of several spot markets. For example, EPEX SPOT operates the electricity market in France, Germany, Great Britain, the Netherlands, Belgium, Austria, Switzerland and Luxembourg.

In the real-time management of an electricity grid, production and consumption must always be balanced to maintain frequency and voltage, otherwise a generalized electrical incident could occur. In countries where the electricity sector is open, there are "wholesale markets" where competing electricity producers sell their electricity production to suppliers. The market price is determined by matching producer offers and consumer demands. It is an equilibrium price resulting from the relationship between supply and demand. An example of market clearing is shown in figure 1.7. The *market clearing price* corresponds to the intersection between the power supply curve and the power demand curve. This price may also be negative, this occurs when a high and inflexible power generation appears simultaneously with low electricity demand.

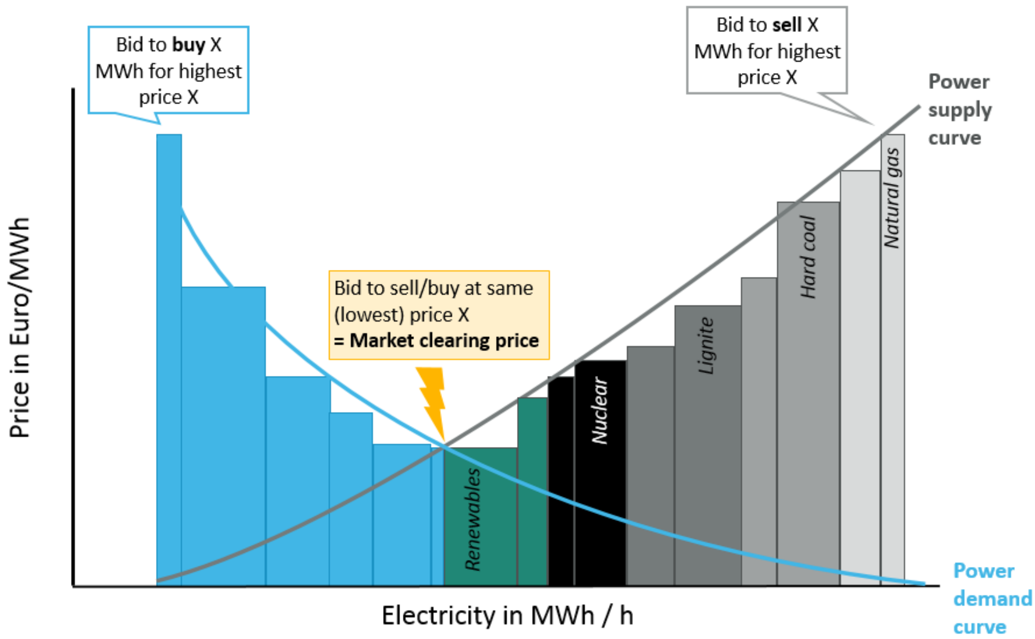

Figure 1.7 — Example of market clearing [27].

More specifically, on the EPEX SPOT market, there are two components: *Day-Ahead* and *Intraday*.

In the *Day-Ahead* market, producers or their traders bid a certain amount of energy at a certain price (which depends mainly on the underlying operational cost of the energy source but also on trading strategies) for the next day in blocks of 30 minutes in France and 15 minutes in Germany. On their side, buyers announce the demand to be met and the desired purchase prices. Based on the requests initiated by the parties, a single auction procedure taking place at a given hour (e.g., at 12 p.m. in France) matches the bids and leads to a single equilibrium price by block. The next day,

- if the producer produces more than expected, he will be able to liquidate his surplus on the Intraday market at a price generally lower than the Day-Ahead price.
- if it produces less than expected, it will have to meet its supply commitments by purchasing additional electricity at a higher price than the Day-Ahead price.

On the *Intraday* market, there is no equilibrium price, the price varies continuously. In addition, the deadlines and temporal resolutions are shorter. In France, the step is 30 minutes and transactions take place until two hours before delivery (these rules are not set by the EU and depend on each country). After this period, if the producer has not been able to honor his commitment, the TSO will supply electricity from its reserve set aside for this purpose. This action is costly, in the form of penalties, for players who are not in balance.

1.4.3 Specificity of renewable energies on the market

Historically, renewable energies were first heavily subsidized to develop the sectors because renewable energy sources are not yet mature enough. The level of subsidy has started to decrease in order to move towards a non-subsidized market system.

Firstly, the *Feed-in tariff* mechanism forced network operators to buy electricity from renewable sources at a fixed price (higher than the market value of electricity). In France, the *Energy Transition for Green Growth Act* [28] came into force in 2018. This law requires new renewable energy installations to sell their electricity on the spot market (or the old installations once the pre-established agreements expire).

In order to keep prices competitive, the *Feed-in premium* scheme has been introduced in which a subsidy is always granted to producers of renewable electricity. This additional remuneration is calculated at the end of each month as the difference between a reference price (depending on the cost of the installation) and the market value of the sector at national level (e.g., French wind energy sector). This market value corresponds to the remuneration in euros per MWh that the national sector would have received for the sale of its hourly production on the Day-Ahead EPEX SPOT auction.

Since there may be a difference between the national market value and the producer's sales revenue benefiting from the additional remuneration (weather decorrelation, periods of unavailability, trading performance, etc.), this mechanism has introduced an incentive for producers to better sell their production on the markets. Since the bonus depends on the entire sector, a producer who can sell better than another will still receive the same additional remuneration.

In order to avoid high volatility of Intraday prices and to prevent penalty, an intermittent energy producer must predict its production in advance. The trading activity is very complex and requires a dedicated team, operational 24/7 for the Intraday market. Consequently this activity is often outsourced to another market player: *aggregators*. These are brokers who manage the energy trading of a portfolio of power plants including different types of production processes (photovoltaic, wind, hydro, but also gas, coal, etc.) selected in order to obtain a more profitable total production and better guaranteed against weather variability.

1.4.4 Grid balancing

The Transmission System Operator (TSO) ensures at all times the balance between electricity production and consumption and resolves congestion on the transmission network.

To this end, it establishes and activates balancing reserves provided by the balancing actors: producers, consumers, other actors likely to inject energy into or to extract energy from the network. In addition, to minimize balancing needs, balance responsible entities (suppliers, producers, etc.) are encouraged to balance their injections and withdrawals on the network in advance.

In France, RTE has three types of reserves to reduce the imbalances between electricity production and consumption [29]:

- ▷ *Primary reserve*: it is activated in a decentralized way at the level of each production group and takes place in 15 to 30 seconds. The primary reserve must be able to cope with the simultaneous loss of the two largest generating units, i.e. 3000 MW at European level. The French system contributes about 540 MW.
- ▷ *Secondary reserve*: it is automatically activated by RTE, in about 400 seconds. The secondary reserve set up in France is between 500 MW and 1180 MW. All producers operating generating units of more than 120 MW in France are required to participate.
- ▷ *Tertiary reserve*: the activation of this reserve is manual, carried out by an RTE dispatcher, through the *adjustment mechanism*. It is used to supplement the secondary reserve if it is exhausted or insufficient to deal with an imbalance, but also to replace the primary and secondary reserves or to anticipate a future imbalance. This reserve takes longer to mobilize than the others but once started it is available for a longer period of time.

When RTE activates an upward adjustment offer, i.e. an offer that makes it possible to resolve a lack of energy (production less than consumption), or conversely a downward adjustment offer, RTE pays the actor who proposed this flexibility. The costs and income related to the activation of adjustment offers are managed by RTE within the "adjustments" account, a management account that is intended to be balanced: the costs of imbalances are allocated to the actors who are responsible for them during the process of calculating and settling the gaps.

The price of the differences is directly linked to the price of the adjustment offers requested by RTE to maintain the balance of the French electricity system. The principles for calculating the *Imbalance Settlement Price* (ISP) of differences make it possible to send balance responsible entities a financial incentive on their imbalances and reflect the operational cost of balancing.

Thus, an electricity producer who has overestimated or underestimated his production will have a financial penalty. The amount of this penalty will depend on the context of the entire power grid. For example, if the electricity grid is "tight" (i.e. supply barely covers demand), an overestimation of production can be expensive.

For controllable electricity production (e.g., by fossil fuels), an incorrect estimate may be due to technical problems (turbine failure or breakdown). For intermittent renewable energies, production is by nature difficult to predict. Since aggregators (or more rarely directly power producers) are financially responsible for any deviation from their commitments, improving power forecasting accuracy enables to reduce the penalties they incur [30]. The functioning of the market and the strong increase in renewable energies are thus pushing the players to constantly improve production forecasts.

The next part will focus on wind power forecasting by introducing all the necessary details on the wind and how wind turbines work. Then the different production forecasting models will be presented.

2 Wind Power Forecasting

As seen in chapter 1, the energy sector is increasingly turning to renewable energy. The variability of these energy sources requires producers to accurately predict their production in order to make it profitable in the electricity market and avoid the penalties of the Imbalance Settlement Price.

This chapter details the case of the wind power forecasting while focusing first on the specificities of this kind of forecast. Then, the different methods currently used are compared with a focus on spatial dependencies.

2.1 Wind energy 24
2.1.1 The wind as energy producer 24
2.1.1.1 Wind definition 24
2.1.1.2 Statistical distribution 26
2.1.1.3 Physical phenomena 27
2.1.2 Wind turbines 28
2.1.2.1 General principles 29
2.1.2.2 Wind turbine types 29
2.1.2.3 Theoretical studies 31
2.1.2.4 Power curve 32
2.1.2.5 The production in practice 33
2.1.2.6 Curtailments 34
2.1.2.7 Effects to be considered 35
2.2 Wind power forecasting 37
2.2.1 Different uses of forecasts 37
2.2.2 Forecast formats 38
2.2.3 Offshore specificity 39
2.3 Physical models 39
2.3.1 Numerical Weather Prediction (NWP) 40

2.3.1.1 The weather: a chaotic system . . . 40
2.3.1.2 Description of the NWP models . . . 40
2.3.1.3 The particularities of the wind forecasting . . . 41
2.3.2 Downscaling method . . . 42
2.4 Statistical models . . . 44
2.4.1 Time series models . . . 44
2.4.2 Machine Learning methods . . . 45
2.4.2.1 Linear methods . . . 46
2.4.2.2 Support Vector Machine (SVM) . . . 46
2.4.2.3 Artificial Neural Networks . . . 46
2.4.2.4 Ensemble methods . . . 47
2.5 The spatial forecasting methods . . . 48
2.5.1 Short scale dependencies . . . 48
2.5.2 Large scale dependencies . . . 49
2.6 Analysis of the current models . . . 50
2.6.1 Analysis criteria . . . 50
2.6.2 Evaluation of the criteria . . . 51
2.6.3 Synthesis . . . 51

2.1 Wind energy

Wind energy is becoming one of the pillars of the energy transition and the forecasting of this intermittent and variable energy will be an important condition for its success. Before talking about wind power production forecasts, we first look at the energy that drives wind turbines: the wind; and wind turbines are then presented in detail.

2.1.1 The wind as energy producer

Wind is a force that everyone encounters daily without always understanding it. This section describes precisely what wind is and shows its temporal and spatial diversity by describing its statistical distribution. Finally, some interesting features of winds that are useful for wind turbines are explained.

2.1.1.1 Wind definition

The wind is the movement within an atmosphere of a mass of gas located on the surface of a planet. The main causes of wind are:

- ▷ the uneven warming on the surface of the planet from stellar radiation, which causes a pressure difference; and

- ▷ the rotation of the Earth (the effect of the Coriolis force) which diverts the air flow.

The wind belts girdling the planet are organized into three cells in each hemisphere: the *Hadley cell*, the *Ferrel cell*, and the *Polar cell*. The vast bulk of the atmospheric motion occurs in the Hadley cell. The high pressure systems acting on the Earth's surface are balanced by the low pressure systems elsewhere. As a result, there is a balance of forces acting on the Earth's surface. These cells are shown in figure 2.1.

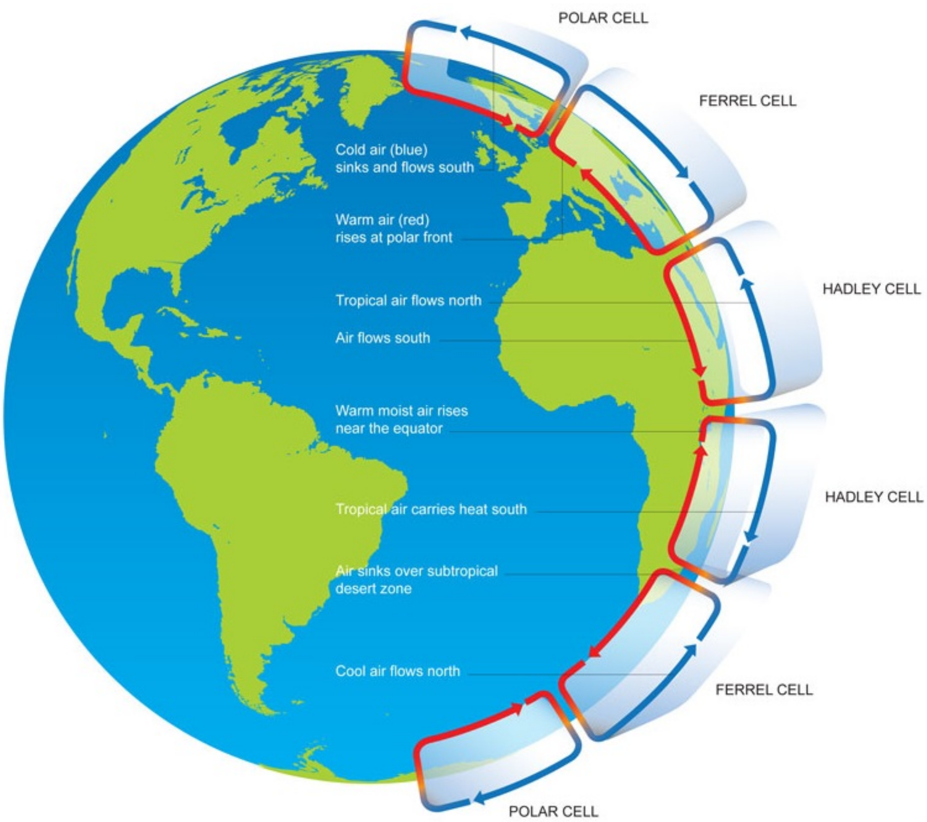

Figure 2.1 — Representation of the three main convective cells: Polar, Ferrel and Hadley cells [31].

The ambient air is a fluid, so its movements can be described by equations 2.1 and 2.2 used in fluid mechanics. They are called the Navier-Stockes equations, where ρ is the density, μ is the viscosity, v is the flow velocity, p is the pressure and t is time. These are non-linear partial differential equations for which no general solution is known, so the problem is solved by approximation. The mathematical existence of solutions of Navier-Stokes equations is not demonstrated, and this is one of the problems of the Clay Mathematical Institute's Millennium Prize [32].

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho v) = 0 \tag{2.1}$$

$$\rho \left(\frac{\partial v}{\partial t} + v \cdot \nabla v \right) = -\nabla p + \mu \nabla^2 v \qquad (2.2)$$

2.1.1.2 Statistical distribution

By measuring wind speed over a year, we can see that, in most parts of the world, extreme winds are very rare, while fresh or moderate winds are quite frequent. Usually, wind variations are described using a Weibull distribution, a mathematical expression which provides a good approximation to many measured wind speed distributions [33].

Figure 2.2 shows the wind observations at a given site for a year and the Weibull distribution that best matches these data. Such a distribution is described by two parameters: the "scale", a parameter which is closely related to the mean wind speed, and the "shape" parameter, which is a measurement of the width of the distribution. This distribution corresponds of course to an average over large periods, at shorter scales the distributions can be very different.

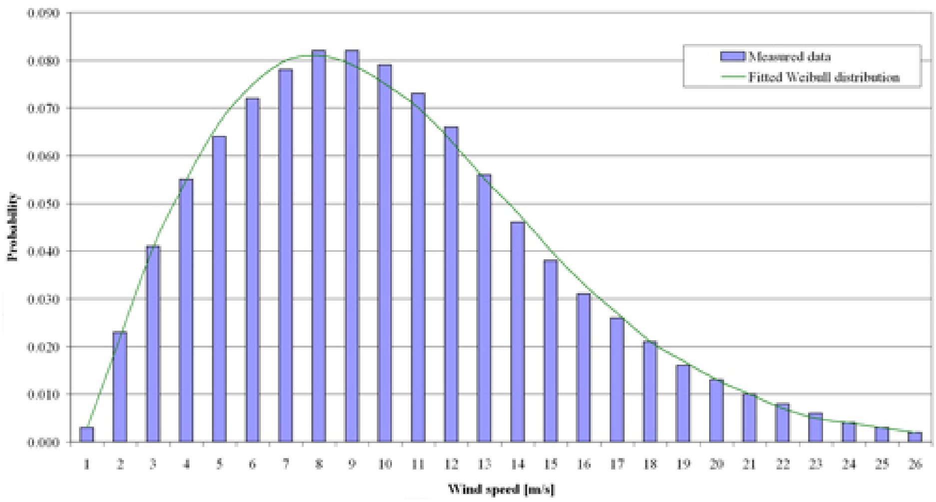

Figure 2.2 — Example of wind speed distribution [34].

Although the wind distribution at a given site can be fairly well represented by a well parameterized Weibull distribution, wind speed is far from being evenly distributed over the Earth. Figure 2.3 represents the world wind speed potential. Some regions have very low winds (e.g., Amazonian forest), while others have high winds (e.g., Greenland). Overall, there is more wind near the coasts.

The wind distribution is also different according to the seasons. Overall, the wind speed is higher in winter than in summer. There is a much greater variation in temperature during the winter. While most summer days are roughly the same temperature, winter temperatures fluctuate dramatically, and this leads to a more rapid flow of air between atmosphere layers.

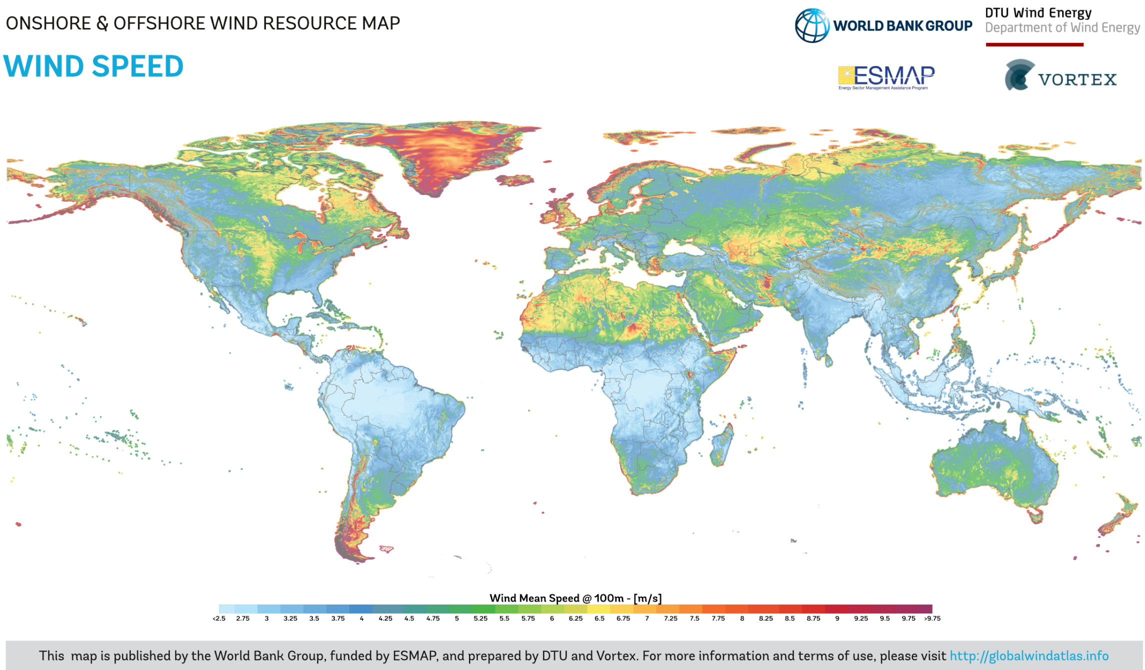

Figure 2.3 — World wind speed potential [35].

2.1.1.3 Physical phenomena

Wind can be characterized on a large scale (as seen in section 2.1.1 with large convective cells) but also on a smaller scale with more local phenomena depending on specific environmental factors. The main phenomena are the following:

Turbulence — In fluid dynamics, a turbulence (or turbulent flow) is a fluid motion characterized by chaotic changes in pressure and flow velocity. It is in contrast to a laminar flow, which occurs when a fluid flows in parallel layers, with no disruption between those layers.

Influence of roughness — The wind is not the same on the ground and at altitude. Indeed, when the wind blows, it is slowed on the ground and adopts a logarithmic profile. This profile can be more or less accentuated or even deformed depending on the ground. Roughness Z_0 is a measure of small-scale variations of amplitude in the height of a surface. The roughness of the landscape and, in particular, the "soft" roughness of the trees (that of forests, groves, savannas, compared to the rocks and buildings that do not move) has an impact on the winds and the turbulence. Figure 2.4 shows the influence of terrain on wind speed at higher altitudes. In a dense urban environment, the wind is strongly slowed down. In addition, as the wind encounters many obstacles, it becomes highly turbulent.

Tunnel effect — The tunnel effect, also known as the Venturi effect, is created at the level of the passes, between two mountains as between two large buildings; the wind is often stronger there. Air is compressed on the windward side of mountains or buildings, in order to keep a constant air flow, so wind speed increases considerably between obstacles. In addition, the wind generally keeps a constant direction.

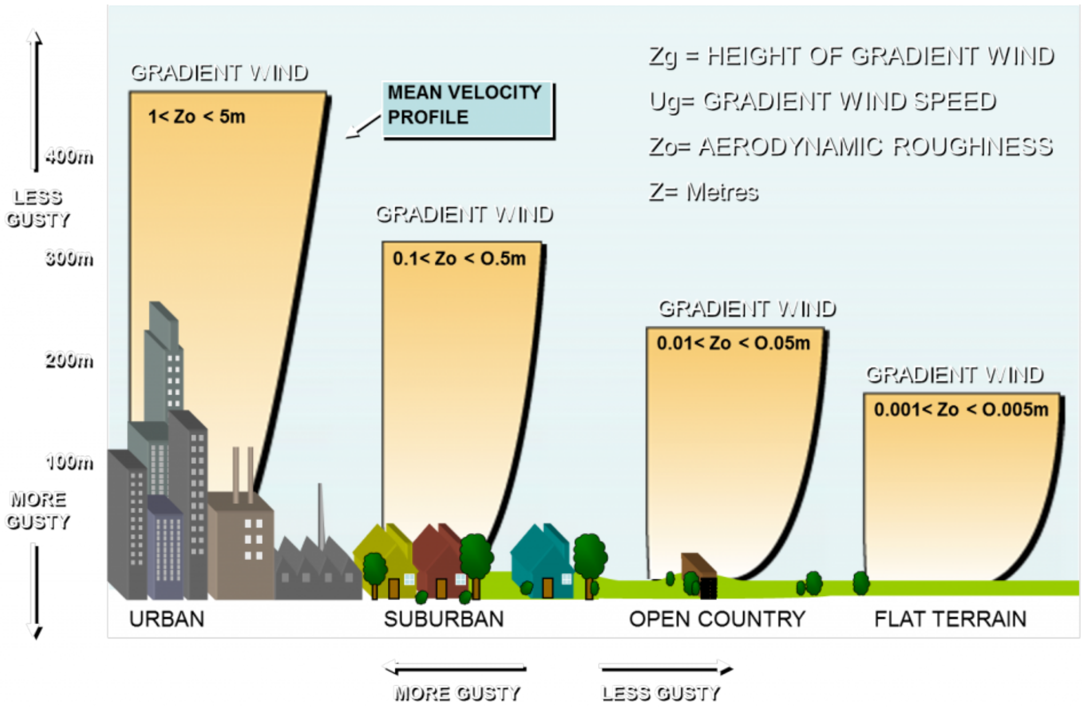

Figure 2.4 — Difference of wind profiles according to topography [36].

Hill effect — On hills, the wind speeds are higher than in the surrounding area. The wind becomes compressed on the windy side of the hill, and once the air reaches the ridge it can expand again as its soars down into the low pressure area on the lee side of the hill. This effect depends on the slope of the hillsides, turbulence can indeed be caused by too large an angle. Figure 2.5 shows the wind profile when the land includes a hill. The speed-up is a function of height above the surface. The height of maximum speed up (l) is related to the geometry of the hill (L).

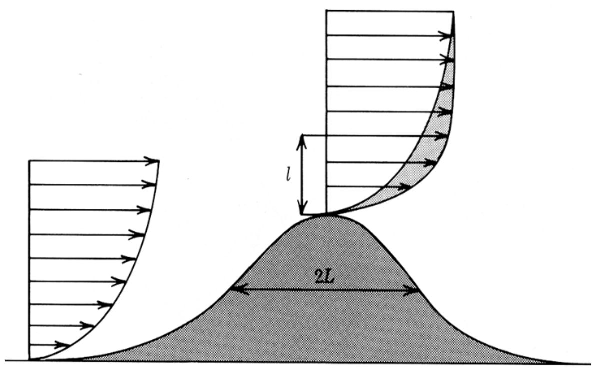

Figure 2.5 — The vertical profile of wind speed upwind and on top of a hill [37].

2.1.2 Wind turbines

From what we have seen above on the wind, we will now have an overview of wind turbines. This section details the functioning of wind turbines, the different types and different theoretical equations. Practical figures are then given and specific effects, such as curtailment, that have to be taken into account when dealing with wind turbines are then described.

2.1.2.1 General principles

A wind turbine is a device that transforms the kinetic energy of the wind into mechanical energy, which is then transformed into electrical energy using an alternator.

A wind turbine is composed of the following elements (see figure 2.7):

- *A mast or tower* — It allows the rotor to be placed at a height sufficient to allow its movement (necessary for horizontal axis wind turbines), or at a height where the wind blows stronger and more evenly than at ground level.
- *A nacelle or hub* — It is mounted at the top of the mast and houses the mechanical, pneumatic, some electrical and electronic components necessary for the machine to operate. The platform connecting it to the mast can rotate to orient the machine in the wind direction.
- *A rotor* — It is composed of the nose of the wind turbine receiving the blades, fixed on a shaft rotating in bearings installed in the nacelle. The rotor, fixed to the blades, is driven by wind energy. It is connected directly or indirectly (via a gearbox speed multiplier) to the mechanical system that uses the collected energy.
- *Several blades* — They are usually three and are driven by the wind. The shape of the blades changes the air flow. At the rear it accelerates the flow and slows it down at the front. The acceleration is accompanied by a decrease in pressure while deceleration results in an increase in pressure. This results in a force divided into a *lift force* perpendicular to the movement and a *drag force* in the direction of flow (See figure 2.6). The lift of a blade depends on the angle of attack. When this angle is too high, a *stall effect* occurs, the lift decreases significantly. The blades usually end with a winglet at the end of the blades to reduce induced drag.
- Other elements such as the transformer or foundations.

2.1.2.2 Wind turbine types

The majority of wind turbines producing electricity have three blades and a horizontal axis parallel to the wind direction. The more blades there are on a wind turbine, the higher will be the torque (the force that creates rotation) and the slower the rotational speed (because of the increased drag caused by wind flow resistance). But turbines used for generating electricity need to operate at high speeds, and actually do not need much torque. So, the fewer the number of blades, the better suited the system is for producing power. A three-bladed turbine represents the best combination of high rotational speed and minimum stress [40].

There are also other types of wind turbines (see figure 2.8):

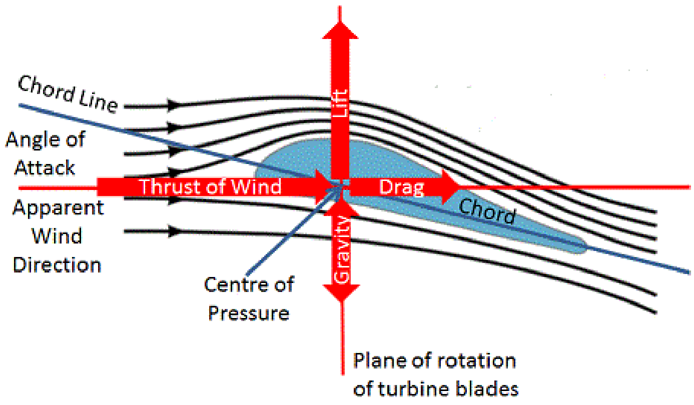

Figure 2.6 — Diagram of the forces acting on a wind turbine blade [38].

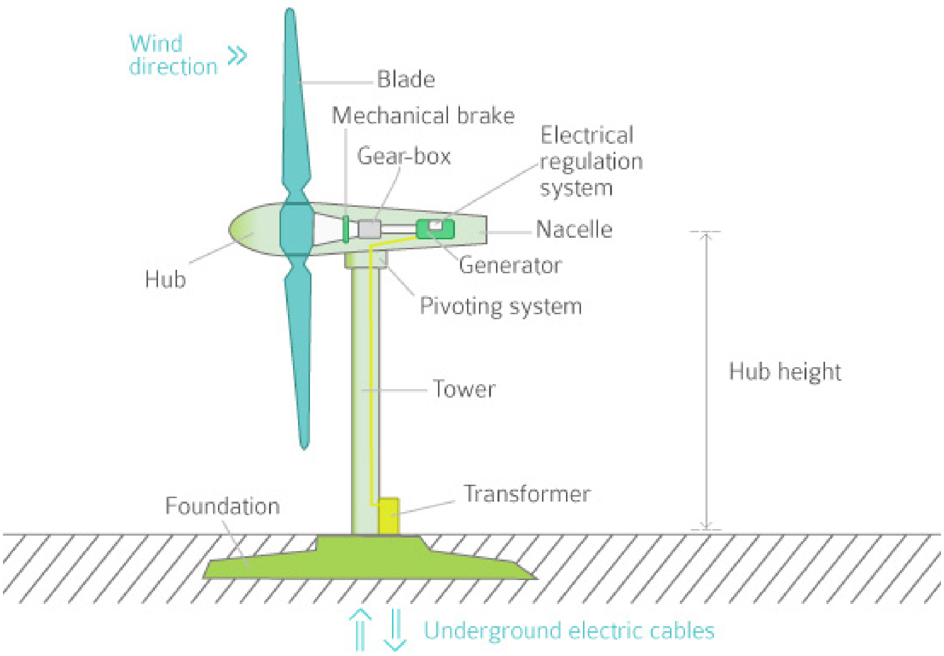

Figure 2.7 — Wind turbine schematic [39].

- ▹ Mills which are not intended to produce electricity(e.g., to grind cereals **(a)**, crush or operate a pump **(b)**). They often have more blades due to a higher torque requirement.
- ▹ Wind turbines with vertical axis which are especially used on a small scale in places where it is not possible to install a wind turbine or when the performance is not important. There are two main ones: *Darrieus* **(c)** which has for defects its low performance and its difficult start. Indeed, the weight of the rotor weighs on its base and generates friction. *Savonius* **(d)** which are also used in cases where cost or reliability are given more importance than performance.
- ▹ *Two-bladed wind turbines* **(e)** which have a noisier and more complex design because they must be equipped with a tilting rotor to avoid strong shocks. The main advantage is that the wind turbine is lighter and this makes it possible to make foldable models in case of

storms. In addition, it saves the price of a blade.

▷ *Offshore wind turbines* **(f)** which follow the same principles as the three-bladed wind turbines with only a different base. In the state of prototypes, there is also the *floating wind turbines* **(g)**. They are mounted on a floating structure that allows the turbine to produce electricity further from the coasts, where the water is much deeper and the winds are stronger and more stable.

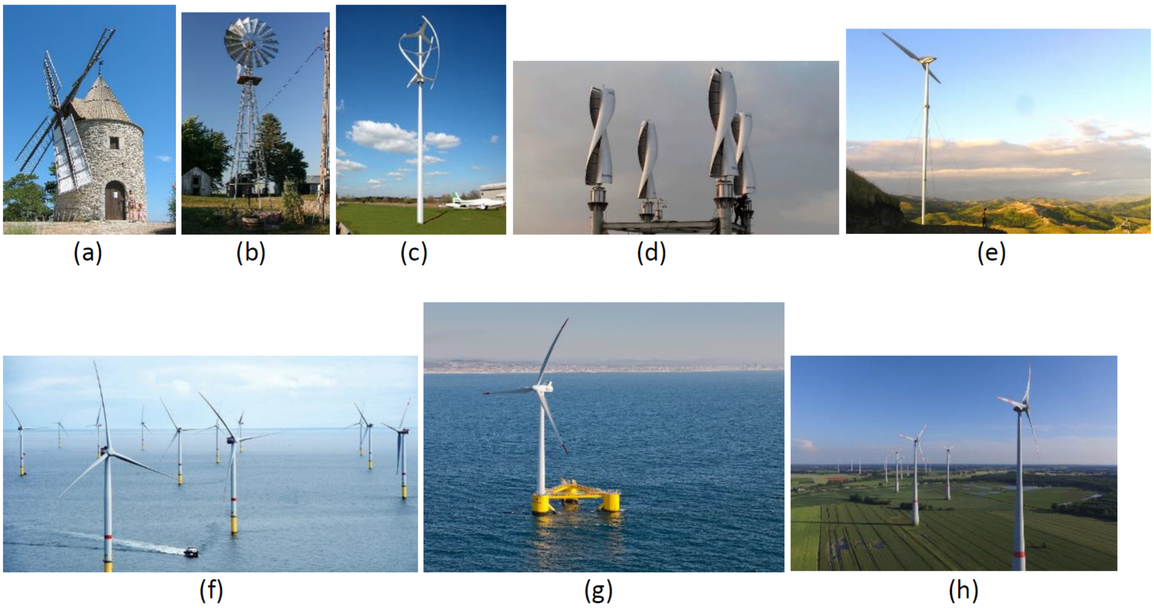

Figure 2.8 — (a) Windmill, (b) Wind water pump, (c) Darrieus, (d) Savonius, (e) Two-blade turbine, (f) Offshore, (g) Floating, (h) Standard [41]–[48].

2.1.2.3 Theoretical studies

The wind is moving air, and like any moving body, it can be associated with kinetic energy E_c (J) given by the following formula:

$$E_c = \frac{1}{2}mv^2 \tag{2.3}$$

where m is the air weight (kg) and v is the instant wind speed (m/s).

Air weight is represented by:

$$m = \rho V \tag{2.4}$$

where V is the volume (m^3) and ρ is the air density (kg/m^3).

By considering a device for recovering this surface energy S and assuming that the wind speed is identical at each point on this surface, the volume of air that passes through this surface in one second is equal to vS. Thus, the theoretically recoverable power $P_{theoretical}$ is equal to:

$$P_{theoretical} = \frac{1}{2}\rho v S v^2 = \frac{1}{2}\rho S v^3 \tag{2.5}$$

This power (in Watts) is a theoretical power and cannot be recovered entirely by a wind turbine (this would be like stopping the wind). In [49], Albert Betz demonstrated that the maximum recoverable power is:

$$P_{max} = \frac{16}{27} \cdot P_{theoretical} \tag{2.6}$$

The theoretical maximum yield of a wind turbine is therefore set at $\frac{16}{27}$, or about 59.3%. This value is called the *Betz limit*. This figure does not take into account energy losses caused by the conversion of mechanical wind energy into electrical energy. In practice, the real power is:

$$P_{real} = C_p \cdot P_{max} \tag{2.7}$$

where C_p is the power coefficient (i.e. the fraction of wind energy that the wind turbine is able to extract). It is the ratio of actual electric power produced by a wind turbine divided by the total wind power flowing into the turbine blades at specific wind speed.

As the energy supplied by the wind turbine is converted from one form to another, this coefficient is therefore affected by all the yields specific to the different transformations such as the blades, the alternator, the transformer, the rectifier, the batteries and the power line losses. The efficiency of each element varies with the operating speed linked to the blades rotation speed, which further reduces the overall efficiency of the device. Generally, C_p does not exceed 70% of the Betz limit.

2.1.2.4 Power curve

From section 2.1.2.3, the power P delivered by a wind turbine follows the equation:

$$P = \frac{1}{2}\rho S C_p v^3 \tag{2.8}$$

where ρ is the air density, S is the rotor surface (the area swept by the blades), C_p is the power coefficient and v is the wind speed.

This function only forms part of the full power curve (i.e. the graph representing the wind speed-production relationship). According to [50], a typical power curve for an operational wind turbine is sketched in figure 2.9 and is made up of 4 parts:

1. below *cut-in wind speed* (typically between 3 and 4 m/s) where the turbine does not operate.
2. between *cut-in* and *rated (or nominal) wind speed* where the power follows equation (2.8).
3. above *rated wind speed* where the power is limited to the turbine *rated power*.

4. above *cut-out wind speed* (usually between 20 and 25 m/s) where the turbine is shut down to prevent damage.

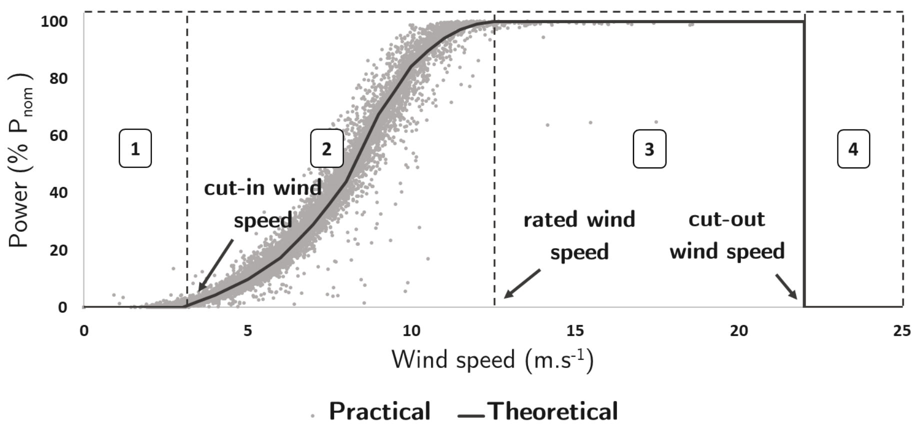

Figure 2.9 — A theoretical power curve compared with the observed production as a function of the wind speed measurement at the nacelle.

The power curve is ideally built from on-site measurements by fixing an anemometer on a mast located near the wind turbine (not directly on the wind turbine itself because it may cause turbulence that will affect the reliability of the measurements). As shown in figure 2.9, the power curve is actually made up of a multitude of points spaced on either side of the line. Indeed, there will always be fluctuations in the wind speed that will make it impossible to accurately measure the air flow through the wind turbine rotor. In practice, the average of the different measurements is taken for each wind speed to build the power curve.

This curve, provided by the turbine manufacturer [51], corresponds to a specific turbine model but there are several types of wind turbines and each has a different power curve. The wind farm developer therefore initially chooses the wind turbine model that meets all its constraints (financial, topographical and storm or cyclone risk) and optimizes energy production according to the wind profile at the wind farm location.

2.1.2.5 The production in practice

The *nominal power* P_{nom}, or "installed power", is a term that indicates the maximum power of a wind turbine. The nominal power of a wind turbine can vary from a few kW to several MW. Currently, the most powerful wind turbine is the *Haliade-X 12 MW* with a nominal power of 12MW [52]. Figure 2.10 shows the number of wind turbines as a function of P_{nom} in 2016 in Europe. We notice that half of them have a power less than or equal to 2MW.

The *Capacity factor* is the ratio between the actual amount of energy produced over a given period and the maximum theoretical production of a wind turbine operating at full power on a full-time basis. In theory, the capacity factor is between 0% and 100%, in practice it will be between 20% and 30%. Wind turbines are not primarily designed to optimize the capacity factor, but to generate as much electricity as possible at a certain wind speed.

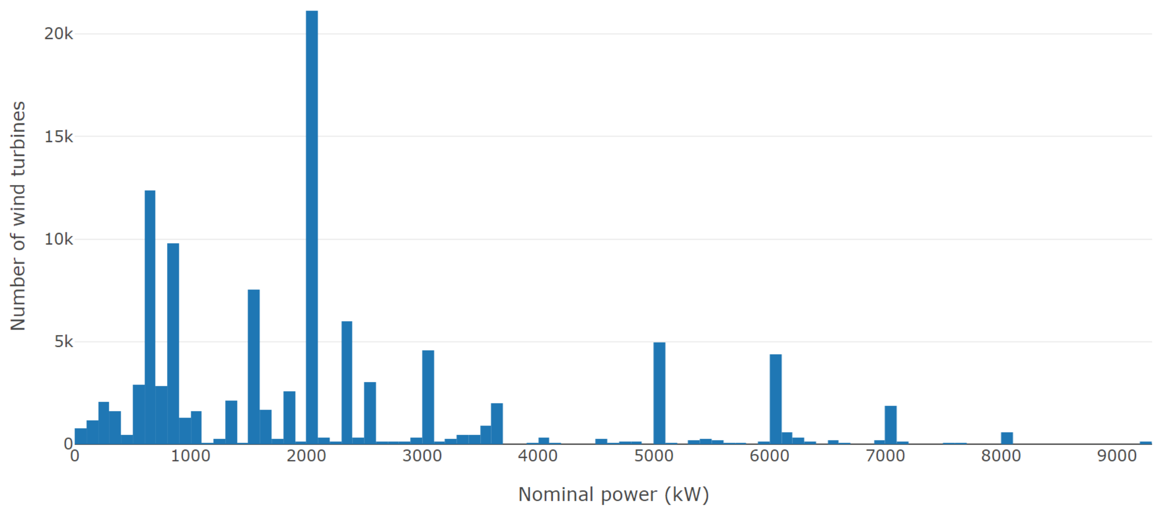

Figure 2.10 — Number of wind turbines as a function of P_{nom} in 2016 in Europe [51].

2.1.2.6 Curtailments

The curtailment consists in limiting or stopping completely the rotation speed of the blades, for several reasons [53]:

- Limitation of noise pollution for neighboring homes.
- Reduction of turbulence between wind turbines that can cause wear and tear (*wind sector management*).
- Wildlife protection (e.g., birds and bats).

The wind operator carries out this clamping manually or automatically according to wind speed, direction and time criteria defined by legislation. There are several technical ways to limit the speed of a wind turbine:

- *Pitch-control* — Some wind turbines (called *pitch-regulated turbine*) allow to change the orientation of the blades in order to change the angle of attack. It is also possible to place the blades in flags (feathering) in relation to the wind direction to reduce the forces exerted to a minimum.
- *Stall-regulation* — Other wind turbines (called *stall-regulated turbine*) have blades rigidly attached to the hub. The blade geometry has been designed to take advantage of the stall effect (see in section 2.1.2.1) at too high wind speeds by causing turbulence on the part of the blade that is not facing the wind. This avoids the installation of moving parts and a complex control system.
- *Yawing* — If the type of the wind turbine allows it, the entire rotor can be rotated according to the wind direction. The rotation can be on a vertical axis for a left-right shift, or on a horizontal axis to tilt the rotor horizontally.

▷ *Mechanical brakes* for a total stop.

The power curves between pitch-regulated and stall-regulated turbine are different because the latter cannot maintain the same power in strong winds, it drops as the blades stall (see figure 2.11). The power curve of a stall-regulated turbine is optimized for a certain wind speed while for a pitch-regulated turbine the maximum power is reached for a whole speed range.

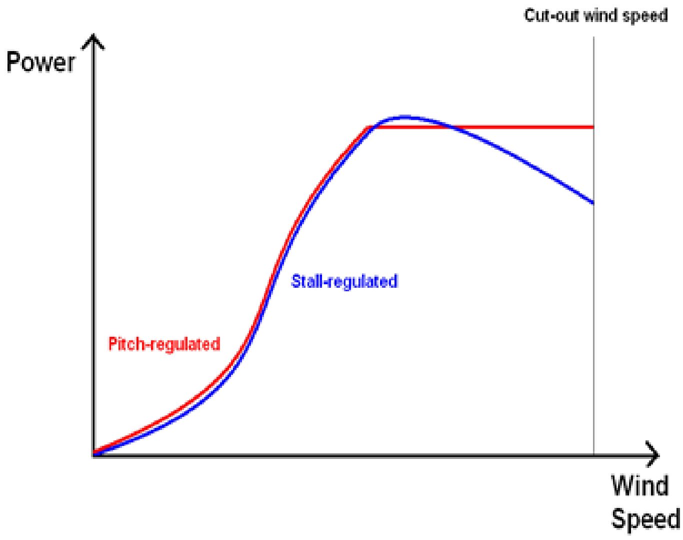

Figure 2.11 — Pitch-regulated and Stall-regulated wind turbine power curves [54].

All the modern megawatt-class wind turbines make use of pitch control to optimize the rotor performance. However, for kilowatt-range machines, stall-regulated solutions are still attractive and largely used for their simplicity and robustness [55].

2.1.2.7 Effects to be considered

Some physical effects have an influence on the production of wind turbines, the main ones are the following:

▷ *Wake effect* — Wind turbines extract energy from the wind and downstream there is a wake from the wind turbine, where wind speed is reduced. As the flow proceeds downstream, there is a spreading of the wake and the wake recovers towards free stream conditions. The wake effect is the aggregated influence on the energy production of the wind farm, which results from the changes in wind speed caused by the impact of the turbines on each other. It is important to consider wake effects from neighboring wind farms and the possible impact of wind farms which will be built in the future [56]. The wake effect induced by wind turbines is particularly noticeable in figure 2.12.

▷ *Hysteresis* — It is the property of a system whose the evolution is different according to whether an external cause increases or decreases. In the case of wind turbines, the production depends on whether the wind is dropping or increasing. To prevent frequent

Figure 2.12 — Example of wake effect [57].

shutdowns and restarts (at the edge of the cut-off), which contribute to turbine fatigue, hysteresis is often applied, so that the wind turbine starts up only when the average wind speed reaches a value lower than the shutdown wind speed, i.e. the turbine will only restart after the wind speed has dropped several m/s below the cut-off speed. This phenomenon is quite rare because wind turbines are placed and selected in such a way as to avoid exceeding the cut-off speed. The phenomenon can also occur at low speeds, in which case the wind turbine only produces energy if the speed decreases (i.e. if the blades are already turning). The effect is illustrated in figure 2.13, from Ws_2 the production is zero and will only increase if the speed drops to Ws_1.

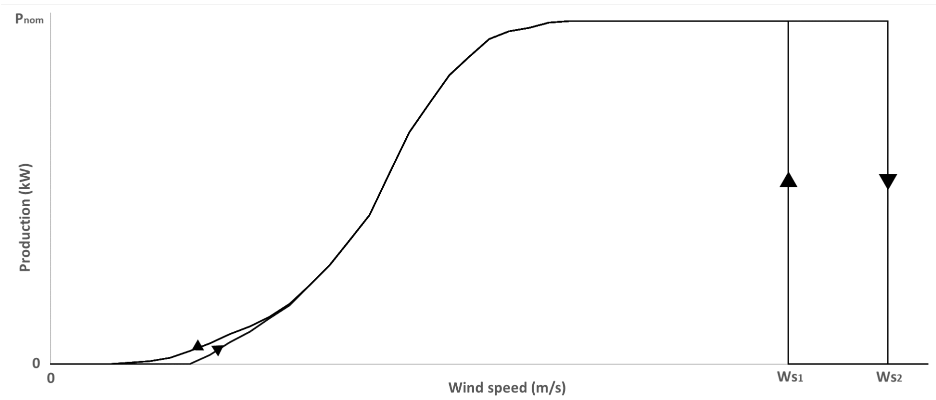

Figure 2.13 — Hysteresis phenomenon applied to wind generation.

▷ *Wind phenomena seen in section 2.1.1.3* — As seen above, hill effect, tunnel effect, roughness or turbulence imply particularities in the choice and placement of wind turbines (micro-sitting). *Hill effect* implies that wind turbines placed on a hill have a rather low mast to take advantage of wind acceleration. The *tunnel effect* mainly has an impact on the small wind turbine located in cities. Due to *roughness*, wind turbines are placed in places far from cities for better efficiency (this also avoids inconvenience for neighboring houses). Finally, *turbulence* increases the fatigue of the mechanical components of the wind turbine. In general, attempts are made to increase the height of the mast to prevent turbulence generated near the ground from affecting the surface swept by the rotor.

2.2 Wind power forecasting

One of the largest challenges of wind power, as compared to conventionally generated electricity, is its dependence on the volatility of the wind. Wind is not as easily "controllable" as fossil fuels and requires to be predicted with the greatest possible accuracy.

This section introduces the different types of wind power forecasts, the different forecast formats and the particularity of offshore wind farm forecasts.

2.2.1 Different uses of forecasts

Wind power forecasts have different uses, the main ones are the following:

- *Wind turbine control* — The forecasts are used to modify the operation of a wind turbine in order to optimize performance and ensure safe operation under all wind conditions. The turbine control is out of the scope of this overview because it involves very short-term (a few seconds) forecasts usually using a lidar (a remote measuring device based on the analysis of the properties of a beam of light returned to its transmitter) in the nose of the turbine, and therefore is qualitatively different from the rest of the approaches mentioned here.

- *Intraday or Day-ahead market trading* — As seen in section 1.4, the producers or their aggregator must provide forecasts for spot markets, therefore forecasts have to be made the day before (Day-ahead) or on the same day (Intraday). An incorrect estimate of production can lead to severe financial penalties.

- *Unit commitment (UC)* and *Economic dispatch (ED)* — If the electricity market is not open, the forecasts may serve for deciding on the use of conventional power plants and for the optimization of the scheduling of these plants.

- *Maintenance planning* — Long time scales are interesting for the maintenance planning of large power plant components, wind turbines, or transmission lines. However, the accuracy of weather predictions decreases strongly looking at 5-7 days in advance. Shorter horizons can also be considered for maintenance to ensure that the crew does not experience too strong winds at the top of the turbine or that they can return safely in the case of offshore wind turbines.

- *Resource assessment* — Before building a wind farm, it is necessary to estimate the average production it will provide during its years of use. These are very long-term forecasts essentially based on climatology.

Table 2.1 presents the different names of the forecasts according to the horizons, the time resolution and the uses. As the time-scale classification of wind power forecasting methods is not expressed clearly in the literature, not everyone uses exactly the same terms. We have chosen the classification used by Jung [58].

Since this thesis was mainly carried out in a company specialized in wind power forecasting, we focused on the medium-term forecasting, more precisely the hourly forecast

Table 2.1 — The different types of wind power forecasts according to time scale.

	Horizon	Resolution	Use
Very-short-term	Few minutes	Seconds, minutes	Wind turbine control
Short-term	30 minutes - 6 hours	15, 30 minutes, 1 hour	Intraday market, UC, ED
Medium-term	6 - 48 hours	1 hour, 3 hours	Day-ahead markets, UC, ED
Long-term	Days, weeks or more	Days, weeks, months	Maintenance planning, resource assessment

for the next day for several possible players in the electricity market. Indeed, since the majority of requests was for this type of forecast, we had access to more data and had points of comparison.

2.2.2 Forecast formats

The production forecasts are mainly provided in two different formats: point and probabilistic forecasts. This section presents these two types of forecasts.

Point forecasts — The point forecasts are the simplest and most familiar type of forecasts. They comprise a single prediction of some future observations, e.g., "the wind speed will be 10m/s one hour from now". The point forecasts, sometimes called deterministic forecasts, are favored by many practitioners because of their ease of use: a non-expert can produce, communicate and interpret point forecasts with relative ease. Most media that provide weather forecasts for public consumption will offer a point forecasts for precisely this reason.

Probabilistic forecasts — The point forecasts are inherently uncertain, and while they offer a "best estimate" of some future quantity, they provide no information as to how confident one can be in that outcome being realized. The probabilistic forecasts offer more information than a point forecast by providing an estimate of the likelihood of a range of possible outcomes, information that is essential for optimal decision-making in many situations. The probabilistic forecasts are the optimal input to decision-making problems with non-symmetric cost functions.

The probabilistic forecasts come in a variety of forms: the quantile forecasts, for instance, estimate the probability that an observation will exceed some value, e.g. "there is a 90% chance that the wind speed will be greater than 5 m/s one hour from now". Similarly, an interval forecast predicts the probability that an observation will fall within some interval. Information pertaining to the full range of possible outcomes is contained in a predictive distribution, where the full probability density function for a future observation is estimated, this may take the form of either a parametric or non-parametric distribution. [59] provides a more detailed review of these techniques.

When multiple connected forecasts are required, such as the wind power generation at several wind farms in the same region, capturing dependence between observations is extremely important. In these situations, scenario forecasts capture both spatial structures and temporal structures necessary for multi-stage decision making problems. Over the past 20 years, there has been a shift from deterministic to probabilistic forecasting in applications from economic and financial risk management to demographic and epidemiological projections [60]. The ability to quantify the confidence of a prediction is extremely valuable to decision makers and is now a common requirement of many forecasting tools, including those designed for the wind power.

However, the probabilistic forecasts are still rarely used in the industry. Since trading is not always automated, there is a lot of information for an operator who does not necessarily know how to interpret forecasts. It is one of the IEA Wind Task 36 Forecasting (Group of experts from different countries and sectors meeting to advance the technological development and global deployment of wind energy technology) objectives to promote the probabilistic forecasting. This group published a scientific overview that allows forecast users to better understand the probabilistic predictions and how to use them [61].

2.2.3 Offshore specificity

The properties of the wind in the offshore environment can be very different from those onshore. The reduced diurnal heating of the surface and the effect of low roughness over vast areas on the atmospheric boundary layer mean that the wind does not exhibit some properties which are familiar onshore [62], [63]. Therefore, authors have proposed methods specifically for offshore wind power forecasting, such as [64], [65].

In [66], authors produced a comparison of prediction accuracy on- and offshore concluding that the performance of offshore forecasting lies somewhere between the onshore sites with simple terrain, which can be forecast with relatively high accuracy, and the onshore sites in complex terrain that are more difficult to forecast.

The wind power forecasting models can be categorized into two parts: the *physical* models, based mainly on fluid dynamics and *statistical* models based on historical data without necessarily understanding meteorological models.

In the next section, we will first present the physical models.

2.3 Physical models

Physical models are the first approach to production forecasting. They are so called because these models are based on physical fluid mechanics equations. These models require a good knowledge of wind physics, unlike statistical methods which are essentially based on data and do not necessarily require to be understood by the forecaster. This section details the weather forecast models and then focuses on downscaling methods to predict wind speed at the wind turbine.

2.3.1 Numerical Weather Prediction (NWP)

The wind production forecast is essentially dependent on a good wind forecast provided by a meteorological model. First, the chaotic nature of the weather is described, then, the characteristics of the different weather models are given before presenting the specificities of wind prediction.

2.3.1.1 The weather: a chaotic system

As Edward Lorenz has noticed, the weather is chaotic, i.e. a dynamic system very sensitive to initial conditions. With his famous question "Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?", he described the impact of a small global change over time [67]. Although atmospheric motions are governed by well-known fluid dynamics laws (deterministic system), inaccuracy in initial conditions makes weather unpredictable (also known as *deterministic chaos*). Indeed, instead of knowing the state (pressure, temperature, etc.) of the atmosphere at any point, they are known only at a certain observation station with a certain precision (the other points are estimated by regression).

Although some of the world's most powerful supercomputers are used for weather forecasting, the chaotic nature of weather limits the spatial and temporal resolution of forecasts. The weather forecasts are therefore typically issued with forecast horizons of between 7 and 10 days with spatial resolution ranging from 5 km to 25 km; and temporal resolution of either 1 or 3 hours. Longer term climate forecasts are made but at much lower resolution. Due to the vast computational expense of weather prediction models, forecasts are typically calculated every 6 or 12 hours [68], [69].

2.3.1.2 Description of the NWP models

Numerical Weather Prediction (NWP) forms the basis of most meteorological forecasts. NWP involves using observations to estimate the current state of the atmosphere and oceans in order to compute their future states. The atmospheric model is initialized and a set of linearized equations describing atmospheric physics, including the Navier-Stokes equation and ideal gas law, are solved on a 3-dimensional grid. Both the initialization of atmospheric parameters and the linearization of the governing equations are critical in producing meaningful forecasts. The main parameters to be determined at all altitudes are pressure, temperature, wind and humidity. They then make it possible to evaluate the meteorological situation (presence of rain, thunderstorms, good weather, high temperatures, etc.)

Since this type of system continuously involves each of the points in the atmosphere, it is also at each of these points that solutions should be found, which is impossible with our current computing capabilities: so the models use a simplified representation of the atmosphere, in which the meteorological fields are known only at the points of a horizontal and vertical grid called the *mesh* or the *grid* [70].

A model can be a *large-mesh* model that can reach several tens of kilometers. These are often *global models* that cover the entire planet. It makes it possible to predict long-term and large scale phenomena (e.g., depressions or high pressure systems) that travel across the globe.

However, a model can be based on a *fine-mesh* with a distance between points of a few kilometers. In this case, the number of calculations to be performed at each time step of the forecast becomes high, due to the large number of grid points distributed within the atmosphere. They are therefore essentially *local models* covering only one part of the planet (e.g., a region or a country). This type of model is often only used for relatively short deadlines (often 36 hours at most), otherwise the calculation time required for each forecast would become excessive. These models are important because a fine mesh makes it possible to take local meteorological phenomena (e.g., in places with reliefs). They are always coupled with a global model, especially for the edges.

A number of NWP models covering different regions of the planet are run in several countries around the world using measurements from weather satellites and radiosondes. Table 2.2 summarizes several main models: the American Global Forecast System (GFS), the European Integrated Forecast System (IFS) and the French "Action de Recherche Petite Echelle Grande Echelle" (ARPEGE) and "Applications de la Recherche à l'Opérationnel à Méso-Echelle" (AROME). The resolution is given in degrees of latitude and longitude. In terms of latitude (the north-south position of a point), 1° always represents about 111.11 km. However, in terms of longitude the conversion depends on the latitude. Thus at the equator 1° represents 111.11 km but only 77.2 km in France located at a latitude of approximately 46° ($111.11 \times cos(46°) \simeq 77.2km$). It should be noted that there are several versions of each model and that the characteristics are not always fixed. For example, the further away the horizon is, the more the time-step increases for IFS and ARPEGE (up to twelve hours).

Table 2.2 — Comparison of several NWP models [71], [72], [73], [74].

Model	**Resolution**	**Time-step**	**Coverage**	**Max horizon**
GFS	0.25° , 0.5° or 1°	3 hours	Global	192 hours
IFS	0.1°	1 hour (live), 3 hours (history) and more	Global	240 hours
ARPEGE	0.1° (France), 0.25° (Europe) and 0.5° (World)	1 hour (France), 3 hours (World) and more	Global	102 hours
AROME	0.025° and 0.01°	1 hour	France	42 hours

2.3.1.3 The particularities of the wind forecasting

Knowing the wind at high altitudes allows to predict large atmospheric movements. On the ground the wind has a strong impact on the temperature felt and can be annoying or even dangerous at high speed. Between the two heights, demand is lower and is not the priority of forecasters. For a wind turbine, it is interesting to have the forecasts at a wind height between 50 and 150m a.g.l (above ground level).

The weather models are regularly recalibrated using the sensor observations. As there is little data available at these heights due to the complexity of the implementation, the forecasts are of low quality. In practice, the models calculate wind speed from the ground speed (the standard is 10 m a.g.l) using a formula to move from 10 m to a higher height.

Two mathematical models or "laws" are generally used to model the vertical profile of wind speed over regions of homogenous, flat terrain [75]. The first approach, the *logarithmic law*, has its origins in boundary layer flow in fluid mechanics and in atmospheric research. It is based on a combination of theoretical and empirical research. Equation 2.9 enables to estimate the mean wind speed at one height z_2 based on that at another z_1 [76].

$$u(z_2) = u(z_1)\frac{\ln\left((z_2 - d)/z_0\right)}{\ln\left((z_1 - d)/z_0\right)} \tag{2.9}$$

where z_0 is the roughness, $u(z_1)$ is the mean wind speed at height z_1 and d the zero-plane displacement (the height in meters above the ground at which zero wind speed is achieved as a result of flow obstacles such as trees or buildings).

The second approach is the *power law*, presented in equation 2.10.

$$u(z_2) = u(z_1)\left(\frac{z_2}{z_1}\right)^{\alpha} \tag{2.10}$$

where the exponent α is an empirically derived coefficient depending the stability of the atmosphere. For neutral stability conditions, α is approximately $\frac{1}{7}$.

Both approaches are subject to uncertainty caused by the variable, complex nature of turbulent flows. The logarithmic wind profile is generally considered to be a more reliable estimator of mean wind speed than the wind profile power law in the lowest 10-20 m of the planetary boundary layer. Between 20m and 100m both methods can produce reasonable predictions of mean wind speed in neutral atmospheric conditions. From 10m to near the top of the atmospheric boundary layer the power law produces more accurate predictions of mean wind speed [77].

The use of these formulas assumes neutral atmospheric stability. This assumption is acceptable when the average wind at 10m height exceeds 10 m/s, so the turbulence mixing outweighs atmospheric instability.

2.3.2 Downscaling method

The physical approach uses the detailed physical description to model the on-site conditions at the location of the wind farm [78]. The basic operation of a physical approach is illustrated in figure 2.14. It carries out the refinement of the Numerical Weather Prediction (NWP) data to take into account the on-site conditions by the downscaling method, which are based on the physics of the lower atmospheric boundary layer. The downscaling method requires the detailed physical descriptions of the wind farms and their surroundings, including: the description of the wind farm (wind farm layout and wind turbine power curve, etc.) and the description of the terrain (orography, roughness, obstacles, etc.). Then, the refined wind speed data at the hub height of the wind turbines is plugged into the corresponding wind power curve to calculate the wind power production. If the on-line data are available, a regression is performed from the last known data on the site (also called *Model Output Statistics (MOS)*) to reduce the error of the forecast. The physical approach does not necessarily require training input from on historical data (which can nevertheless be used to calibrate the parameters). Acquiring the physical data is one of the main drawbacks of the approach.

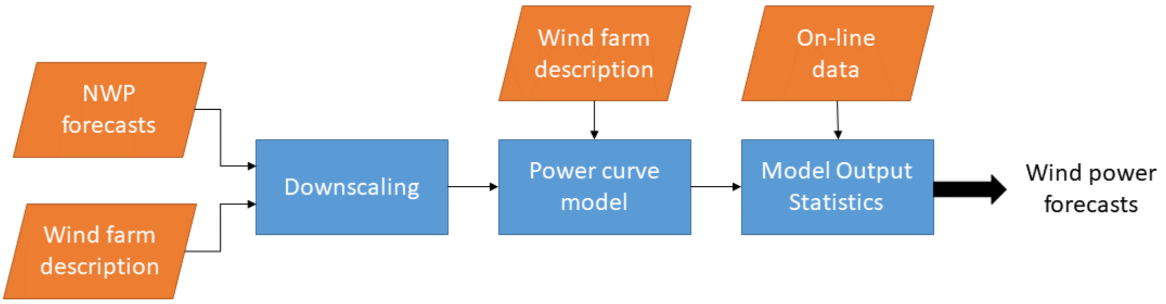

Figure 2.14 — Physical approach process for the wind power forecasting.

A number of physical approaches have been introduced [78]–[82]. The Prediktor is developed by the Risoe National Laboratory in Denmark. It uses Wind Atlas Analysis and Application Program (WAsP) and PARK program to take the local conditions into account by using the NWP forecast from High Resolution Limited Area Model (HIRLAM) [83]. The Previento, developed by the University of Oldenburg in Germany has a similar physical approach but uses a different NWP forecast from Lakelmodell of the German Weather Service [84]. The LocalPred is developed by CENER - National Renewable Energy Centre in Spain. It involves adaptive optimization of the NWP forecast, time series modeling, meso-scale modeling with MM5, and power curve modeling [85]. The eWind, developed by AWS TrueWind Inc. in the USA, has a similar physical approach to Prediktor but uses a high-resolution boundary layer model (ForeWind) as a numerical weather model to take the local conditions into account [86].

Physical approaches are based on the models using the fundamental physical principles for conservation of mass, momentum, and energy in air flows. These models address Computational Fluid Dynamics (CFD) for simulating the atmosphere. Although there are many CFD models available, they are all based on the same basic physical principles. They differ in how the grids are structured and scaled, and how the numerical computations are performed.

The physical approaches are currently mainly used for the very short-term and short-term horizons (up to six hours). In these cases the use of physical approaches provide good results due to the influence of atmospheric dynamics.

In the next section, we will present another type of approach: the statistical approaches which produce good forecasts at all horizons and are increasingly being used.

2.4 Statistical models

The *statistical models* represent the other main category of approaches for the wind power forecasting. Instead of using knowledge about the wind physics or the description of the wind farm, the models are based solely on data. This is done by the use of gray- or black-box statistical models that are able to combine inputs, such as NWP of speed, direction, temperature etc., of various model levels together with the online measurements of wind power. Unlike the process of physical approaches, seen in figure 2.14, the process of statistical approaches is a direct transformation of the input variables to wind power.

The advantage of statistical methods is that they theoretically make it possible to model additional effects or phenomena based on empirical observations. For example, effect such as hysteresis can be taken into account in the data while it is physically complex to model.

This type of method (especially machine learning) has been developed in all applications of artificial intelligence for years [87]. It requires less specification and knowledge of the field but a sufficiently large amount of data, representative of the reality, is needed. Their current popularity is due in particular to the performance of algorithms in many applications and the ease of access and use and large volume of data.

This section presents the two main categories of statistical methods: the models based on time series analysis and the machine learning models including linear methods, support vector machine, artificial neural networks and ensemble methods.

2.4.1 Time series models

The wind speed is a time series which is a series of numerical values representing the evolution of a specific quantity over time. This temporal characteristic implies certain relationships and coherence between the different values of this series. Based on the past values of a time series, it is possible to predict future values.

In the case of wind forecasting, if the horizon is short (less than six hours), the forecast can be performed by time series analysis methods without using a weather model. Direct time series models are models that use recent observed values of wind and other variables to predict the future wind speed or wind power. The temporal analysis methods are commonly referred to as conventional statistical approaches.

The Box-Jenkins method [88] is a standard approach for short-term forecasting. The method is divided into four main steps to make a mathematical model of the problem including model identification, model estimation, model diagnostics checking, and forecasting. Several types of time series model may be considered, including Auto-Regressive model (AR), Moving Average model (MA), Auto-Regressive Moving Average model (ARMA), and Auto-Regressive Integrated Moving Average model (ARIMA). The general form of the model is given in the following equation:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{2.11}$$

where X_t represents the forecasting parameter at time t, φ_i is the autoregressive parameter, θ_i is the moving average parameter, c is a constant, and the random variable ε_t is the white noise.

This model called ARMA(p,q) represents the ARMA model having the autoregressive model of order p and the moving average model of order q. If the order of the moving average model q is zero, it represents the autoregressive model of order p (AR(p)). If the order of the autoregressive model p is zero, it represents the moving average model of order q (MA(q)). The ARIMA model is a generalization of an ARMA model.

In summary, the conventional statistical approaches are based on classical linear statistical models such as AR, MA, ARMA, and the Box-Jenkins approach. These models are easy to formulate and are capable of providing timely forecasts. Other methods can also be found in the literature, such as Markov Switching AutoRegressive model (MSAR) [89] or wavelet transform [90].

These approaches are mostly aimed at very short-term and short-term (less than six hours) and are suitable for data at short intervals (between 5 and 10 minutes). In the case of medium-term wind power forecasting, with time steps of one hour, the temporal dependencies are less important. In this case, the methods based on weather forecasts and machine learning are the most popular.

2.4.2 Machine Learning methods

Another statistical approach is to use Machine Learning (ML) algorithms. The NWP and potentially other variables are transformed into the wind power by ML which has been trained by the large sets of historical data in order to learn the dependence of the output on input variables. From the historical data, the model learns a function that links weather forecasts and production that is then applied to live weather forecasts (see figure 2.15). This approach does not require explicit mathematical expressions to be used in the field of application such as those used in the physical approach discussed above.

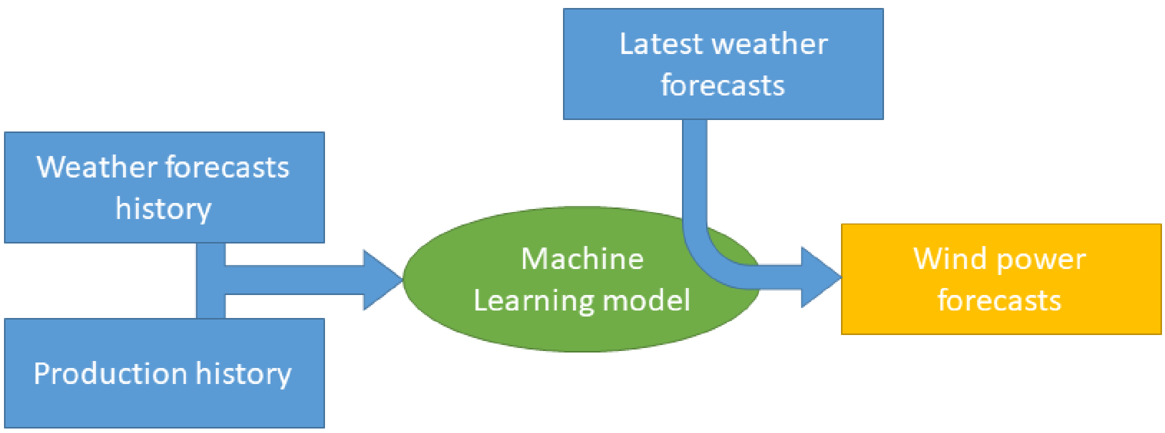

Figure 2.15 — General process of using a Machine Learning model for wind power forecasting.

In recent years, the ML applications have increased significantly. The most well-known algorithms are based on methods that are decades old. It is essentially the good results obtained in several application areas, ease of use and data access that have made ML so popular. In the field of wind power forecasting, there is a lot of work testing different ML methods on data sets. The main ones are presented in the following sections.

2.4.2.1 Linear methods

Multi-Linear Regression (MLR) is a technique used to model the relationship between multiple input variables (feature variables) and an output dependent variable. The model remains linear in that the output is a linear combination of the input variables.

There is most general case called Polynomial Regression where the relationship is modeled as an nth degree polynomial. This however requires knowledge of how the data relates to the output.

The linear regression is simple to understand which can be very valuable for business decisions. For non-linear data, the polynomial regression can be quite challenging to design, as one must have some information about the structure of the data and relationship between feature variables. As a result of the above, these models are not as good as others when it comes to highly complex data.

2.4.2.2 Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a machine learning technique for classification [91]. In this method, classifications are done by finding an Optimum Separating Hyperplane (OSH) which maximizes the minimum distance between the classes. In particular, when an OSH cannot be found in the original space, a nonlinear and high-dimensional mapping should be applied. The notion of OSH can be extended to regression problems, in this case the model is called Support Vector Regression (SVR).

This method is widely used in wind power forecasting [92]–[94] and more generally in any machine learning application. The SVR models perform well and are robust to outliers. However, they are sensitive to parameters (notably C a penalty constant for regularization and ε which controls the width of the ε-insensitive zone) and require precise tuning.

2.4.2.3 Artificial Neural Networks

An Artificial Neural Network (ANN) consists of an interconnected group of nodes called neurons. The input feature variables from the data are passed to these neurons as a multi-variable linear combination, where the values multiplied by each feature variable are known as weights. A non-linearity is then applied to this linear combination which gives the neural network the ability to model complex non-linear relationships. A neural network can have multiple layers where the output of one layer is passed to the next one in the same way. At the output, there is generally no non-linearity applied. The neural networks are trained using Stochastic Gradient Descent (SGD) and the backpropagation algorithm [95].

Since neural networks can have many layers (and thus parameters) with non-linearities, they are very effective at modeling highly complex non-linear relationships. However, they can be quite challenging and computationally intensive to train, requiring careful hyper-parameter tuning and setting of the learning rate schedule. Moreover, they require a lot of data to achieve high performance and are generally outperformed by other ML algorithms in "small data" cases.

The basic network is the Multi-Layer Perceptrons (MLP) [96] which is a feedforward neural network, i.e. the connections between the nodes do not form a cycle. The network often has between 1 and 3 hidden layers (which are neither the input nor the output). When the number of hidden layers is large, the method is called *Deep learning*. The model is thus more abstract and takes longer to train, but it can model more complex relationships. Several neural network architectures exist, the main ones are the following:

- *Convolutional Neural Networks* (CNN) are a classic network to which convolution and pooling layers are added in order to pre-process the information. These layers aim to limit the number of entries while maintaining the strong spatially local correlation of natural images. Usable on large images unlike MLP it is mainly used for image and video recognition (see [97], [98] for some examples).
- *Long/Short Term Memory Networks* (LSTM) are Recurrent Neural Networks (RNN), i.e. networks wherein connections between the nodes form at least one cycle. While basic neural networks use "independent" inputs, LSTM uses data inputs where time correlations (at short and long term) are important. LSTM is suitable for time series analysis, for example in speech recognition or translation (see [99]–[104] for some examples).
- *AutoEncoders* are a type of network used to learn efficient data encoding in an unsupervised manner (i.e. without knowing the output). In this type of network, the input is equal to the output and the goal is to obtain a hidden layer of smaller size in order to compress the signal without loss of information. In [105], AutoEncoders are used as a pre-processing tool in conjunction with other machine learning methods.

2.4.2.4 Ensemble methods

Instead of using a single machine learning model, it is possible to aggregate thousands of models with divergent opinions but each of which can be specialized on parts of the data given. These methods are called ensemble methods and most often give better results. The general principle is to convert several weak learners into a strong learner. A "weak learner" is defined to be a model that is only slightly correlated with the true value (it can forecast examples better than random guessing). In contrast, a strong learner is a model that is well-correlated with the true value. According to [106], there are several ensemble method approaches, the main ones are the following:

- *Bagging*: the training set is randomly sampled into several sub-sets uniformly and with replacement (bootstraping). Then, models are fitted using the above bootstrap samples and combined by averaging the output (for regression) or voting (for classification) [107].

Since weak learners are independent, the fitting can be done in a totally parallel way. These algorithms have good performance. The most commonly used basic model is the decision tree, in this case the associated ensemble model is called Random Forest [108], [109].

- *Boosting*: the idea of boosting is to train weak learners sequentially, each trying to correct its predecessor. Unlike bagging, where weak learners have the same weight, they are weighted according to their accuracy. At each step, after a weak learner is added, the data weights are readjusted, known as "re-weighting". Mispredicted input data gain a higher weight and examples that are predicted correctly lose weight. Thus, the future weak learners focus more on the examples that the previous weak learners mispredicted. Some of the most famous boosting algorithms include AdaBoost [110] and Gradient boosting [111]. (see [112]–[116] for some examples)

- *Stacking*: also called meta-ensembling, Stacking is an ensemble method used to combine information from multiple predictive models (not necessarily weak) to generate a new model [117], [118]. Several learners with different algorithms are fitted on all data by cross-validation. Then, another model (often a linear model, for example elasticnet) is driven from the forecasts of the first models. This is a time-consuming method that requires testing many combinations. It is mainly used in data science challenges to gain a few tenth of percentage of error on a particular set but it not cost-effective for most businesses.

The ensemble methods are widely used in artificial intelligence because they are robust to outliers and avoid overfitting (when the model is too specialized to the learning set).

2.5 The spatial forecasting methods

The previous section presented the different standard forecasting methods in which the forecast is generally performed at a single wind farm based on the total production history of the farm and a single grid point of the weather forecast model. However, the wind has a strong spatial dependence because it is a continuous physical phenomenon. It is therefore consistent to integrate these spatial dependencies into the wind power forecasting models. This section presents the two types of interdependencies: short and large scale interdependencies.

2.5.1 Short scale dependencies

Due to the low resolution of weather models or a lack of data, the wind power forecasts are usually provided at the farm scale (more precisely to the injection value on the grid because it is the reference value for penalties). In other cases, a production is forecasted for every wind turbine independently and the farm production is simply obtained by the sum of these forecasts. However, for a single farm, wind turbines productions are very correlated with each other, especially between close turbines. Figure 2.16 compares the production at the same time between two near (dark color) and far (light color) wind turbines from the same farm. For a better visualization, the graph shows the statistical distribution of the values with twenty boxplots. If all the wind turbine productions were equal, we would obtain identity functions. Overall, productions follow the same trends, but a greater distance between the wind turbines leads to a larger production difference.

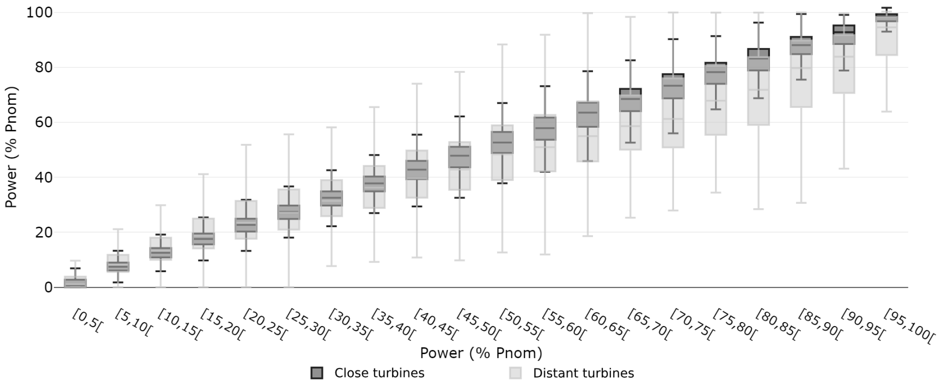

Figure 2.16 — Comparison between the production of two close and distant wind turbines in the same wind farm.

Moreover, since a wind turbine generates electricity from the energy in the wind, the wind leaving the turbine has a lower energy content than the wind arriving in front of the turbine. A wind turbine thus interferes with its neighbors and can cause a production decrease on the turbines located behind it downwind. This phenomenon is the wake effect seen in section 2.1.2.7. This additional information has to be taken into account in the forecast process with the aim of improving the prediction accuracy. Therefore, the problem is to forecast the production at wind farm level by considering local constraints between turbines.

This problem is most often studied by a physical approach via Large-Eddy Simulations (LES) [119]–[124]. Although useful for the initial placement of turbines, simulating wind turbines and their effects with LES is computationally expensive, making wind-farm-scale simulations unreasonable in an operational context.

In [125], authors proposes to solve this problem with a statistical approach. A first forecast is done independently at turbine-level with machine learning algorithms. The farm production is then computed by a weighted sum of the forecasts where the weights are determined by linear regression. An improvement can be observed by firstly dividing data by the wind direction and then determining the weights. However this approach solves the problem in a global way, considering all the turbines together. It does not take into account the local constraints between close turbines.

2.5.2 Large scale dependencies

As with a single farm, a set of farms may have correlations in their production if they are in similar wind regimes (i.e. a regional wind with the same characteristics over a large area). It is therefore interesting to forecast the production for the entire region directly, on a "large scale".

The large scale wind power forecasts have been the subject of several studies [126], [127]. The most straightforward method to predict the power generated over an entire area is to sum the individual forecasts of each power plant (with a single point from a grid of NWP). To do that, the locations and characteristics of each plant must be known. As the weather models are not uniformly effective, independent wind power forecasting of farms does not lead to better global results (although more data are available).

Another alternative method was proposed in [128] and [129]. In these papers, the authors enter a large number of grid points from the weather model covering an entire region. The data is pre-processed via a Principal Component Analysis (PCA) in order to reduce the size of the input without loss of information. Then, a machine learning model (Among neural network, analog encoder and gradient boosting) is trained from the total production of the region. These methods are particularly useful if farm data are not accessible.

In this thesis, we assume that we have access to the wind turbine and wind farm production data and we want to improve the short and large scale forecasting with these additional data. In the literature, these data are often not available or ignored.

2.6 Analysis of the current models

This thesis focuses on improving forecasting models by taking into account spatial dependencies, more specifically the interdependencies between wind turbines productions. This section compares and analyzes the different standard methods presented before, in order to evaluate their relevance for our problem. This comparison is carried out by defining evaluation criteria, comparing the main approaches using them and finally providing a synthesis of this chapter.

2.6.1 Analysis criteria

In order to compare the models, we chose several criteria corresponding to constraints common to the forecast models and others more specific to the production forecast; these are the following:

Ease of implementation — Depending on the method, the model implementation could be easy or not. This criterion aims to evaluate the necessary effort required to adapt a general concept in order to develop a specialized method.

Data dependency — Historical production data are not always available and weather forecasts do not always have the same spatial resolution depending on the forecast location. Data dependency can be a critical criterion for some wind farms.

Consideration of spatial correlation — Since wind is a continuous physical phenomenon, this criterion aims to determine whether spatial correlations are taken into account in the forecast.

Interdependencies — The forecast of a wind farm can be done independently or in conjunction with other wind farms or with wind turbine data, taking into account the interdependencies between the different entities. This consistency between several forecasts makes the model more robust in case of a forecast error and can provide information to the model. This criterion aims to assess whether the consistency of forecasts and the interdependencies between wind turbines and wind farms are used to improve the forecast.

Ability to handle dynamics — The production of a wind farm is constantly evolving according to weather conditions but can also be modified in the long term (modification of the environment, wear on wind turbines, etc.). A forecasting model must adapt to these changes without external intervention.

Execution time — It is important to obtain the forecast in a timely manner. For day-ahead forecasting this criterion is flexible because the forecaster has several hours to send all the forecasts for the next day. The execution time can be divided into model training time and forecast time.

Forecast accuracy — The main role of a forecasting model is to provide the most accurate forecasts possible. This criterion corresponds to the forecast error obtained by the method.

2.6.2 Evaluation of the criteria

As seen in sections 2.3 and 2.4, there are two main classes of forecast models: the physical models that simulate the physical effects of wind and the statistical models that learn from historical data. This section compares these two approaches with the criteria described above. The evaluation of the criteria is commented in tables 2.3 and 2.4.

Table 2.3 — Assessment of the criteria for physical models.

Criteria	Score	Comments
Ease of implementation	- -	Models and algorithms are not easily accessible and require specific meteorological skills.
Data dependency	-	Physical models require accurate information on terrain topography, farm characteristics and wind sensor data. However, models do not require a learning phase on a data set.
Spatial correlation	+ +	Models are directly a simulation of the physical wind, local phenomena and spatial correlations are taken into account.
Interdependencies	+	In some cases, the effects of wake between wind turbines are taken into account in the simulation.
Ability to handle dynamics	- -	Physical models do not adapt to the specificities of the farm or to changes without external modification of the model.
Execution time	-	The forecast time can be long depending on the quality of the simulation. However, models do not need to be trained for long periods of time.
Forecast accuracy	-	Physical models represent the historical approach to forecasting. However, for day-ahead forecasts, they perform less well than statistical models.

2.6.3 Synthesis

The physical and statistical approaches were evaluated according to our criteria in the previous section. Table 2.5 summarizes the scores obtained for each criterion.

We observe that, overall, the statistical approach is better rated according to our criteria. It is simple to implement in any farm, can withstand changes, is not time-consuming to make

Table 2.4 — Assessment of the criteria for statistical models.

Criteria	Score	Comments
Ease of implementation	+ +	There are many open source libraries implementing statistical algorithms. Apart from data recovery and formatting, there is not much difference between predicting on one farm or another.
Data dependency	-	Weather forecasts and production data history, in sufficient quantity, are almost mandatory to obtain good results. However, it is not necessary to know the topography of the land or any specific information about the wind farm.
Spatial correlation	- -	In most cases, a single grid point in the weather model is used to predict an entire wind farm.
Interdependencies	-	In most cases, the forecast is made directly for the entire wind farm or possibly by making an independent forecast for each wind turbine.
Ability to handle dynamics	+	Statistical models reflect the characteristics, even complex ones (e.g., curtailments, hysteresis, etc.), that can be observed in the data. To handle dynamics, the algorithm must be re-trained with more recent data.
Execution time	+	Although the learning phase of statistical approaches can be very long depending on the complexity of the algorithms and the size of the data, the forecast phase is often almost instantaneous.
Forecast accuracy	+	Statistical models perform well in terms of error, they tend to exceed physical models because they are more studied nowadays.

Table 2.5 — Wind power forecasting methods summary.

Criteria	Physical methods	Statistical methods
Ease of implementation	- -	+ +
Data dependency	-	-
Consideration of spatial correlation	+ +	- -
Interdependencies	+	-
Ability to handle dynamics	- -	+
Execution time	-	+
Forecast accuracy	-	+

forecasts and provides accurate forecasts (in relation to the literature). However, statistical methods rarely take into account spatial correlations and interdependencies between wind turbines and wind farms. These criteria are points on which physical methods focus. However, the latter are less and less used because they are difficult to implement, require specific data, do not adapt to changes and produce forecasts that are now less accurate than statistical methods. These two approaches are limited on some criteria and the purpose of this thesis is to study a new approach that takes up the main current challenge of the field of application: fulfilling the enumerated criteria while improving forecast accuracy.

The main research focus of this thesis is the consideration of spatial correlations and interdependencies between wind turbines and wind farms, which are the particularly missing criteria of the state of the art for statistical methods. The scientific track followed is to consider a wind farm as a complex system, i.e. a non-linear dynamic system composed of interacting entities. Complex systems are systems whose behavior is intrinsically difficult to model due to the dependencies, competitions, relationships, or other types of interactions between their parts or between a given system and its environment. A wind farm (or a set of wind farms) is a complex system because its entities interact with each other through the wind due to wake effects. The farm itself evolves in another complex and chaotic system, the atmosphere, whose evolution is calculable but unpredictable.

In view of these observations, in the next chapter, we will study Multi-Agent Systems (MAS) because MAS are potentially suitable for modeling complex systems and interdependencies between entities. We will precisely formulate the problems studied, present the theory of MAS and the subcategory of Adaptive MAS and explore how an approach based on these paradigms may be relevant to our problems.

Adaptive Multi-Agent Systems for Wind Power Forecasting

Contributions

3 Integration of Interdependencies in Forecasting through Cooperation

We have seen the limitations of current forecasting models, especially when considering spatial constraints. This chapter details the problems addressed and briefly presents the MAS (Multi-Agent System) paradigm. Finally, it introduces the AMAS (Adaptive Multi-Agent System) paradigm, which represents a potential solution to our problem.

3.1 Interdependencies in wind power forecasting 58
3.1.1 Short scale 58
3.1.2 Large scale 59
3.2 Multi-Agent Systems 61
3.2.1 Agent 61
3.2.2 Environment 63
3.2.3 Properties of MAS 63
3.2.4 MAS for energy and forecasting 64
3.3 AMAS: Adaptive Multi-Agent Systems 64
3.3.1 Introduction 65
3.3.1.1 Complex systems 65
3.3.1.2 Emergence 65
3.3.2 Interaction and cooperation 66
3.3.3 Functional adequacy 67
3.3.4 Adaptation 68
3.3.5 Applications of AMAS 69
3.3.6 Relevance of the AMAS approach to the problem 70

3.1 Interdependencies in wind power forecasting

This section details the problem addressed in this thesis: how to take into account the interdependencies between entities involved in forecasting the production of one or more wind farms. The interdependencies are studied at two levels: short and large scale.

3.1.1 Short scale

For technical and economical reasons, the wind turbines in the same wind farm are placed in restricted areas and are physically close. The cost of the land is cheaper, maintenance is simpler and the installation of wind turbines is regulated (especially due to noise pollution). Figure 3.1 shows a French wind farm with 15 wind turbines arranged in a line. Data is provided by Boralex and will be presented in section 6.2.

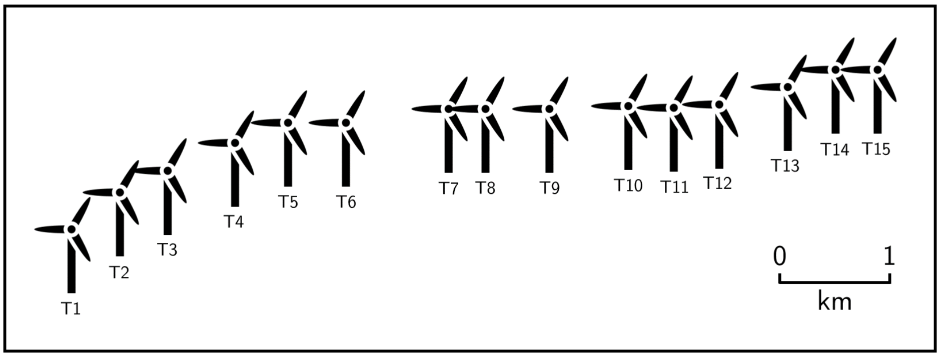

Figure 3.1 — The layout of a wind farm consisting of 15 wind turbines.

All the wind turbines being in the same area (about 8 km separate the most distant wind turbines in this example), a simple assumption is that production will be very close between the wind turbines. However, by observing the correlation between the production of two nearby (T1 and T2) or distant (T1 and T15) wind turbines, there are significant differences. These production comparisons are shown in figure 3.2. An equal production would be shown by the curve $y = x$. Excluding the cases of breakdowns and maintenance (the points on the x and y axis), we can see that the points follow this trend. However, instead of a line, there is a large interval that increases with the distance between the wind turbines.

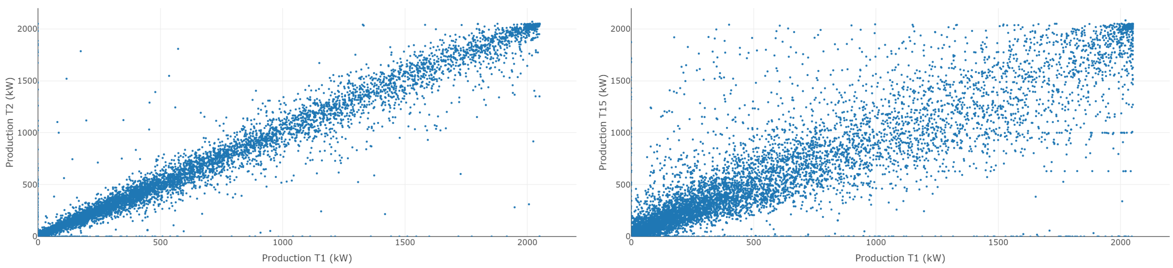

Figure 3.2 — Comparison of the production between two close (left) and distant turbines (right).

As seen in section 2.5.1, these differences are related to local phenomena. The wind and therefore the production can be different between two wind turbines that seem to be close and located in the same wind regime.

More generally, figure 3.3 shows the Pearson correlation coefficient between the production of each wind turbine over a three-year history. The nearest wind turbines have a correlation factor of about 96% while the most distant ones have a correlation factor of 87%. This decrease is gradual and produces a color gradient in the table. Failures have been filtered from the data.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
T1	100,0%	96,9%	96,6%	95,4%	95,6%	94,6%	93,0%	92,4%	92,7%	91,8%	91,8%	91,3%	91,7%	91,1%	87,5%
T2	96,9%	100,0%	97,5%	96,1%	95,8%	94,3%	92,6%	91,9%	92,2%	91,1%	91,0%	90,5%	91,1%	91,0%	87,2%
T3	96,6%	97,5%	100,0%	96,9%	96,5%	94,9%	93,2%	92,3%	93,0%	91,5%	91,5%	90,8%	91,3%	91,2%	87,8%
T4	95,4%	96,1%	96,9%	100,0%	97,0%	95,0%	93,5%	92,9%	93,1%	91,7%	91,3%	90,5%	91,2%	91,0%	87,1%
T5	95,6%	95,8%	96,5%	97,0%	100,0%	96,9%	94,9%	94,2%	94,8%	93,7%	93,0%	91,8%	92,5%	92,4%	88,4%
T6	94,6%	94,3%	94,9%	95,0%	96,9%	100,0%	95,1%	94,6%	95,2%	93,5%	93,8%	92,7%	92,8%	92,6%	89,1%
T7	93,0%	92,6%	93,2%	93,5%	94,9%	95,1%	100,0%	95,9%	96,0%	94,2%	94,4%	93,6%	93,1%	92,5%	89,0%
T8	92,4%	91,9%	92,3%	92,9%	94,2%	94,6%	95,9%	100,0%	96,7%	95,0%	94,5%	93,6%	93,2%	92,7%	89,9%
T9	92,7%	92,2%	93,0%	93,1%	94,8%	95,2%	96,0%	96,7%	100,0%	97,0%	96,5%	95,4%	95,2%	94,6%	91,9%
T10	91,8%	91,1%	91,5%	91,7%	93,7%	93,5%	94,2%	95,0%	97,0%	100,0%	96,4%	95,8%	95,8%	95,1%	92,0%
T11	91,8%	91,0%	91,5%	91,3%	93,0%	93,8%	94,4%	94,5%	96,5%	96,4%	100,0%	96,4%	96,4%	96,0%	92,0%
T12	91,3%	90,5%	90,8%	90,5%	91,8%	92,7%	93,6%	93,6%	95,4%	95,8%	96,4%	100,0%	97,2%	95,6%	92,7%
T13	91,7%	91,1%	91,3%	91,2%	92,5%	92,8%	93,1%	93,2%	95,2%	95,8%	96,4%	97,2%	100,0%	97,4%	93,5%
T14	91,1%	91,0%	91,2%	91,0%	92,4%	92,6%	92,5%	92,7%	94,6%	95,1%	96,0%	95,6%	97,4%	100,0%	94,0%
T15	87,5%	87,2%	87,8%	87,1%	88,4%	89,1%	89,0%	89,9%	91,9%	92,0%	92,0%	92,7%	93,5%	94,0%	100,0%

Figure 3.3 — Pearson correlation coefficient between the production of each turbine in a wind farm over three years of data.

In addition, even nearby wind turbines can produce a different amount of energy at the same time. In particular wind conditions, one wind turbine can interfere with another and reduce its production. Figure 3.4 shows the production of two wind turbines T1 and T2 for 5 days. It can be seen that the production of these two close wind turbines may be different on certain dates (e.g., on May 26 or on May 29), while it may be almost identical on other dates (e.g., on May 27). These differences, that are furthermore not regular over time, may be explained by some of the effects seen in section 2.1.2.7. Here, a wake effect certainly explains these differences.

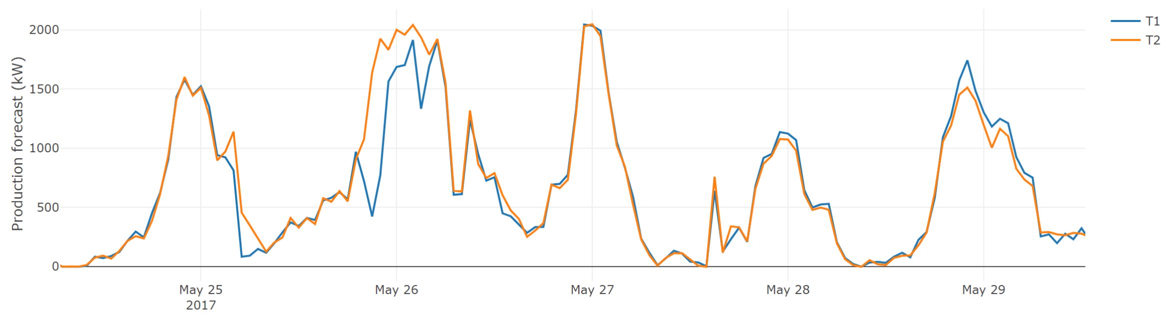

Figure 3.4 — Comparison of production between two neighboring turbines T1 and T2.

3.1.2 Large scale

On a large scale, there are also correlations between wind farms. This occurs when farms are located in an area where the wind has most of the same characteristics (they are said to have the same wind regime). Different large French regional winds are represented in figure 3.5 by wind roses (the lines represent the distribution of the wind origin). Mistral is for example a fresh wind from a north and northwest direction present in the Rhone Valley. These regional winds seem very local but are all linked by global phenomena.

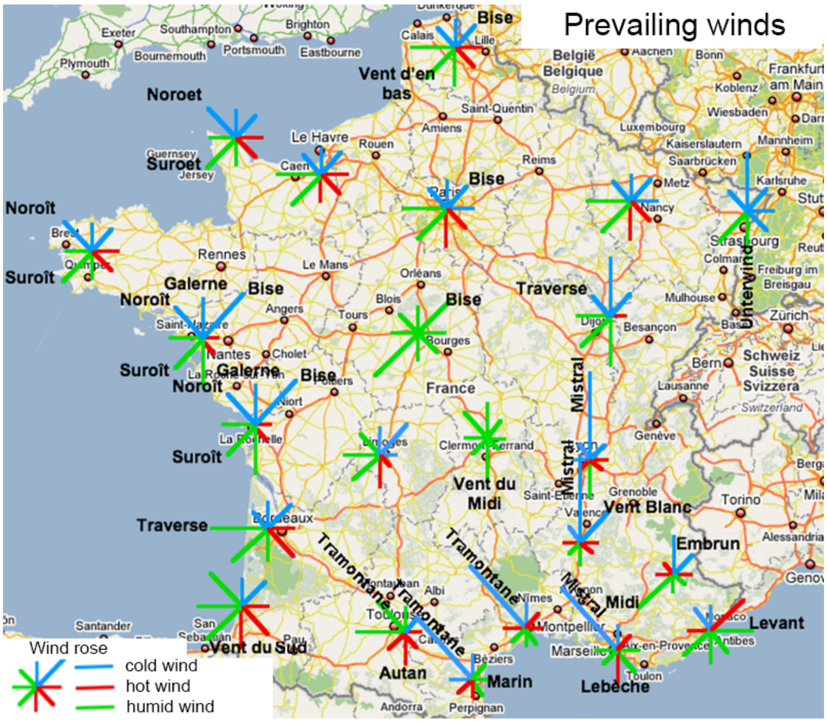

Figure 3.5 — Prevailing winds in France [130].

Five French wind farms are studied in this thesis, their locations are shown in figure 3.6. Figure 3.7 shows the correlation coefficients between the production of the farms. We can see significant correlations between farms A, B and C (between 44 and 57%) while farms D and E are more independent from each other and from others (between -3 and 19%). Note that distance is not always a good indicator of correlation (for example, farms A and C have more correlated production than D and E).

On a large scale, the farms have no (or little) impact on each other because the wake effect is very local. The problem here is how to use our knowledge at the farm level to improve the overall (regional) forecast.

These two problems at short scale and large scale concern systems made up of entities interacting with each other in an environment. The aim is to model this system as well as possible in order to predict its next states, in particular wind turbine production.

The standard approaches presented in section 2.4 do not generally take into account dependencies (neither short nor large scale). Another approach is therefore necessary and the purpose of this thesis is to study a different method from those commonly used in industry. From this perspective, in relation to our initial problem, we are interested in approaches capable of modeling systems in which the parties can interact with each other. A good representative of these approaches are Multi-Agent Systems (MAS) and therefore the following section aims to ensure that this type of approach is appropriate to the problem.

Figure 3.6 — Location of the five wind farms studied.

	A	B	C	D	E
A	100,00%	57,58%	44,11%	19,42%	15,41%
B	57,58%	100,00%	50,79%	3,71%	8,68%
C	44,11%	50,79%	100,00%	14,56%	-3,10%
D	19,42%	3,71%	14,56%	100,00%	16,22%
E	15,41%	8,68%	-3,10%	16,22%	100,00%

Figure 3.7 — Pearson correlation coefficient between the production of the five French wind farms over three years of data.

3.2 Multi-Agent Systems

Multi-Agent Systems (MAS) are systems composed of multiple interacting and autonomous entities, the agents, within a common environment. Each agent has only a partial view of its environment. MAS offer a methodological way to study complex systems with a bottom-up approach. MAS are used in many different domains, from collective problems solving to the study of collective behaviors. The MAS paradigm proposes to focus on the design of agents and their collective behaviors leading to the realization of a particular task. This distribution of tasks inside a MAS makes them highly suitable to overcome a greater complexity than the complexity apprehended by conventional methods. In this section, we present the key concepts of Multi-Agent Systems [131].

3.2.1 Agent

The term "agent" is generic and many paradigms use it. A commonly accepted definition is that "an agent is a computer system that is situated in some environment, and that is capable of autonomous actions in this environment in order to meet its design objectives" [132].

The definition has been enriched by [131], notably by adding a locality criterion. Ferber defines an agent as "an autonomous physical or virtual entity able to act (or communicate) in a given environment given local perceptions and partial knowledge. An agent acts in order to reach a local objective given its local competence".

This definition highlights the fundamental properties of an agent:

- ▹ An agent is *autonomous*, which means that it is the only one to control its behavior. This implies that the choice to act or not is only driven by the agent own behavior. The agent capacity to say "no" (to choose not to act) makes a concrete differentiation between an agent and a sub-program (e.g., an object in object-oriented programming).
- ▹ An agent evolves in an *environment* (physical or virtual) in which it is able to locally perceive information and locally act. An intuitive definition would be that the environment of an agent is everything that is external to the agent and which can be perceived by the agent, including the other agents. This environment acts as the interaction medium.
- ▹ An agent is able to *interact* and *communicate* with other agents either directly or through the environment.
- ▹ An agent possesses a *partial* knowledge of this environment.
- ▹ An agent possesses its own *resources* and *skills*.

The agent behavior is ruled by a three-step life-cycle of "Perception-Decision-Action":

- ▹ *Perception* is the process during which the agent acquires information from its environment and updates its internal state.
- ▹ *Decision* is the process during which the agent decides of actions to perform. This decision is based on its local perceptions, its internal knowledge and its own objectives.
- ▹ *Action* is the process during which the agent performs the actions.

Beside those common properties, agents possess other characteristics that enable their differentiation [133]. An agent is said *reactive* when its actions are triggered by events that occur in its environment as a reflex behavior. The trigger rules are dependent of the agent perceptions and its internal state. Those kinds of agents have generally few or no memory. On the opposite there are *proactive* agents which are able to modify their objectives and create new ones. They are also refereed as *cognitive* agents as they often involve complex reasoning or learning algorithms. There is no concrete frontier between reactive and proactive agents. The reactive agents are less complex, and so they are usually numerous, each agent focusing on a simple task. The system is said to have a fine-grained granularity. On contrary, proactive agents can process more complex tasks, involving that the system needs less agents to reach its objectives. The system is said to have a coarse-grained granularity.

Other characteristics can be mentioned like *situated* agents which interact with other agents through the environment or *communicating* agents which interact with other agents by sending and receiving messages.

3.2.2 Environment

A MAS is a system that is located in an *environment* which is not only a source of information, but also the medium by which the agents act and interact. While being a key component of MAS, the environment lacks of formal definition which reaches a consensus inside the MAS community [134]. Intuitively, the environment of an entity can be described as everything which is not this entity. Depending of the adopted point of view, different environments can be identified. In a MAS, we can either adopt the MAS point of view (the system is viewed at its macro-level) or an agent point of view (the system is viewed at its micro-level).

From the system point of view, the environment is everything that is outside of the system. From the agent point of view, the environment is not only a part of the MAS environment, but also the other agents. For example, if we consider a school of fish as a MAS where each agent represents a fish from the school, the environment of the school corresponds to reefs, plankton, predators, etc. The environment of a fish corresponds to the environment of the school and other fish.

In the literature, the environment is often characterized using the following properties [135], [136]:

- ▷ *Accessible/Inaccessible*: the agent environment is accessible by an agent if the agent is able to perceive all the information required for its task.
- ▷ *Discrete/Continuous*: the agent environment is discrete if it possesses a finite number of distinct states.
- ▷ *Deterministic/Non deterministic*: the agent environment is deterministic if its evolution consecutive to an action is only dependent of its current state.
- ▷ *Dynamic/Static*: the agent environment is dynamic if it evolves despite the agent inactivity or during its deliberation.

3.2.3 Properties of MAS

In MAS, each agent has incomplete information or capabilities for solving the problem and, thus, has a *limited viewpoint*. However, all the required knowledge and skills required for solving the problem are still present, distributed among the system. Thanks to this *distribution*, the MAS paradigm seems particularly suited to problems with a natural distribution.

A MAS is *open* if agents can appear or disappear during the system lifetime. On the opposite case, the system is said to be *closed*. The appearance of an agent is most of the time the result of the decision of an existing agent while its disappearance can be the decision of an agent (which then commits a form of suicide), or initiated by the environment.

Another property is the absence of external or global control system. The control is distributed inside each agent, and each agent is the only one to be responsible of its behavior.

3.2.4 MAS for energy and forecasting

MAS are a highly studied field of research that has had concrete applications in many fields. While MAS have been used in the field of energy, they are mainly used for grid energy management [137]–[142], i.e. the optimal management of the distribution and the storage of the energy produced in the energy grid. Other examples of applications exist such as the optimization of energy consumption at home [143], the management of the maintenance of a wind farm [144], [145], the realization of wind farm diagnoses [146] or the forecast of the electricity consumption for self-consumption [147].

As the electricity network is composed of many heterogeneous and interconnected actors, a distributed and decentralized solution is often relevant for problems related to this field. MAS are particularly suitable for distributed problems because the distribution of control is one of their main properties. For wind power forecasting taking into account the interdependencies between wind turbines and wind farms, MAS are relevant due to the distributed nature of the problem.

However, the other important point of our problem is to take into account the interdependencies between the entities of our system and the fact that these entities are the central point in the problem-solving process. The overall behavior of the system will be modified by the relationships between the entities and thus cause the system to adapt to changes (variability of weather conditions as well as changes in terrain over long periods of time).

Although MAS can somehow adapt or manage interdependencies, they are not specifically designed for this purpose. We want adaptation and interdependencies to be the driving force in solving the problem. Therefore, we will study a particular type of MAS: Adaptive Multi-Agent System (AMAS). The following section presents the AMAS theory and discusses whether it is relevant to the issues addressed in this thesis.

3.3 AMAS: Adaptive Multi-Agent Systems

In the previous section, agents and Multi-Agent Systems (MAS) have been defined. Modeling a wind farm (or a set of wind farms) by a MAS is relevant because the problem is naturally distributed. However, MAS are not specifically suited to other characteristics of the problem (the need to adapt and take into account interdependencies between entities). We therefore propose to study a particular type of MAS, Adaptive Multi-Agent Systems (AMAS). This section presents the theory of AMAS by introducing the concepts of complex systems and emergence then by detailing the notion of cooperation and functional adequacy. Then, examples of work with AMAS in areas related to energy and forecasting are presented. The different characteristics presented are finally confronted with the constraints related to our problem in order to discuss the relevance of AMAS.

3.3.1 Introduction

AMAS are a particular type of multi-agent systems where agents have a specific behavior known as “cooperative”. They are particularly used to model complex systems and emergence, concepts that we will detail in this section.

3.3.1.1 Complex systems

The real world and most of the problems associated with it can be considered *complex*. Complexity must not be mistaken with complication. Where a complicated problem has many different and well-defined parts with well-known behaviors and can be reduced to simpler problems (a big puzzle, for example, can be divided into smaller and easier puzzles, the final solution being the sum of all these parts put side by side), a complex problem is defined by an important number of interacting little parts and it is their interactions that produce the global behavior. The parts of a complex system are guided by simple and individual rules and the behavior of the system cannot be predicted from individual rules.

Complex systems are generally defined by the following characteristics [148], [149]:

- ▷ A large number of heterogeneous entities.
- ▷ Multiple objectives which are possibly conflicting.
- ▷ A high degree of connectivity between variables, i.e a change in one variable may affect many others (feedback loops and non-linearity).
- ▷ A lot of feasible actions, with different effects and consequences that cannot be determined a priori (indeterminism).
- ▷ An environment subjected to dynamic evolving, those changes can be spontaneous making the situation less predictable (dynamics).

Those complex systems can be found in plurality of domains and science: from social to neural sciences. The weather, seen in section 2.3.1.1, is for example a complex system. Modeling such systems and solving related problems are important subjects in computer science.

3.3.1.2 Emergence

Traditionally, the way to tackle a problem or to design a software, is to adopt a top-down approach. The initial problem is divided into simpler sub-problems until the point where they can all be solved. This approach needs to know the system finality: each part bringing something to this finality, the latter must be a priori known. Such an approach is qualified as *reductionist*: a part can always be divided into other parts. Note that the reductionism, besides system design, can be applied to several fields of study of philosophy [150].

This approach is quite good for complicated systems, where the global behavior is well defined. Thus, such systems can be easily decomposed into parts. But the reductionism cannot handle complex systems. In these systems, many little parts are interacting and each has its own rules and objective [151]. Top-down approaches cannot be applied to study complex systems, the approach must be reversed: instead of considering the whole, it is necessary to study and design the little parts and their simple behavior and interactions.

In a complex system, the final function cannot be predicted from the knowledge of the local parts. As Aristotle has written: "the whole is greater than the sum of the parts". This property is called *emergence*. Even if a global definition does not exist for this notion, an emergent phenomenon is something that is perceptible from a macro-level and produced by sub-level interactions (the micro-level) [133], [152]. In the school of fish example, it is the interactions between fish (the sub-level) that produce the school (the macro-level), and so the school cannot be predicted from the fish point of view. As the emergent property is unpredictable from the micro-level, it brings something new to the system. As it is new and unpredictable, the emergent property is decentralized: no single part controls it, but interactions from a lower level produce it.

Emergence cannot be only reduced to the interactions of local parts. The emergent property appears while the system is evolving: emergence is linked to dynamic properties. Moreover, the emergent property is not one-time but it remains during a certain period. At first, some fish interact with one another and then, when there are enough to swim in the same way, the school emerges. Fish keep this configuration during a long time, and the school evolves: new fish, new patterns (to avoid a predator for example), etc.

3.3.2 Interaction and cooperation

A Multi-Agent System can often be considered as a complex system. Indeed, a MAS is composed of many interacting parts, the agents, with simple and local behaviors. Each agent is autonomous and is not controlled from a macro-level. Moreover, no entity controls the whole system. The interactions between agents produce the emergent property: the function of the system. Thus, to design MAS that produce the right function, the key is to find correct local interactions and implement agents with those rules. The Adaptive Multi-Agent Systems (AMAS) theory [153], proposes a theoretical framework to design such systems. This approach is based upon the interactions between agents, and the *cooperation* notion.

Jennings in [154] distinguishes three kinds of interactions: *antinomic*, *neutral* and *cooperative*. An entity has an antinomic interaction if its action disturbs another entity in the accomplishment of its activity. If the action does not disturb but does not favor either, the interaction is neutral. Finally, if the behavior of an entity favors the behavior of another, the interaction between them is cooperative. If all the interactions between the system and its environment are cooperative, the system is in a cooperative state, else if the interactions are neutral or antinomic, the system is in a non cooperative state.

From an agent point of view, cooperation is defined as their ability to work together in order to realize their objectives. Thus, four properties must be satisfied to ensure cooperative interactions [155]:

- ▷ *Sincerity*: an agent is sincere and therefore never lies.

- ▷ *Willingness*: a request is always satisfied if it is coherent with the agent state and if it has skills to perform it.
- ▷ *Reciprocity*: all agents know those properties and respect them.
- ▷ *Fairness*: when it is possible, the agent with the lowest level of non-satisfaction degree is favored to be satisfied.

In order to measure and compare this notion of *fairness* at the local level, an agent needs a criterion to question its behavior and relationships. This criterion is based on its cooperative social attitude and its degree of dissatisfaction with respect to its local goal. That is why the notion of "criticality" has been introduced and can be defined as follows [156]:

Definition 1. *The **criticality** of an agent represents the state of dissatisfaction of it regarding its local goal.*

The cooperative social attitude of an agent consists in always helping the most critical agent in its (limited) neighborhood (without being altruistic i.e. without becoming the most critical agent). The criticality value is domain-dependent and has to be normalized in order to be compared between agents. The actions of the agents aim to minimize the criticality of all agents in the system without the need for global knowledge.

The notion of cooperation is at the heart of the AMAS problem-solving process. The following section presents the functional adequacy theorem that links the local cooperative behavior of agents to the function performed by the entire system.

3.3.3 Functional adequacy

A key notion of the AMAS approach is the functional adequacy. An artificial system is designed to perform a function and intuitively, the system is functionally adequate when it performs the function for which it was designed. Usually, the evaluation of the functional adequacy is determined by an external entity which observes the system activity. However, with a MAS, this evaluation has to be performed by the inner agents which have no clue on the global task. This must be realized by agents with self-observation capacities, evaluating only local criteria. A theorem in the AMAS theory stipulates that a system in which all the agents are in a cooperative state is functionally adequate [157], [158]. The AMAS approach proposes a definition of the functional adequacy based on the categorization of the interactions between a system and its environment.

Definition 2. *A system is functionally adequate if it has no antinomic activity on its environment.*

Reciprocally, a cooperative system, which has only beneficial activities with its environment, is functionally adequate.

Given this definition, [155] expresses the theorem of functional adequacy as:

Theorem 3.3.1 (Theorem of functional adequacy)**.** *Given a functionally adequate system, there exists at least one cooperative internal medium system that fulfills an equivalent function in the same environment.*

A cooperative internal medium is a system in which all the interactions between its constituting parts are cooperative. For more information on the demonstration of the theorem, the reader can refer to [159].

Thus, for each problem where a solution is effectively calculable, there exists a MAS where all the agents are in a cooperative state that solves this problem. The design of a functionally adequate system can be made with a focus on the design of local cooperative interactions between the constituting parts.

Methodologies exist to help in the design of an AMAS such as ADELFE [160]–[162] (French acronym for Atelier de DEveloppement de Logiciels à Fonctionnalite Emergente, Toolkit for Designing Software with Emergent Functionalities). The goal of this methodology is to guide the development of AMAS, through five work definitions, from preliminary requirements to design and fast prototyping.

3.3.4 Adaptation

As it has been said earlier, a MAS is coupled with its environment. As soon as a change in the environment occurs, the system functionality may not be in functional adequacy anymore. The more the system is complex, the more difficult it is to reach and maintain a functionally adequate state. As expressed by the AMAS approach, the non adequacy of the system comes from the existence of non cooperative interactions within the system. In order to repair the functionality of the system and reach a functionally adequate state, agents within the system must locally detect failures in cooperation and modify their behavior accordingly. Self-organization of an AMAS rests on the self-observation capacities of its agents to detect, anticipate and repair non cooperative situations.

An agent is in a Non Cooperative Situation (NCS) when there is a failure in its perception, decision or action process resulting in non cooperative interactions. Seven types of NCS have been identified [163]:

- *Incomprehension*: the agent does not understand the message it has received.
- *Ambiguity*: a single message can be understood in different ways.
- *Incompetence*: the agent has no skill to process the information it has perceived.
- *Unproductiveness*: the agent cannot propose an action to do during the decision.
- *Concurrence*: the agent perceives another agent which is acting to reach the same world state.
- *Conflict*: the agent believes that the transformation it is going to operate on the world is incompatible with the activity of another agent.
- *Uselessness*: the agent believes that its action cannot change the world state or it believes that the results for its action are not interesting for the other agents.

To solve a NCS, an agent has to locally adjust its behavior. In order to do so, the agent disposes of three means [164]:

▷ *Tuning*: the agent adjusts its internal parameters.

▷ *Reorganization*: the agent changes the way it interacts with its neighborhood, i.e. it stops interacting with a given neighbor, or it starts interacting with a new neighbor, or it updates the confidence given to its existing neighbors.

▷ *Openness*: the agent creates one or several other agents, or deletes itself.

The behavior of an agent can be split in two parts:

▷ the *nominal* behavior which ensures the functional adequacy when the agent is in a cooperative state.

▷ the *cooperative* behavior which enables the agent to reach its nominal behavior.

The cooperation in an AMAS is assured by mechanisms which either anticipate or resolve NCS. This task is devolved to the system designer who has to identify the NCS and to propose the adequate mechanism. By resolving NCS at the local level of agents, the overall behavior of the system adapts to changes in the environment.

3.3.5 Applications of AMAS

Currently, there is no theoretical tool powerful enough to model a dynamic system such as an AMAS in the general case. Hence, the applications of the theory plays an important role in the validation of the approach.

AMAS have already been applied to wind power forecasting [165] using AMOEBA (Agnostic MOdEl Builder by self-Adaptation), a generic learning tool based on AMAS [166]. This work was not applied to spatial interdependencies but was a first proof of concept to test the use of AMAS for wind power forecasting. A major issue identified by this study is AMOEBA sensitivity to weather forecast errors, the latters being largely the difficulty of the problem.

Since its conceptualization, the AMAS approach has also been applied to different energy or forecasting problems:

▷ Flood forecast [167].

▷ Analysis of power flux and state estimation for voltage regulation in an electricity grid [168].

▷ Frequency regulation of the electricity grid by using electric vehicles fleet [169].

▷ Optimization of photovoltaic panel production by water cooling [170].

▷ Estimation of missing information for smart cities with a limited number of sensors [171].

These application examples all include some kind of prediction or estimation of values based on interactions between several entities. However, they are based on sensor data considered to be mostly reliable. In the case of medium-term production forecasting (up to several days ahead), weather model forecasts are unreliable and largely inaccurate. Although an AMAS allows learning a function from data in application fields similar to that of this thesis, an important lock is that this learning is done on data that is a priori inaccurate.

3.3.6 Relevance of the AMAS approach to the problem

As we saw before, a wind farm is comparable to a mechanical system because it is composed of several elements that interact with each other via physical laws. One of the missing points in the state of the art is the lack of consistency between high and low level forecasts (e.g., the production forecast of a whole wind farm without using data from the wind turbines). Considering the wind farm as a system composed of interacting elements (and not as a single entity whose total production must be known) is a first step towards taking into account this consistency between the forecasts. Moreover, a wind farm consists of several spatially naturally distributed entities (their positions are fixed but possibly mobile for floating wind turbines). The problem must be solved locally because the behavior of the wind turbines (i.e. the evolution of their production) will depend on their position. In addition, a global method for calculating all interactions is more expensive in terms of computation time than a local method. This is a significant problem for large wind farms, some of which can reach several hundred of wind turbines.

The decentralization and the distribution of control of MAS make them relevant to the problem.

In addition, the problem can be considered as open and dynamic because a wind farm has a lifespan of several decades, in which the wind farm can evolve directly (addition/withdrawal of wind turbines) or indirectly (wear of wind turbines that change production). The external environment is also likely to change directly (addition of buildings, deforestation, etc.) or indirectly (climate change, change of flora). An effective forecasting system must also take these changes into account. In addition, the data is numerous and can be heterogeneous (e.g., a wind farm can be composed of several types of wind turbines operating differently and two wind farms can be of very different sizes).

As seen through the examples of applications presented in section 3.3.5, AMAS are relevant for problems where the model must adapt to the data. Therefore, all the needs for adaptation encountered in our problem make AMAS relevant to solve it in a dynamical an autonomous way. Although the examples of applications given were not specifically designed to use highly erroneous input data, the fact that AMAS are able to learn from heterogeneous and incomplete data may suggest that AMAS are also able to deal with such inputs.

In conclusion, the need for consistency between forecasts, the natural distribution of the problem and the need to adapt the model to inaccurate and heterogeneous data made us identify AMAS as a potential approach for a wind power forecasting method taking into account the interdependencies between wind turbines and wind farms.

Therefore, the next chapters present the two systems – AMAWind-Turbine and AMAWind-Farm – that are based on this technology to deal with wind power forecasting at short and large scales.

4 AMAWind-Turbine: a MAS for Wind Power Forecasting at Turbine-Level

In this chapter we present the first Adaptive Multi-Agent System designed in this thesis: AMAWind-Turbine (**A**daptive **M**ulti-**A**gent system for **Wind** power forecasting at **Turbine**-level). It is dedicated to forecasting the production of a wind farm by taking into account the short-scale interdependencies between wind turbines.

The objectives of AMAWind-Turbine are presented before describing its environment. Entities that make it up are identified with a focus on the cooperative behavior of the agents which are responsible for the emergent solving of our forecasting problem. At the level of the implementation, the criticality of the agents are designed and then discussed. Finally, some technical choices are discussed.

4.1 Objectives 74
4.2 Environment 74
4.3 Entities 75
4.3.1 Grid Point entity 77
4.3.2 Grid Point Hour entity 77
4.3.3 Wind Turbine 78
4.4 A cooperative agent: Wind Turbine Hour agent 78
4.4.1 Description 78
4.4.2 Local goal 79
4.4.3 Behavior 79
4.4.4 Non Cooperative Situations 80
4.4.4.1 Uselessness 81
4.4.4.2 Incompetence 81
4.5 Criticality implementation 82
4.5.1 Percentiles computation 83

4.5.2 Generic criticality function 84
4.5.3 Subcriticalities computation 85
4.5.4 Final criticality 86
4.6 Technical choices 87
4.6.1 Software architecture 87
4.6.2 Agent initialization 88
4.6.2.1 Choice of initial forecasts 89
4.6.2.2 Increment choice 89
4.6.3 Agent scheduling and stop condition 89
4.7 Conclusion 90

4.1 Objectives

As seen in section 2.4, current statistical machine learning methods are able to predict wind power satisfactorily but do not take into account the interdependencies between wind turbines production. In addition, the study on the Adaptive Multi-Agent System (AMAS) approach in section 3.3.6 showed that the AMAS paradigm is a relevant solution because the problem is naturally distributed, deals with interdependencies and requires adaptation.

The main objective of AMAWind-Turbine is to forecast the production of a wind farm by integrating the interdependencies between wind turbines with the constraint of not degrading the quality of the forecasts compared to current state-of-the-art methods. We will likely agentify parts of the wind farm in order to learn these interdependencies locally. The aim will be to adjust the forecast of a wind turbine to take into account both the production history and weather conditions, but also the current forecast of other wind turbines.

4.2 Environment

AMAWind-Turbine is a system based on the representation of a wind farm, which is a physical entity of the real world. The MAS environment is, in this case, the physical environment of the wind farm, i.e. the surrounding air acting on the wind turbines. More generally, the environment is anything that can change air circulation, such as the rugosity or the topography of the land. The description of AMAWind-Turbine environment can be made using the characteristics defined by [135]:

- The environment is *dynamic*: the topography of the land, and therefore the airflow, can be modified by the addition/removal of buildings or the growth of forests. The wind farm can also be modified by adding, removing or replacing a wind turbine and by wear over time (reduction in production efficiency).
- The environment is *continuous*: it is based on physical events in the real world.

- ▷ The environment appears to be *non-deterministic*: the environment is partially observable, the consequences of performed actions in the real world could not be determined in advance with certainty.
- ▷ The environment is *non-accessible*: not all information that could be used is available to the system.

This analysis reinforces the idea that the problem is complex and that the Adaptive Multi-Agent System approach is adequate. The following section presents the different entities and agents that make up the system.

4.3 Entities

As seen in section 3.3, the design of an AMAS is usually done by a bottom-up approach. The entities and agents making up the system are identified and their relationships defined. A differentiation is made between *active entities*, which have their own dynamics and can initiate the activity even in the absence of external stimuli, and *passive entities*, which have no dynamics of their own and can only be perceived and potentially altered by active entities and the system itself (e.g., a passive entity may correspond to data that can be accessed). Finally, an *agent* is an active entity that has a purpose, the ability to perceive its environment and to interact with other active entities autonomously. This agent becomes a *cooperative agent* if, in its interactions with others, it may find itself confronted with Non Cooperative Situations (NCS) presented in section 3.3.4 and has to solve them to come back on a cooperative state. To define the entities and then the agents of the system, we will start from the data that can be used for the problem.

A wind farm consists of a set of wind turbines. We have their characteristics (mainly their geographical position, maximum power, type and power curve) and production history. We also have a weather forecast history corresponding to certain grid points of the Numerical Weather Prediction (NWP) model. This statement reveals two entities corresponding respectively to a wind turbine and a grid point. However, wind power forecasting has two stages: learning from historical data and then forecasting from live data. We can therefore deduce two passive entities that will only serve as databases: the Wind Turbine (WT) entity and the Grid Point (GP) entity. Then, we deduce two active entities in charge of providing the live forecasts: the Wind Turbine Hour (WTH) entity and the Grid Point Hour (GPH) entity. Although one WTH entity and one GPH could be created for each date and time we need to forecast (e.g., a WTH entity which would be in charge of the forecast for the fourth wind turbine on 23/08/2019 at 8:00 am.), we have chosen to limit the number of WTH and GPH entities to 24. Therefore, each WTH and GPH is responsible for one hour of a specific date, chosen during the initialization of the entities. This choice is relevant because 24 agents are sufficient to manage the problem.

Now that the entities are defined, we need to determine the possible links that may exist between these entities. Firstly, WTH entities are connected to WT entities representing the same wind turbine while GPH entities are connected to GP entities representing the same grid point. These links exist because the entities represent the same physical "element". In addition, since the entities are geographically located, it is relevant to connect those which are closest physically. Therefore, WT entities are connected with other WT entities and GP entities, this relationship is also transmitted to the 24 related WTH entities and GPH entities. Figure 4.1 illustrates an example in which WT #1 is related to WT #2 and GP #1. In this case, all WTH entities connected to WT #1 will be related to the WTH entities related to WT #2 corresponding at the same hour as well as the GPH entities related to GP #1 corresponding at the same hour. Several choices of neighborhoods are possible in terms of relations between wind turbines, or between wind turbines and grid points. These choices are discussed in section 4.4.1.

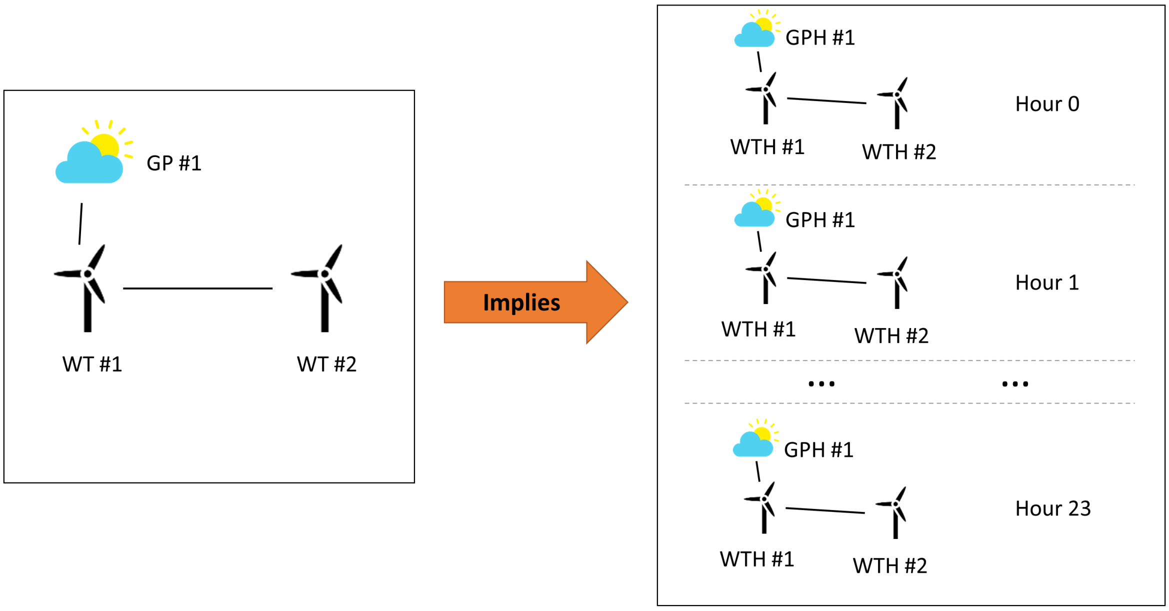

Figure 4.1 — Links between WT/GP and WTH/GPH entities.

While the GPH entity has no possible action on the weather forecast (it is provided by an external numerical weather prediction model), the WTH entity can modify its production forecast. Indeed, since this latter is precisely the desired output of the system, the WTH must ensure that this value is consistent with its neighborhood (i.e. the history of WT, the GPH forecast and other WTH entities). The WTH entity is therefore an agent, and it is by modifying its forecast according to its perceptions that the system will achieve its objective.

To summarize, the analysis stage of the ADELFE methodology has led us to identify several types of entities:

- Passive entity: *Grid Point* (GP) entity.
- Active entity: *Grid Point Hour* (GPH) entity.
- Passive entity: *Wind Turbine* (WT) entity.
- Cooperative agent: *Wind Turbine Hour* (WTH) agent.

The resulting architecture of the AMAWind-Turbine system is given in figure 4.2 and the features and relationships of these entities are detailed below. The WT and GP entities are also linked to each other (in the same way as WTH agents and GPH entities) but the links are not displayed so as not to overload the graph.

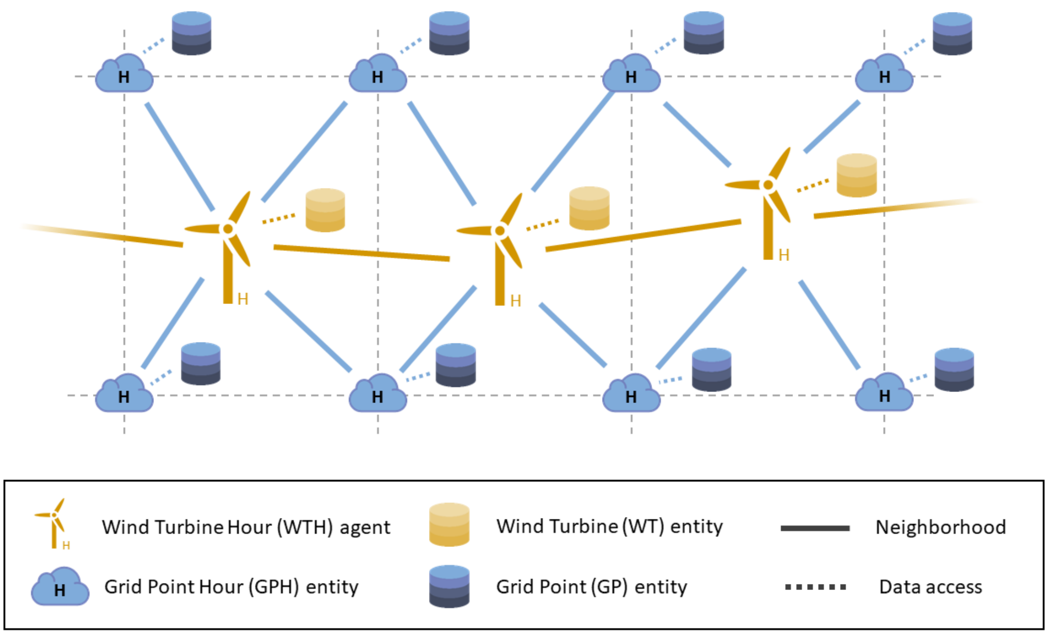

Figure 4.2 — Architecture of the AMAWind-Turbine system and relationships between entities and agents.

4.3.1 Grid Point entity

A weather model provides forecasts on specific coordinates called grid points. For each wind turbine the forecasts of the nearest grid point are used. Since farms can be spread over long distances, several grid points are used per farm.

A grid point entity corresponds to a physical grid point. It therefore has a specific location and contains all the historical weather forecast data at this location. It is a passive entity, it cannot act directly on the system. In computer terms, it is only a database. The GPH entity, which we will present after, can transmit requests to it.

4.3.2 Grid Point Hour entity

A Grid Point Hour (GPH) entity is a grid point associated with an hour. For each grid point, there are 24 GPH, corresponding to the 24 hours of the day.

Its role is to provide the live weather forecast of a grid point at a given hour. Unlike weather forecast history, which are accessible via GP entities and used to train the model, these forecasts are used to compute the wind power forecast.

4.3.3 Wind Turbine

A Wind Turbine (WT) entity corresponds to a wind turbine. It has a location, characteristics (e.g., nominal power, power curve, etc.) and a production history.

Like the GP entity, the WT entity is passive and only corresponds to a database that the Wind Turbine Hour (WTH) agent can access. The latter is presented in more detail in the next section.

4.4 A cooperative agent: Wind Turbine Hour agent

AMAWind-Turbine embeds only a single type of cooperative agent called Wind Turbine Hour (WTH) agent. This section describes this agent, its local goal and nominal behavior before detailing the non cooperative situations with which it may be confronted and its behavior in these cases.

4.4.1 Description

A Wind Turbine Hour (WTH) agent is responsible of the forecast of a wind turbine production at a given hour (e.g., an agent charged with forecasting for the wind turbine 4 at 08:00 a.m.). As for Grid Point Hour entities, there are 24 WTH per wind turbine.

The neighborhood of a WTH agent is based on physical closeness: at a given hour, a WTH agent is related to, at most, the two closest WTH agents. An example is illustrated in figure 4.3, all agents linked to wind turbine 2 will have in their neighborhood the agents corresponding to the same hour and linked to wind turbines 1 and 3. This choice remains arbitrary and other configurations could be considered. In this thesis, since we worked on wind farms with few wind turbines and a simple structure, this choice is consistent. Large farms with several lines can be found, especially in offshore, in which case the neighborhood could be defined in a different way. WTH agents corresponding to the same wind turbines and different hours are not neighbors.

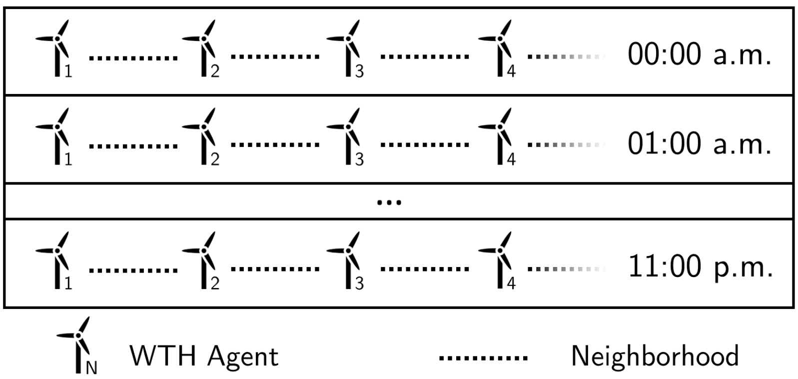

Figure 4.3 — Neighborhood between WTH agents for an example of a layout.

In addition, each WTH agent has access to historical and live weather forecasts and production history through the GP, GPH and WT entities with which it is associated.

4.4.2 Local goal

The local goal of a WTH agent is to adjust a forecast related to a wind turbine and a specific hour. The forecast must be consistent with the wind turbine production history but also with the production history of neighboring wind turbines. For example, if an GPH entity reports that the wind will be strong, the value of the forecast will also be high. However, this is not the only constraint taken into account: if a turbine always produces more than another older turbine, the forecast of the first will always be higher than the other turbine forecast. The consistency of the forecast of an agent therefore corresponds to constraints on the production and on the difference in production with the neighboring agents.

As seen in section 3.3.2, the distance to the local goal of an agent is represented by a value called criticality ranging from 0 (non-critical state towards the goal) to 100 (very critical state). Since the quality of the forecast made by a WTH agent depends on the consistency of its forecast with both its own past productions and the neighboring agents forecast, its criticality is therefore expressed by considering these two factors seen as two subcriticalities: *Local* and *Neighboring* criticalities. How these latter are expressed and combined to implement the criticality function of a WTH agent will be detailed in section 4.5.

4.4.3 Behavior

Since a WTH agent is a cooperative agent, its purpose is to determine the forecast that minimizes its criticality while avoiding aggravating the criticality of its neighbors. As an agent, its behavior follows the classical *Perception-Decision-Action* lifecycle (described in section 3.2.1) during which it may decide to modify its forecast to be as cooperative as possible. After having started with an initial forecast which can be random or chosen to optimize the time of the resolution, the nominal behavior of a WTH agent is the following:

▷ *Perception* — The agent perceives the current forecast and criticality of its neighbors. It also knows its own forecast and criticality.

▷ *Decision* — The agent decides how it will change its forecast. These changes are decided on the basis of previously perceived information. The agent has two possibilities: to increase or to decrease its forecast. The increment is fixed at a constant value, this choice will be explained in section 4.6.2. It simulates these different cases and performs the action minimizing the maximal criticality of its neighbors and itself. Even if this action increases its criticality, as seen in section 3.3, the agent acts in a cooperative way by helping the most critical agent while avoiding to become more critical than this latter. In the event of equal criticality between the two possibilities, the second most critical agent is taken into account (then the third if the second maximum criticalities are equal).

▷ *Action* — The agent finally performs the decided action and therefore possibly changes its forecast. Its criticality is consequently updated at the end of the cycle.

Figure 4.4 shows an example where a WTH agent (#2) has to decide how to modify its forecast considering the criticality perceived from its two neighbors (WTH agents #1 and

#3). As seen before, this agent has two possible choices: increase or reduce its forecast. If it increases it, it increases also its criticality but reduces that of others. If it reduces its criticality, the criticality of the WTH agent #1 increases drastically. Increasing the criticality of a neighbor which is already more critical is not cooperative. Therefore, the solution chosen by the WTH agent #2 is the one that minimizes the highest criticality among its own and that of its neighbors, i.e. increasing its forecast.

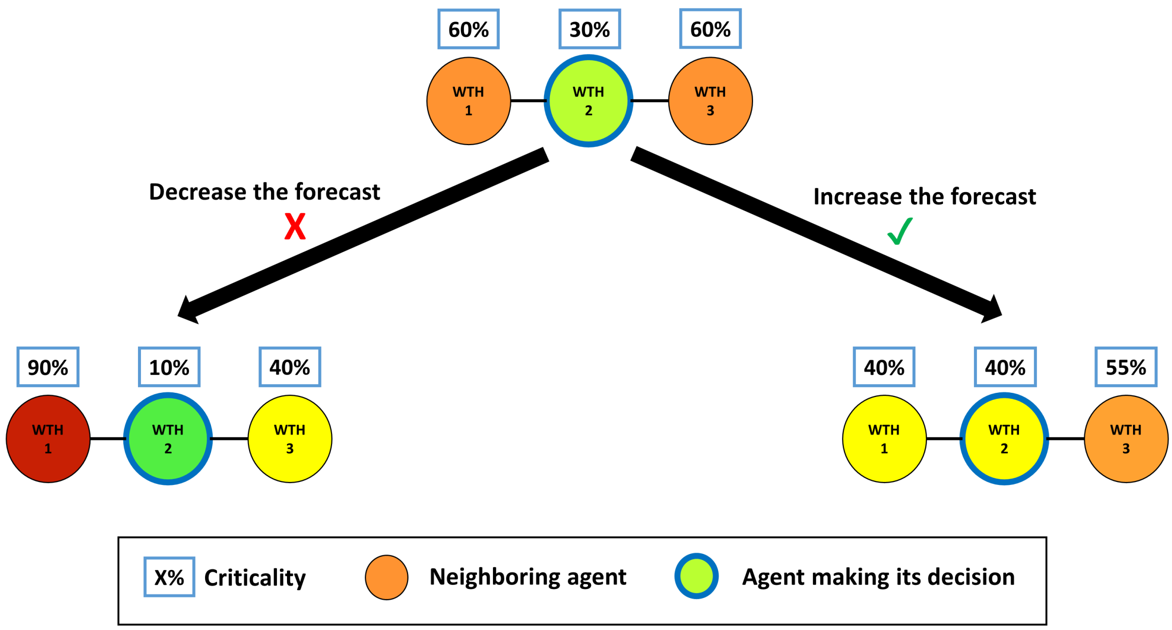

Figure 4.4 — Example of cooperative behavior of a WTH agent.

The nominal behavior of an agent is summed up in Algorithm 1. This behavior is only valid for one date, each WTH agent being assigned to one hour of that day. In order to make a forecast on a new day (in the case of a forecast over a period of several days for example), each WTH agent is reset with a new criticality function that will depend on the weather forecast for that day.

Algorithm 1 Nominal behavior of a WTH agent for a specific date

repeat
  Perceive: Store its own forecast and criticality and those of its neighbors
  Decide: Compute the criticality of each possible action (increase or decrease the forecast) and decide cooperatively the action that minimizes the highest criticality of its neighbors and its own
  Act: Perform the decided action and inform its neighbors of its new criticality
until Global criticality convergence *or* limit number of cycles exceeded

4.4.4 Non Cooperative Situations

WTH agents have a nominal behavior detailed in Algorithm 1. In case of Non Cooperative Situation (NCS), they may deviate from this nominal behavior to resolve these situations.

4.4.4.1 Uselessness

Problem description — The agent possible actions (i.e. reducing or increasing its forecast) worsen the situation in its neighborhood (see figure 4.5).

Type — *Uselessness*: it means that the agent believes that its action cannot change the world state or it believes that the results for its action are not interesting for the other agents.

Detection — This situation is detected during the agent decision phase. When the agent calculates its criticality and those of its neighbors by considering that it raises or lowers its forecast, it considers that any action will worsen the highest criticality.

Resolution — The agent does not take any action to avoid worsening its criticality and that of neighboring agents. As long as none of its actions can reduce the worst criticality, it remains in this state. This action does not further constrain neighboring agents.

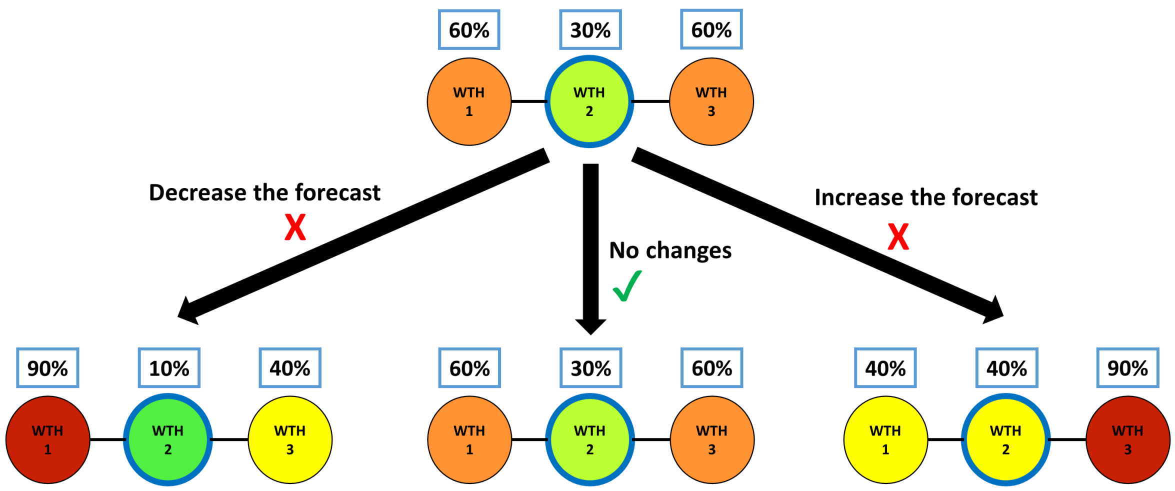

Figure 4.5 — Example of a non cooperative situation of type *Uselessness*.

4.4.4.2 Incompetence

Problem description — The information perceived by the agent does not allow it to decide on the action to be taken (see figure 4.6).

Type — *Incompetence*: it means that the agent has no skill to treat the information it has perceived.

Detection — The situation is detected when the maximum criticalities obtained by considering a decrease, increase or keeping of the forecast are equal. Whatever the agent decision, it will have no impact on its own criticality and that of its neighbors.

Resolution — Whatever the agent action, it will not worsen the agent local situation. However, a random choice would risk causing criticality degradation to agents not directly in the neighborhood. Consequently, the agent does not change its forecast as long as the situation is detected, waiting for a neighbor to unblock the situation.

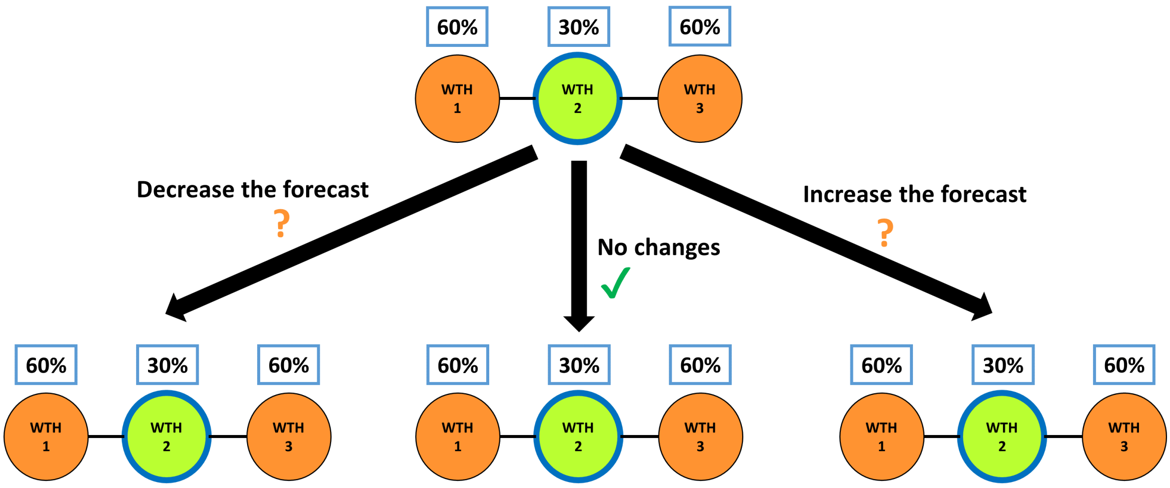

Figure 4.6 — Example of a non cooperative situation of type *Incompetence*.

4.5 Criticality implementation

The behavior of a cooperative agent is guided by the value of its criticality which dictates its choices (as well as those of its neighbors). It must therefore best reflect the deviation from the agent local goal because a wrong definition of criticality will lead to a wrong local solution and therefore a wrong global resolution of the problem. The local goal of a WTH agent is to adjust a forecast for a specific wind turbine and time so that it is consistent with the wind turbine production history, but also with the production history of neighboring wind turbines. When an agent considers that it has achieved or is approaching this goal, i.e. that it has achieved the right forecast, its criticality will be at a minimum.

The criticality function corresponds to a WTH agent for a particular date, it is calculated at the initialization of the agent, it depends on the agent forecast and that of its neighbors. As an agent must be consistent with both its own history and that of its neighbors, we have decided to divide the criticality into two subcriticalities: *local* and *neighboring* criticality. The two subcriticalities are obtained from a production probability density forecast, obtained by a model trained on the weather forecast and production data history. The principle of using a predicted probability density of the wind power as the basis for the criticality function is relevant because it enables to delineate confidence intervals of production value in which the agent can modify its forecast (e.g., to help a more critical neighboring agent). The predicted percentiles are then transformed via a generic criticality function into two subcriticalities functions. The final criticality function corresponds to the maximum of these two functions for all values.

This section presents the construction of the criticality function through the calculation of the percentiles, the application of the generic criticality function and finally the decomposition into two subcriticalities.

4.5.1 Percentiles computation

Quantiles are cut points dividing the range of a probability distribution into continuous intervals with equal probabilities. The quantile function Q is defined by the equation:

$$Q(p) \ = \ \inf\{x \in \mathbb{R} : p \leq F_X(x)\}$$

where F_X is the cumulative distribution function defined below:

$$F_X(x) = P(X \leq x)$$

When the range is divided into one hundred equal-sized groups, the quantiles are called percentiles. The i^{th} percentile is written P_i. There are 99 of them, from the 1^{st} percentile to the 99^{th} percentile. For example, if the 25^{th} percentile is equal to 300 kW, the prediction algorithm estimates that there is a 25% probability that the production is less than 300 kW and therefore a 75% probability that it is greater. We have chosen to use percentiles in order to obtain a better precision on the criticality function (compared to deciles for example). The percentile function P_i for the ith percentile corresponds to:

$$P_i = Q(i \div 100)$$

By combining two percentiles, a prediction interval I_j is obtained if the difference between the percentiles is equal to j. For example, I_{10} corresponds to the pair $\{P_{45}, P_{55}\}$ and I_{60} to the pair $\{P_{20}, P_{80}\}$, there is respectively a 10% and 60% probability of being inside.

In order to obtain the quantiles, a Gradient Boosting Model (GBM) [111] is used. This method was already applied to probabilistic wind power forecasting with success in [172]. As seen in section 2.4.2.4, it is an ensemble model composed of a set of weak prediction models, typically decision trees. It is a very popular algorithm in many areas because it is robust to outliers and shows good results without having to over-tune the model. It is mainly used to solve classification or regression problems (single value forecast) but can also be used to make probabilistic forecasts by independently predicting all quantiles. Indeed, the Friedman's gradient calculation (see figure 4.7), used in GBM, allows for the use of arbitrary loss functions for adjusting the targets of each consecutive learner (steps 1 and 3), without requiring a change to the regression model itself (step 2). It is therefore well-suited for probabilistic predictions.

Separate GBM models were used to produce each of the 99 percentiles. Each percentile is calculated independently, this allows the parallelization of model training but yields to 99 independent predictions. This can produce sets of predictions where a lower percentile received a higher prediction (e.g., if the 75^{th} percentile corresponds to a higher production value than the 74^{th} percentile). Before using these percentiles, sorting them provides a consistent probability distribution.

Although decision trees are typically bounded by the target range of the input, which is between 0 and the nominal power, the additive nature of gradient boosted machines can occasionally yield results that are outside the input range. To improve the accuracy, any

Initialize $\hat{f}(\mathbf{x})$ as a constant, $\hat{f}(\mathbf{x}) = \arg\min_\rho \sum_{i=1}^N \Psi(y_i, \rho)$.
For t in $1, \ldots, T$,

1. compute the negative gradient as the working response

$$z_i = -\frac{\partial}{\partial f(\mathbf{x}_i)} \Psi(y_i, f(\mathbf{x}_i)) \Big|_{f(\mathbf{x}_i)=\hat{f}(\mathbf{x}_i)}. \tag{1}$$

2. fit a regression model $g(\mathbf{x})$ for predicting z_i from the covariates $\mathbf{x}_i$.
3. choose a gradient descent step size as

$$\rho = \arg\min_\rho \sum_{i=1}^N \Psi(y_i, \hat{f}(\mathbf{x}_i) + \rho g(\mathbf{x}_i)). \tag{2}$$

4. update the estimate of $f(\mathbf{x})$ as

$$\hat{f}(\mathbf{x}) \leftarrow \hat{f}(\mathbf{x}) + \rho g(\mathbf{x}). \tag{3}$$

Figure 4.7 — Friedman's gradient boosting algorithm [173].

results outside those bounds were replaced with the corresponding boundary value.

Figure 4.8 shows an example of a wind power probabilistic divided into percentiles forecast on our data. Each band represents an interval from I_2 to I_{98} (respectively dark purple to light pink) and the dark curve is the median P_{50}. The red dots represent the real production values.

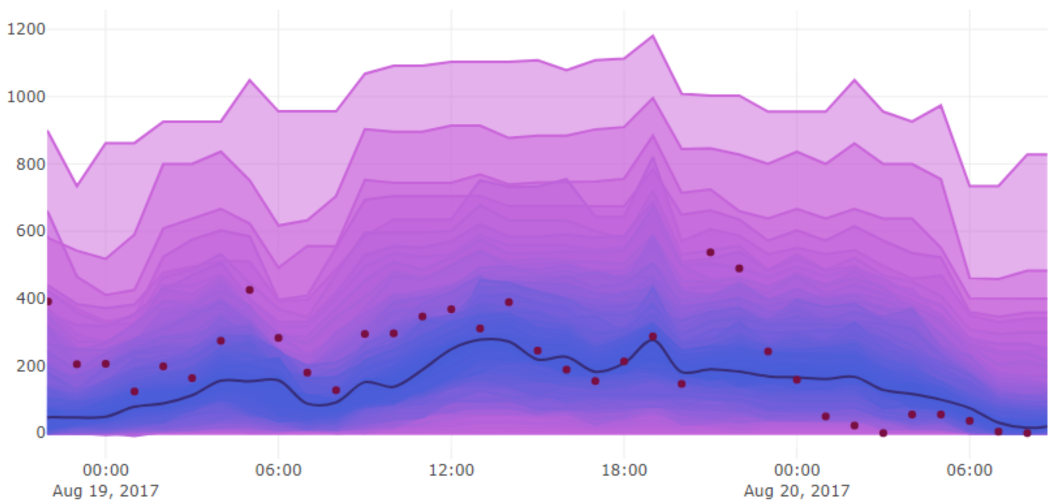

Figure 4.8 — Example of probabilistic wind power forecast divided into percentiles.

4.5.2 Generic criticality function

We want the criticality of the agent to be directly related to the probability interval to which the agent current forecast belongs. The criticality is low for a forecast close to the median and increases as it moves away from the median. However, the final criticality function being the maximum of several subcriticality functions, the minimum criticality will not necessarily correspond to one of the medians.

Then, the criticality function is built according to the equation:

$$\text{Criticality} = \{P_i \mapsto |100 - 2i|\} \tag{4.1}$$

$$\text{where } i \in \mathbb{N} \mid 0 < i < 50 \text{ and } i \neq 25$$

Intuitively, the construction of this function corresponds to the selection of a single horizontal slice from figure 4.8, the inversion of the axes and the application of a low criticality value to the dark parts and a high value to the light parts.

The median P_{50} is removed to obtain a minimum over an interval of values instead of a single value. This leads to greater flexibility in the action of the agents (e.g., to temporarily deviate from the minimum in order to help a neighboring agent). In addition, for a better interpretation of the criticality, the value is also re-scaled between 0 and 100 (instead of being between 2 and 98). Finally, the missing values are calculated by a linear interpolation.

4.5.3 Subcriticalities computation

The quality of the forecast made by a WTH agent depends on the consistency of its forecast with both its own past productions and the neighboring agents forecast. The criticality is then expressed by considering these two factors and combines two kinds of subcriticalities:

▷ **Local criticality**

The forecast made by an agent has to be consistent with the productions observed with a similar weather situation in its history (e.g., the wind turbine rarely produces energy when the weather model forecasts a very light wind). The corresponding criticality function is then built based on equation (4.1). Figure 4.9 shows an example of the local criticality function of a wind turbine at a given date. The x-axis corresponds to the value of the forecast made by the agent and the y-axis corresponds to its local criticality. In this example, the least critical situation is obtained when the forecast is equal to about 65% of the maximum power.

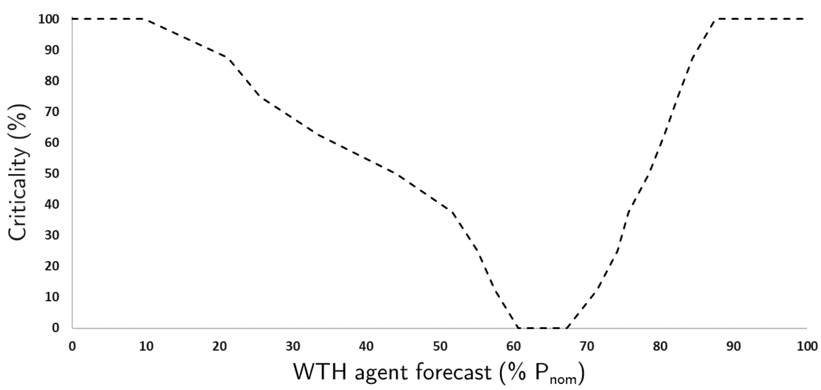

Figure 4.9 — Local criticality function example.

▷ **Neighboring criticality**

The forecast has also to be consistent with the neighboring agents forecasts (e.g., if a wind turbine always produces more amount of power than its neighbor, this constraint has to be taken into account). Indeed, a turbine slows down the wind behind it due to the wake effect and thus interferes with its neighbors. Therefore, a difference appearing in past observations between two wind turbines will be included in the resolution through this criticality. The neighboring criticality is built in the same way as the local criticality, except that the forecasting model learns from the history of production differences between two wind turbines (instead of the production history of a single turbine). Figure 4.10 shows probabilistic forecasting applied to the difference in production between two nearby wind turbines. The forecasts are centered on 0 kW because the productions are close.

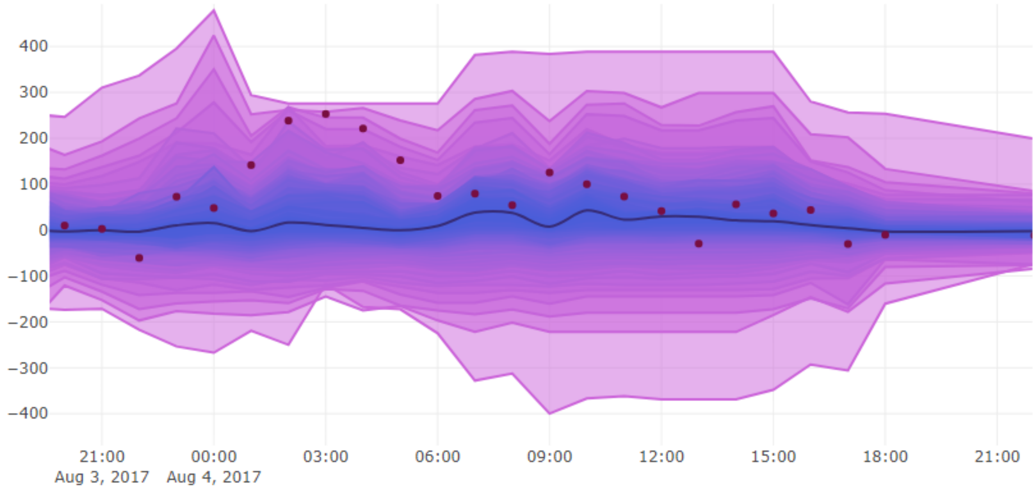

Figure 4.10 — Example of probabilistic wind power forecast difference between two wind turbines divided into percentiles.

For one agent, one criticality function is built for each neighbor of the agent based on equation (4.1). Figure 4.11 shows an example of the neighboring criticality functions of an agent with two neighbors at a given date. The x-axis corresponds to the value of the forecast made by the agent and the y-axis corresponds to its neighboring criticalities. In this example, with the two neighbors, the least critical situations are obtained if the agent has a slightly lower forecast than them.

4.5.4 Final criticality

The final criticality of a WTH agent corresponds to the maximum between its local criticality and each neighboring criticality. This choice enables not to give an advantage to one criticality over the other, they are considered equivalent.

Figure 4.12 shows an example of a final criticality function as a heat map. For a given WTH agent, the x-axis corresponds to its forecast, the y-axis corresponds to the difference between its forecast and that of its neighbor and finally the color corresponds to the criticality (yellow for very critical and dark blue for not critical). In this example, the WTH agent

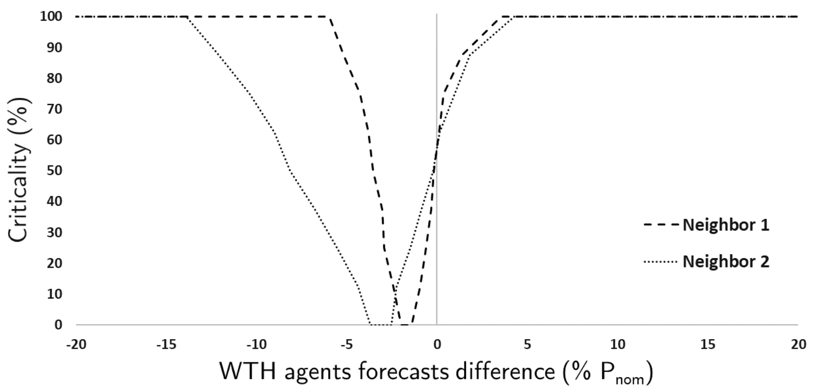

Figure 4.11 — Neighboring criticality functions example.

whose criticality we calculate has only one neighbor (a wind turbine located at one end of the wind farm) to simplify the visualization. Otherwise, an additional dimension should have been added to the graph for each neighbor. The graph is deliberately zoomed to the global minimum of the function but the search space is larger. Indeed the forecast can vary between 0 and the nominal power P_{nom} (which is equal to 2050 kW in this example) while the forecast difference can possibly vary between $-P_{nom}$ and P_{nom}.

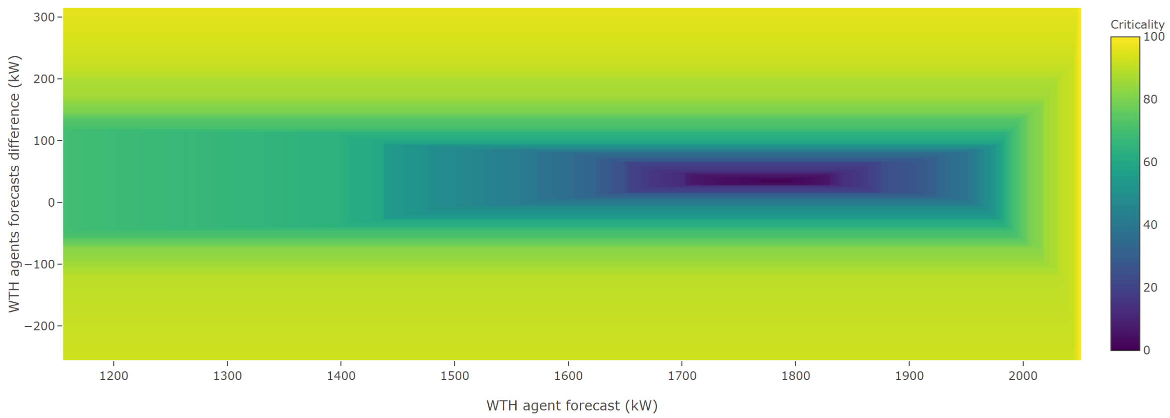

Figure 4.12 — Example of a final criticality function for a WTH agent with a single neighbor.

4.6 Technical choices

Since a Multi-Agent System is a computerized system, the transition from theory to practice includes specific technical issues that we will solve in this section. This section presents the software architecture, the agent initialization, the agent scheduling and the system stop condition.

4.6.1 Software architecture

We have developed a structure largely inspired by AMAK [174], a framework for developing Adaptive Multi-Agent Systems (AMAS). It provides a basic foundation and allows us to focus essentially on the behavior of agents. The three main abstract classes, represented in figure 4.13, are the following:

- ▷ **Amas**: The main class of the system that manages the scheduling of agents (discussed in section 4.6.3).
- ▷ **Agent**: Each agent inherits from this class. In order to follow the Perception-Decision-Action life cycle, the abstract class Agent mainly contains three overridable methods: *perceive*, *decide* and *act*. It also contains a *compute_criticality* method necessary to calculate the criticality of the agent.
- ▷ **Environment**: The class Environment is abstract and must be extended by the environment of the AMAS. The extended class must provide direct access to any information the agents may require. Each agent belonging to the AMAS has a pointer to this environment and uses it to perceive its part of the environment.

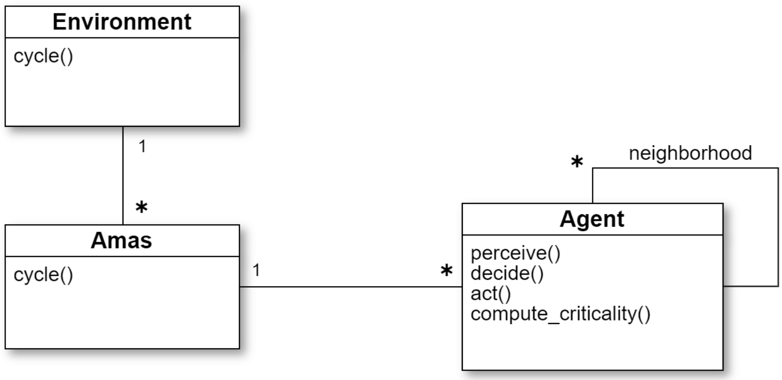

Figure 4.13 — Global class diagram based on AMAK.

Figure 4.14 represents the AMAWind-Turbine class diagram to understand the relationships between entities and agents. In AMAWind-Turbine, only Wind Turbine Hour (WTH) agents inherit the Agent class. As described above, for each Wind Turbine (WT) and Grid Point (GP) entity correspond respectively 24 WTH agents and 24 Grid Point Hour (GPH) entities. Each WT entity is connected to a single GP entity, but each GP entity can correspond to several different WT entities (and similarly for WTH agents and GPH entities).

The system was developed in Python because many libraries developed in this language are highly popular in the field of data science [175]. This simplifies comparison between models and allows the use of practical data processing tools. In addition, the company **SWIFT* which co-supervised this thesis works mainly in Python. This language was also chosen to facilitate the integration of the work with the company tools.

4.6.2 Agent initialization

When the agent is initialized, an initial forecast and an increment value are set. This section discusses the choices of these values to allow the system to work properly.

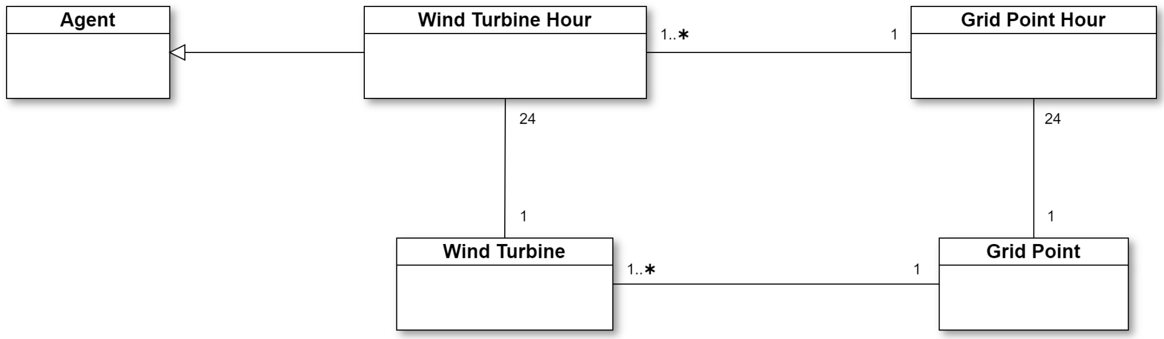

Figure 4.14 — Class diagram representing the links between entities and agents in AMAWind-Turbine.

4.6.2.1 Choice of initial forecasts

As described in section 4.4.3, each agent starts with an initial forecast that it can modify at each cycle. In chapter 7, the evaluation part will test the three following possible choices for the initial forecast of an agent:

- The average production value of the wind turbine.
- A random value within the possible production interval $[0, P_{nom}]$ of the wind turbine.
- A "consistent" value predicted by a machine learning algorithm.

The main criterion we want to verify is the impact on computation time which may be different because an initial forecast far from the final forecast will require a larger number of cycles from the system.

4.6.2.2 Increment choice

As seen in section 4.4.3, at each cycle, a WTH agent decides to raise or lower its forecast by a fixed value, it may also not change it in case of NCS.

This increment has been set to 1 kW. This value has been chosen arbitrarily, the lower it is and the more accurate the results can be. However, the smaller the value, the larger the number of cycles required to converge, so a compromise must be found between computation time and accuracy. Since the maximum power of a wind turbine is several MW and the mean absolute error of forecast is about 10%, the forecast will not be degraded by an increment value representing less than 0.1% of the maximal production value. In addition, we will see in chapter 7 that the computation times are acceptable when the increment value is equal to 1 kW.

4.6.3 Agent scheduling and stop condition

For each date, the agents are initialized with a certain forecast value and a criticality function, and then perform actions at each cycle. At each cycle the agents are run sequentially in a random order to avoid biasing the simulation by a predetermined order. The forecast is carried out independently for each date, so the calculation can be run in parallel.

The whole system can converge towards a final solution, i.e. all agents consider that their actions will worsen the highest criticality between their own and their neighbors. In this case, this forecast is chosen as the final solution. However, if the system diverges and does not find a solution, the simulation is arbitrarily stopped at a predetermined number of cycles. This value is chosen experimentally, it must be high enough not to stop a system that was going to converge but not too high to slow down the entire forecast because of a blocking situation.

4.7 Conclusion

In this chapter, we have presented the AMAWind-Turbine system. It is an adaptive multi-agent system for predicting the production of a wind farm in a way that is consistent with the wind turbines that compose it.

In AMAWind-Turbine, cooperation between agents is guided by the criticality value which is defined by a function built at the agent initialization. The final forecast is the result of the cooperative local behavior of Wind Turbine Hour (WTH) agents. Unlike other forecasting methods, our objective is at the macro-level (the wind farm) while the system resolution is at the micro-level (the wind turbines).

Returning to the criteria presented in section 2.6.1 allows to handle dynamics and is dependent on the presence and the quality of data. The reason is that a part of the learning is included in the construction of the criticality function and is performed by a statistical method, therefore the evaluation of these criteria is the same. Moreover, our method is simple to implement, it does not require specific knowledge in fluid mechanics, only data at the wind turbine scale. The advantage of AMAWind-Turbine is that the system takes into account the interdependencies between wind turbines and spatial correlations in the forecast. Concerning the execution time and forecast accuracy, it will be necessary to check these points experimentally. Therefore, the evaluation of the system on real data will be presented in chapter 7.

In the next chapter we will present the AMAWind-Farm system, a system similar to AMAWind-Turbine but designed for forecasting on several wind farms.

5 AMAWind-Farm: a MAS for Wind Power Forecasting at Farm-Level

In this chapter we present the second Adaptive Multi-Agent System designed in this thesis: AMAWind-Farm (**A**daptive **M**ulti-**A**gent system for **Wind** power forecasting at **Farm**-level). It is dedicated to forecasting the production of a region (i.e. several wind farms) by taking into account the large-scale interdependencies between wind farms. AMAWind-Farm has an architecture and operation similar to AMAWind-Turbine. This shorter chapter will therefore make many references to the previous one while focusing on the differences between the two systems.

5.1 Objectives . . . 91
5.2 Entities . . . 92
5.2.1 Grid Point and Grid Point Hour entities . . . 93
5.2.2 Farm . . . 93
5.3 A cooperative agent: Farm Hour agent . . . 94
5.3.1 Description . . . 94
5.3.2 Local goal . . . 95
5.3.3 Behavior . . . 95
5.4 Implementation differences . . . 95
5.4.1 A criticality adapted to heterogeneous data . . . 96
5.4.2 A variable increment . . . 97
5.5 Conclusion . . . 97

5.1 Objectives

While between two wind turbines there may be a wake effect (i.e. the turbulence caused by the blades of the turbine reduces the production of neighboring wind turbines), this is not a priori the case between wind farms because the distances are too large. However, as mentioned in section 3.1.2, the production of two wind farms, even distant ones, can be correlated if they are located in similar wind zones.

Therefore, the objective of AMAWind-Farm is to forecast the production of a region, i.e. several wind farms, by integrating the interdependencies between wind farms. As with AMAWind-Turbine, the objective related to the field of application is also to reduce the forecast error compared to standard methods.

5.2 Entities

In order to determine the different entities of the system and their links, we follow the same approach as in section 4.3. A "region" consists of a set of wind farms for which their characteristics (mainly their geographical position and maximum power) and their production history are known. We also have a weather forecast history corresponding to certain grid points of the numerical weather prediction model.

This statement reveals two entities corresponding respectively to a wind farm and a grid point. Similarly to AMAWind-Turbine, we can deduce two passive entities that will only serve as databases: the Farm (F) entity and the Grid Point (GP) entity.

Then, we deduce two active entities in charge of providing the live forecasts: the Farm Hour (FH) entity and the Grid Point Hour (GPH) entity. Each FH and GPH is responsible for one hour for a specific date, chosen during the initialization of the entities.

In the same way that in AMAWind-Turbine, FH entities are connected to F entities representing the same farm while GPH entities are connected to GP entities representing the same grid point. F entities are connected with other F entities and GP entities, this relationship is also transmitted to the 24 related FH entities and GPH entities. Figure 5.1 illustrates an example in which F #1 is related to F #2 and GP #1. In this case, all FH entities connected to F #1 will be related to the FH entities related to F #2 corresponding at the same hour (e.g., FH #1 and FH #2 at hour 1 are connected) as well as the GPH entities related to GP #1 corresponding at the same hour.

While a GPH entity has no possible action on the weather forecast (it is provided by an external numerical weather prediction model), an FH entity can modify its production forecast. Indeed, since this latter is precisely the desired output of the system, the FH must ensure that this value is consistent with its neighborhood (i.e. the history of F, the GPH forecast and other FH entities). The FH entity is therefore an agent, and it is by modifying its forecast according to its perceptions that the system will achieve its objective.

To summarize, the analysis stage of the ADELFE methodology has led us to identify several types of entities:

- ▷ Passive entity: *Grid Point* (GP) entity.
- ▷ Active entity: *Grid Point Hour* (GPH) entity.
- ▷ Passive entity: *Farm* (F) entity.

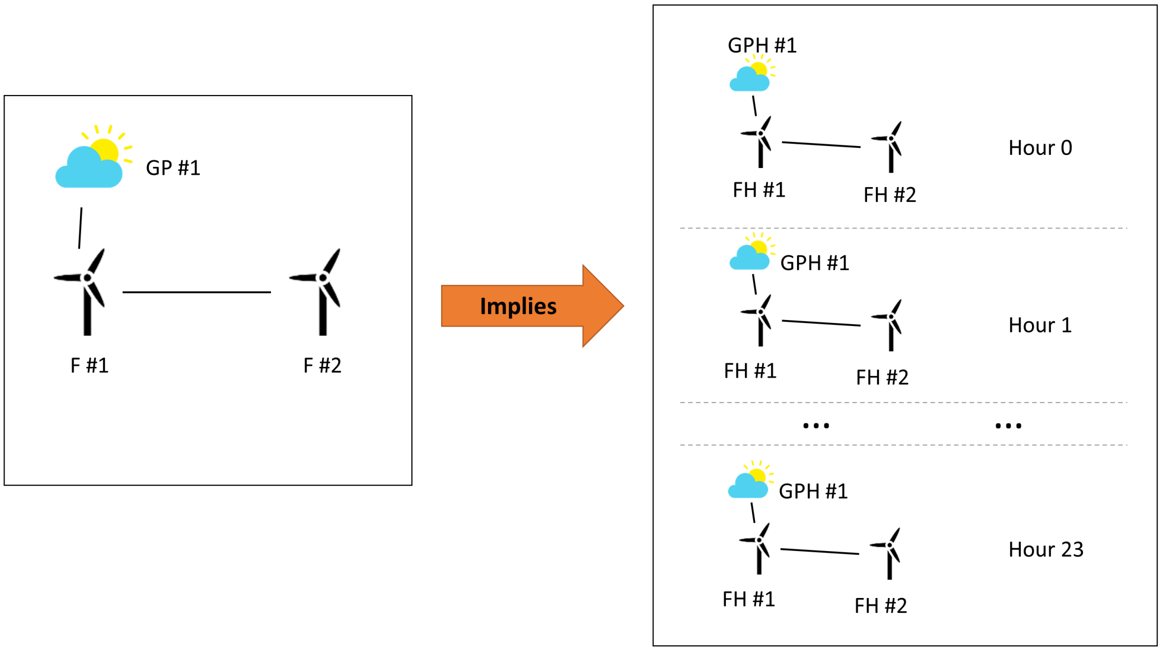

Figure 5.1 — Links between F/GP and FH/GPH entities.

▷ Cooperative agent: *Farm Hour* (FH) agent.

The resulting architecture of the AMAWind-Farm system is given in figure 5.2 and the features and relationships of these entities are detailed below. The F and GP entities are also linked to each other (in the same way as FH agents and GPH entities) but the links are not displayed so as not to overload the graph.

5.2.1 Grid Point and Grid Point Hour entities

The Grid Point (GP) and Grid Point Hour (GPH) entities have the same characteristics and roles as in the AMAWind-Turbine system (presented in section 4.3.1 and section 4.3.2 respectively). For each farm, the forecasts from the most central grid point are used.

5.2.2 Farm

A farm can be seen as a set of wind turbines which themselves have a particular location and characteristics. In this problem, however, we consider that we do not have the data for each wind turbine and that a wind farm is a single entity. Therefore, the farm has a total power, a total power curve and an approximate location (which represents the center of the farm).

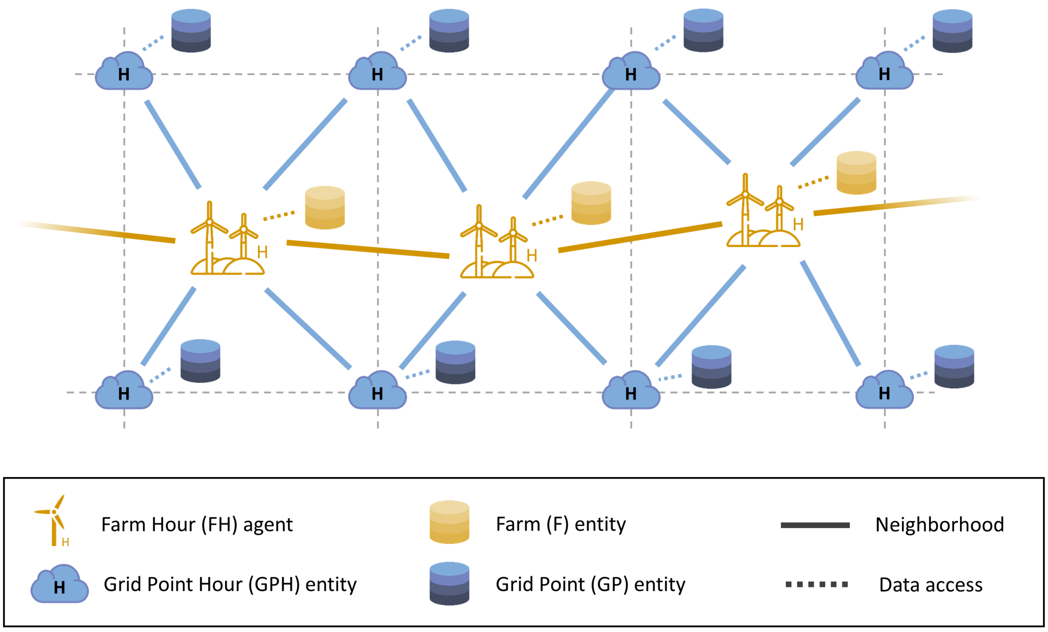

Figure 5.2 — Architecture of the AMAWind-Farm system and relationships between entities and agents.

5.3 A cooperative agent: Farm Hour agent

5.3.1 Description

A Farm Hour (FH) agent is responsible of the forecast of a farm production at a given hour (e.g., an agent in charge of forecasting the production of farm A at 08:00 a.m.). As for Grid Point Hour entities, there are 24 FH per farm.

The neighborhood of an FH agent could be based on physical proximity, but the distance between farms is often so large that it does not necessarily mean a high correlation criterion. In section 3.1.2, we observed this with an example between two nearby farms with less correlated production than the others. We have decided to link all the farms together because we work with a small number of farms.

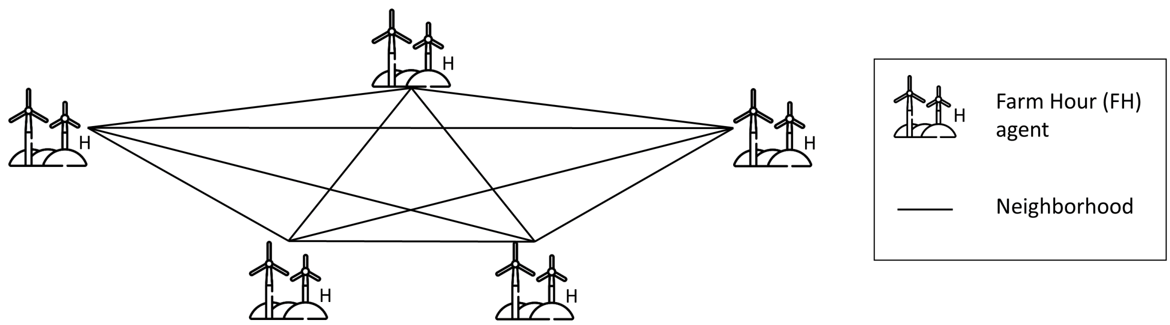

Figure 5.3 — Neighborhood of Farm Hour (FH) agents.

5.3.2 Local goal

The local goal of an FH agent is to adjust a forecast related to a wind farm at a specific hour. The forecast must be consistent with the wind farm production history but also with the production history of the neighboring wind farms. Therefore, the consistency of the forecast of an agent corresponds to some constraints on the production and on the difference in production with the neighboring agents.

Similarly to AMAWind-Turbine, this local goal is also represented by a criticality expressed by the combination of two subcriticalities: *Local* and *Neighboring* criticalities.

5.3.3 Behavior

Since an FH agent is a cooperative agent, its purpose is to determine the forecast that minimizes its criticality while avoiding aggravating the criticality of its neighbors. FH agents follow the same Perception-Decision-Action lifecycle as WTH agents presented in section 4.4.3:

- ▷ *Perception* — The FH agent perceives the current forecast and criticality of its neighbors. It also knows its own forecast and criticality.
- ▷ *Decision* — The agent decides how it will change its forecast. These changes are decided on the basis of previously perceived information. The agent has two possibilities: to increase or to decrease its forecast. The choice of the increment is discussed in section 5.4.2. The agent simulates these different cases and performs the action minimizing the maximal criticality of its neighbors and itself.
- ▷ *Action* — The FH agent finally performs the decided action and therefore possibly changes its forecast. Its criticality is consequently updated at the end of the cycle.

Compared to the AMAWind system, the Non Cooperative Situations (NCS) presented in section 4.4.4 are also present in this system. The two possible NCS are:

- ▷ *Uselessness*: The FH agent can only make the situation worse (i.e. increase the maximum criticality).
- ▷ *Incompetence*: The FH agent cannot decide on the best action to take.

5.4 Implementation differences

The implementation of AMAWind-Farm follows the same principles as AMAWind-Turbine. In particular, the software architectures are similar, as shown in figure 5.4. However, there are two specific points of divergence detailed hereafter: the criticality normalization and the increment variability.

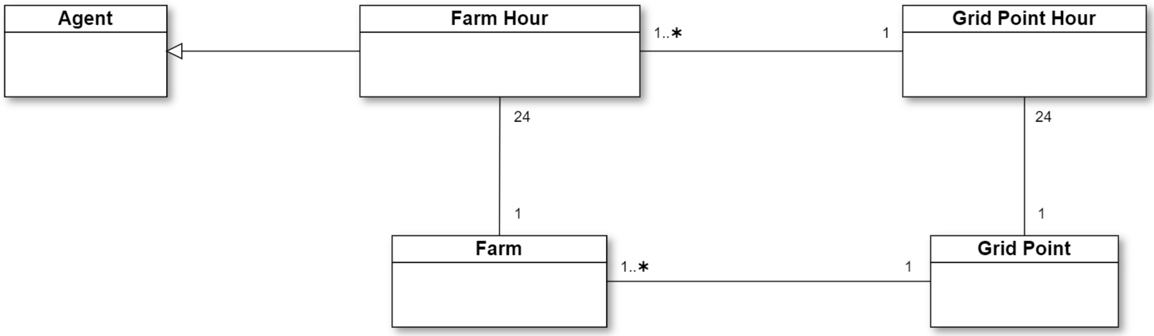

Figure 5.4 — Class diagram representing the links between entities and agents in AMAWind-Farm.

5.4.1 A criticality adapted to heterogeneous data

In the AMAWind-Turbine system, for a given wind farm, all wind turbines have very similar data. The reason is that they correspond to the same type of turbine (in our data and in most cases), have the same nominal power, are controlled in the same way and are positioned in a limited area.

For wind farms, each farm is a priori different (in size, location and use). With equivalent criticalities, agents would have an unbalanced relationship between them. Indeed, a large wind farm will have more weight in the final forecasts than a smaller wind farm. For example, an extreme case would be to predict a region in which there is a large 100 MW wind farm and another wind farm of a few kW. With such a difference in scale, the smallest farm will have a small impact on solving the problem and the agents bound to that farm will follow the behavior of those bound to the largest farm. In this case, it seems consistent not to consider the two wind farms on the same scale in order to encourage cooperative behavior by agents.

In order to penalize or favor certain agents, we decide to normalize the criticality function by applying a multiplicative factor. This factor is fixed at the initialization of the agents and depends on the nominal power of the different farms. For a farm with a nominal power P, the factor associated with its criticality φ will be:

$$\varphi = 1 - \frac{P - P_{\text{min}}}{P_{\text{max}}} \tag{5.1}$$

where P_{min} and P_{max} are the minimum and maximum nominal powers of all farms respectively. For the smallest farm $\varphi = 1$ and for the largest farm $\varphi = \frac{P_{\text{min}}}{P_{\text{max}}}$. For the others they will have a factor proportional to their nominal power and located between these two extreme values. Figure 5.5 shows the normalization of a criticality function by a factor $\varphi = 0.5$. The agent with this criticality will be less constrained, because it is less critical, than an agent with a higher criticality factor.

This factor is applied directly to all values of the generic criticality function presented in section 4.5.2. However, since the final criticality is the combination of two subcriticalities, we have chosen to test different possibilities and apply the normalization on:

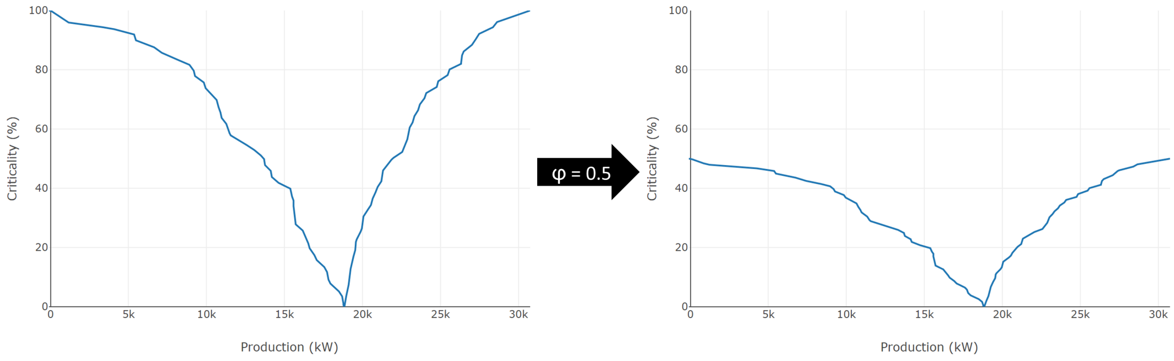

Figure 5.5 — Example of the normalization of a criticality function with $\varphi = 0.5$.

- ▹ none of the subcriticalities.
- ▹ only local subcriticality.
- ▹ only neighboring subcriticality.
- ▹ both subcriticalities.

These different scenarios will be tested in the system evaluation in section 8.2.2.

5.4.2 A variable increment

At each cycle, an FH agent decides to raise or lower its forecast by a value, it may also not change it. In AMAWind-Turbine, this increment value is fixed. Although the farms have different nominal powers, this does not change the resolution because the farms are not related.

In AMAWind-Farm, the farms solve the problem together. A fixed increment value for all farms would disadvantage the larger farms that would require more cycles to achieve their objective. This would allow a smaller farm to reach its goal faster and at an advantage.

For a given Farm entity F, we therefore chose an increment value $i_F = \frac{\mathrm{P_{nom}}(F)}{\mathrm{N_{rate}}}$ depending on the nominal power $\mathrm{P_{nom}}(F)$ of the farm and a fixed rate $\mathrm{N_{rate}}$ chosen experimentally. The higher the nominal power of the farm, the higher the increment value will also be. For example, if $\mathrm{N_{rate}} = 1000$ and $\mathrm{P_{nom(\mathit{F})}} = 2000$ kW, at each cycle the forecast of an FH agent will be lowered or increased by $i = \frac{2000}{1000} = 2$ kW. The value of $\mathrm{N_{rate}}$ will be discussed in the chapter dedicated to the evaluation of the system in section 8.1.

5.5 Conclusion

In this chapter, we have presented the AMAWind-Farm system. It is an adaptive multi-agent system for predicting the production of a region (i.e. several wind farms) in a way that is consistent with the wind farms that compose it.

AMAWind-Farm works very similarly to AMAWind-Turbine because it also uses interdependencies between entities in a forecast model. The architecture of both system as well as the behavior of the agents are equivalent, there is only a difference of scale. While AMAWind-Turbine aimed to forecast the production of an entire wind farm using the wind turbine data, in this case AMAWind-Farm aims to predict the production of an entire region using the wind farm data.

The main problem is the heterogeneity of the different agents. While the wind turbines were all the same for a given wind farm with very similar data and operation, the wind farms are all very different. This has led us to modify the criticality of agents in order to avoid unbalanced exchanges between agents.

Although this system has the same characteristics as AMAWind-Turbine, it is also necessary to evaluate it because the data used are different and we must also check the proper functioning of the new criticality function. Therefore, the evaluation of the system on real data will be presented in chapter 8.

Experimental Evaluation

6 Methodology of the Experiments

The last chapters presented the two systems developed to predict wind production. In order to validate and evaluate these models, we carried out experiments with data from real wind farms. This chapter describes the conditions for experimentation, i.e. the data used, the data preprocessing, the evaluation criteria and the validation method.

6.1 General description of the experiments 101
6.2 Data description 102
6.2.1 Production data 102
6.2.2 Weather forecasts 103
6.3 Data preprocessing 105
6.3.1 Filtering of errors 105
6.3.2 Filtering of maintenance or failure cases 105
6.4 Evaluation criteria 106
6.4.1 Criteria related to the system validation 106
6.4.2 Criteria related to wind power forecasting 107
6.5 Model validation by cross-validation 108

6.1 General description of the experiments

The experiments consist of replaying forecasts from data history (weather forecasts and wind production history) on several wind turbines and wind farms. In the evaluation phase, the system starts with initial forecasts that evolve through interactions between agents. Once the system has converged on solutions, the forecasts can be compared with historical data.

The experiments carried out during this thesis have two main purposes:

- ▷ The validation of the functioning of our systems and analysis of their global behavior. In order to do this, we will study the evolution of the local behavior of agents and the global behavior of the system. As we are using a new forecasting method, it is important to understand whether the system is working and how it is evolving in real conditions.

In our case, we want essentially to verify that a collective resolution leads to a decrease in global criticality.

▷ The assessment of the accuracy of our forecasts and comparisons with other state-of-the-art methods applied to the same data. As with any forecasting method, we want to ensure that it leads to at least more accurate forecasts than other conventional methods.

6.2 Data description

The experiments are based on production data as well as weather forecast data. This section describes the characteristics of these data.

6.2.1 Production data

This thesis being co-supervised by **SWIFT*, we were able to access real wind farm data as part of real wind power forecast requests. The data were provided by *Boralex*, a Canadian electricity producer specialized in renewable energy. We thank them for allowing us to use this data for our research.

The study was carried out on five wind farms in France, their geographical location is available in section 3.1.2. Table 6.1 specifies for each wind farm the code used in this thesis, the name, the wind turbine model, the number of wind turbines, the nominal power P_{nom} and the hub height of a wind turbine and the capacity factor C_f (the ratio between actual electrical energy production and maximum electrical energy production, which was computed on the available data).

Table 6.1 — Characteristics of the wind farms studied.

Code	**Name**	**Model**	**Size**	$\mathbf{P_{nom}}$	**Hub height**	$\mathbf{C_f}$
A	Vallée de l'Arce	Senvion MM92	15	2050 kW	100 m	25.93%
B	Chasse Marée	Enercon E82	4	2300 kW	78 m	27.22%
C	Plouguin	Enercon E66	4	2000 kW	85 m	28.40%
D	Avignonet 1	Nordex N50	10	800 kW	50 m	24.86%
E	Cham-Longe 1	GE GE1.5s	12	1500 kW	65 m	38.16%

It is important to compare forecasts under several conditions in order to have an overall view of the model accuracy and to avoid that a model only improves a particular case. Therefore, we chose these farms because they are located in different wind zones (e.g., on average, the production of farm E reaches its minimum at 01:00 p.m. and its maximum at 12:00 a.m., while for farm C it is the opposite), and on land with different topographies (e.g., farm E is located on a hill). The layouts of the wind turbines in the five wind farms, shown in figure 6.1, are also different. We also notice in figure 6.2 that the power curves of these turbines have different shapes, wind amplitudes and maximum powers. Finally, the types and the sizes of the wind turbines are not the same for each wind farm.

The distribution of the measured wind and production for one wind turbine in each farm is shown in figure 6.3. The x-axis represents the wind speed in m/s on the left and the

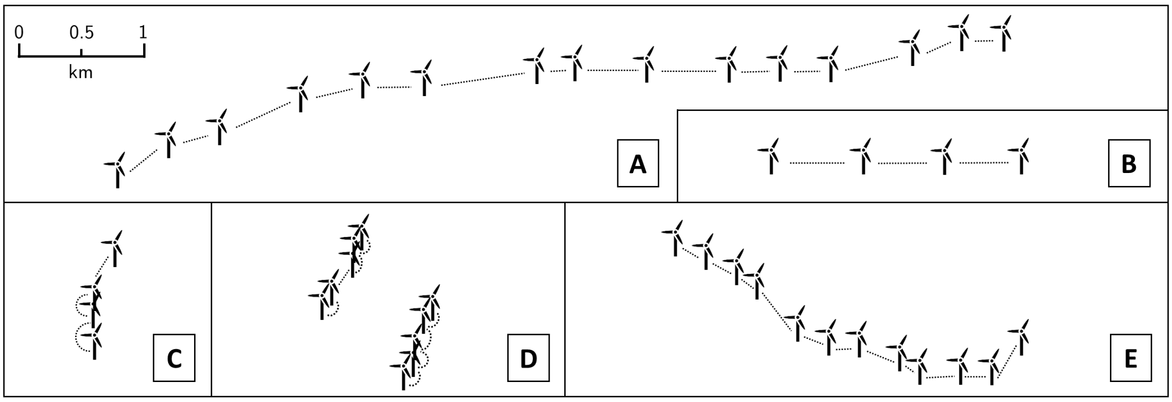

Figure 6.1 — The layout of the five studied wind farms at the same scale.

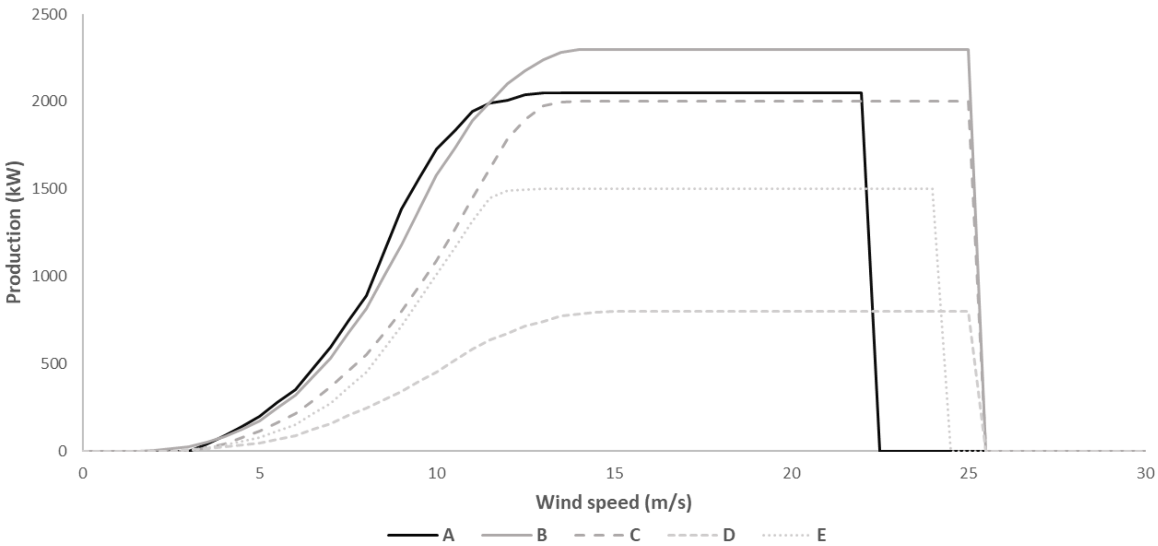

Figure 6.2 — Power curves of the wind turbines of the five wind farms studied.

production in kW on the right, while the y-axis represents the statistical distribution of data. We notice differences in distribution: for example the wind speeds at wind farm D are, on average, lower than at the wind farm C due to a lower hub height (wind speed increases with altitude). For the distribution of the production, the general shape is the same, marked by two peaks at minimum and maximum power. This is due to the shape of the power curve, with two plateaus at low and high wind (and possibly the third plateau after the cut-out but it is very rarely reached). In the production graphs, the y-axis is truncated for more visibility because there are many values approaching 0 kW.

6.2.2 Weather forecasts

Weather forecasts are provided by the Météo-France AROME [176] high-resolution forecast model for the entire next day (with time horizon from 21 to 45 hours). The grid points are separated by 0.025° of longitude and latitude. This corresponds to a resolution of approximately 2.7 km in North/South direction and 1.9 km in West/East direction in

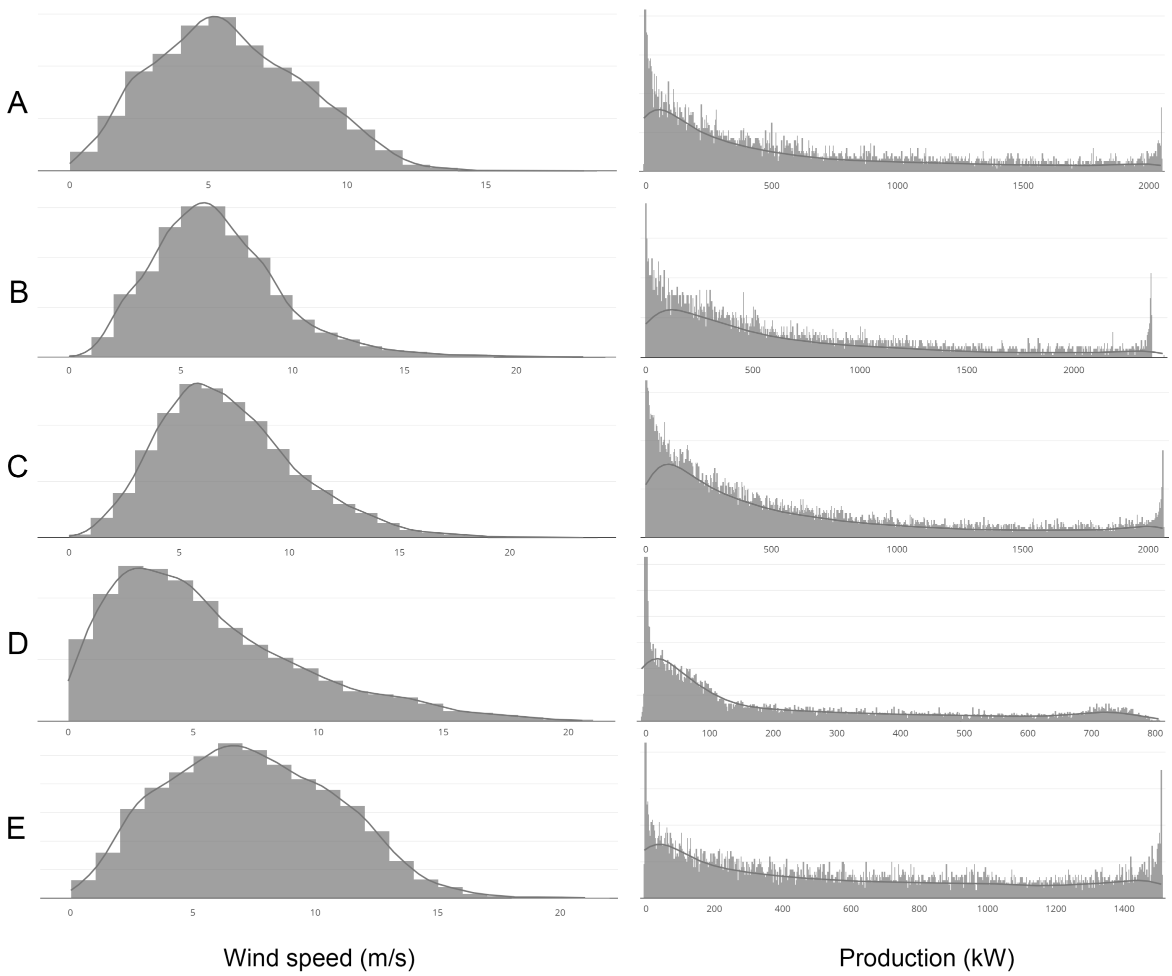

Figure 6.3 — Distribution of the measured wind and production of one wind turbine in each farm.

metropolitan France. The experiment covers a large period thanks to a nearly three-year history of wind power and weather forecasts from 01/2014 to 09/2017. However, as we do not have a complete history for all the farms, we only have about 15000 entries for farm A and about 20000 entries for farms B, C, D and E. These amounts are however large enough to cover a significant number of weather conditions. In terms of data size, the weather forecast and production history corresponds to about 300 MB, which is low and easy to handle.

For each wind turbine, the nearest available grid point is used. For a large wind farm, several points can therefore be used. In order to obtain the quantiles, the GBM is trained with the parameters at 100m related to equation 2.8 presented in section 2.1.2.4:

- W_s, the wind speed.
- W_d, the wind direction.
- T, the temperature.
- P, the pressure.
- R_h, the relative humidity.

The last three parameters appear indirectly in the theoretical formula (2.8) by the air density ρ. The wind direction W_d may indicate if the wind has been disturbed by a neighboring wind turbine.

Although the current weather forecast data are freely available on the website of Météo-France [177], the access to forecasting history was made possible through a partnership with the CNRM (French acronym for Centre National de Recherches Météorologiques, French Weather Research Center) on the METEOSWIFT project [1]. The data are provided in GRIB2 format (GRIdded Binary 2) allowing all forecasts for all grid points in a geographical area to be compressed into a single file. For ease of use, we selected grid points with the necessary parameters and we generated CSV (Comma-Separated Values) files.

6.3 Data preprocessing

The wind farms on which we are conducting our experiments are currently operational. Wind turbines are therefore potentially subjected to maintenance operations and breakdowns. There may also be errors in the values measured by the sensors. Although failure prediction is an existing field of research [178], it is outside the scope of this thesis. We have therefore decided to filter out outliers in the data history to have results that are not dependent on the quality of production history. This concerns both data errors and maintenance or failure cases.

6.3.1 Filtering of errors

We consider as erroneous the values that are not consistent with the rest of the history. For example, production values that are largely negative or much higher than the nominal power of the wind turbine are not physically possible.

There are also whole periods in the data when the production is constant. This is because the data were not recorded and the last known value was chosen. We therefore filter the productions if for a minimum period of 3 consecutive hours the production is stricly equal with an accuracy of 6 decimal places.

6.3.2 Filtering of maintenance or failure cases

The filtering of maintenance or failure cases is more subtle because these cases are often only detectable by looking at production data. Indeed, when a wind turbine does not produce energy when wind has been predicted, either the wind forecast was incorrect or the turbine was undergoing maintenance, curtailment or failure that must be filtered.

On the wind farms studied, we have data from wind sensors located on the hub. Although these data can also be subject to errors, they are a good approximation for filtering extreme cases. Thus, when the measured wind exceeds a certain threshold and the corresponding production is low or null, the record is filtered.

1. The METEOSWIFT project (2016-2018) was funded by the European Regional Development Fund of the European Union and the French Occitanie Region and involved *SWIFT, CNRM and IRIT.

Table 6.2 — Size of wind farm production history before and after filtering.

Farm	Raw	Error	Maintenance	Final
A	14903	-437	-1825	12641
B	14375	-842	-97	13436
C	30336	-4552	-148	25636
D	15270	-622	-974	13674
E	15048	-252	-3218	11578

Table 6.2 shows the number of deleted entries for each farm by each filter type. Filtering is common to all wind turbines in a wind farm, i.e. if for a given date the data are filtered for one wind turbine, they are filtered for all wind turbines in the farm. It is necessary in our case because the cooperative resolution method predicts the production of all wind turbines at the same time. Figure 6.4 shows the observed power curve of wind farm C by separating the final and filtered records. Most of the 4000 filtered inputs are located on the abscissa axis, i.e. when production is zero.

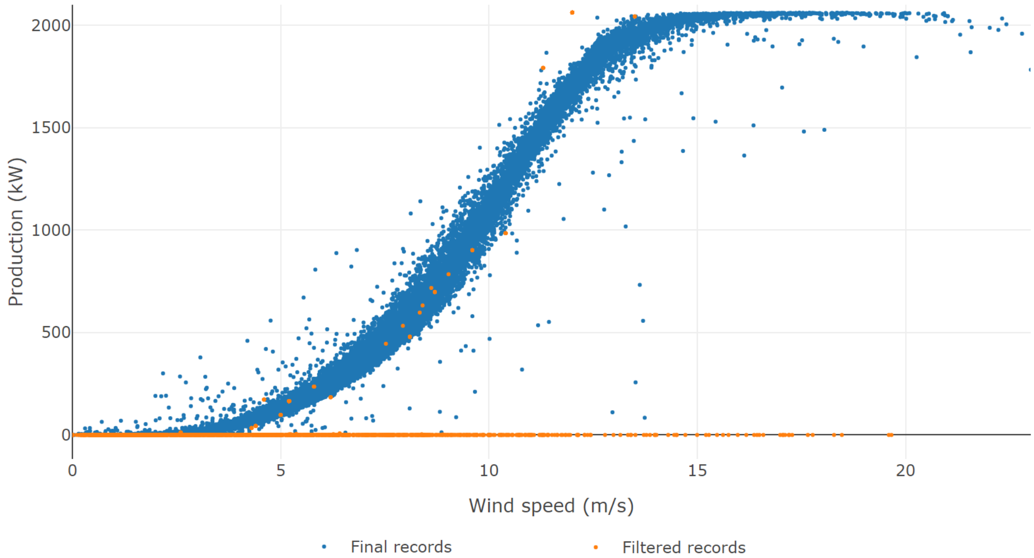

Figure 6.4 — Power curve of the wind farm C with final and filtered records.

6.4 Evaluation criteria

This section presents the evaluation criteria. They can be divided into two categories: those related to the multi-agent system validation and those related to the results (i.e. forecast accuracy).

6.4.1 Criteria related to the system validation

The first criteria are specific to the systems we have developed. They make it possible to validate the system and better understand how it works. These different criteria are as follows:

- ▷ The **computation time** taken by the system to provide the forecasts. Indeed, the system must be able to calculate forecasts without exceeding the time limits imposed. For our field of application, which is wind power forecasting for the day-ahead market, time constraints are rather low because we often have several hours to predict the production of only one day.
- ▷ The **global decrease in criticality** because it is the objective of our system to achieve a reduction in overall criticality through the cooperative behavior of agents.
- ▷ The **correlation between criticality improvement and error improvement**. This makes it possible to verify that the system is moving towards the application goal of our system, which is to reduce the forecast error.
- ▷ The **convergence rate**. As detailed in section 4.6.3, either the system converges to a final forecast or it diverges and stops after a certain number of cycles. A divergent case means that the system fails to achieve a minimum criticality and that the forecast performed is probably not good. Therefore, it is important to know the number of cycles required by the system to converge and to verify that there are no cases of non-convergence.

During this validation, we will also discuss some of the issues raised earlier:

- ▷ The **choice of the initial forecast** described in section 4.6.2.1.
- ▷ The **choice between the different normalization cases** for AMAWind-Farm described in section 5.4.1.
- ▷ The **evaluation of AMAWind-Farm with farms with independent (or slightly correlated) productions**.

6.4.2 Criteria related to wind power forecasting

In our field of application, when the forecast is used for the market, the real evaluation criterion should be the financial gain obtained by clients through the forecasts. However, this gain depends on the choices made by the client based on forecasts but also on the other players in the electricity market. Therefore, the best way to evaluate our forecast is to measure its accuracy against actual production.

To evaluate forecasts performance two standard measures were used: the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE). While for the MAE each error has the same weight in the final error, for the RMSE larger errors have a disproportionately large effect. For production forecasting in the context of electricity grid management, this makes sense because larger forecast errors are more difficult to compensate.

These metrics are given by the equations 6.1 and 6.2 where $\hat{y}_i$ is the forecast and y_i the observation.

$$\mathrm{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i| \tag{6.1}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{6.2}$$

Errors are normalized by the nominal power to obtain percentages, more convenient to interpret. This makes it easier to compare errors between wind farms. The new metrics are called NMAE and NRMSE (Normalized MAE and Normalized NRMSE). These metrics can also be calculated on subdivisions of the data by hour, season, year or horizon to reveal specific improvements or deteriorations.

6.5 Model validation by cross-validation

The model is validated by a *k*-fold cross-validation. The original sample is partitioned into *k* equal sized subsamples. Of the *k* subsamples, a single subsample is retained as the validation data for testing the model, and the remaining *k* - 1 subsamples are used as training data [179]. Although this method does not represent the real conditions (learning from more recent data than the validated ones), it enables to give an insight on how the model will generalize to an independent data set. This is a standard method to verify that the model does not overfit, i.e. it is too specialized on the data on which it learns and not general enough.

Ideally, *k* is chosen to be as large as possible (the extreme case, when *k* is equal to the number of inputs is called leave-one-out cross-validation). However, each learning can be very long and a compromise is necessary. In this thesis, we chose a 10-fold which corresponds to a validation set of approximately three months (only two months for the farm A, due to fewer entries). It is a good compromise between the calculation time required and the quality of the evaluation.

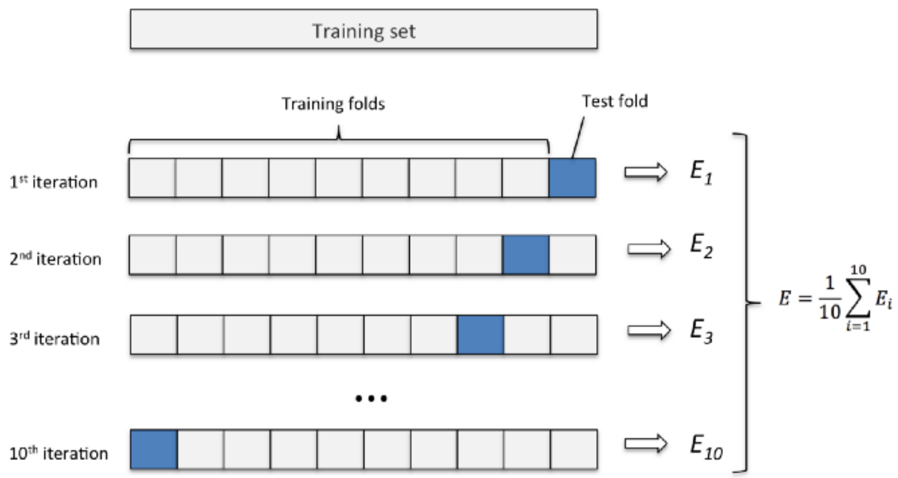

Figure 6.5 — 10-fold cross validation process [180].

Figure 6.5 summarizes the 10-fold process, each iteration allows to compute a forecast error on each subset. The final error is the average of all these errors.

After presenting the methodology, we will proceed with the experiments. The next two chapters are respectively the experiments carried out on AMAWind-Turbine and AMAWind-Farm.

7 AMAWind-Turbine Experiments

This chapter presents the experiments carried out on AMAWind-Turbine and discusses the results obtained. The description of the experiments is first detailed and then the evaluation is divided into two parts: the criteria related to the validation of the functioning of the system and those related to the evaluation of the results. A general synthesis finally points out the properties and limitations of AMAWind-Turbine.

7.1 Description 111
7.2 System validation 113
7.2.1 Impact of the initial forecast 113
7.2.2 Computation times 114
7.2.3 Criticality evolution 115
7.2.4 Relationship between criticality and forecast error 116
7.2.5 Number of cycles and convergence 117
7.3 Evaluation of forecast accuracy 119
7.3.1 Reference forecasting methods 119
7.3.1.1 Choice of a machine learning algorithm 119
7.3.1.2 Method integrating interdependencies 121
7.3.2 Forecast errors 122
7.3.3 Detailed results 123
7.4 General synthesis 124

7.1 Description

While the generic framework of the experiments carried out in this thesis has been presented in section 6, this section details the AMAWind-Turbine specific choices.

First, as explained in section 4.6.2.1, it is necessary to decide on the initialization value of the forecast and the increment value. In this evaluation, we use the "consistent" value, which is a forecast with good accuracy compared to the state of the art. The impact of the initial forecast value is studied in section 7.2.1. The increment value presented in section 4.6.2.2 has been set to 1 kW, i.e. at each cycle an agent can increase or decrease its forecast of only 1 kW.

The experiments are carried out on the five wind farms independently, i.e. in AMAWind-Turbine the forecast of one farm does not influence another. The Wind Turbine Hour (WTH) neighborhood is the same as the one shown in figure 6.1, i.e. each turbine has the two nearest wind turbines as neighbors. If it is at the end of the line it has only one. In addition, each WTH is connected to the nearest Grid Point Hour (GPH) geographically in order to obtain the weather forecasts at the nearest location. Two nearby wind turbines can be connected to the same grid point.

As a reminder, for a given date and time, the various WTH agents in the system start with an initial forecast that they modify during a certain number of cycles until the system stops (due to convergence or exceeding the maximum number of cycles). An example of a forecast calculation for wind farm A at a given date and time is shown in figure 7.1. The graph plots the forecasts and criticalities according to the cycle number. Each agent modifies its forecast at each cycle, leading to a change in criticality.

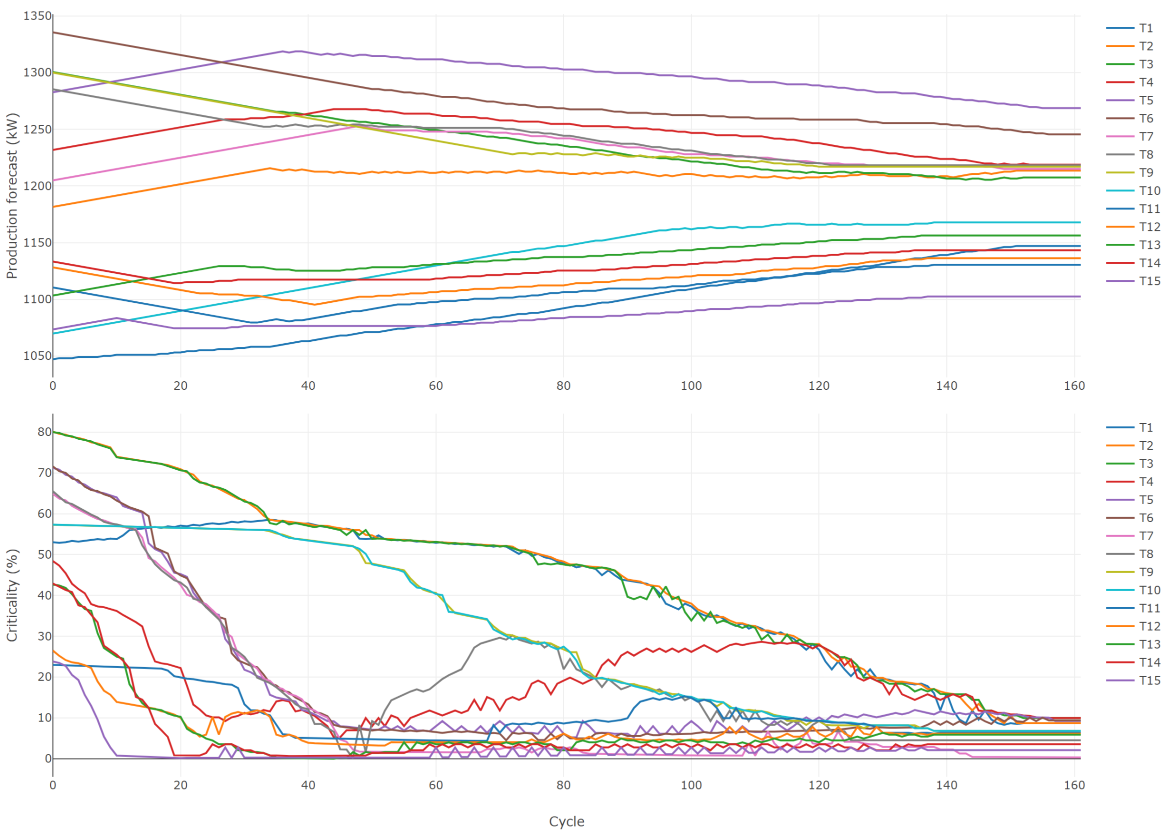

Figure 7.1 — Forecast and criticality evolution, as functions of the cycle number, on the fifteen turbines of the farm A.

Once the system execution is complete, the forecasts for each wind turbine are summed to obtain a global forecast for the wind farm. An example of a forecast over several days is shown in figure 7.2.

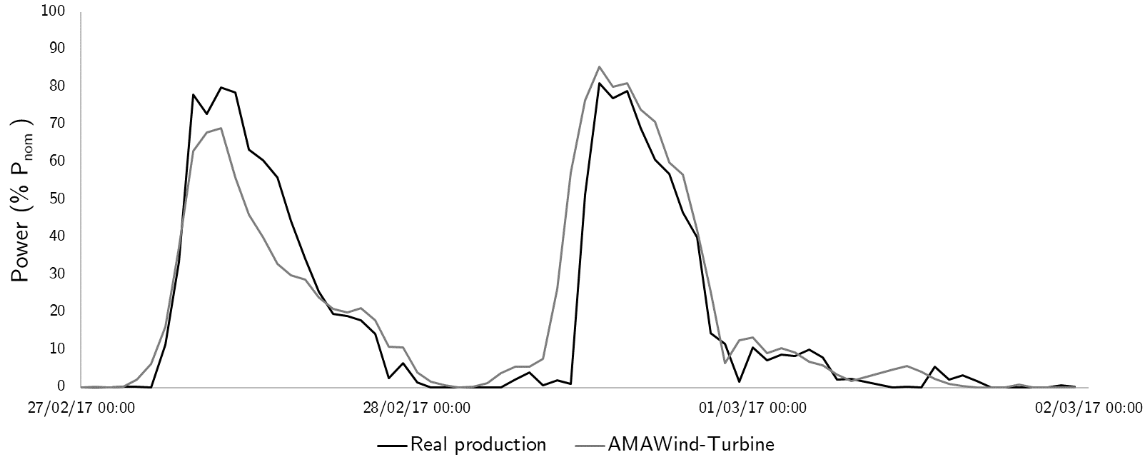

Figure 7.2 — Example of wind power forecasts provided by AMAWind-Turbine over a set of days.

7.2 System validation

AMAWind-Turbine is a system that has been designed from the bottom up, first by defining the entities, their behaviors and then how they can interact with each other. The system validation consists in studying the overall behavior of the system in order to verify its consistency with its objectives and compliance with the constraints imposed.

7.2.1 Impact of the initial forecast

As described in section 4.6.2.1, each agent starts with an initial forecast that it can modify at each cycle. We present a discussion on the choice of the initial forecast to decide which value will be used in the experiments. We have chosen the three following possible choices for the initial forecast of an agent:

- The average production of the entire available history of the wind turbine.
- A random value within the possible production interval $[0, P_{nom}]$ of the wind turbine.
- A "consistent" value predicted by a machine learning algorithm, which corresponds to the median P_{50} of the probabilistic forecasts. Unlike the first two choices, this forecast is already the result of a model that has learned from historical data and therefore achieves good forecast accuracy.

We tested the three choices on AMAWind-Turbine. Table 7.1 summarizes the computation times obtained. When the initial forecast was random or average, performance has been degraded. Indeed, since the initial forecast is further away from the final forecast, more cycles are required. For farms A and E, we stopped the experiments in progress because they exceeded 24 hours. The results of the computation times are detailed and discussed more precisely in section 7.2.2.

Table 7.1 — Comparison of computation time according to the initial forecast used (HH:MM).

Farm	N records	N turbines	Random	Average	Consistent
A	12599	15	+24 hours	07:11	**04:39**
B	16522	4	08:15	02:28	**01:32**
C	16499	4	09:01	02:17	**01:38**
D	17917	10	22:49	06:28	**03:58**
E	18638	12	+24 hours	10:17	**05:29**

We therefore choose the consistent value as the initial forecast for all the next experiments in this chapter.

7.2.2 Computation times

In the previous section we briefly presented the calculation times. This section discusses this simple but important measure in more detail. Indeed, the system must be able to compute forecasts without exceeding the time limits imposed. Table 7.2 shows the computation times required for each farm. *N records* represents the number of records in the history, *N turbines* is the number of turbines in the farm, *Time (HH:MM)* is the calculation time in hours and minutes, and *Time/WT/N* is the calculation time divided by the number of wind turbines and the number of records. The durations seem long (e.g., more than 5 hours for farm E) but it should be remembered that this is a replay of a 3-year forecast of historical data. This length of history is necessary to accurately evaluate the model, but this operation is not often performed. In addition, we can observe that the large time differences are due to the differences between the number of records and the number of wind turbines.

Table 7.2 — Summary of computation times for each farm on AMAWind-Turbine.

Farm	A	B	C	D	E
N records	12599	16522	16499	17917	18638
N turbines	15	4	4	10	12
Time (HH:MM)	04:39	01:32	01:38	03:58	05:29
Time/WT/N	88.58 ms	83.52 ms	89.91 ms	79.89 ms	88.33 ms

As described in section 4.5, our system uses probabilistic forecasts as inputs to calculate the criticality functions of agents. Although these probabilistic forecasts also take a few milliseconds of calculations, the model used must also be trained for several hours on the historical data. This time is not included in the calculation times presented in the table but this training should only be done once.

By calculating the Time/WT/N ratio, we obtain an average of about 80 ms per forecast. Even if AMAWind-Turbine is applied on several farms, the forecasts will be made in a reasonable time (i.e. few seconds), especially since the calculations can be made in parallel. For our field of application, which is wind power forecasting for the day-ahead market, several hours are available (depending on the run time of the weather model) to make forecasts for only one day. We can therefore conclude that in terms of calculation time, our system is applicable in real conditions.

7.2.3 Criticality evolution

The criticality, which is the distance of an agent from its local purpose, is the central point of the functioning of an adaptive multi-agent system. Since the agents are cooperative, i.e. their behaviors are aimed at reducing the worst criticality of their neighborhood, a criticality decrease should be observed over the cycles.

Figure 7.3 shows the evolution of the average criticality for the five farms by box plots. As a reminder, in a box plot the ends of the box represent the first and third quartiles (Q1 and Q3) while the second quartile (Q2) is marked by a line inside the box. The length of the whiskers (i.e. the lines extending vertically from the box) corresponds to 1.5 times the interquartile range (Q3-Q1). For each wind farm, for each execution (representing an hour and a date), the average of the criticalities of all WTH agents in the farm is calculated at each cycle. However, the number of cycles completed is different between each execution and between each farm. We therefore normalized the data in the graph so that cycle #0 always corresponds to the beginning and cycle #1000 always to the end (even if in reality the execution required less cycle time). To facilitate visualization, the boxes represent intervals of 50 cycles.

Overall in this graph, over the cycles, a significant and progressive decrease in criticality can be observed in the five farms. Note that the size of the boxes and whiskers allows to estimate the variability of the criticality. For example, the most left-handed boxes are often large, meaning that the initial criticality is very variable. This is because, depending on the initial situations, the constraints between the wind turbines can almost already be respected, which leads to a low average criticality. There are differences between the farms. Indeed, the boxes are higher on wind farms with more wind turbines (especially A and E), meaning that the criticality is generally higher.

Table 7.3 summarizes the average criticality in the first and last cycles. On average, the criticality decrease is between 27 and 36%. It can also be seen from this table that the initial and final criticalities are higher in farms A and E. One explanation is that the number of wind turbines logically increases the number of constraints to be respected, which are themselves represented in criticality.

Table 7.3 — Average first and last cycle criticality for each farm.

Average criticality \ Farm	A	B	C	D	E
First cycle	49.63%	35.62%	41.27%	36.12%	46.44%
Final cycle	13.55%	7.38%	10.72%	8.11%	12.99%
Difference	-36.08%	-28.24%	-30.55%	-27.51%	-33.45%

We can conclude that the cooperative local behavior of agents leads to an overall decrease in criticality. Their local objective, which is to modify the forecast by taking into account interdependencies with neighboring wind turbines, is therefore respected.

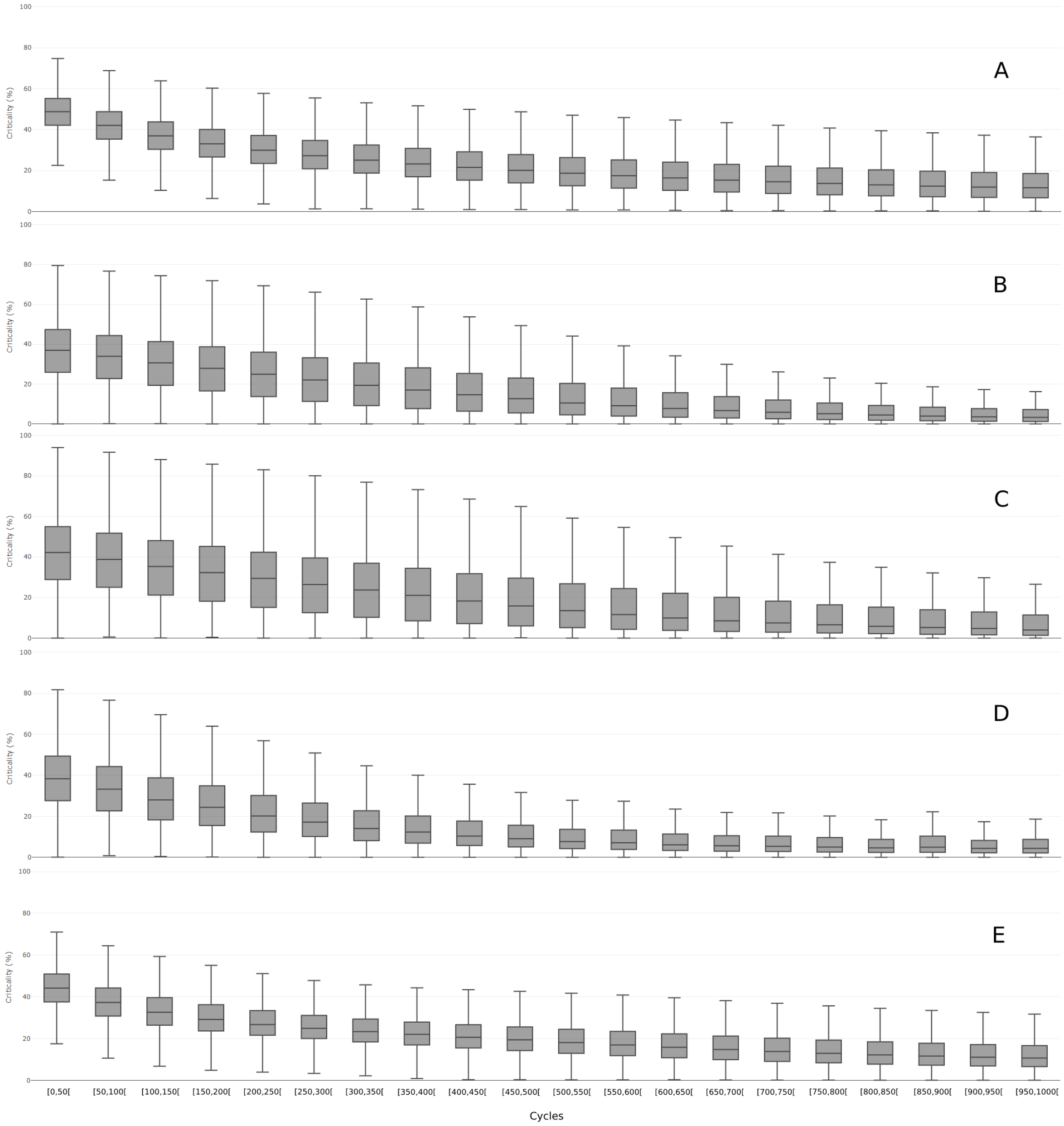

Figure 7.3 **—** Criticality evolution.

7.2.4 Relationship between criticality and forecast error

We have previously observed that the behavior of agents has led to an overall decrease in criticality. However, the overall goal of the system is also that the change in the forecast should lead to a decrease in error. It is therefore important to study the link between criticality and forecast error.

Figure 7.4 shows the error improvement as a function of the criticality improvement for the five farms. For better visibility, the points represent the average of 5% criticality improvement intervals. The relationship between error and criticality improvements is globally increasing on the five farms. This is particularly noticeable in all farms for an

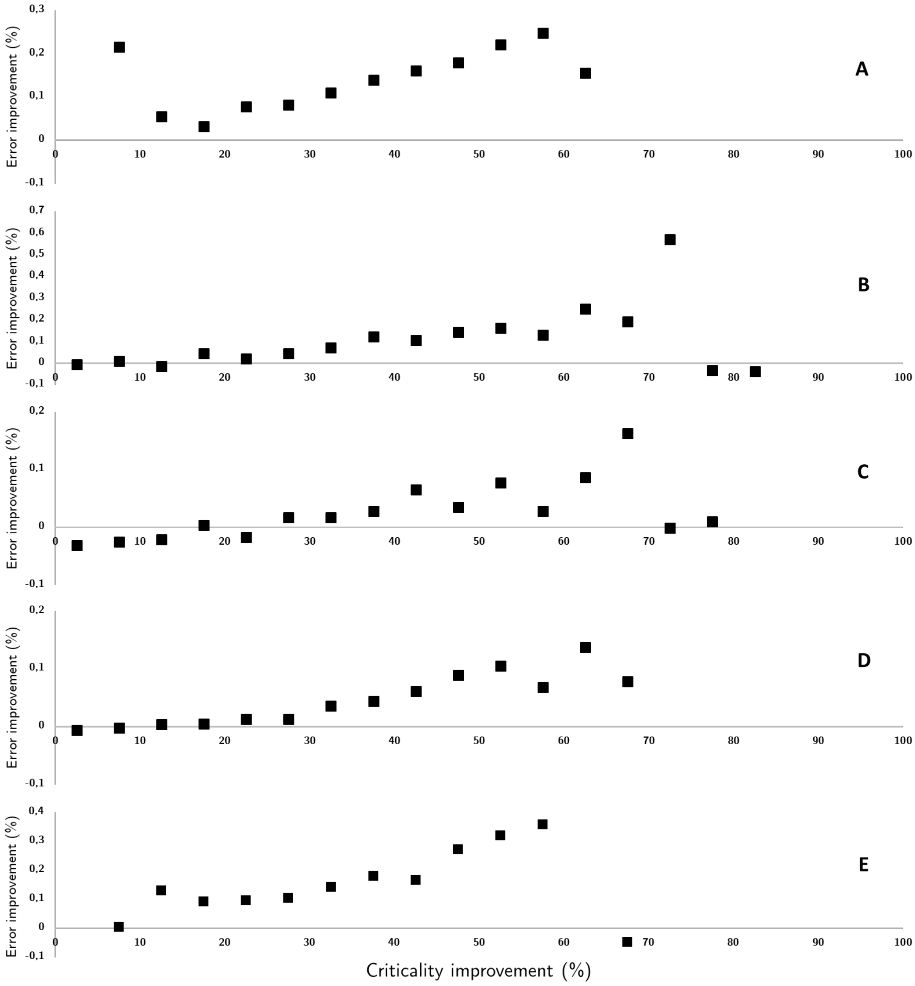

Figure 7.4 — Relationship between error improvement and criticality improvement (average over 5% intervals).

improvement in criticality between 20% and 60%.

Error degradations can be observed for farms B, C and E for major criticality improvements (greater than 70%). In fact, these averages are based on very few examples (about 0.05% of the forecasts) and are not representative. Moreover, it can be seen that when criticality is slightly improved, the error also varies slightly. In fact, little change in criticality often means that the forecast has not changed much, so the error improvement (or degradation) is small.

From this, it can be concluded that there is a global decrease of the error and that this improvement is rather correlated, with some exceptions, to the improvement in criticality.

7.2.5 Number of cycles and convergence

The system is considered convergent if its agents have converged on a forecast and no longer modify it for 5 cycles. If this is not the case, the system has not converged and is stopped at 3000 cycles.

We calculated the total number of cycles required for each forecast performed and then plotted the distribution histogram in figure 7.5. The scale of the y-axis is not displayed so as not to disturb the visualization, it is the shape of the curve in relation to the number of cycles that interests us. The x-axis is limited between 0 and 500 cycles because this represents the vast majority of the data and the curves continue to decrease afterwards. In farms B, C, D and more marginally in A, distributions of similar shape with a different amplitude are observed. On these farms, we observe that the majority of executions end in a small number of cycles (between 30 and 100 depending on the farms). The shape of the distribution histogram of farm E has similar characteristics as the others but is globally flatter.

There is a large difference in "horizontal amplitude" between the farms. For example, farm D rarely ends after 100 cycles while farm E exceeds 500. This is partly due to the difference in the nominal power of the farms (e.g., 800 kW for D and 2300 kW for B). Since the increment value is fixed, a higher power leads to a greater number of cycles because there is more "distance" to cover to reach the objective. A large number of wind turbines can also lead to a larger number of necessary cycles because the system only stops when all the agents have converged. Indeed, the highest number of cycles are observed on farms A and E which are respectively composed of 15 and 12 wind turbines.

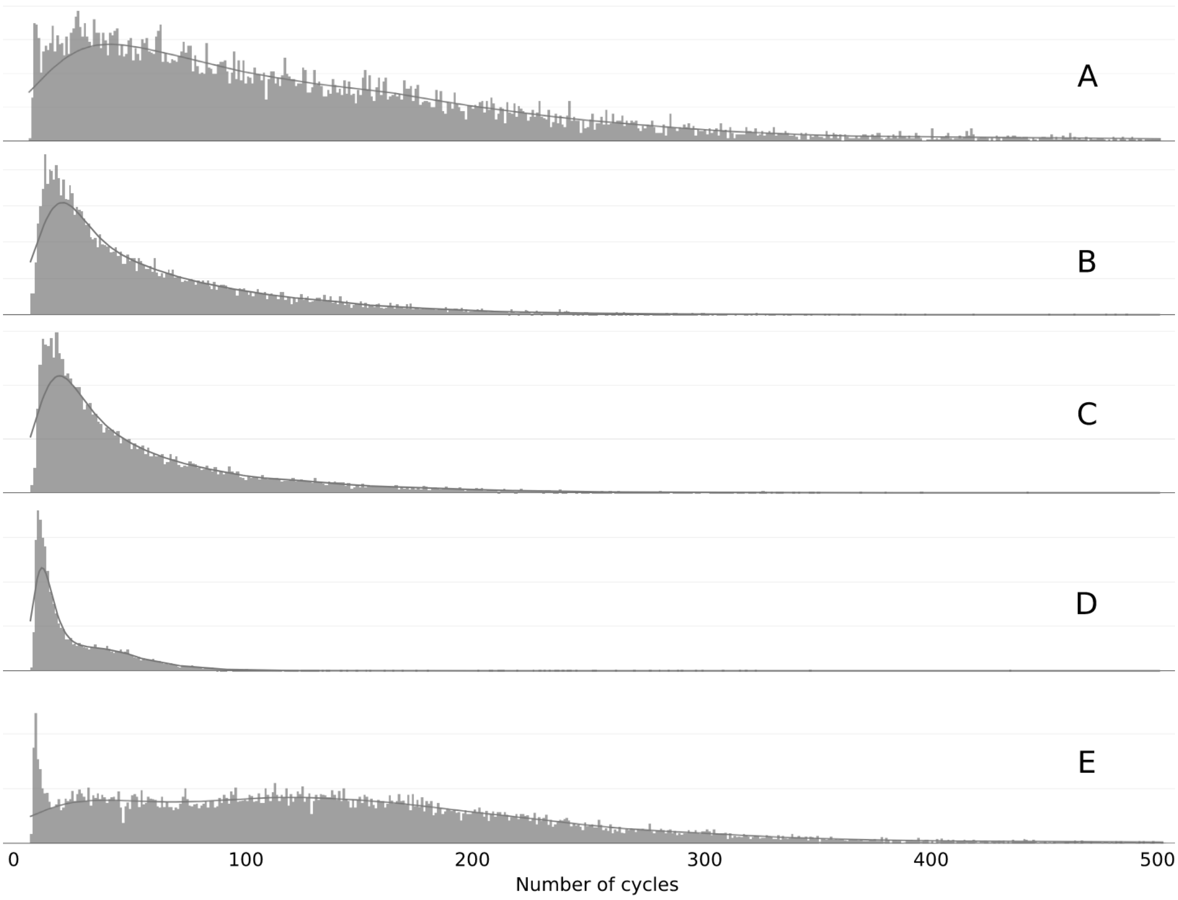

Figure 7.5 — Distribution histograms of number of cycles on AMAWind-Turbine.

Concerning convergence, the maximum number of cycles reached is 1875 for farm E and

1067 for farm A. As the limit was 3000 cycles, the system has always converged and never had to be stopped before convergence.

7.3 Evaluation of forecast accuracy

The previous section has shown that the agents behavior led to a decrease in criticality and a decrease in error in comparison to the initial situation of the system. However, it is important to compare the error obtained with other state-of-the-art methods to truly assess the quality of our forecasting system. In this section we first select and present three forecasting methods as a reference, present the errors obtained on the five farms with these methods, and then detail the results.

7.3.1 Reference forecasting methods

In order to evaluate the forecasts provided by AMAWind-Turbine we must compare it with methods that have already proven their worth. To do this, we compared the results obtained with AMAWind-Turbine with those provided by three methods:

- The first method is a naive forecasting method which gives an overview of the error obtained with little knowledge. It is based on the power curve, i.e. the graph representing the wind speed-production relationship, which is provided by the turbine manufacturer. With a wind speed forecast at wind turbine height, we obtain a production forecast without having to train the model on a production history. The forecasts have a low accuracy but this method gives an overview of the error obtained with almost no knowledge.
- The second method involves learning on data but does not use production data from wind turbines. We use a standard machine learning method to predict production. In order to compare ourselves to only one method of this type, while there are many algorithms available, we experimentally select the method most suitable for our data in the following section (7.3.1.1).
- Finally, the last method also aims to integrate the interdependencies between the production of wind turbines in the production forecast of a wind farm. This method is described in section 7.3.1.2.

7.3.1.1 Choice of a machine learning algorithm

Several machine learning methods have already been introduced in section 2.4.2 and we have decided to select the most efficient one on our data to serve as a comparison method. The first step was to pre-select five popular algorithms that are implemented in the *scikit-learn* library [181], and therefore are easy to use. Then, the parameters of these models were chosen by grid-search i.e. by evaluating all the combinations of parameters within a predefined set. Finally, we trained the five models on the weather forecasting and production history of one wind turbine in each farm for evaluating the error on the entire history by a 10-fold cross-validation.

Below, we briefly present each model with the parameters by default, specified in the documentation [182]–[186], or the parameters chosen when different from the former. Then the errors are compared to choose the method that will be used to serve as a comparison with AMAWind-Turbine.

▷ **Linear Regression (LR)**: linear regression attempts to model the relationship between input variables and an output variable by fitting a linear equation to observed data. As seen in equation 2.8 of section 2.1.2.4, there is a cubic relationship between production and wind. Therefore, with regard to the weather parameters presented in section 6.2.2, we add the cube of the wind speed to the data on which the model is training.

▷ **Multi-Layer Perceptron (MLP)**: an MLP is a class of feed-forward artificial neural networks composed of at least one hidden layer.
Parameters: the solver for weight optimization is an optimizer in the family of quasi-Newton methods (*lbfgs*), the maximum number of iterations is set at 100 and there is one hidden layer of 50 neurons.

▷ **Random Forest (RF)**: random forest is a bagging method that consists in building a multitude of independent decision trees at the time of training and calculating the average forecast of individual trees.
Parameters: the number of trees in the forest is set at 150.

▷ **Gradient Boosting Machine (GBM)**: gradient boosting is a boosting method which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. It builds the model in stages unlike bagging methods and generalizes them by allowing the optimization of an arbitrary loss function.
Parameters: the loss function to be optimized is the least absolute deviation (*lad*) and the number of boosting stages to perform is set at 125.

▷ **Support Vector Machine (SVM)**: in SVM, classifications are done by finding an Optimum Separating Hyperplane (OSH) which maximizes the minimum distance between the classes. In particular, when an OSH cannot be found in the original space, a nonlinear and high-dimensional mapping should be applied. The notion of OSH can be extended to regression problems.
Parameters: the penalty parameter C is set at 110 and the kernel coefficient is set at 0.004.

After the training of the five models on the weather forecasting and production history of one wind turbine in each farm, the NMAE obtained were those presented in table 7.4. According to them, the Gradient Boosting Machine method outperforms all other approaches on all farms. Although a few percent improvement seems little, the difference in error is considerable if the forecast is applied to large wind farms. For example, wind power production in France was 27800 GWh in 2018 [21], therefore 1% error gives 278 GWh corresponding to about 18500 households (assuming that a house of 100 m^2 with 4 people consumes 15 MWh per year). It should be noted that there are significant differences between

Table 7.4 — NMAE of the different machine learning methods for the five wind farms.

	GBM	LR	MLP	RF	SVM
A	**9.64**	13.25	11.31	10.49	10.79
B	**8.31**	11.48	9.8	8.88	9.12
C	**8.68**	11.67	10.43	9.19	9.27
D	**7.84**	11.62	10.03	8.65	8.18
E	**12.83**	16.29	14.88	13.66	13.75
Mean	**9.46**	12.862	11.29	10.17	10.22

the farms, between D and E for example. This is explained by the complexity of the terrain making weather forecasting more difficult and by the capacity factor (a farm that produces less will have an artificially reduced error due to normalization by the nominal power of the farm).

Therefore, only the Gradient Boosting Machine method will be used in the following experiments. In addition, it is generally known to require little parameter tuning and to be robust to erroneous data and overfitting.

7.3.1.2 Method integrating interdependencies

The method proposed by [125], and already briefly presented in section 2.5.1, is a work close to that of this thesis. The authors propose, as we do, to predict the production at the wind turbine scale. However, the final wind farm forecast is a linear combination of these forecasts performed by the LASSO algorithm.

This method is a stacking method, i.e. a first prediction is made with a model and a second model weights the first predictions. Their first model is GBM because it is the model that worked best in the paper and the stacking method is the LASSO (Least Absolute Shrinkage and Selection Operator) linear regression algorithm. As the second model trains from the outputs of the first model, the wind turbine forecasts are obtained for the entire history by cross-validation.

A linear model for the quantity y_t we are attempting to forecast is the weighted sum of input features $X_t = [x_{1,t}, ..., x_{p,t}]$ plus an error term ϵ_t:

$$y_t = \sum_{i=1}^{p} \beta_i x_{i,t} + \epsilon_t \tag{7.1}$$

where the weights $\beta = [\beta_1, ..., \beta_p]^\top$ are unknown parameters to be estimated by the LASSO algorithm.

The paper compares two weighting methods based on linear models: WS (Weighted Sum) and CWS (Conditional Weighted Sum). The particularity of the CWS method is that it takes into account the direction of the wind θ to express y_t:

$$y_t = \sum_{i=1}^{p} \omega_i(\theta) x_{i,t} + \epsilon_t \tag{7.2}$$

where $\omega_i(\theta)$ are the weights that are dependent on θ.

Instead of simply summing the forecasts of each wind turbine to obtain the farm forecast, the CWS method makes a weighted sum that depends on the predicted wind direction. In the paper, the CWS method gives better results than the WS method, therefore we chose to adopt it as weighting method.

7.3.2 Forecast errors

The results obtained are shown in table 7.5 and are divided into NMAE (Normalized Mean Absolute Error) and NRMSE (Normalized Root Mean Square Error). As a reminder, the *PC* method is based on theoretical power curves of wind turbines, the *GBM* method predicts production directly to the entire wind farm by gradient boosting machine, the *CWS* method predicts separately the production of wind turbines and aggregates them by a weighted sum, and finally the *AMAW-T* method corresponds to the forecasts obtained by operating the AMAWind-Turbine system. These results are obtained by cross-validation, i.e. the model learns and predicts on two disjoint subsets from the data history and the operation is repeated until forecasts are obtained on the entire data history. NMAE and NRMSE are calculated directly on the entire history and the lowest errors for each farm are shown in bold.

Table 7.5 **—** Results in terms of NMAE and NRMSE for the five farms.

	Method \ Farm	A	B	C	D	E	Mean
NMAE (%)	PC	10.32	11.71	10.54	11.33	16.92	12.16
	GBM	9.09	7.92	8.11	7.81	12.02	8.99
	CWS	8.93	**7.86**	**8.06**	7.76	11.08	8.74
	AMAW-T	**8.86**	7.87	8.08	**7.65**	**10.92**	**8.68**
NRMSE (%)	PC	15.88	17.08	16.28	18.17	24.41	18.37
	GBM	13.90	11.82	12.29	13.24	16.82	13.62
	CWS	13.52	**11.76**	**12.25**	13.13	15.77	13.29
	AMAW-T	**13.48**	11.78	12.3	**12.87**	**15.47**	**13.18**

The *PC* method obtains results that are 3 and 5% higher in terms of NMAE and NRMSE than other methods. This is not surprising because this forecast does not involve the learning of a model from data.

For the other three methods, it can be seen that both the *CWS* and *AMAW-T* methods, using production data at turbine level, obtain significantly lower errors than *GBM* (on average 0.28% of NMAE and 0.39% of NRMSE). By linking these results to the characteristics of the wind farms in table 6.1, a relationship can be observed between the improvement and the number of wind turbines in the farms. Indeed, the greatest improvements are obtained for wind farms E, A and D, which are composed of 12, 15 and 10 wind turbines respectively. These results reflect a real contribution of data at the wind turbine scale to the wind power forecasting at farm level.

Concerning the two methods *CWS* and *AMAW-T*, on average the *AMAW-T* method obtains the lowest errors, closely followed by the *CWS* method (the difference is 0.06% of NMAE and 0.11% of NRMSE). Concerning the results at the farm level, we observe that

AMAW-T obtains the lowest errors for farms A, D and E while *CWS* has the best results for B and C by a very small percentage (on average 0.015% of NMAE and 0.035% of NRMSE).

We can conclude that methods that integrate interdependencies between wind turbines reduce the forecasting error. In addition, *AMAW-T* reduces the error compared to *CWS* on average, but this improvement is small.

7.3.3 Detailed results

We have studied above, with table 7.5, the error computed on the whole history. As the forecasts are made by 10-fold cross-validation, it is also interesting to study the results obtained for each of the subsets. Tables 7.6 and 7.7 show the NMAE and NRMSE performed on the 10 subsets for each wind farm. The interest of these detailed results is to make it possible to determine whether the improvement obtained on average is homogeneously distributed or corresponds to a significant improvement in specific cases. This allows, for example, to use a model only for certain situations where it is more efficient.

For better readability, only the results of the *GBM*, *CWS* and *AMAW-T* methods are displayed, the *PC* method obtaining much higher errors for each subset. As the farm data do not cover exactly the same dates, the subset does not necessarily correspond to the same periods between each farm. Thus, the five main rows of the table must be interpreted independently. Note that for NMAE, the "Mean" column is equivalent to the farm error in table 7.5 because the error calculation is linear (i.e the average of the errors obtained on the 10 subsets is equal to the average of the errors on the whole set). However, for NRMSE, the means are different due to the non-linearity of the quadratic error.

Table 7.6 — Detailed results for each subset of the 10-fold cross-validation (NMAE).

Farm	Subset / Method	1	2	3	4	5	6	7	8	9	10	Mean
A	GBM	9.46	7.21	7.92	11.32	11.29	10.96	6.06	10.21	7.9	8.52	9.09
	CWS	9.2	7.37	7.81	**11.16**	11.09	10.82	5.92	9.92	**7.63**	8.35	8.93
	AMAW-T	**9.17**	**7.19**	**7.78**	11.17	**11.05**	**10.77**	**5.79**	**9.76**	**7.63**	**8.34**	**8.86**
B	GBM	8.74	6.29	8.15	7.71	7.6	10.1	**8.3**	8.68	6.09	7.56	7.92
	CWS	**8.69**	6.15	**8.13**	**7.58**	**7.47**	**10.09**	8.31	8.57	6.06	7.54	**7.86**
	AMAW-T	8.74	**6.11**	8.28	7.6	7.48	10.16	8.31	**8.56**	**6.02**	**7.46**	7.87
C	GBM	8.84	7.43	9.54	6.16	7.5	11.32	8.56	7.21	7.94	6.64	8.11
	CWS	8.77	7.42	**9.41**	**6.14**	7.49	11.26	**8.5**	7.17	**7.83**	6.61	**8.06**
	AMAW-T	**8.75**	**7.3**	9.49	6.26	**7.43**	**11.24**	8.66	**7.14**	7.88	**6.6**	8.08
D	GBM	10.04	6.31	7.74	6.12	6.88	9.28	9.66	7.73	7.23	7.1	7.81
	CWS	10.0	6.31	**7.58**	6.06	6.88	9.24	9.66	7.63	7.15	7.12	7.76
	AMAW-T	**9.85**	**6.28**	**7.58**	**5.94**	**6.76**	**9.15**	**9.44**	**7.5**	**7.06**	**6.93**	**7.65**
E	GBM	14.95	13.86	13.7	9.71	11.44	13.39	12.56	10.19	10.36	10.08	12.02
	CWS	14.0	**12.05**	13.31	**8.65**	**10.48**	12.53	11.74	9.33	9.51	9.19	11.08
	AMAW-T	**13.39**	12.11	**13.26**	9.05	**10.48**	**12.1**	**11.53**	**9.15**	**9.14**	**9.01**	**10.92**

Concerning NMAE, in table 7.6, it can be seen that for farms A and D, where the *AMAW-T* method achieves better results on average, the majority of subsets obtain lower or almost equal errors. This is also true for farm E where there is however a subset where the *CWS* method is more accurate. For farms B and C, where the *CWS* and *AMAW-T* methods obtain similar average errors, the two methods obtain the best results in about half of the subsets each.

Table 7.7 — Detailed results for each subset of the 10-fold cross validation (NRMSE).

Farm	Method \ Subset	1	2	3	4	5	6	7	8	9	10	Mean
A	GBM	14.4	**11.41**	12.79	**15.78**	16.57	16.12	9.57	15.54	12.01	13.03	13.72
	CWS	**13.69**	11.66	**12.07**	15.8	16.2	**15.58**	9.3	14.79	**11.52**	12.89	13.35
	AMAW-T	13.72	11.45	12.11	15.81	**16.19**	15.6	**9.12**	**14.63**	11.54	**12.86**	**13.3**
B	GBM	12.98	9.33	**12.23**	11.18	11.39	**14.31**	**12.64**	12.87	9.0	11.23	11.72
	CWS	**12.88**	9.16	12.3	**10.98**	**11.21**	**14.31**	**12.64**	12.77	8.97	11.28	**11.65**
	AMAW-T	12.93	**9.14**	12.48	11.02	11.24	14.35	12.65	**12.73**	**8.94**	**11.21**	11.67
C	GBM	13.68	11.52	14.13	9.65	**11.28**	15.88	12.5	11.08	11.87	**9.97**	12.16
	CWS	13.62	11.52	**14.01**	**9.6**	11.3	15.88	**12.38**	11.1	**11.78**	10.0	**12.12**
	AMAW-T	**13.6**	**11.34**	14.12	9.83	**11.28**	**15.86**	12.65	**11.06**	11.9	10.04	12.17
D	GBM	16.26	11.39	12.09	10.17	12.5	14.63	15.88	13.69	12.17	12.33	13.11
	CWS	16.16	11.38	11.86	10.06	12.42	14.51	15.78	13.53	11.97	12.33	13.0
	AMAW-T	**15.78**	**11.19**	**11.76**	**9.81**	**12.1**	**14.25**	**15.42**	**13.25**	**11.81**	**12.02**	**12.74**
E	GBM	20.68	18.75	18.79	13.62	16.0	17.96	17.16	14.42	15.15	14.19	16.67
	CWS	19.13	**16.89**	18.29	**12.27**	14.76	17.04	16.48	13.75	14.4	13.2	15.62
	AMAW-T	**18.12**	17.08	**17.89**	12.77	**14.66**	**16.32**	**16.36**	**13.51**	**13.83**	**12.94**	**15.35**

In terms of NRMSE, in table 7.7, the results follow the same trend but are less marked. For farms D and E, the *AMAW-T* method confirms that it achieves the best results on almost all subsets. For farms B and C, the three methods are very similar and even the *GBM* method, which predicts directly production to the farms, gets errors slightly less than or equal to the others. This observation is also visible for farm A where the differences with the *GBM* method are however higher.

Therefore, our study leads us to conclude that the improvement obtained by the *AMAW-T* method on farms A, D and E is fairly homogeneous because the majority of the subsets of the 10-fold cross-validation have lower errors than the other methods. However, it also confirms that the *AMAW-T* method leads to results equivalent to the reference methods for smaller farms.

7.4 General synthesis

In chapter 4, we designed AMAWind-Turbine which is an Adaptive Multi-Agent System (AMAS) designed to forecast the production of a wind farm by taking into account the short-scale interdependencies between wind turbines. In this system, the wind turbine forecasts are all built at the same time and depend on each other. In this chapter, we conducted experiments to verify the proper functioning of the system, to study various proposed hypotheses and to evaluate the forecast error. The experiments were conducted on real historical data provided by *Boralex*, a customer and partner of **SWIFT*. The wind farms were chosen to represent several possible situations in terms of environmental conditions as well as the number, layout or type of wind turbines.

In section 7.2, we first assessed the impact of different initial forecasts and concluded that the best choice was an initial forecast called "consistent" which is a value predicted by a machine learning algorithm. The interest is to converge more quickly towards the final solution. We then measured AMAWind-Turbine computation times and concluded that the system generates forecasts in an acceptable time and can therefore be used in an industrial context. The study of the evolution of criticality then showed that the overall criticality of the system decreases. This confirms that the cooperative behavior of the selected agents leads to a proper functioning of the system. We have also shown that a decrease in criticality leads to a decrease in error. This allows to validate the choice of the criticality function based on probabilistic forecasts. Finally, we studied the number of cycles required to converge agents towards a final forecast. We concluded that in most cases, few cycles were enough to converge and that the system always converges according to our criteria.

Section 7.3 presented different forecasting methods and compared the results with those obtained by AMAWind-Turbine in terms of NMAE and NRMSE. The integration of interdependencies between wind turbine production, whether via *CWS* or *AMAW-T* methods, showed an improvement in results. This confirms that we can use this type of data for wind farm forecasting and that it can improve forecast accuracy. The study also showed that the improvement seems to be correlated with the number of wind turbines. This link seems logical because the more wind turbines there are, the more interdependencies there are between them (the extreme case being a farm with only one wind turbine where there is therefore no interdependence).

However, the study of forecast errors showed some weaknesses in AMAWind-Turbine. Indeed, as mentioned above, for small wind farms such as B and C, the accuracy of forecasts is very little improved compared to a method, such as *GBM*, that directly forecasts production for the entire wind farm. In addition, when the *AMAW-T* method obtains the lowest error among the methods tested, for example for farms A, D and E, the results are very similar to those of the *CWS* method.

A question to be asked, especially in an industrial context, is whether the slight improvement obtained thanks to AMAWind-Turbine justifies the means implemented. From our point of view, AMAWind-Turbine remains an original forecasting method that has shown satisfactory results in terms of error. Moreover, it is a fairly simple system to set up, without any special knowledge required in wind power forecasting.

8 AMAWind-Farm Experiments

This chapter presents the experiments carried out on AMAWind-Farm and discusses the results obtained. It follows the model of chapter 7, therefore the common points are not explained again in detail and only the differences are given. First, the experimental conditions are described and then the evaluation of the system is divided into two parts: validation of its functioning, and analysis of the results obtained. A general synthesis finally points out the properties and limitations of AMAWind-Farm.

8.1 Description 127
8.2 System validation 129
8.2.1 Wind farm selection 129
8.2.2 Impact of the normalization 129
8.2.3 Impact of the initial forecast 131
8.2.4 Computation times 131
8.2.5 Criticality evolution 132
8.2.6 Relationship between criticality and forecast error 132
8.2.7 Number of cycles and convergence 133
8.3 Evaluation of forecast accuracy 133
8.3.1 Reference forecasting methods 134
8.3.2 Forecast errors 134
8.3.3 Detailed results 135
8.4 General synthesis 136

8.1 Description

As indicated in section 6.2.1, data are available for five wind farms. However, as the farms are heterogeneous, a first study is carried out in section 8.2.1 in order to select the farms on which we will carry out the experiments. A set of farms forms a “region”.

As mentioned in section 5.3, all the farms are neighboring each other. The system is only started once and predicts production for all farms simultaneously. Each farm is connected to a single central grid point.

The impacts of the different normalization possibilities, presented in section 5.4.1, are analyzed in section 8.2.2. To obtain the probability forecasts used to calculate the neighboring criticality (corresponding to two neighboring farms), the Gradient Boosting Machine models are trained with the weather forecasts of the two neighboring farms.

The N_{rate} value, described in section 5.4.2, defines the rate of the nominal power that corresponds to the increment. We decide to set it at 2000, i.e. it will require a minimum of 2000 cycles to decrease or increase the initial forecast of the nominal power of the farm. This choice is large enough to obtain a precise forecast while avoiding excessive computation time. To avoid stopping the system before it converges, the maximum number of iterations is set at 3000.

An example of a forecast calculation for the region at a given date and time (on 02/04/2015 at 05:00) is shown in figure 8.1. The graph plots the forecast and criticality of each Farm Hour agent according to the cycle number. Each agent modifies its forecast at each cycle, and this makes its criticality change. In this example, an agent may increase or decrease its forecast by $\frac{P_{nom}}{N_{rate}}$ (e.g., for farm A, this corresponds to a value of 15kW), making the forecast vary little; however, we can see that the criticalities of the five agents decrease until they converge to a value between 3 and 5%.

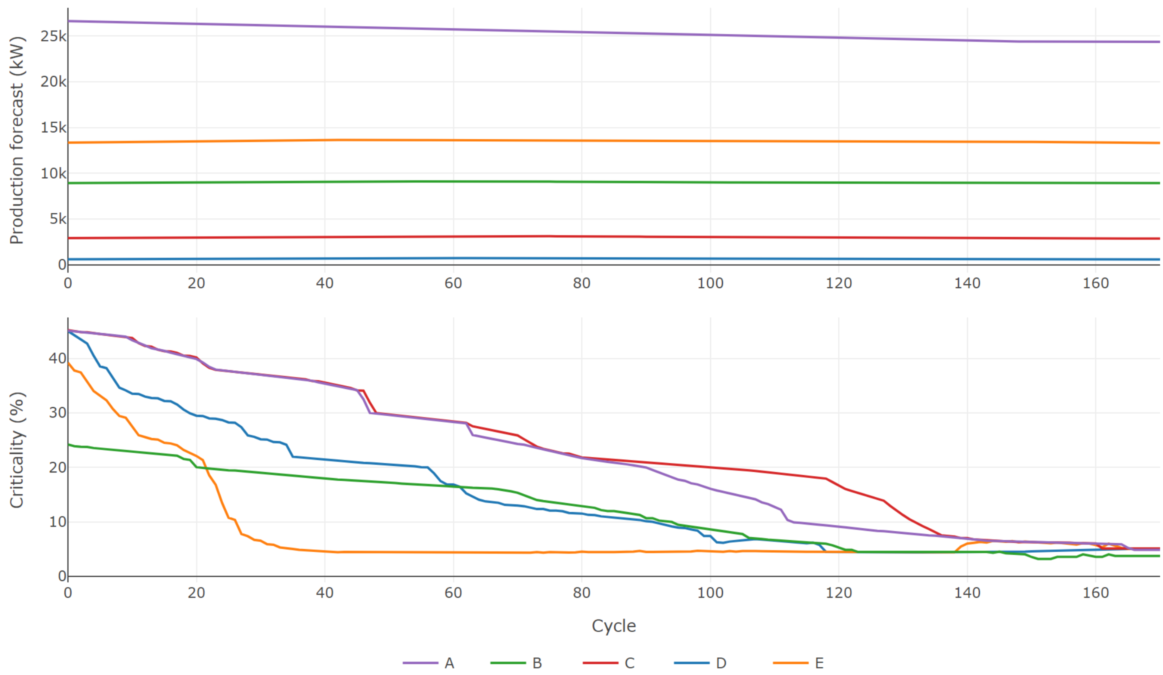

Figure 8.1 — Criticality and forecast evolution as a function of the cycles on five Farm Hour agents.

8.2 System validation

AMAWind-Farm is a system that has also been designed from the bottom up, first by defining the entities, their behaviors and then how they can interact with each other. The system validation consists in studying the overall behavior of the system in order to verify its consistency with its objectives.

This section presents criteria specific to AMAWind-Farm: the wind farm selection as well as the study of the different cases of normalization of the criticality function. Then, it follows the same path as section 7.2 by studying the impact of an initial forecast change, the computation times, the evolution of criticality, the relationship of criticality with the forecast error and the convergence of the system.

8.2.1 Wind farm selection

In the AMAWind-Turbine system, all wind turbines have highly correlated productions because they are located very closely and controlled in the same way. In AMAWind-Farm, this is not the case and we have seen in figure 3.7 that the correlation coefficients are very different between the farms. Since our system is based on the cooperative behavior of the agents representing the farms, we have the intuition that the resolution will not work (or not so well) for farms that are too different.

By testing the AMAWind-Farm system on all five farms together, we obtained an increase in error (compared to the initial value) of 0.23%. This means that our system, by trying to solve the constraints of production difference between the farms, has degraded the quality of the forecast.

We therefore decided to choose the farms A, B and C, which are the three most correlated of the five farms. On this set of farms, AMAWind-Farm almost does not reduce the error but does not degrade it (a decrease of 0.01% of NMAE). However, this study was conducted without the normalization described in section 5.4.1. The different normalization scenarios are tested in the next section.

8.2.2 Impact of the normalization

In order to deal with the strong heterogeneity in size of the farms, we have decided to modify the criticality function of the agents. In this section we test four scenarios of normalization:

- **L-N** (standard **L**ocal – standard **N**eighboring): no normalization.
- **NL-N** (**N**ormalized **L**ocal – standard **N**eighboring): normalization only on local subcriticality.
- **L-NN** (standard **L**ocal – **N**ormalized **N**eighboring): normalization only on neighboring subcriticality.

▷ **NL-NN** (**N**ormalized **L**ocal – **N**ormalized **N**eighboring): normalization on both local and neighboring subcriticalities.

Figure 8.2 summarizes these four possible cases of normalization by presenting the different factors φ used for each farm based on the equation 5.1 seen in section 5.4.1. Farms A, B and C have a nominal power of 30750 kW, 9200 kW and 8000 kW respectively. Their factors φ will therefore be 0.26, 0.96 and 1.

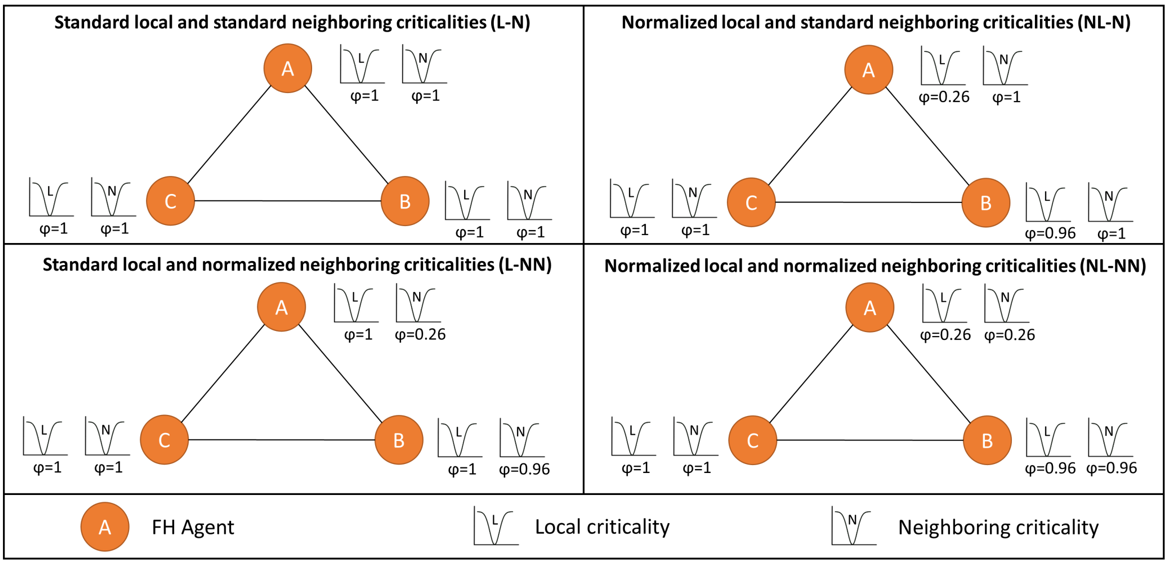

Figure 8.2 — The four scenarios tested and the factors φ applied to subcriticalities.

Table 8.1 summarizes the different criteria already studied (computation times, convergence rate, criticality decrease and error decrease) for the four different scenarios.

Table 8.1 — Validation indicators according to the normalization applied.

Case	L-N	NL-N	L-NN	NL-NN
Computation times (HH:MM)	00:43	01:05	00:40	00:40
Convergence rate	100%	100%	100%	100%
Average first cycle	25.21%	25.08%	19.61%	19.02%
Average final cycle	11.08%	9.47%	9.8%	8.34%
Difference first/final	14.13%	15.62%	9.81%	10.67%
Error decrease	-0.01%	-0.19%	-0.08%	-0.01%

In terms of performance, the NL-N normalization is slower (one hour instead of about 40 minutes) than the others. All experiments achieved a convergence rate of 100%, regardless of the normalization scenario.

The average initial criticality is lower when the neighborhood subcriticality is normalized (cases L-NN and NL-NN). This implies that it is often the neighborhood subcriticality that increases the total criticality in the first place. This behavior is consistent because the initial forecast is chosen to satisfy local criticality first (consistency with weather forecasts at farm location) and not neighborhood criticality.

The NL-N scenario has the best results in terms of error. The decrease in error is almost negligible for the L-N and NL-NN normalizations. It can be noted that in terms of error, the L-N and NL-NN normalizations (no normalization or both subcriticalities are normalized) obtain the same result. By normalizing the two subcriticalities, the agents behavior leads to the same final choices. Between the two scenarios, only the order of decisions will be different. Indeed, in NL-NN, the FH agents linked to the farm C will a priori be the most critical and will therefore be the first ones to see their criticality decrease.

Therefore, in the rest of the chapter, all experiments will be conducted with NL-N normalization.

8.2.3 Impact of the initial forecast

As in section 7.2.1 of the previous chapter, we discuss the choice of the initial forecast. We have similarly chosen the three following possible choices for the initial forecast of an agent:

- The average production of the entire available history of the wind farm.
- A random value within the possible production interval $[0, P_{nom}]$ of the wind farm (which is equal to the nominal power of each wind turbine in the farm times the number of wind turbines).
- A "consistent" value predicted by a machine learning algorithm, which corresponds to the median P_{50} of the probabilistic forecasts. Unlike the first two choices, this forecast is already the result of a model that has learned from historical data and therefore achieves good forecast accuracy.

Table 8.2 summarizes the computation times obtained with each possible initial forecast. Compared to a "consistent" value, we can see that the performance is degraded when a random or an average value are chosen for this initial forecast. Therefore, a "consistent" value will be used in the following experiments.

Table 8.2 — Comparison of computation times on AMAWind-Farm according to the initial forecast used (HH:MM).

N records	N farms	Random	Average	Consistent
9046	3	02:17	01:51	01:05

8.2.4 Computation times

In this section we detail the calculation times already partially presented previously. We can see in table 8.2 that the forecast was performed on 9046 records representing a total of 3 years of history spaced by periods without data. Indeed less data were accessible for AMAWind-Turbine because data from all the three farms had to be available at the same time. On these data, the system takes 1 hour and 5 minutes to run. The evaluation is carried out on the three wind farms selected, therefore each forecast takes 143 ms per farm for a given date and time. This time is short by putting it back into an operational context where only a few dozen records are needed each day.

8.2.5 Criticality evolution

The evolution of the average criticality for the three farms is plotted in figure 8.3 by box plots of 50-cycle steps. A single graph is plotted because our evaluation is carried out over a single region made up of several farms. There is a significant decrease in criticality over the cycles.

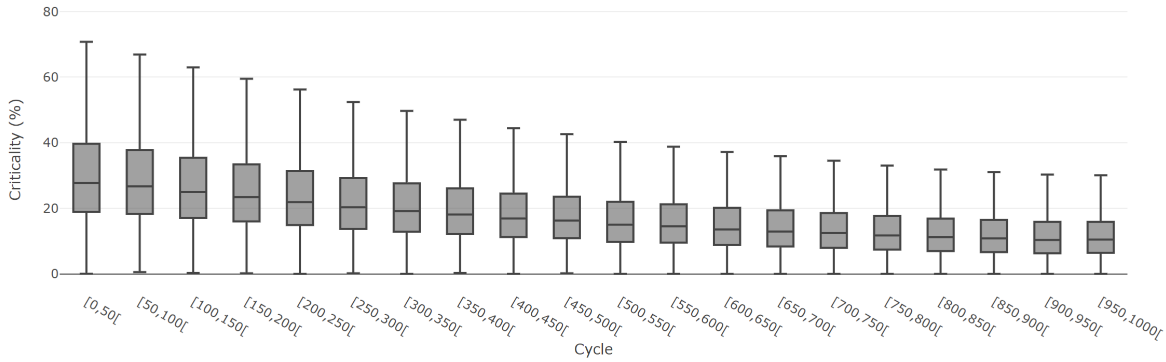

Figure 8.3 — Criticality evolution during the operation of the AMAWind-Farm system.

Table 8.3 summarizes the average criticality in the first and last cycles. We observe that the overall criticality is decreasing. Compared to the results obtained with AMAWind-Turbine, the shape of the curves are similar and the final criticalities are equivalent with an average value of about 10%. However, here the criticality at the beginning is lower (25% against an average of 41% previously). This is because the differences in production between wind farms are more diffuse and less pronounced than the differences between the production of wind turbines. This has the impact of making probabilistic predictions "wider" with larger distances between quantiles. The consequence is that neighborhood criticalities are less constraining (i.e. the criticality function is more extensive and does not form a peak).

Table 8.3 — Average first and last cycle criticality on AMAWind-Farm.

Average criticality		
First cycle	**Final cycle**	**Difference**
25.08%	9.47%	-15.62%

8.2.6 Relationship between criticality and forecast error

The error improvement as a function of the criticality improvement is plotted in figure 8.4 with errors averaged per 5% criticality improvement intervals. The relationship between error and criticality improvements is globally increasing. A slight error deterioration can be observed around 55% criticality improvement. This decrease is offset by error improvements in other cases. Therefore, in AMAWind-Farm, there is also a correlation between criticality and error improvement.

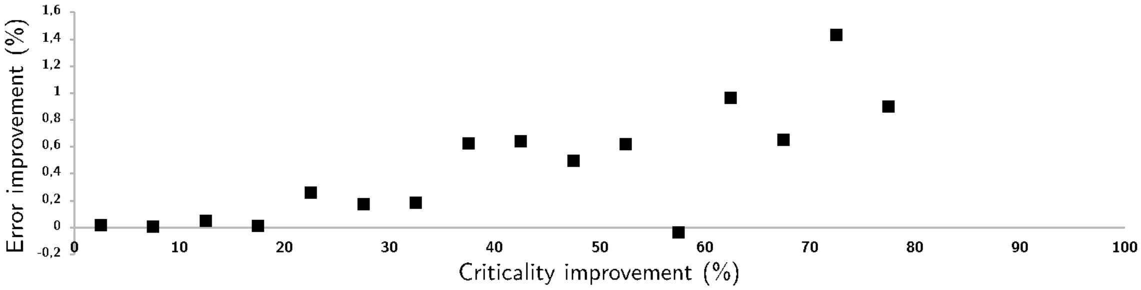

Figure 8.4 — Relation between error improvement and criticality improvement on AMAWind-Farm (average over 5% intervals).

8.2.7 Number of cycles and convergence

The system is considered convergent if all agents have converged on a forecast and no longer modify it for 5 cycles. If this is not the case, the system has not converged and is stopped at 3000 cycles. We calculated the total number of cycles required for each forecast performed and then plotted the distribution histogram in figure 8.5.

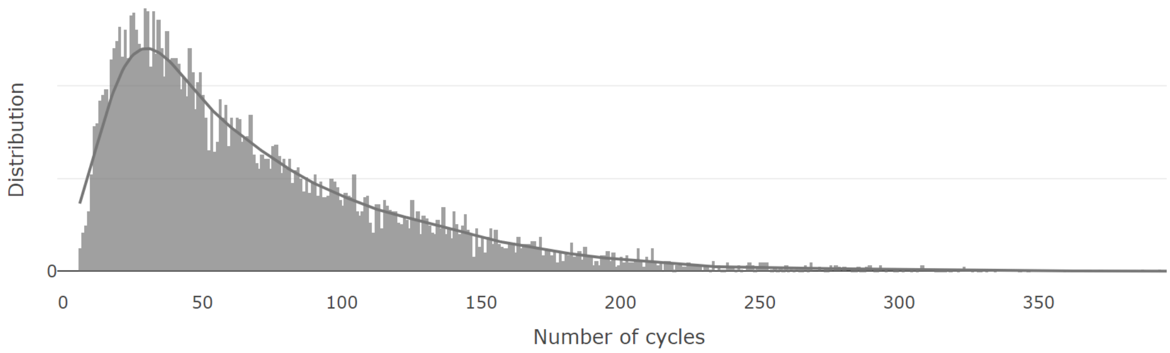

Figure 8.5 — Distribution histograms of number of cycles required on AMAWind-Farm.

The shape of the distribution is similar to those obtained with AMAWind-Turbine shown in figure 7.5. The system requires a small number of cycles to reach the end of the resolution, the most frequent number of cycles required is about 30. The system converges in all cases because the maximum number of cycles reached is 408, while the maximum we have allowed is 3000.

8.3 Evaluation of forecast accuracy

The previous section showed that the agents behavior led to a decrease in criticality and a decrease in error. In this section, the error is compared to those obtained by other standard methods. The different results obtained in terms of forecast accuracy are presented here.

8.3.1 Reference forecasting methods

The Power Curve (*PC*) method corresponds to the forecast obtained using the theoretical power curve function (linking production to wind speed) provided by the wind turbine manufacturer. We use the sum of the power curves of all the wind turbines in the farm and take as input the wind forecast at 100m from a central grid point at the wind farm.

We have shown in section 7.3.1.1 that Gradient Boosting Machine (GBM) is the most suitable machine learning method for our problem. We also use it as a reference method in this section. We are building two models: *GBM-F* (GBM-Farm) which is trained for each wind farm with data from a central grid point and *GBM-R* (GBM-Region) which is trained directly for the region with data from the grid points of each farm.

Finally the *AMAW-F* method refers to the results obtained by operating the AMAWind-Farm system.

8.3.2 Forecast errors

The results obtained are summarized in table 8.4 and are divided into NMAE (Normalized Mean Absolute Error) and NRMSE (Normalized Root Mean Square Error). For each farm, the lowest error is in bold. The errors of the region do not correspond to the average of the errors because it is the sum of the forecasts compared to the total production. Thus, errors may compensate each other: e.g., if farm A produces less than expected and farm B produces more than expected, the total error may be zero when the model has been wrong twice. The *GBM-R* method provides production forecasts directly at the regional level, therefore there is no result corresponding at farms level in the table.

Table 8.4 — Results summary on AMAWind-Farm.

	Method \ Farm	A	B	C	Region
NMAE (%)	PC	11.02	11.88	10.41	8.15
	GBM-R	–	–	–	7.16
	GBM-F	9.66	8.23	**8.08**	6.94
	AMAW-F	**9.51**	**8.11**	8.27	**6.85**
NRMSE (%)	PC	16.61	17.34	16.00	11.70
	GBM-R	–	–	–	10.28
	GBM-F	14.42	12.24	**12.25**	10.07
	AMAW-F	**14.10**	**11.98**	12.42	**9.92**

As for the experiments conducted for AMAWind-Turbine, the *PC* method based on the theoretical power curve obtains significantly higher errors than the other methods, both at the farm and regional level.

In addition, both the *GBM-F* and *AMAW-F* methods obtain significantly lower errors than *GBM-R* (on average 0.27% of NMAE and 0.29% of NRMSE). As these two methods use production data at farm level, these results reflect a real contribution of data at the wind farm scale to the regional wind power forecast.

Concerning the two methods *GBM-F* and *AMAW-F*, on average the *AMAW-F* method obtains the lowest errors, closely followed by the *GBM-F* method (the difference is 0.09% of NMAE and 0.15% of NRMSE). On average, these two methods can be considered equivalent in terms of forecast accuracy. However, concerning the results at the farms level, we observe that *AMAW-F* obtains the lowest errors for farms A and B while *GBM-F* has the best results for C. For the three farms, the differences in error between the two models are significant (on average 0.15% of NMAE and 0.25% of NRMSE).

Although the *AMAW-F* method does not always obtain the best results, the fact that both methods are significantly better on particular farms is still interesting. Indeed, if it is possible to distinguish situations where one method is better than another, it is sufficient to create a meta-model selecting the right method at the right time (such as the stacking methods presented in section 2.4.2.4). In our example, this corresponds to the use of *AMAW-F* to predict the production of farms A and B, and *GBM-F* for farm C. We have tested this meta-model, it obtained a regional error of 6.81% of NMAE and 9.85% of NRMSE (respectively a decrease of 0.04% and 0.07%) and therefore allowed the error to be reduced slightly further.

8.3.3 Detailed results

Similar to the previous chapter, we also detail errors for each subset of the 10-fold cross-validation. Tables 8.5 and 8.6 present the errors in terms of NMAE and NRMSE respectively. For better readability, only the results of the *GBM-R*, *GBM-F* and *AMAW-F* methods are displayed, the *PC* method obtaining much higher errors for each subset. Since the *GBM-R* method predicts production directly at the regional level, errors can only be computed for the region.

Concerning NMAE, in table 8.5, the detailed results confirm that the *AMAW-F* method obtains lower errors for farms A and B than *GBM-F* on a majority of subsets. The opposite is also true for farm C. At the regional level, the *AMAW-F* method also achieves the best results on a majority of subsets.

Table 8.5 — Detailed results on AMAWind-Farm in terms of NMAE (%).

Farm	Method \ Subset	1	2	3	4	5	6	7	8	9	10	Mean
A	GBM-F	9.54	7.81	9.23	11.46	12.64	**11.2**	6.02	11.75	**8.42**	9.45	9.75
	AMAW-F	**9.48**	**7.7**	**9.03**	**11.14**	**11.78**	11.32	**5.86**	**10.96**	8.5	**9.34**	**9.51**
B	GBM-F	7.66	7.24	8.01	10.15	10.12	8.97	**6.73**	9.65	6.7	7.12	8.24
	AMAW-F	**7.47**	**7.12**	**7.96**	**9.96**	**9.95**	**8.87**	6.76	**9.34**	**6.64**	**7.04**	**8.11**
C	GBM-F	**8.08**	**5.9**	**7.19**	**10.97**	**11.59**	**7.6**	**5.45**	**10.77**	**7.47**	**5.8**	**8.08**
	AMAW-F	8.18	5.92	7.41	11.5	11.88	7.7	**5.45**	10.94	7.59	6.13	8.27
Region	GBM-R	7.07	5.85	6.87	8.5	9.02	8.1	4.73	8.52	6.25	6.66	7.16
	GBM-F	6.72	5.59	6.71	8.18	9.07	**8.0**	4.55	8.0	**6.11**	**6.5**	6.94
	AMAW-F	**6.7**	**5.51**	**6.6**	**7.98**	**8.59**	8.05	**4.45**	**7.92**	6.14	6.52	**6.85**

In terms of NRMSE, in table 8.6, the results follow the same trend. For farms A and B, the *AMAW-F* method confirms that it achieves the best results on almost all subsets. For farm C, it is the same, but for the *GBM-F* method. For the region, the *AMAW-F* method achieves the best results on a majority of subsets.

Therefore, our study leads us to conclude that the improvement obtained by the *AMAW-F*

Table 8.6 — Detailed results on AMAWind-Farm in terms of NRMSE (%).

Farm	Method \ Subset	1	2	3	4	5	6	7	8	9	10	Mean
A	GBM-F	**13.81**	12.14	14.2	15.72	17.65	15.81	9.19	16.71	12.71	14.29	14.22
	AMAW-F	13.87	**11.89**	**13.86**	**15.15**	**16.87**	**15.8**	**8.92**	**16.17**	**12.51**	**14.17**	**13.92**
B	GBM-F	11.33	10.73	12.09	13.97	14.84	13.02	**10.03**	14.43	9.93	10.77	12.11
	AMAW-F	**11.09**	**10.51**	**11.88**	**13.7**	**14.17**	**12.88**	10.21	**13.92**	**9.75**	**10.63**	**11.87**
C	GBM-F	**11.69**	9.32	**11.04**	**15.24**	**16.67**	11.42	8.62	**15.12**	**11.28**	**9.21**	**11.96**
	AMAW-F	11.94	**9.3**	11.23	15.53	16.85	**11.41**	**8.58**	15.21	11.48	9.78	12.13
Region	GBM-R	10.32	8.64	10.23	11.27	12.36	11.08	7.06	12.04	8.85	9.71	10.16
	GBM-F	**9.68**	8.51	10.16	10.93	12.53	**10.95**	6.73	12.23	**8.6**	9.7	10.0
	AMAW-F	9.73	**8.39**	**9.92**	**10.72**	**11.98**	11.01	**6.59**	**11.35**	8.71	**9.59**	**9.8**

method on farms A and B is fairly homogeneous because the majority of the subsets of the 10-fold cross-validation have lower errors than the *GBM-F* method. It is also true that the *GBM-F* method has lower errors on the majority of subsets for farm C.

8.4 General synthesis

In chapter 5, we designed AMAWind-Farm which is an Adaptive Multi-Agent System (AMAS) designed to forecast the production of a region (i.e. several wind farms) by taking into account the large-scale interdependencies between wind farms. In this system, the wind farm forecasts are all built at the same time and depend on each other. In this chapter, we conducted experiments to verify the proper functioning of the system, to study various proposed hypotheses and to evaluate the forecast error.

In section 8.2, some experiments were carried out only on AMAWind-Farm. Firstly, we tested the system on the five farms and on the three most correlated farms, this led to a deterioration and a slight improvement in the error rate respectively. Therefore, we continued the evaluations with only the three most correlated farms so that farms with independent productions would not degrade the predictions of the others. Then, we evaluated the impact of a change in the criticality function of agents with several normalization scenarios with a factor φ depending on the nominal powers of the farms. The study concluded that the best results were obtained for the *NL-N* scenario, where only the local subcriticality is normalized. To avoid measuring correlations and automatically choosing the best farms to be included in a forecast, one possible automatic solution would be for the system alone to ignore a farm that degrades the overall forecast.

Then, as in the previous chapter, we assessed the impact of different initial forecasts, we measured AMAWind-Farm computation times, we studied the evolution of criticality and the relationship between criticality and forecast error, and we finally assessed the convergence of the system. These experiments led to the same conclusions as for the AMAWind-Turbine system, which is the proper functioning of the system.

In section 8.3, we presented different forecasting methods and compared the results with those obtained by AMAWind-Farm in terms of NMAE and NRMSE. The integration of interdependencies between wind farm production, whether via *GBM-F* or *AMAW-F* methods, showed an improvement in results. This confirms that we can use this type of data for regional forecasting and that it can improve forecast accuracy. However by detailing the errors obtained per wind farm, it became clear that the result was uneven. Indeed, although the *AMAW-F* method achieves the best results for farms A and B, it is the *GBM-F* method that reduces the error the most for farm C. Since the results differ significantly from one farm to another, this counter-performance can be compensated by a meta-model that chooses the method to be used according to the situation.

In conclusion, we can raise the same question as in section 7.4 for AMAWind-Turbine concerning the justification of such a system for the relatively small improvement achieved. The improvements obtained being of the same order of magnitude, we can also answer that this method is first successful in reducing the error, even slightly, and then, that its originality opens the way for forecasting methods based on a decentralized approach.

Conclusion and Perspectives

This thesis was devoted to the study of wind power forecasting methods. With the strong growth of renewable energies, this challenge, which was a minor problem, has come to the forefront. In this last chapter, we summarize the different points discussed throughout this thesis. We then detail the contributions to wind power forecasting and to AMAS. We conclude by offering some interesting work perspectives.

General Conclusion

The study on the energy sector showed a heavy dependence on fossil fuels available in limited quantities and a threat of a major climate change. These two points are leading to an increase in renewable energy, particularly wind energy, which has the disadvantage of being intermittent and variable. As supply is not controllable, changes in demand can lead to an energy surplus or shortage. In the absence of sufficiently mature, cost-effective and available technologies to store energy, improving wind power forecasting models seems to be a necessity to support the integration of renewable energy.

Chapter 2 provided a theoretical study on wind, the functioning of wind turbines and the specificity of the wind power forecasting problem. It described the two main prediction approaches: physical models, based mainly on fluid dynamics and statistical models based on historical data without necessarily understanding meteorological models. Then, it highlighted the strong spatial dependencies between the production of wind turbines from the same wind farm or between the production of wind farms from the same region. An analysis of the forecasting models commonly used by companies was conducted using multiple criteria including the ability to handle dynamics, forecast accuracy or the integration of spatial interdependencies. This study showed that, on the one hand, statistical approaches meet many criteria but rarely take spatial interdependencies into account in forecasts. On the other hand, interdependencies are taken into account by physical approaches, which nevertheless tend to be less precise and are gradually being neglected.

The distributed, open and dynamic nature of our problem has led us to become interested in Multi-Agent Systems (MAS). More specifically, in chapter 3, we studied Adaptive Multi-Agent Systems (AMAS) which are a specific type of MAS that enables a system to perform complex behavior by emergence from the local cooperative behavior of agents with simple logic. The need for consistency between forecasts, the natural distribution of the problem and the need to adapt the model to inaccurate and heterogeneous data made us identify AMAS as a potential approach for a wind power forecasting method taking into account the interdependencies between wind turbines and wind farms.

In chapter 4, we presented AMAWind-Turbine which is an AMAS dedicated to predicting the production of a wind farm by taking into account the interdependencies between wind turbines. In this system, wind turbines are agents that act cooperatively with their neighbors, i.e. the nearest wind turbines, to provide consistent forecasts between them. One of the particularities of this system is that the criticality, the value that guides agents actions towards their own objectives, is built from probabilistic production forecasts. The criticality of each agent is divided into local and neighboring subcriticalities which respectively allow to take into account in the forecast the wind turbine own production history as well as the history of production differences with neighboring wind turbines.

Chapter 5 described AMAWind-Farm, a similar system to AMAWind-Turbine applied to predicting the production of a region, i.e. a set of wind farms, based on farm data. The main difference is a change of scale because the agents are, in this case, wind farms that cooperate to predict the production of a region. In addition, wind farms are different in terms of location, size and operation, unlike wind turbines in the same farm. This heterogeneity of agents has led to a change in criticality that requires normalization depending on the size of the farms.

In chapter 6, production and weather forecast data have been described in detail. In order not to disturb the results with parameters that are outside the scope of this thesis, erroneous records or records corresponding to maintenance or failure cases have been filtered. The evaluation methodology has been described by the presentation of the 10-fold cross-validation as well as by the description of the different criteria. These latter concerned both the proper functioning of the systems and the accuracy of the forecasts obtained in comparison with other methods. As this thesis is co-supervised by **SWIFT*, this last point is very important because our research addresses a very practical need.

For the AMAWind-Turbine system, the experiments conducted in chapter 7 first discussed the choice of the initial forecast. Then, they showed that computation times were acceptable in an industrial context, that the cooperative behavior of the agents led to a decrease in overall criticality and that this decrease was correlated with a decrease in forecast error. All these elements have validated the proper functioning of the system. In the error study, it was observed that methods that take into account the interdependencies between wind turbines obtained better results and even more so for farms with a high number of wind turbines. Compared to *CWS*, another method that also takes interdependencies into account, AMAWind-Turbine obtains slightly lower errors on the largest farms and almost similar results for the smallest farms.

In chapter 8, we carried out experiments on the AMAWind-Farm system and the same conclusions were made regarding the validation of the proper functioning of the system. The study to test AMAWind-Farm on all wind farms showed that the system was not robust to farms with low correlation with other wind farms. Therefore, the evaluations focused on the three farms with the most correlated production. Then, the different ways of changing the criticality computation to compensate for the heterogeneity of the farms, and then the agents, was studied and this led to the choice of a normalization on only the local subcriticality. In terms of error, we observed an improvement in the results for methods that take into account the interdependencies between farms. Compared to *GBM-F*, another method that also takes into account interdependencies, AMAWind-Farm obtained significantly lower errors for two farms. For the last farm, it was the opposite. Since the results differ significantly from one farm to another, we have shown that this counter-performance of AMAWind-Farm can nevertheless be compensated by a meta-model that chooses the method to be used according to the situation.

Contribution

Contribution to Wind Power Forecasting

This thesis contributed to wind power forecasting, the field of application, by proposing a new method for predicting wind production. The originality of this method is that it takes into account in the large-scale production forecast (a farm or a region) the interdependencies between finer-scale production (respectively a turbine or a farm). These interdependencies are induced by the strong spatial correlation of the data and provide new information to the forecast model. This method has been applied in two systems: AMAWind-Turbine and AMAWind-Farm dedicated respectively to forecasting a wind farm using wind turbine data and forecasting a region using wind farm data.

Following the experiments conducted at both scales, the methods integrating the interdependencies of wind turbines or wind farms obtained errors significantly lower than the standard methods. The use of finer-scale data has therefore improved the forecast accuracy. However, these data are not widely used in forecasting models because they are not always provided by aggregators or producers, as they sometimes require more complex infrastructure or additional resources. This work shows, that this additional information can improve forecast accuracy. This thesis therefore contributes to justify a democratization of these data for wind power forecasting.

In addition, on the largest wind farms and on average, AMAWind-Turbine obtained slightly better results in terms of error than *CWS*, a method that also takes into account the interdependencies between wind turbines. For its part, AMAWind-Farm obtained significantly better results than *GBM-F* for two of the three farms evaluated. Therefore, with methods using equivalent data, the results are either slightly improved or significantly improved or degraded depending on the farm. Although these results do not revolutionize the field of forecasting, they are encouraging. The results can be put into perspective because, in practice, the methods used as references in this thesis already obtain lower errors on many forecasters on benchmarks organized by aggregators and producers.

Contribution to AMAS

This thesis provides another validation of the AMAS approach in a new domain and confirms its ability to manage complexity. The overall principle of the work presented is to use data at finer scales and to take into account the interdependencies between them in order to make forecasts on a larger scale. In this thesis we apply this concept to wind power forecasting, at two different levels, but it is general enough to apply it to other situations and scope. In fact, the principle of our systems is sufficiently generic to apply to any hierarchical time series forecasting problem in which time series collections must be added together in a consistent manner.

Although learning and forecasting issues have already been processed by AMAS, in this thesis, we have to use mostly false data. Indeed, the quality of medium-term weather forecasts is highly variable, particularly for wind, due to the chaotic nature of the weather. This lock has been highlighted during a preliminary study on the use of AMAS for wind power forecasting [165]. This issue was solved by delegating part of the learning to a statistical model adapted to this type of highly noisy data.

Another point studied in this thesis is the behavior that agents must have when the agents of a system are heterogeneous. A similar problem has been addressed in a thesis [187] where the author has to deal with an AMAS representing a system of system. Since the systems are not designed by the same people, criticalities are not always comparable between them. The proposed solution is that an agent can normalize its perception of a criticality of a neighbor (by applying a multiplicative factor) if the neighbor does not have the same criticality standards. In AMAWind-Farm, a similar problem of heterogeneity between farms was solved by normalizing the criticality function by a factor depending on the size (more precisely the nominal power) of the wind farms. Since the two methods are similar and have improved the results, our work confirms the idea that criticality normalization is effective in addressing agent heterogeneity issues and this enriches the AMAS theory.

Perspectives

Even though AMAWind-Turbine and AMAWind-Farm are ready to use and have shown conclusive results during the experiments, several improvements have appeared to us during this research work.

Concerning the heterogeneity of agents, a problem raised during this thesis is the integration of farms with different productions in AMAWind-Farm. Indeed, during the evaluation with five farms, we noticed that some farms tended to disrupt the cooperative resolution and eventually led to a higher error. More formally, a first perspective of research would be to detect disruptive agents locally and, then, managing these non cooperation situations in order to come back to a global cooperative behavior.

Another possible follow-up to this thesis work would be to merge the two systems AMAWind-Turbine and AMAWind-Farm in order to obtain a single system that adapts to the data received. Communication between the wind farms and the wind turbines, each with their own objectives, represents a challenge in the continuity of the problem of heterogeneity between the Farm Hour agents of the AMAWind-Farm system.

The two systems have been evaluated on five wind farms and this has led to a decrease in forecasting error. However, this improvement should be put into perspective because it remains low compared to the means used in this thesis. In addition, we experimentally observed that AMAWind-Turbine obtained better results for the wind farms with the highest number of wind turbines. The largest wind farm evaluated in this thesis consists of 15 wind turbines while there are many wind farms with several hundred wind turbines. We therefore hypothesize that repeating the experiment on larger farms would lead to greater improvements. Indeed, taking into account the interdependencies between wind turbines makes more sense for large wind farms, with strong wake effects, than for a small wind farm with four wind turbines. This research track would raise the question of the choice of the agents neighborhoods (when the layout of the farm is not just a line).

Finally, the systems could be extended to take into account temporal interdependencies. The initial design of the systems has already been thought in this direction by the decomposition of the Wind Turbine Hour and Farm Hour agents into 24 hours. By following the current method, the principle is to add in the neighborhood of an agent, the agents forecasting the production of the same wind turbine at the previous and following hours. The agents behavior should be adapted to this new neighborhood.

Bibliography

[1] F. Berna, P. Goldberg, L. K. Horwitz, J. Brink, S. Holt, M. Bamford, and M. Chazan, "Microstratigraphic evidence of in situ fire in the acheulean strata of wonderwerk cave, northern cape province, south africa", *Proceedings of the National Academy of Sciences*, vol. 109, no. 20, E1215–E1220, 2012.

[2] K. H. Fasol, "A short history of hydropower control", *IEEE Control systems magazine*, vol. 22, no. 4, pp. 68–76, 2002.

[3] D. Le Gourieres, *Wind power plants: theory and design*. Elsevier, 2014.

[4] C. Zou, Q. Zhao, G. Zhang, and B. Xiong, "Energy revolution: from a fossil energy era to a new energy era", *Natural Gas Industry B*, vol. 3, no. 1, pp. 1–11, 2016.

[5] J. Rifkin, *The third industrial revolution*, 2011.

[6] R. P. Feynman, R. B. Leighton, and M. Sands, "The feynman lectures on physics; vol. i", *American Journal of Physics*, vol. 33, no. 9, pp. 750–752, 1965.

[7] G. R. Slemon, *Magnetoelectric devices: transducers, transformers, and machines*. Wiley, 1966.

[8] International Energy Agency, *Global energy & climate outlook*, 2018.

[9] M. Sato, "Thermochemistry of the formation of fossil fuels", *Spec Publ-Geochemical Soc*, vol. 2, pp. 271–283, 1990.

[10] D. Sarkar, *Thermal power plant: design and operation*. Elsevier, 2015.

[11] International Energy Agency, *World energy outlook*, 2018.

[12] BP, *Bp statistical review of world energy*, 2019. [Online]. Available: `https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2019-full-report.pdf`.

[13] Intergovernmental Panel on Climate Change, *Summary for policymakers*, 2018.

[14] World Nuclear Association, *Supply of uranium*, 2018. [Online]. Available: `http://world-nuclear.org/information-library/nuclear-fuel-cycle/uranium-resources/supply-of-uranium.aspx`.

[15] H. Ritchie and M. Roser, *Co2 and other greenhouse gas emissions*, 2019. [Online]. Available: `https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions`.

[16] negaWatt, *Scenario negawatt 2017-2050*, 2017.

[17] Central Intelligence Agency, *The world factbook*, 2019. [Online]. Available: https://www.cia.gov/library/publications/resources/the-world-factbook/index.html.

[18] International Renewable Energy Agency, *Capacity and generation*, 2018. [Online]. Available: http://resourceirena.irena.org/gateway/dashboard/?topic=4&subTopic=54.

[19] M. Combe, *Energies renouvelables: des investissements en baisse!*, 2018. [Online]. Available: http://www.natura-sciences.com/energie/energies-renouvelables-investissements-baisse.html.

[20] International Renewable Energy Agency, *Finance and investiment*, 2018. [Online]. Available: https://www.irena.org/financeinvestment.

[21] Réseau de transport d'électricité, *Bilan electrique 2018*, 2018. [Online]. Available: https://www.rte-france.com/sites/default/files/be_pdf_2018v3.pdf.

[22] M. Aneke and M. Wang, "Energy storage technologies and real life applications – a state of the art review", *Applied Energy*, vol. 179, pp. 350–377, 2016.

[23] A. Barbaux, *Pourquoi edf n'investit pas plus dans les step pour le stockage des énergies renouvelables*, 2017. [Online]. Available: https://www.usinenouvelle.com/article/pourquoi-edf-n-investit-pas-plus-dans-les-step-pour-le-stockage-des-energies-renouvelables.N620998.

[24] ARES, *Grid scale energy storage*, 2019. [Online]. Available: https://www.aresnorthamerica.com/grid-scale-energy-storage.

[25] Energy Vault, *Energy vault*, 2019. [Online]. Available: https://energyvault.com/.

[26] Commission de Régulation de l'Energie, *Observatoire des marchés de détail du 4e trimestre 2018*, 2019. [Online]. Available: https://www.cre.fr/Documents/Publications/Observatoire-des-marches/Observatoire-des-marches-de-detail-du-4e-trimestre-2018.

[27] S. Amelang and K. Appunn, *The causes and effects of negative power prices*, 2018. [Online]. Available: https://www.cleanenergywire.org/factsheets/why-power-prices-turn-negative.

[28] Legifrance, *Loi n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte*, 2015. [Online]. Available: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031044385&categorieLien=id.

[29] Commission de régulation de l'énergie, *Services système et mécanisme d'ajustement*, 2018. [Online]. Available: https://www.cre.fr/Electricite/Reseaux-d-electricite/Services-systeme-et-mecanisme-d-ajustement.

[30] P. Pinson, "Wind energy: forecasting challenges for its operational management", *Statistical Science*, pp. 564–585, 2013.

[31] Met Office, *Global cells*, 2019. [Online]. Available: https://www.metoffice.gov.uk/binaries/content/gallery/mohippo/images/migrated-image/f/figure-4-global-cells%28edit%292.jpg.

[32] Clay Mathematics Institute, *Navier–stokes equation*, 2019. [Online]. Available: https://www.claymath.org/millennium-problems/navier%5C%E2%5C%80%5C%93stokes-equation.

[33] S. E. Tuller and A. C. Brett, "The characteristics of wind velocity that favor the fitting of a weibull distribution in wind speed analysis", *Journal of Climate and Applied Meteorology*, vol. 23, no. 1, pp. 124–134, 1984.

[34] Wind Energy - The Facts, *The annual variability of wind speed*, 2019. [Online]. Available: https://www.wind-energy-the-facts.org/the-annual-variability-of-wind-speed.html.

[35] Wind Energy Department of the Technical University of Denmark, *Global wind atlas*, 2019. [Online]. Available: https://globalwindatlas.info/.

[36] Cermak Peterka Petersen, *Wind profile characterization*, 2015. [Online]. Available: https://www.cppwind.com/blogs/wind-profile-characterization.

[37] Wind Energy Department of the Technical University of Denmark, *Global wind atlas methodology*, 2019. [Online]. Available: https://globalwindatlas.info/about/method.

[38] Multi Body, *Wind turbine blade*, 2016. [Online]. Available: http://www.multibody.net/teaching/msms/teachingmsmsstudents-projects-2016/pregrasso-matteo/.

[39] Futuren, *L'énergie éolienne, comment ça marche ?*, 2019. [Online]. Available: http://www.futuren-group.com/lenergie-eolienne/comment-ca-marche.

[40] W. Tong, *Wind power generation and wind turbine design*. WIT press, 2010.

[41] Wikipedia, *Moulin saint-elzéar de montfuron*, 2019. [Online]. Available: https://fr.wikipedia.org/wiki/Moulin_Saint-Elz%5C%C3%5C%A9ar_de_Montfuron.

[42] ——, *Windpump*, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Windpump.

[43] Quiet Revolution, *Photos*, 2019. [Online]. Available: https://www.quietrevolution.com/photos/.

[44] Solutions alternatives, *Éolienne savonius*, 2019. [Online]. Available: https://solutions-alternatives.org/base-de-connaissances/eolienne-savonius/.

[45] Vergnet, *Gev mp c*, 2019. [Online]. Available: http://www.vergnet.com/wp-content/uploads/2016/04/dc-11-00-01-fr_gev_mp-c_275_kw.pdf.

[46] Orsted, *World's largest offshore wind farm officially unveiled*, 2019. [Online]. Available: https://walneyextension.co.uk/News/2018/09/Worlds-largest-offshore-wind-farm-officially-unveiled.

[47] Wikipedia, *Floating wind turbine*, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Floating_wind_turbine.

[48] Sean Gallup, *Germany invests in renewable energy sources*, 2019.

[49] A. Betz, *Introduction to the theory of flow machines*. Oxford: Pergamon Press, 1966.

[50] T. Burton, N. Jenkins, D. Sharpe, and E. Bossanyi, *Wind energy handbook*. John Wiley & Sons, 2011.

[51] The Wind Power, *Database*, 2016. [Online]. Available: https://www.thewindpower.net/index.php.

[52] General Electric Renewable Energy, *Haliade-x offshore wind turbine platform*, 2019. [Online]. Available: https://www.ge.com/renewableenergy/wind-energy/offshore-wind/haliade-x-offshore-turbine.

[53] L. Bird, J. Cochran, and X. Wang, *Wind and solar energy curtailment: Experience and practices in the United States*. National Renewable Energy Laboratory, 2014.

[54] Research Hubs, *Pitch-regulated and stall-regulated wind turbine*, 2015. [Online]. Available: http://researchhubs.com/post/engineering/wind-energy/pitch-regulated-and-stall-regulated-wind-turbine.html.

[55] F. Grasso, D. Coiro, N. Bizzarrini, and G. Calise, "Design of advanced airfoil for stall-regulated wind turbines", in *Journal of Physics: Conference Series*, IOP Publishing, vol. 753, 2016, p. 022 008.

[56] N. G. Nygaard, "Wakes in very large wind farms and the effect of neighbouring wind farms", *Journal of Physics: Conference Series*, vol. 524, no. 1, 2014.

[57] Vattenfall, *Horns rev 1, systerpark till horns rev 3*, 2016. [Online]. Available: https://news.cision.com/vattenfall/i/horns-rev-1--systerpark-till-horns-rev-3,c1912031.

[58] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting", *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762–777, 2014.

[59] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation", *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[60] T. Gneiting, "Probabilistic forecasting", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 171, no. 2, pp. 319–321, 2008.

[61] R. Bessa, C. Möhrlen, V. Fundel, M. Siefert, J. Browell, S. Haglund El Gaidi, B.-M. Hodge, U. Cali, and G. Kariniotakis, "Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry", *Energies*, vol. 10, no. 9, p. 1402, 2017.

[62] C. Graham, "The parameterisation and prediction of wave height and wind speed persistence statistics for oil industry operational planning purposes", *Coastal Engineering*, vol. 6, no. 4, pp. 303–329, 1982.

[63] S. Kuwashima and N. Hogben, "The estimation of wave height and wind speed persistence statistics from cumulative probability distributions", *Coastal Engineering*, vol. 9, no. 6, pp. 563–590, 1986.

[64] P. Pinson and H. Madsen, "Ensemble-based probabilistic forecasting at horns rev", *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, vol. 12, no. 2, pp. 137–155, 2009.

[65] ——, "Adaptive modelling and forecasting of offshore wind power fluctuations with markov-switching autoregressive models", *Journal of forecasting*, vol. 31, no. 4, pp. 281–313, 2012.

[66] K. Rogers, J. Collins, J. Parkes, and L. Landberg, "Wind power forecasting offshore, more or less accurate than onshore?", 2012.

[67] E. N. Lorenz, "Deterministic nonperiodic flow", *Journal of the atmospheric sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[68] J. Coiffier, *Fundamentals of numerical weather prediction*. Cambridge University Press, 2011.

[69] P. Bauer, A. Thorpe, and G. Brunet, "The quiet revolution of numerical weather prediction", *Nature*, vol. 525, no. 7567, p. 47, 2015.

[70] Météo France, *Glossaire: maille*, 2019. [Online]. Available: `http://www.meteofrance.fr/publications/glossaire/152616-maille`.

[71] NOAA (National centers for environmental information), *Global forecast system (gfs)*, 2019. [Online]. Available: `https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs`.

[72] ECMWF (European Centre for Medium-Range Weather Forecasts), *Documentation and support*, 2019. [Online]. Available: `https://www.ecmwf.int/en/forecasts/documentation-and-support`.

[73] Météo France, *Données de modèle atmosphérique global*, 2019. [Online]. Available: `https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=130&id_rubrique=51`.

[74] ——, *Données de modèle atmosphérique à aire limitée à haute résolution*, 2019. [Online]. Available: `https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=131&id_rubrique=51`.

[75] J. F. Manwell, J. G. McGowan, and A. L. Rogers, *Wind energy explained: theory, design and application*. John Wiley & Sons, 2010.

[76] J. D. Holmes, *Wind loading of structures*. CRC press, 2018.

[77] T. R. Oke, *Boundary layer climates*. Routledge, 2002.

[78] G. Kariniotakis, P. Pinson, N. Siebert, G. Giebel, R. Barthelmie, *et al.*, "The state of the art in short-term prediction of wind power-from an offshore perspective", in *Proceedings of*, 2004, pp. 20–21.

[79] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power", *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.

[80] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction", *Renewable and Sustainable Energy Reviews*, vol. 12, no. 6, pp. 1725–1744, 2008.

[81] L. Landberg, G. Giebel, H. A. Nielsen, T. Nielsen, and H. Madsen, "Short-term prediction—an overview", *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, vol. 6, no. 3, pp. 273–280, 2003.

[82] N. Siebert, "Development of methods for regional wind power forecasting", PhD thesis, École Nationale Supérieure des Mines de Paris, 2008.

[83] L. Landberg, "Short-term prediction of the power production from wind farms", *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 80, no. 1-2, pp. 207–220, 1999.

[84] U. Focken, M. Lange, and H.-P. Waldl, "Previento-a wind power prediction system with an innovative upscaling algorithm", in *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark*, Citeseer, vol. 276, 2001.

[85] I. Martı, D. Cabezón, J. Villanueva, M. J. Sanisidro, Y. Loureiro, E. Cantero, and J. Sanz, "Localpred and regiopred. advanced tools for wind energy prediction in complex terrain", in *Proceedings of the European Wind Energy Conference EWEC'03*, 2003, pp. 16–19.

[86] J. Zack, M. Brower, and B. Bailey, "Validating of the forewind model in wind forecasting applications", in *Talk on the EUWEC Special Topic Conference Wind Power for the 21st Century, Kassel, Germany*, 2000, pp. 25–27.

[87] E. Alpaydin, *Introduction to machine learning*. MIT press, 2009.

[88] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[89] P. Ailliot and V. Monbet, "Markov-switching autoregressive models for wind time series", *Environmental Modelling & Software*, vol. 30, pp. 92–101, 2012.

[90] J. P. d. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Short-term wind power forecasting in portugal by neural networks and wavelet transform", *Renewable energy*, vol. 36, no. 4, pp. 1245–1251, 2011.

[91] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.

[92] J. Zeng and W. Qiao, "Support vector machine-based short-term wind power forecasting", in *2011 IEEE/PES Power Systems Conference and Exposition*, IEEE, 2011, pp. 1–8.

[93] J. Zhou, J. Shi, and G. Li, "Fine tuning support vector machines for short-term wind speed forecasting", *Energy Conversion and Management*, vol. 52, no. 4, pp. 1990–1998, 2011.

[94] G. Santamarıa-Bonfil, A. Reyes-Ballesteros, and C. Gershenson, "Wind speed forecasting for wind farms: a method based on support vector regression", *Renewable Energy*, vol. 85, pp. 790–809, 2016.

[95] S. S. Haykin *et al.*, *Neural networks and learning machines/Simon Haykin*. New York: Prentice Hall, 2009.

[96] M. L. Minsky and S. A. Papert, "Perceptrons: expanded edition", 1988.

[97] D. Dıaz, A. Torres, and J. R. Dorronsoro, "Deep neural networks for wind energy prediction", in *International Work-Conference on Artificial Neural Networks*, Springer, 2015, pp. 430–443.

[98] H.-z. Wang, G.-q. Li, G.-b. Wang, J.-c. Peng, H. Jiang, and Y.-t. Liu, "Deep learning based ensemble approach for probabilistic wind power forecasting", *Applied energy*, vol. 188, pp. 56–70, 2017.

[99] A. Ghaderi, B. M. Sanandaji, and F. Ghaderi, "Deep forecast: deep learning-based spatio-temporal forecasting", *arXiv preprint arXiv:1707.08110*, 2017.

[100] Q. Cao, B. T. Ewing, and M. A. Thompson, "Forecasting wind speed with recurrent neural networks", *European Journal of Operational Research*, vol. 221, no. 1, pp. 148–154, 2012.

[101] Z. Liu, W. Gao, Y.-H. Wan, and E. Muljadi, "Wind power plant prediction by using neural networks", in *2012 IEEE energy conversion congress and exposition (ECCE)*, IEEE, 2012, pp. 3154–3160.

[102] S. Balluff, J. Bendfeld, and S. Krauter, "Short term wind and energy prediction for offshore wind farms using neural networks", in *2015 International Conference on Renewable Energy Research and Applications (ICRERA)*, IEEE, 2015, pp. 379–382.

[103] Z. O. Olaofe, "A 5-day wind speed & power forecasts using a layer recurrent neural network (lrnn)", *Sustainable Energy Technologies and Assessments*, vol. 6, pp. 1–24, 2014.

[104] E. López, C. Valle, H. Allende, E. Gil, and H. Madsen, "Wind power forecasting based on echo state networks and long short-term memory", *Energies*, vol. 11, no. 3, p. 526, 2018.

[105] A. S. Qureshi, A. Khan, A. Zameer, and A. Usman, "Wind power prediction using deep neural network based meta regression and transfer learning", *Applied Soft Computing*, vol. 58, pp. 742–755, 2017.

[106] Z.-H. Zhou, *Ensemble methods: foundations and algorithms*. Chapman and Hall/CRC, 2012.

[107] L. Breiman, "Bagging predictors", *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[108] A. Lahouar and J. B. H. Slama, "Hour-ahead wind power forecast based on random forests", *Renewable energy*, vol. 109, pp. 529–541, 2017.

[109] T. K. Ho, "Random decision forests", in *Proceedings of 3rd international conference on document analysis and recognition*, IEEE, vol. 1, 1995, pp. 278–282.

[110] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[111] J. H. Friedman, "Greedy function approximation: a gradient boosting machine", *Annals of statistics*, pp. 1189–1232, 2001.

[112] J. Wu, B. Zhang, and K. Wang, "Application of adaboost-based bp neural network for short-term wind speed forecast", *Power System Technology*, vol. 36, no. 9, pp. 221–225, 2012.

[113] Y. Ren, X. Qiu, and P. N. Suganthan, "Empirical mode decomposition based adaboost-backpropagation neural network method for wind speed forecasting", in *2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*, IEEE, 2014, pp. 1–6.

[114] J. P. Heinermann, "Wind power prediction with machine learning ensembles", PhD thesis, Carl von Ossietzky Universität Oldenburg, 2016.

[115] M. Landry, T. P. Erlinger, D. Patschke, and C. Varrichio, "Probabilistic gradient boosting machines for gefcom2014 wind forecasting", *International Journal of Forecasting*, vol. 32, no. 3, pp. 1061–1066, 2016.

[116] T. Hong, P. Pinson, and S. Fan, *Global energy forecasting competition 2012*, 2014.

[117] B. Gorman, *A kaggler's guide to model stacking in practice*, 2016. [Online]. Available: `http://blog.kaggle.com/2016/12/27/a-kagglers-guide-to-model-stacking-in-practice/`.

[118] B. Himmetoglu, *Stacking models for improved predictions*, 2016. [Online]. Available: `https://burakhimmetoglu.com/2016/12/01/stacking-models-for-improved-predictions/`.

[119] M. Abkar and F. Porté-Agel, "A new wind-farm parameterization for large-scale atmospheric models", *Journal of Renewable and Sustainable Energy*, vol. 7, no. 1, p. 013 121, 2015.

[120] M. L. Aitken, B. Kosović, J. D. Mirocha, and J. K. Lundquist, "Large eddy simulation of wind turbine wake dynamics in the stable boundary layer using the weather research and forecasting model", *Journal of Renewable and Sustainable Energy*, vol. 6, no. 3, p. 033 137, 2014.

[121] M. Calaf, C. Meneveau, and J. Meyers, "Large eddy simulation study of fully developed wind-turbine array boundary layers", *Physics of fluids*, vol. 22, no. 1, p. 015 110, 2010.

[122] M. J. Churchfield, S. Lee, J. Michalakes, and P. J. Moriarty, "A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics", *Journal of turbulence*, no. 13, N14, 2012.

[123] V. Sharma, M. Calaf, M. Lehning, and M. Parlange, "Time-adaptive wind turbine model for an les framework", *Wind Energy*, vol. 19, no. 5, pp. 939–952, 2016.

[124] A. Jimenez, A. Crespo, E. Migoya, and J. Garcıa, "Advances in large-eddy simulation of a wind turbine wake", in *Journal of Physics: Conference Series*, IOP Publishing, vol. 75, 2007, p. 012 041.

[125] J. Browell, C. Gilbert, and D. McMillan, "Use of turbine-level data for improved wind power forecasting", in *2017 IEEE Manchester PowerTech*, 2017, pp. 1–6.

[126] R. Girard, K. Laquaine, and G. Kariniotakis, "Assessment of wind power predictability as a decision factor in the investment phase of wind farms", *Applied Energy*, vol. 101, pp. 609–617, 2013.

[127] P.-J. Trombe, P. Pinson, and H. Madsen, "A general probabilistic forecasting framework for offshore wind power fluctuations", *Energies*, vol. 5, no. 3, pp. 621–657, 2012.

[128] F. Davò, S. Alessandrini, S. Sperati, L. Delle Monache, D. Airoldi, and M. T. Vespucci, "Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting", *Solar Energy*, vol. 134, pp. 327–338, 2016.

[129] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions", *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1571–1580, 2017.

[130] Météo Chailles, *Vents dominants*, 2019. [Online]. Available: `https://jmj41.com/meteo/climat_vent_dominant.php`.

[131] J. Ferber, *Multi-agent systems: an introduction to distributed artificial intelligence*. Addison-Wesley Reading, 1999, vol. 1.

[132] G. Weiss, *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press, 1999.

[133] G. D. M. Serugendo, M.-P. Gleizes, and A. Karageorgos, *Self-organising Software*. Springer, 2011.

[134] D. Weyns, H. V. D. Parunak, F. Michel, T. Holvoet, and J. Ferber, "Environments for multiagent systems state-of-the-art and research challenges", in *International Workshop on Environments for Multi-Agent Systems*, Springer, 2004, pp. 1–47.

[135] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited, 2016.

[136] J. Lind, *Iterative software engineering for multiagent systems: the MASSIVE method*. Springer-Verlag, 2001.

[137] C. Paniah, "Approche multi-agents pour la gestion des fermes éoliennes offshore", PhD thesis, Paris 11, 2015.

[138] R. Roche, F. Lauri, B. Blunier, A. Mirao, and A. Koukam, "Chapter 16 multi-agent technology for power system control", 2012.

[139] R. Roche, B. Blunier, A. Miraoui, V. Hilaire, and A. Koukam, "Multi-agent systems for grid energy management: a short review", in *IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society*, IEEE, 2010, pp. 3341–3346.

[140] L. Hernández, C. Baladron, J. M. Aguiar, B. Carro, A. Sanchez-Esguevillas, J. Lloret, D. Chinarro, J. J. Gomez-Sanz, and D. Cook, "A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants", *IEEE Communications Magazine*, vol. 51, no. 1, pp. 106–113, 2013.

[141] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings, "Agent-based micro-storage management for the smart grid", in *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1-Volume 1*, International Foundation for Autonomous Agents and Multiagent Systems, 2010, pp. 39–46.

[142] S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. Jennings, "Agent-based control for decentralised demand side management in the smart grid", in *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, International Foundation for Autonomous Agents and Multiagent Systems, 2011, pp. 5–12.

[143] A. González-Briones, P. Chamoso, F. De La Prieta, Y. Demazeau, and J. Corchado, "Agreement technologies for energy optimization at home", *Sensors*, vol. 18, no. 5, p. 1633, 2018.

[144] M. Sahnoun, D. Baudry, N. Mustafee, A. Louis, P. A. Smart, P. Godsiff, and B. Mazari, "Modelling and simulation of operation and maintenance strategy for offshore wind farms based on multi-agent system", *Journal of Intelligent Manufacturing*, pp. 1–17, 2015.

[145] M. Kpakpo, M. Itmi, and A. Cardon, "A MAS model approach to a wind farm maintenance strategy", in *Proceedings of the 10th International Conference on Agents and Artificial Intelligence, ICAART 2018, Volume 1, Funchal, Madeira, Portugal, January 16-18, 2018.*, 2018, pp. 159–167. DOI: 10.5220/0006554501590167. [Online]. Available: https://doi.org/10.5220/0006554501590167.

[146] A. Zaher and S. McArthur, "A multi-agent fault detection system for wind turbine defect recognition and diagnosis", in *2007 IEEE Lausanne Power Tech*, IEEE, 2007, pp. 22–27.

[147] J. Albouys-Perrois, N. Sabouret, Y. Haradji, M. Schumann, and C. Inard, "Simulation multi-agent de l'autoconsommation collective en relation avec l'activité des foyers", in *Journées Francophones sur les Systèmes Multi-Agents*, 2019, pp. 116–125.

[148] P. A. Frensch and J. Funke, *Definitions, traditions, and a general framework for understanding complex problem solving*. Universitätsbibliothek der Universität Heidelberg, 1995.

[149] J. Funke, "Complex problem solving: a case for complex cognition?", *Cognitive processing*, vol. 11, no. 2, pp. 133–142, 2010.

[150] J. C. Polkinghorne, "Reductionism", *Interdisciplinary Encyclopedia of Religion and Science*, p. 10, 2002.

[151] J. Ladyman, J. Lambert, and K. Wiesner, "What is a complex system?", *European Journal for Philosophy of Science*, vol. 3, no. 1, pp. 33–67, 2013.

[152] T. De Wolf and T. Holvoet, "Emergence and self-organisation: a statement of similarities and differences", in *Proceedings of the International Workshop on Engineering Self-Organising Applications 2004*, 2004, pp. 96–110.

[153] V. Camps, M.-P. Gleizes, and P. Glize, "A self-organization process based on cooperation theory for adaptive artificial systems", in *Problems of Evolution in Real and Virtual Systems: Proceedings of the First International Conference on Philosophy and Computer Science*, 1998, p. 10.

[154] N. R. Jennings, "Agent-oriented software engineering", in *European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, Springer, 1999, pp. 1–7.

[155] P. Glize, "L'adaptation des systèmes à fonctionnalité émergente par auto-organisation coopérative", *Hdr, Université Paul Sabatier, Toulouse III*, 2001.

[156] V. Noël and F. Zambonelli, "Methodological guidelines for engineering self-organization and emergence", in *Software Engineering for Collective Autonomic Systems*, Springer, 2015, pp. 355–378.

[157] J.-P. Georgé, "Résolution de problèmes par émergence: étude d'un environnement de programmation émergente", PhD thesis, Toulouse 3, 2004.

[158] M.-P. Gleizes, V. Camps, and P. Glize, "A theory of emergent computation based on cooperative self-organization for adaptive artificial systems", in *Fourth European Congress of Systems Science*, 1999, pp. 20–24.

[159] J.-P. Georgé, B. Edmonds, and P. Glize, "Making self-organising adaptive multiagent systems work", in *Methodologies and Software Engineering for Agent Systems*, Springer, 2004, pp. 321–340.

[160] N. Bonjean, W. Mefteh, M. P. Gleizes, C. Maurel, and F. Migeon, "Adelfe 2.0", in *Handbook on agent-oriented design processes*, Springer, 2014, pp. 19–63.

[161] G. Picard, "Méthodologie de développement de systèmes multi-agents adaptatifs et conception de logiciels à fonctionnalité émergente", PhD thesis, 2004.

[162] N. Bonjean, *Ingénierie des systèmes multi-agents adaptatifs: vers un guide pour la conception du comportement d'agent coopératif*, 2009.

[163] M.-P. Gleizes, "Self-adaptive complex systems", in *European Workshop on Multi-Agent Systems*, Springer, 2011, pp. 114–128.

[164] D. Capera, "Systèmes multi-agents adaptatifs pour la résolution de problãmes application à la conception de mécanismes", PhD thesis, Ph. D. thesis, Université de Toulouse, Toulouse, France, 2005.

[165] J. Nigon, E. Glize, D. Dupas, F. Crasnier, and J. Boes, "Use cases of pervasive artificial intelligence for smart cities challenges", in *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)*, IEEE, 2016, pp. 1021–1027.

[166] J. Nigon, "Apprentissage artificiel adapté aux systèmes complexes par auto-organisation coopérative de systèmes multi-agents", PhD thesis, Université Paul Sabatier de Toulouse III, 2017.

[167] J.-P. Georgé, M.-P. Gleizes, P. Glize, and C. Régis, "Real-time simulation for flood forecast: an adaptive multi-agent system staff", in *Proceedings of the AISB*, vol. 3, 2003, pp. 109–114.

[168] A. Perles, "An adaptive multi-agent system for the distribution of intelligence in electrical distribution networks: state estimation", PhD thesis, Paul Sabatier University, Toulouse, France, 2017.

[169] J.-B. Blanc-Rouchossé, A. Blavette, G. Camilleri, and M.-P. Gleizes, "Electric vehicles fleet for frequency regulation using a multi-agent system", in *International Conference on Practical Applications of Agents and Multi-Agent Systems*, Springer, 2018, pp. 84–96.

[170] V. Guivarch, C. Bernon, and M. P. Gleizes, "Power optimization by cooling photovoltaic plants as a dynamic self-adaptive regulation problem", in *ICAART*, 2018, pp. 276–281.

[171] D. Guastella, V. Camps, and M.-P. Gleizes, "Multi-agent systems for estimating missing information in smart cities", Feb. 2019. DOI: `10.5220/0007381902140223`.

[172] M. Landry, T. P. Erlinger, D. Patschke, and C. Varrichio, “Probabilistic gradient boosting machines for gefcom2014 wind forecasting”, *International Journal of Forecasting*, vol. 32, no. 3, pp. 1061–1066, 2016.

[173] G. Ridgeway, “The state of boosting”, *Computing Science and Statistics*, pp. 172–181, 1999.

[174] A. Perles, F. Crasnier, and J.-P. Georgé, “Amak-a framework for developing robust and open adaptive multi-agent systems”, in *International Conference on Practical Applications of Agents and Multi-Agent Systems*, Springer, 2018, pp. 468–479.

[175] J. VanderPlas, *Python data science handbook: essential tools for working with data*. " O’Reilly Media, Inc.", 2016.

[176] Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson, “The arome-france convective-scale operational model”, *Monthly Weather Review*, vol. 139, no. 3, pp. 976–991, 2011.

[177] Météo-France, *Données publiques*, 2019. [Online]. Available: `https://donneespubliques.meteofrance.fr/`.

[178] B. Lu, Y. Li, X. Wu, and Z. Yang, “A review of recent advances in wind turbine condition monitoring and fault diagnosis”, in *2009 IEEE Power Electronics and Machines in Wind Applications*, IEEE, 2009, pp. 1–7.

[179] T. Hastie, R. Tibshirani, J. Friedman, and J. Franklin, “The elements of statistical learning: data mining, inference and prediction”, *The Mathematical Intelligencer*, vol. 27, no. 2, pp. 83–85, 2005.

[180] K. Rosaen, *K-fold cross-validation*, 2016. [Online]. Available: `http://karlrosaen.com/ml/learning-log/2016-06-20/`.

[181] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, “Scikit-learn: machine learning in python”, *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[182] scikit-learn, *Linear regression documentation*, 2019. [Online]. Available: `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html`.

[183] ——, *Mlp regressor documentation*, 2019. [Online]. Available: `https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html`.

[184] ——, *Random forest regressor documentation*, 2019. [Online]. Available: `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html`.

[185] ——, *Gradient boosting regressor documentation*, 2019. [Online]. Available: `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html`.

[186] ——, *Svr documentation*, 2019. [Online]. Available: `https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html`.

[187] T. Bouziat, "A cooperative architecting procedure for systems of systems based on self-adaptive multi-agent systems", PhD thesis, Université de Toulouse, 2017.

List of Figures

1.1 World electricity generation mix in 2018 [8]. 12
1.2 Atmospheric CO_2 concentration estimated from gas microbubbles trapped in ice cores [15]. 13
1.3 IEA World electricity generation scenarios. 14
1.4 French primary energy balance of the negaWatt scenario [16]. 14
1.5 Percentage of renewable energy in electricity generation by country in 2018 [18]. 16
1.6 Global trends in renewable energy investment [20]. 16
1.7 Example of market clearing [27]. 20

2.1 Representation of the three main convective cells: Polar, Ferrel and Hadley cells [31]. 25
2.2 Example of wind speed distribution [34]. 26
2.3 World wind speed potential [35]. 27
2.4 Difference of wind profiles according to topography [36]. 28
2.5 The vertical profile of wind speed upwind and on top of a hill [37]. 28
2.6 Diagram of the forces acting on a wind turbine blade [38]. 30
2.7 Wind turbine schematic [39]. 30
2.8 (a) Windmill, (b) Wind water pump, (c) Darrieus, (d) Savonius, (e) Two-blade turbine, (f) Offshore, (g) Floating, (h) Standard [41]–[48]. 31
2.9 A theoretical power curve compared with the observed production as a function of the wind speed measurement at the nacelle. 33
2.10 Number of wind turbines as a function of P_{nom} in 2016 in Europe [51]. 34
2.11 Pitch-regulated and Stall-regulated wind turbine power curves [54]. 35
2.12 Example of wake effect [57]. 36
2.13 Hysteresis phenomenon applied to wind generation. 36
2.14 Physical approach process for the wind power forecasting. 43
2.15 General process of using a Machine Learning model for wind power forecasting. 45

2.16 Comparison between the production of two close and distant wind turbines in the same wind farm. . . . 49

3.1 The layout of a wind farm consisting of 15 wind turbines. . . . 58
3.2 Comparison of the production between two close (left) and distant turbines (right). . . . 58
3.3 Pearson correlation coefficient between the production of each turbine in a wind farm over three years of data. . . . 59
3.4 Comparison of production between two neighboring turbines T1 and T2. . . . 59
3.5 Prevailing winds in France [130]. . . . 60
3.6 Location of the five wind farms studied. . . . 61
3.7 Pearson correlation coefficient between the production of the five French wind farms over three years of data. . . . 61

4.1 Links between WT/GP and WTH/GPH entities. . . . 76
4.2 Architecture of the AMAWind-Turbine system and relationships between entities and agents. . . . 77
4.3 Neighborhood between WTH agents for an example of a layout. . . . 78
4.4 Example of cooperative behavior of a WTH agent. . . . 80
4.5 Example of a non cooperative situation of type *Uselessness*. . . . 81
4.6 Example of a non cooperative situation of type *Incompetence*. . . . 82
4.7 Friedman's gradient boosting algorithm [173]. . . . 84
4.8 Example of probabilistic wind power forecast divided into percentiles. . . . 84
4.9 Local criticality function example. . . . 85
4.10 Example of probabilistic wind power forecast difference between two wind turbines divided into percentiles. . . . 86
4.11 Neighboring criticality functions example. . . . 87
4.12 Example of a final criticality function for a WTH agent with a single neighbor. 87
4.13 Global class diagram based on AMAK. . . . 88
4.14 Class diagram representing the links between entities and agents in AMAWind-Turbine. . . . 89

5.1 Links between F/GP and FH/GPH entities. . . . 93
5.2 Architecture of the AMAWind-Farm system and relationships between entities and agents. . . . 94
5.3 Neighborhood of Farm Hour (FH) agents. . . . 94
5.4 Class diagram representing the links between entities and agents in AMAWind-Farm. . . . 96

5.5 Example of the normalization of a criticality function with $\varphi = 0.5$ 97

6.1 The layout of the five studied wind farms at the same scale. 103
6.2 Power curves of the wind turbines of the five wind farms studied. 103
6.3 Distribution of the measured wind and production of one wind turbine in each farm. 104
6.4 Power curve of the wind farm C with final and filtered records. 106
6.5 10-fold cross validation process [180]. 108

7.1 Forecast and criticality evolution, as functions of the cycle number, on the fifteen turbines of the farm A. 112
7.2 Example of wind power forecasts provided by AMAWind-Turbine over a set of days. 113
7.3 Criticality evolution. 116
7.4 Relationship between error improvement and criticality improvement (average over 5% intervals). 117
7.5 Distribution histograms of number of cycles on AMAWind-Turbine. 118

8.1 Criticality and forecast evolution as a function of the cycles on five Farm Hour agents. 128
8.2 The four scenarios tested and the factors φ applied to subcriticalities. 130
8.3 Criticality evolution during the operation of the AMAWind-Farm system. . . 132
8.4 Relation between error improvement and criticality improvement on AMAWind-Farm (average over 5% intervals). 133
8.5 Distribution histograms of number of cycles required on AMAWind-Farm. . . 133

List of Tables

1.1 Estimated number of years of production at this rate by source [12], [14]. . . . 12
1.2 World electricity production by renewable source [12]. 15

2.1 The different types of wind power forecasts according to time scale. 38
2.2 Comparison of several NWP models [71], [72], [73], [74]. 41
2.3 Assessment of the criteria for physical models. 51
2.4 Assessment of the criteria for statistical models. 52
2.5 Wind power forecasting methods summary. 52

6.1 Characteristics of the wind farms studied. 102
6.2 Size of wind farm production history before and after filtering. 106

7.1 Comparison of computation time according to the initial forecast used (HH:MM). 114
7.2 Summary of computation times for each farm on AMAWind-Turbine. 114
7.3 Average first and last cycle criticality for each farm. 115
7.4 NMAE of the different machine learning methods for the five wind farms. . . 121
7.5 Results in terms of NMAE and NRMSE for the five farms. 122
7.6 Detailed results for each subset of the 10-fold cross-validation (NMAE). . . . 123
7.7 Detailed results for each subset of the 10-fold cross validation (NRMSE). . . . 124

8.1 Validation indicators according to the normalization applied. 130
8.2 Comparison of computation times on AMAWind-Farm according to the initial forecast used (HH:MM). 131
8.3 Average first and last cycle criticality on AMAWind-Farm. 132
8.4 Results summary on AMAWind-Farm. 134
8.5 Detailed results on AMAWind-Farm in terms of NMAE (%). 135
8.6 Detailed results on AMAWind-Farm in terms of NRMSE (%). 136

Glossary

AMAS – *Adaptive Multi-Agent System*

AMAW-F – *refers to AMAWind-Farm*

AMAW-T – *refers to AMAWind-Turbine*

AMAWind-Farm – *Adaptive Multi-Agent system for Wind power forecasting at Farm-level*

AMAWind-Turbine – *Adaptive Multi-Agent system for Wind power forecasting at Turbine-level*

CWS – *Conditional Weighted Sum*

DSO – *Distribution System Operator*

F – *Farm entity*

FH – *Farm Hour agent*

GBM – *Gradient Boosting Machine*

GBM-F – *GBM at Farm level*

GBM-R – *GBM at Regional level*

GP – *Grid Point entity*

GPH – *Grid Point Hour entity*

ISP– *Imbalance Settlement Price*

NCS – *Non Cooperative Situation*

NMAE – *Normalized Mean Absolute Error*

NRMSE – *Normalized Root Mean Squared Error*

NWP – *Numerical Weather Prediction*

MAS – *Multi-Agent System*

PC – *Power Curve*

TSO – *Transmission System Operator*

WT – *Wind Turbine entity*

WTH – *Wind Turbine Hour agent*