SafeWind	Collaborative project funded by the European Commission under the 7th Framework Program, Theme 2007-2.3.2: Energy
"Multi-scale data assimilation, advanced wind modelling & forecasting with emphasis to extreme weather situations for a safe large-scale wind power integration"	Grant Agreement N°: 213740

"Dp-7.3 : Study of the economic impact of the level of predictability on selected wind farms"

DOCUMENT TYPE	deliverable
DOCUMENT NAME:	Swind_deliverableDp-7.3_ARMINES.pdf
VERSION:	V1.1
DATE:	2012-01-11
CLASSIFICATION:	R0: Public
STATUS:	Reviewed

Abstract: **This document is the contribution of ARMINES to Task 7.1 and Deliverable Dp7.3. The impact of predictability is evaluated through the day ahead market participation.**

AUTHORS[1], REVIEWERS				
MAIN AUTHOR/EDITOR:	K. Laquaine			
AFFILIATION:	ARMINES			
ADDRESS:	B.P. No 207, F-06904 Sophia-Antipolis Cedex, France.			
TEL.:				
EMAIL:	robin.girard@mines-paristech.fr			
FURTHER AUTHORS:	R.Girard, G. Kariniotakis			
PEER REVIEWERS:	R. Sanz			
REVIEW APPROVAL:	Approved :		Rejected (improve as indicated below) :	
SUGGESTED IMPROVEMENTS:	*For a long list of remarks make reference to another document*			

VERSION HISTORY			
VERSION[2]:	DATE:	COMMENTS, CHANGES, STATUS:	PERSON(S):
1.0	26.10.2011	First version to be reviewed	R. Girard
1.1	11.01.2012	Adaptation of the document according to Javier's review	R. Girard

STATUS, CONFIDENTIALITY, ACCESSIBILITY										
STATUS:				CONFIDENTIALITY:				ACCESSIBILITY:		
S0	Approved/Released			**R0**	General public	**X**		Private web site		
S1	Reviewed	**X**		**R1**	Restricted to project members			Public web site	**X**	
S2	Pending for review			**R2**	Restricted to European Commission			Paper copy		
S3	Draft for comments			**R3**	Restricted to WP members + PL					
S4	Under preparation			**R4**	Restricted to Task members +WPL+PL					

1 **The authors of this document are solely responsible for its content, which does not represent the opinion of the European Community and the European Community is not responsible for any use that might be made of data appearing therein.**

2 **VERSION NAMING : V0.x** draft before peer-review approval, **V1.0** at the approval, **V1.x** minor revisions, **V2.0** major revision

Contents

Introduction **1**

1 State of the art in wind ressource assessment and benefit from market participation **4**
1.1 Wind resource assessment versus predictability 4
1.1.1 The wind ressource assessment process 4
1.1.2 Wind resource assessment and wind power predictability . 6
1.2 The value of predictability related to electricity market participation 8
1.2.1 Electricity markets 8
1.2.2 The value of predictability related to day-ahead market participation 10
1.2.3 Advanced strategies for hedging imbalance costs 11
1.3 The stakes for integrating predictability in wind resource assessment 12
1.3.1 The single case : "got no wind farm, want 1" 12
1.3.2 The aggregated case : "got n wind farm, want k" 13
1.4 Conclusion 13

2 Understanding the value of predictability : a case study over Denmark **14**
2.1 The value of predictability when installing a new wind farm . . . 14
2.1.1 Formulation of the decision-making problem 14
2.1.2 Scope of the study 16
2.1.3 Results and analysis 17
2.2 Investigating the influence of hedging solutions on the value of predictability 22
2.2.1 Formulation of the imbalance penalty resulting from hedging solutions 22
2.2.2 Results and analysis 24
2.3 Conclusion 26

3 The economical value of predictability in case of aggregated wind farms **28**
3.1 The spatial smoothing effect and the reduction of imbalance penalties 28
3.1.1 The spatial smoothing effect from aggregating wind farms 28
3.1.2 The portfolio effect : benefits from the market participation of aggregated wind farms 30

3.2 The economical value of predictability when aggregating several wind farms 30
3.2.1 Identifying the most relevant parameters of aggregation . 30
3.2.2 Spatial analysis of the effect of aggregation on predictability and market revenue 34
3.2.3 Results and analysis 36
3.3 Strategies for the aggregation of wind farms 37
3.3.1 Definition of different aggregation strategies 37
3.3.2 Results and analysis 40
3.4 Conclusion 41

Conclusion and Perspectives **43**

A Estimating predictability **45**
A.1 Methodology for estimating predictability 45
A.2 Results for a French wind farm 46

B Pearson's and Spearman's correlation formulas **51**
B.1 Linear correlation coefficient 51
B.2 Rank coefficient correlation 51

Introduction

In the upcoming years, renewable energy sources (RES) are expected to play an increasing role in the energy mix by replacing higher shares of conventional generation. As regards Europe, the European Commission has set a range of targets that are to be fulfilled by 2020, that is to say a 20% reduction of carbon emission, a 20% increase in energetic efficiency and an increase in the share of renewables in the final energy consumption up to 20%. However, those targets translate into a share of about 30%-40% of renewables in electricity generation, as far as the other energy sectors give only few growth perspectives. In particular, wind power is the fastest growing renewable energy source in Europe in terms of MW and EU wind power capacity is expected to reach at least 230 GW[1] in 2020, while present capacity equals 84,278 GW (see Table 1).

Such a large-scale integration of wind energy in power systems raises the challenge of managing large shares of wind power and increase the need for accurate wind power output prediction. It also entails that a significant number of new wind farms will be built across Europe. As the most windy sites are already occupied in certain countries (e.g. in Spain [18] and Germany [39]), it will lead to deal with more and more complex terrains, which often have lower predictability. That highlights the need for accurate wind resource assessment. Wind resource assessment refers to the use of numerical tools and experimental techniques in order to produce the wind field over a specific site topography, with the final aim of estimating both technical and economical feasibility of a new wind farm installation. Techniques for wind resource assessment experienced significant developments during the last few years, due to the growing use of ground-based remote sensing devices as a supplement to mast-based instrumentations and thanks to the development of mesoscale and microscale models.

In the early stages of wind power integration, wind power producers may benefit from a feed-in tariff system, which purchase cost-based prices, long-term contracts, guaranteed grid access and no need to care about predictability. But wind generation is intrinsically intermittent, highly variable and has limited dispatchability. When operating under market conditions, that results in penalizing power producers because electricity markets have been originally designed for power producers that can control their power production, which in fact is not the case of RES. Typically, power producers can participate in day-ahead markets, where they propose their quantity-price bids prior to the delivery time while having imperfect knowledge about what the future wind generation will be. Market clearing price is determined by the interaction of the aggregated supply and demand curves. Wind power producers are balance responsible par-

[1] $1GW = 10^3 MW = 10^6 kW$.

ties, as a result of which, they are charged for any imbalance between contracted and delivered energy. This can lead to significant amounts of imbalance costs that reduce wind power producers' income.

EU-27 wind power capacity (GW)	2010	2020 (low)	2020 (high)
Onshore	81.332	190	210
Offshore	2.946	40	55
Total	84.278	230	265

Table 1: Evolution of wind power installed capacity in EU countries. *Source*: EWEA, 2011.

Such a burden makes particularly challenging the estimation of predictability right from the resource assessment phase, when predictability is understood as the degree to which the wind power output can be given an accurate estimation. The question is : having a set of pre-selected sites wind farms with similar estimated yield, is it relevant to decide between those farms on the bases of their predictability? How much is the additional benefit for high predictability? Until now, the economical part of wind resource assessment processes has focused on wind energy potential as the main criterion when deciding to set up a new wind farm (especially the so-called Annual Energy Production criterion). When taking into account wind predictability, one might better assess not only the wind power potential, but also the market benefits that depend on asymmetric effects due to the regulation state of the market. Predictability can thus be considered as a system design parameter linked to the resource assessment phase, where the aim is to take optimal decisions for the installation of a new wind farm. This implies being able to provide rules for decision-makers.

The first objective of this report is to quantify, in terms of increased power producer's income, the effect of integrating predictability in the wind resource assessment phase, when power producers participate in electricity markets. We work with real-world production data from 182 Danish wind farms thus knowing both the wind power production output and the value for predictability, which in fact is not the case at the time of resource assessment. To validate this approach, it has to be assessed whether the pre-construction estimation of predictability is close to the predictability that is stated when the wind farm is operational. In other words, we ought to know if they are of the same magnitude the difference between the actual wind power and the predicted wind power from the moment the wind farm is producing, on one hand, and between the estimated wind power and the predicted wind power before the wind farm is constructed, on the other hand. For the rest of this work, prediction refers to the forecast of some variables using past power data, whereas estimation is only linked to the wind resource assessment process.

The second objective is to define rules that one could refer to when facing to a wind farm site choice decision-making problem, derived from the above-mentioned relations that were found to be stable enough. The aim is to be able to compose optimal portfolios that take advantage of the smoothing effect that arises from aggregation in case of the simultaneous installation of multiple wind farms. The present report is organized as follows:

- Chapter 1 gives an overview of the state of the art in wind resource assessment. It secondly explains the market framework and the value of

predictability in terms of market revenue. Then, two distinct cases of the installation of wind farms decision-making problem are defined.

- Chapter 2 investigates the relation between predictability and market revenue when installing a new wind farm. In particular, the influence of hedging solutions is highlighted. The results are from a set of wind farms spread over Denmark.
- Chapter 3 is about evaluating the economical value of predictability when aggregating several wind farms. Both the spatial smoothing effect and the portfolio effect that arises from the aggregation of wind farms are introduced. Aggregation is characterized by a set of parameters. Their influence is investigated, with emphasis on the spatial parameters, in order to derive optimal and sufficiently generic aggregation strategies.
- Finally, the general conclusions of this work are drawn as well as some perspectives.

Chapter 1

State of the art in wind ressource assessment and benefit from market participation

This chapter investigates the merits of integrating wind power predictability in the wind resource assessment phase. It first provides an overview of the literature linked with wind resource assessment and the benefit from the participation of wind farms in power exchange. Then the market pattern is concisely introduced and both ordinary and advanced strategies for the participation in electricity markets are described. The chapter finally distinguishes the case of installing a single wind farm when no other wind farm from the case of installing several wind farms when having already n in a specific pool.

1.1 Wind resource assessment versus predictability

First, this section gives a short summary of the literature related to the wind resource assessment process. Then it gives an overview of the papers related to the integration of wind power predictability in this particular process. For a far more exhaustive state of the art in wind resource assessment, we refer to [36]. In [27], a typology of different wind resource assessment methods is given, ranging from *Folklore* to advanced mesoscale and microscale modeling of wind speed.

1.1.1 The wind ressource assessment process

Wind resource assessment is a matter of estimation under conditions of uncertainty. It aims at giving an estimation of wind energy potential at a specific site, more precisely the long-term kinetic energy that wind holds. The target variable to define is the so-called Annual Energy Production (AEP), which measures the

expected annual wind energy production (e.g. in GWh) of the future power plant. This estimation, as well as the associated uncertainty, is then used as input in the study of economical feasibility of the project, which aims to assess whether the investment is valuable in terms of expected discounted income.

Once a site has been selected based on site availability and regional wind climate, wind resource assessment phase can take place. This process can be divided into several stages, according to [7]:

Monitoring towers are installed and are fitted with cup anemometers or wind vanes to conduct on-site measurements of wind speed. In addition to that, ground-based remote sensing devices can be used, such as SODAR or LIDAR[1]. They are especially well-suited for flat terrain.

Then data are **validated and corrections** are made if needed. As the measurement campaign at the site is quite short for economical reasons, it is necessary to extrapolate wind speed time series for longer period that matches with the life duration of a wind turbine, that is to say approximately 20 years at onshore locations. This is done by correlating data with climatologically representative time series, usually provided by a reference site (often a meteorological station) standing nearby the measurement site. Then the **long-term average wind speed** is extrapolated. This method is called Measure-Correlate-Predict (MCP) and gives satisfactory results. Another way is to use more advanced methods such as Numerical Weather Prediction (NWP) tools and then downscaling the results at the measurement site location. A review of the use of NWP models in wind resource assessment is given by [37].

Wind shear extrapolation aims at taking into account of the difference in wind speed between the wind speed measurement devices and the hub-height of future wind turbines. This shearing effect normally results in an increase of wind velocity at higher vertical locations (there can also be negative gain due to complex terrain). Vertical extrapolation becomes more and more challenging as taller masts are being installed.

Then the wind resource has to be simulated over the site of interest using numerical models. **Wind flow modeling** includes the study of wake effects and effective turbulence across the wind farm, taking into account local topography. The most popular linear model is WAsP (Wind Atlas Application and Analysis Program), which was developed at RisøNational Laboratory, DK. It is used for resource assessment in flat to medium complex terrain, [20]. Computational Fluid Dynamics (CFD) non-linear models are well-suited for modeling complex terrain flows, whereas classical methods express their limits. Wind flow softwares fit input wind speed data with parametric distribution. In [11], the suitability of the Weibull and Rayleigh functions in modeling the wind speed distribution is tested. Both are empirically performing and thus are widely used among wind data analysts. A neural based approach is considered in [29], which appears to overperform the Weibull approximation. The authors of [17] enumerate several alternative techniques.

We can mention an additional stage, which consists in the determination of extreme wind speed, where extreme wind speed is defined as the value of the average highest wind speed over 10 min with a return period of 50 years. It is a design parameter and it is calculated for design compliance purposes, using either the so-called EWTSII method or extreme value analysis depending on

[1] respectively Sonic Detection And Ranging and Light Detection And Ranging

the amount of available data [36].

Most of the ongoing researches in the field of wind resource assessment are focusing on developing methodologies and improving wind speed modeling in case of non-classical terrain types. For example, a methodology for offshore wind resource assessment is proposed in [2], taking into account that roughness length increases with wind speed at offshore locations. The most recent developments in wind flow modeling in case of hilly and woody terrains are presented in [5].

Finally, at the **energy estimation stage**, power production is predicted using the wind turbine power curve, which provides the associated power rate as a function of wind speed. The energy yield is then estimated by taking into account the annual availability rate of wind turbine as well as the loss energy rate. Once the mean energy yield has been estimated, the associated uncertainty has to be quantified too.

1.1.2 Wind resource assessment and wind power predictability

Here, and for the rest of this study, predictability is the degree to which a correct prediction or estimation of wind and power data can be made. Prediction concerns the mean behavior of wind power, as well as the occurrence of extreme events. Predictability of market price also affects the income of wind power producers but it has nothing to do with wind resource assessment, so it is out of the scope of this work. A recent survey [36] pointed out that the integration of wind power output predictability is a stake for 7% of researchers; although it remains quite unconsidered among developers, manufacturers and consultants. In order to understand why predictability is considered to be important, we will have to clarify the links between wind variability, wind resource assessment uncertainty and, at last, predictability.

First, variability refers to the degree to which a variable experiences changes in its value, in time or within a set of data. Here (Fig. 1.1) are the standard deviations of wind power in percentage of nominal power at different time scales over year 2008, ranging from hour to month, for 182 transformation stations grouping operational Danish wind farms. The value for 20-year inter-annual wind power variability is also plotted. It is set to 0.1, which is consistent with the value found in the few empirical study that were conducted, see [38].

Most of the wind farms have an hourly variability ranging from 25% to 30% of their nominal power. Variability decreases as the time scale increases, thus illustrating a temporal smoothing effect. The consequence is that the level of future wind energy production is marred by uncertainty, at all horizons.

When estimating the wind resource of a candidate site, the literature as considered as particularly relevant to analyze the uncertainty associated with the annual level of wind energy production. In [25], a framework to assess the wind resource of a wind turbine using uncertainty analysis is presented. An overall 11% level of uncertainty associated with wind resource assessment is found, which is consistent with the value we can find when glancing through the literature (from 8% to 12% in average, often twofold for highly complex terrain). Monte Carlo simulation is used in [3] to estimate the quantiles of AEP for different time horizons. The authors find that the annual capacity factor will lie within 1.50% of his long-term predicted average. Based on large cross-country data, it is found in [4] that the computation of wind energy potential

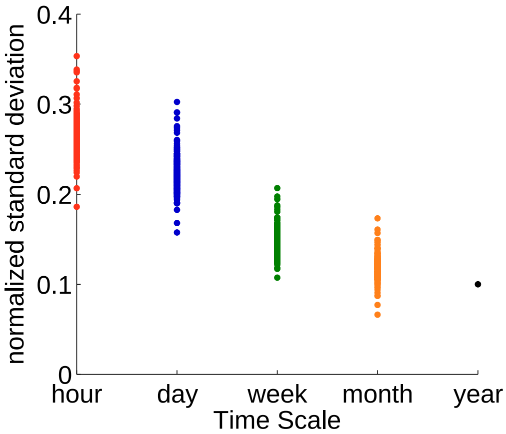

Figure 1.1: Wind power variability for different time scales over year 2008, for 182 transformation stations grouping operational Danish wind farms.

tends to underestimate the effective wind energy production by 6-7%.

According to [30], there exist four sources of errors in wind energy production estimation: (1) wind speed measurement, (2) long-term resource estimation, (3) wind resource variability and (4) site assessment. Each of the above-mentioned wind resource assessment stages belongs to one of the (1), (2) and (4) uncertainty sources. The authors of [7] gives interesting estimation of the uncertainty related to some stages of wind resource assessment, as well as the associated costs, based on their experience. The authors estimate that cumulating those errors can result in a 24% error in energy yield estimation, which leads in his turn to a 2.5 million € error in annual revenue estimation for a 50 MW sample European wind project, when assuming a 0.08€/kWh price. In [30], sensitivity factors are used to combine these independent sources of uncertainty and to provide more accurate estimation of the uncertainty in the site assessment process. Unless it is not in the scope of this work to discuss the uncertainties associated to one or another stage of the wind assessment process, knowing their magnitude could help in assessing how important is to estimate predictability.

In the pre-cited work [7], tower sheltering effect is shown to potentially induce a 1.8% error in energy yield estimation, while top anemometer speed-up leads to a 7% error. Another main source of uncertainty is the wind speed modeling which results in even higher uncertainty when translated into a power production profile. The Weibull approximation is found to result in a 1.5% error. A simplistic wind shear extrapolation could also induce a 1.8% error. The most dramatic source of error is the equipment problems that can cause an 11% error. Another interesting work is the one from [1], which investigate the effect of missing data on wind resource estimation, mainly caused by icing of equipment that occurs in colder months. The authors found that missing data decrease predictability and result in significant bias for expected economic revenue in case of low electricity price variability. They develop a seasonality-based method to help remove this bias. It appears that the uncertainty is caused either by the quality of the data, the number of years involved in estimation.

Apart from type (3), the uncertainty sources are partially reducible by improving equipment selection, monitoring and modeling. E.g. it is noticed in [30] that the use of ground-based remote sensing devices could potentially reduce the uncertainty in wind power estimation of about 25%.

However, the estimation of annual energy production might be not sufficient to assess whether a wind farm is economically valuable or better than another one. One could consider as well the short-term predictability, which is at the root of the economic valorization of wind energy. In fact, it is the discrepancy between short-term wind power output and the associated prediction that determine the producers' income through market participation. The economic value of accurate prediction is shown in [16] and [33], among other works. In a wind resource assessment perspective, the idea is that a candidate site could have huge energy yield potential but limited short-term predictability, and thus not so good economic value. It highlights the need for short-term uncertainty quantification (and thus predictability) from the resource assessment stage, in full accordance with the market time. In order to consider this economic value of predictability, we will have to introduce the market framework under which power producers operate.

1.2 The value of predictability related to electricity market participation

This section explains how the predictability of wind generation decisively impacts the income of power producers when operating under market conditions. It first introduces the characteristics of the market framework, before describing the participation in day-ahead market, which is seen as the reference case, as well as more advanced strategies for hedging imbalance settlements.

1.2.1 Electricity markets

Due to the ongoing liberalization process[2], electricity has become a commodity that is usually traded on physical markets where the demand from consumers meets the offer from power producers. However, electricity has very special characteristics, which may differentiate it from other tradable goods. Electricity generation has to meet demand at all times in order to ensure system stability despite being hardly subject to storage. On the side of demand, consumers may react very smoothly to changes in price, which amounts to saying that demand for electricity is quite inelastic. This is mainly due to the lack of any substitutable good.

When trading their production at the market, wind power producers, as well as conventional power producers, are economically responsible for any imbalance between contracted and delivered energy. Market participants are said to be balance responsible parties (BRPs). To avoid these regulations costs, power producer have to make suitable decisions about their energy contracts before delivery, while having no certainty about the future wind speed intensity and

[2]To get an overview on what are the pros and cons of liberalization, look at [10]. To know how power market design could be shifted in order to fit onto the large-scale integration of renewable energy sources in the electricity system, [22, 9]. About the Nordic Power Market, [35]

thus about wind power production. In other words, the limited predictability of wind generation makes it particularly sensitive to regulation costs. Those costs translates the wish of the Transmission System Operator (TSO)[3] to provide incentives to power producers for minimizing imbalances so the equilibrium between generation and load can be maintained at all times and the system is kept stable.

There exist bilateral markets, where buyers and sellers trade directly with each other, and pool markets, where buyers and seller submit respectively their offer and demand bids during the period between the gate opening and the gate closure time. Bids consist in an amount of energy for a given price. Day-ahead markets and intraday market are such markets. They also belong to the physical markets category, as they permit to trade energy, which is a physical commodity. On day-ahead markets, producers submit their quantity-price bids on the day before delivery and a single market price is defined by matching the aggregated offer and demand curves (see Figure 1.2). E.g. In the Nord Pool Elspot day-ahead market, bids are to be settled before noon for every hour of the following day.

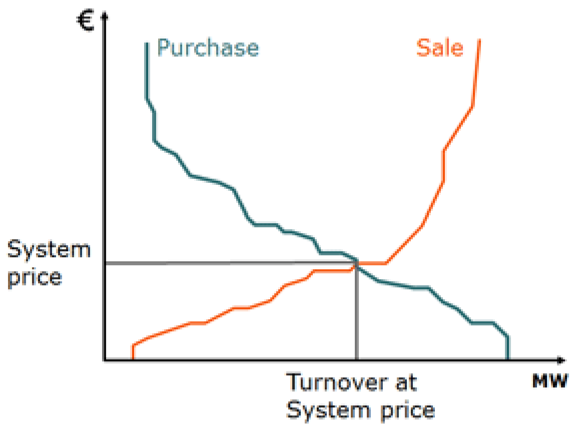

Figure 1.2: Single market clearing process. *Source*: Nord Pool website, May 2011.

At the time of imbalance settlement, a posteriori to production and consumption, the imbalance volume is measured and the imbalance penalty is obtained by multiplying it with the imbalance price. The imbalance volume is the amount of energy between delivered energy and net contracted energy. Positive energy imbalance occurs in the case where delivered energy exceeds contracted energy, while negative energy imbalance occurs when delivered energy is lower than contracted energy. The sum of imbalance volumes from all the balance responsible parties determines the regulation state on the real-time market. The TSO ensures the equilibrium between production and consumption at all times by applying secondary reserves and tertiary reserves. If the real-time overall imbalance is positive, we talk about down-regulation. If the opposite occurs,

[3]The TSO is mainly responsible for the secure and efficient operation of the network. It manages transmission networks.

we talk about up-regulation. The imbalance prices are either fixed or market-based, depending on the market. Sometimes, balance responsible parties are only charged for imbalances opposite to the regulation state. When no overall imbalance occurs, both up regulation and down regulation prices are equal to the spot price(see Figure 1.3).

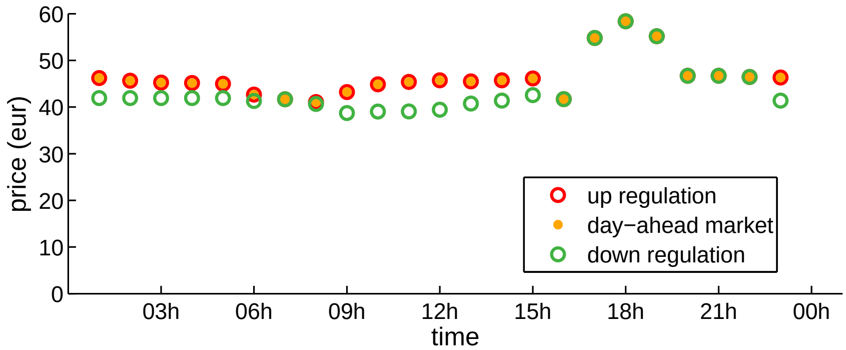

Figure 1.3: Day-ahead electricity market price in the Elspot market for the Western Denmark area, the 01/01/2008, and the associated up and down regulation price for the same area.

1.2.2 The value of predictability related to day-ahead market participation

Once the market mechanism has been introduced, it is obvious that increasing predictability results in lower imbalance penalties and thus increases wind producer's income. We already said that power producers have to propose bids before delivery time, thus having no precise idea about what the future power production will be. The longer is the delay between gate closure time and delivery time, the lower is the predictability and the higher are the imbalance volumes and penalties.

A lot of researches have been carried out on the field of the participation of wind power producers in day-ahead markets. They all attach great value to predictability as it permits to make optimal bids on the day-ahead market. The authors of [19] use a Markov Process to model wind generation and simulate the participation in the day-ahead market. A bidding strategy based on stochastic programming is developed in [24]. In [34], the authors aim at demonstrate the benefits of a quantile-based bidding strategy, while the authors of [15] adopt a risk-based approach. [8] is a case-study for an Italian wind farm operating in the day-ahead market.

Finally, [6] formalizes the participation in the day-ahead market and proposes a generic formulation for the loss function of a power producer, including the possibility for hedging solutions. We will use this framework for the purpose of this work.

1.2.3 Advanced strategies for hedging imbalance costs

Many solutions exist to manage imbalance penalties related to renewable generation in electricity markets. There are *physical* solutions, such as storage and combining with conventional energy generators. Additional *financial* solutions like participation in intraday markets can help reduce the imbalance penalties. All those solutions are given a short description in this subsection. Note that an exhaustive description of the strategies for hedging imbalance costs is given in [6]. We assume that wind producers do not control generation, that is to say they do not operate strategic reductions of wind power output.

Storage

Combining wind turbines with energy storage device (ESD) is a way to improve dispatchability and to manage imbalances. The basic mechanism is the following one: if wind generation is greater than contracted ESDs can store the excess energy and make it available to manage latter shortage situations. If wind generation is lower than contracted, ESDs can deliver the energy that has previously been stored, if any. However, storage entails energy losses and both investment and operational costs. According to [21], the installation of storage facilities is in the order of 40% of investment over 20 years for electricity production systems and only become profitable if a certain quantity of energy is supplied.

A simulation for the participation of a virtual power plant including wind turbines and energy storage facilities is presented in [28]. The author of [6] estimates that this situation leads to a reduction of 10%-15% of the imbalance penalties if operating at the day-ahead market and assuming a 30 MW capacity for the ESD and a round-trip efficiency of 0.85. In [13], a risk-based strategy is developed, including strategic coordination of pumped-hydro storage and the participation in the intraday market.

Conventional generation

Another solution is to combine the wind power turbines with dispatchable conventional generation units, such as gas turbines, diesel engines or tidal generators. Unlike storage, combining with conventional generation can only reduce negative imbalances. If wind generation is lower than contracted, these units can be switched on in order to fill the lack of wind generation.

In [32], an optimal bidding curve is developed for the participation of a mixed portfolio made of RES and conventional fossil-fuelled plants in a dual-pricing imbalance environment. In [23], the joint operation of hydro and wind power on electricity spot market is considered. However, those works do not take into account the additional cost for installing back-up conventional generation units. It is done in [6], assuming a linear variation of the marginal cost for any variation of the delivered energy. The sensitivity analysis reveals that the imbalance penalty is at his lowest level when the marginal cost is low and the conventional generation unit nominal power is 1.35 MW (for a 18 MW wind farm).

Participation in intraday markets

The participation in intraday markets is a financial hedging solution but has the same effects as physical solutions. The intraday market permits to trade energy between the day-ahead gate closure time and the delivery time. Intraday markets gate closure and opening times changes depending on the national or international market that is considered. They permit to take advantage of more recent (and thus more accurate) predictions to revise their production schedule and to trade the difference between the day-ahead bid and the revised schedule.

The market clearing process is often different from the spot market one. A pay-as-bid mechanism is used, that is to say market power producers are remunerated at their bid price and the transaction occurs on a "first come first serve" basis : the first matching offer to a bid is rewarded and fixed into two bilateral transactions between the seller and the buyer.

Such an advanced hedging strategy with sequential participation in both day-ahead and intraday market is described in [12]. The author finds that the participation in intraday market can help reduce the imbalance penalties by up to 18%. For the moment, very few actors are trading on the intraday markets and the volume is quite low. E.g. For the Nordic Elbas intraday market and year 2007, only 1.6 TWh were traded on this market, while the Elspot day-ahead market and the regulation market traded volumes reached respectively 291 TWh and 13 TWh.

1.3 The stakes for integrating predictability in wind resource assessment

This section formulates the two distinct problems that our work will cope with. The case of the installation of a single wind farm is presented first. Then the case of the simultaneous installation of multiple wind farms is presented. This last case is more complicated, as it requires analyzing spatial effects.

1.3.1 The single case : "got no wind farm, want 1"

A basic situation is the one of a new Independent Power Producer (IPP) entering the market, who wishes to build a single wind farm or, it is the same for us, an IPP who already owns one or several wind farms but does not take account of the information provided by those wind farms in the installation choice. If energy yield estimation is available for several candidate sites where a wind resource assessment process was carried out, it could be useful not only to compare those sites on the basis of energy yield, but also to consider short-term predictability of wind power, which exposes producers to imbalance penalties. It is a basic decision-making problem, which can involve a trade-off between wind power yield and wind power predictability. If a higher energy yield results in lower wind predictability, it is necessary to make the balance and to estimate what the producer's income could have been if the wind farms had been producing energy. To our knowledge, no paper in the literature has dealed with the so defined problem.

This situation, where only the installation of a single wind farm is considered is denoted as the *single case*. This basic case is studied first in order to

investigate the relations between variability, predictability and market revenue that *does not* rely on spatial effect. It can be coarsely resumed by the following statement: "got no wind farms, want 1". It is the main object of chapter 2.

1.3.2 The aggregated case : "got *n* wind farm, want *k*"

A more complex situation is the one of an Independent Power Producer who wants to settle several wind farms in an optimal way. This IPP can already have either several wind farms or none. If any, the initial number of wind farms is assumed to be known and is thus considered as a parameter, whereas the number of new wind farms to be settled is a variable we want to fix in an optimal way, or at list being given boundaries. More generally, a power producer owning wind turbines of a given capacity and spread over several locations wants to install k more turbines of a given capacity into present wind farms (*repowering*) or at new locations (*installation* strictly speaking).

The study of such a problem entails the investigation of the spatial and temporal effects of the wind farms when being strewed. For example, the distance between wind farms, and the power capacity dispersion notably affects predictability and the value from the market. This situation, where considering the output and market revenue of aggregated wind farms is denoted as the *aggregated case*. It can be coarsely resumed by the following statement: "got n wind farms, want k". It is the main object of chapter 3.

1.4 Conclusion

It appears that wind resource assessment is a multi-step process, which involves many sources of uncertainty. Wind short-term variability is particularly high-stake and has a dramatic impact on producers' income as it affects wind predictability. Indeed, producers operate under market conditions and face imbalance penalties when they make wrong predictions about the expected available short-term wind power. Having an idea of wind power short-term predictability at the stages of wind resource assessment could therefore be useful to take optimal decisions when one wishes to install a new wind farm. We distinguish two cases, the single case and the aggregated case, which implies spatial and temporal effects. For each of them, we seek to evaluate the value of having an idea of what predictability is before a wind farm is constructed.

Chapter 2

Understanding the value of predictability : a case study over Denmark

This chapter makes explicit the value of predictability when looking at the installation of a single wind farm. It first formulates the decision-making problem and brings to light a set of results for a case study over Denmark. Then the influence on predictability and market revenue of hedging solutions is considered and results are given for a set of Danish wind power plants.

2.1 The value of predictability when installing a new wind farm

This section formulates the decision-making problem and the market penalization functions. The data are also introduced.

2.1.1 Formulation of the decision-making problem

Given a finite set of pre-selected prospecting sites $\mathcal{F} = \{A, B, C, \ldots\}$, the power producer has to choose the most appropriate one to build a wind farm ω^* which will maximize his revenue R_{tot} over a given period :

$$\omega^* = argmax_{\omega \in \mathcal{F}} R_{tot}^{\omega}$$

Prospecting sites are characterized by their location and area coverage and we assume that sufficient wind data is available in order to estimate the wind energy yield and the predictability of wind power generation at this location (see Appendix A for estimating predictability).

We keep only the transformation stations that represent more than 1 MW of installed power. Smaller ones often have a very specific behavior (very high capacity factor[1], low predictability, high prediction bias) and thus greatly influ-

[1]By capacity factor, we mean the actual energy output of a wind farm divided by the theoretical maximum output, if the machine were running at its maximum rated power during all the year.

ence the results.

As we already said, the value of predictability is related to market participation. Consequently, we do need a model to express market participation and characterize the penalization of energy imbalances. As we said, market participants are charged for the imbalance between their energy bids and their actual wind production.

For each T_i market time, the revenue (also referred to as the income) R_{T_i} of a market participant bidding an amount of energy $E^C_{T_i}$ but actually generating $\tilde{E}_{T_i}$ can be formulated as the combination of the income from selling the actual wind generation $\tilde{E}_{T_i}$ at the spot price $\pi^C_{T_i}$ minus the cost $\delta_{T_i}(d_{T_i})$ for regulation[2] of imbalance d_{T_i}:

$$R_{T_i} = \pi^C_{T_i}.\tilde{E}_{T_i} - \delta_{T_i}(d_{T_i}), \quad \forall i = 1, \ldots, N$$

As the market participant is assumed to be price-taker and the actual wind generation $\tilde{E}_{T_i}$ is not controllable (i.e. wind input depends on meteorological conditions and no strategic reduction of electricity generation occurs), the decision problem can be rewritten as follows:

$$\omega^* = argmin_{\omega \in \mathcal{F}} \delta^\omega_{tot}, \quad \delta^\omega_{tot} = \sum_{i=1}^{N} \delta^\omega_{T_i}(d_{T_i}),$$

where $\delta_{T_i}(d_{T_i})$ is the penalization function that given the imbalance cost knowing the amount of imbalance energy d_{T_i}at time T_i :

$$\delta_{T_i}(d_{T_i}) = \begin{cases} \pi^{*,+}_{T_i}.d_{T_i}, & \mathrm{d}_{T_i} \geq 0 \\ -\pi^{*,-}_{T_i}.d_{T_i}, & \mathrm{d}_{T_i} < 0 \end{cases}$$

$$d_{T_i} = \tilde{E}_{T_i} - E^C_{T_i},$$

where $\pi^{*,+}_{T_i}$ and $\pi^{*,-}_{T_i}$ are the regulation unit costs for positive and negative imbalances, and are given by :

$$\pi^{*,+}_{T_i} = \pi^C_{T_i} - \pi^{C,+}_{T_i}$$
$$\pi^{*,-}_{T_i} = \pi^{C,-}_{T_i} - \pi^C_{T_i},$$

with $\pi^C_{T_i}$ being the spot price, $\pi^{C,+}_{T_i}$ and $\pi^{C,-}_{T_i}$ being, respectively, the positive and negative imbalance prices. Finally, the market participant is supposed to bid the amount of predicted energy for each hour. The quantity-price bid is supposed to be accepted and is then equal to the contracted energy $E^C_{T_i}$, which is determined by the power prediction $\hat{P}_{T_i}$ multiplied by the time step $\Delta\tau$ at time T_i :

$$E^C_{T_i} = \hat{P}_{T_i} \times \Delta\tau, \quad \hat{P}_{T_i} \in [0, P_{nom}], \quad \forall i = 1, \ldots, N$$

The aim of the next section is to evaluate by how wind power predictability can be a valuable choice criterion when having such decision to make. In fact the rest of the study will lie on real world data and so predictability will by

[2] We neglect the subscript ω in the rest of this section, as the market design is the same for all wind farms

associated to a predictive model that takes real production data as input to a series of measured prediction error. To draw conclusions with a resource assessment point of view, it has to be verified that predictability estimation is possible and give good enough results. This is done in Appendix A for a French wind farm.

2.1.2 Scope of the study

The dataset

In the present study, we deal with Western Denmark data. Western Denmark refers to the country part that is situated from the western side of the Great Belt.

This area is characterized by high wind potential and flat landscape. It has the world's higher share of wind power in its electricity mix and therefore represents a particularly interesting case study as it is advanced in the ongoing process of wind power penetration in electricity systems.

Indeed, we do not deal with wind farms data directly. The data are grouped by the nearest transformation station. That is to say that power generation data are aggregated and that some variables in the database, like hub height, are nothing but an average among wind farms. Among the 565 transformations stations available in the data, 361 were removed for having too poor data availability or for having experienced a flagrant change in their installed power during the period. 22 more were removed for representing too low installed capacity (below 1 MW). The remaining 182 transformation stations still form a substantial dataset and may be able to get us some meaningful results. From now on, we will talk about wind farms even if data are from transformation stations.

All power data were provided by the Danish TSO Energinet.dk, according to the EU project Safewind's data sharing program. Market data are available on `http://www.Energinet.dk`. Power data contain the hourly rated power of each transformation station as well as their features (identifier, coordinates, hub height, installed capacity).

Wind power production forecasts were made using the *Random Forest* non-parametric estimation procedure, implemented in the *RandomForest* R package. However it is outperformed by some state-of-the-art models, it gives satisfactory results.

The Nord Pool power Exchange

Danish power producers operate under Nord Pool power Exchange rules. This market gathers all power producers from the Nordic Country and is divided into several areas. Western Denmark corresponds to the DK-1 territorial subdivision of Nord Pool power Exchange.

The market participation simulation is conducted according to this market's specific rules. In particular, power producers are not charged with penalties if their energy imbalance is opposite to market regulation state (dual pricing mechanism). Typically, a power producer operating on this market and actually producing less than contracted will be rewarded his energy at the spot price if market is down-regulating, that is to say if the total amount of energy generated among market participants is greater than consumption.

We also make several other assumptions. Market participants are price-takers, they are not big enough to influence the market price level. Only selling bids are considered.

Elspot is the name of the day-ahead market, and Elbas is the one for the intraday market. The last one was opened for trading on the first of March 2007 for the Western Denmark (DK-1) area.

2.1.3 Results and analysis

We are now able to run a simulation of the participation of those wind farms to the market for the period ranging from the 18/11/2007 to the 29/09/2009, which corresponds to 682 days of market participation. This period is determined according to the availability of both production and market data and is quite large (almost two years). In this reference case, only the participation in day-ahead market is considered. The simulation are made through a platform developed at Ecole des Mines de Paris, Center for Energy and Processes and implemented under MATLAB©software. It is described in [6].

Predictability is measured by the Normalized Mean Absolute Error (NMAE for the rest of the study), where P^{ω}_{nom} is the nominal power capacity of wind farm ω:

$$NMAE(Err^{\omega}) = \frac{1}{24N} \sum_{i=1}^{N} \sum_{h=12}^{H=35} \frac{|Err^{\omega}_{i,h}|}{P^{\omega}_{nom}}, \quad \forall i = 1, \ldots, N$$

$$Err^{\omega}_{i,h} = P^{\omega}_{i,h} - \hat{P}^{\omega}_{i,h}$$

Knowing the historical wind power production data for 182 transformation stations, having computed predictions, we can now simulate the participation of those production units in the power market and being given an idea of what are the magnitude of predictability and market revenue as well as the relation between them. Table 2.1 summarizes the results for the dataset, reported for a 365-day year.

	min	q1	q2	q3	max
Nominal power (MW)	1.030	3.044	5.662	10.950	34.121
NMAE(Err) (%)	9.82	11.24	11.85	12.56	16.20
Perfect revenue ($\times 1000€/MW$)	73.1	82.8	87.4	97.6	147.4
Revenue ($\times 1000€/MW$)	69.9	79.0	83.1	93.1	142.1
Positive $\delta(d)$ ($\times 1000€/MW$)	1.653	2.686	2.912	3.085	5.306
Negative $\delta(d)$ ($\times 1000€/MW$)	0.737	1.287	1.396	1.524	2.074
Imbalance penalty ($\times 1000€/MW$)	3.170	4.002	4.303	4.550	6.043

Table 2.1: Simulation results. Min, q1, q2, q3 and max respectively refer to the minimum, the first quartile, the median, the third quartile and the maximum that were found for each indicator.

The annual income for each wind farm through market participation lies between 69,900 and 142,100€ per MW of installed power. That means that the

wind farms that achieve the best results overperform by a factor of two those who achieve poor results. The annual imbalance penalty lies between 3,170 and 6,043€ per MW of installed power. The total of imbalance penalties resulting from positive energy imbalances is twice bigger than those resulting from negative energy imbalances. This is explain by the dissymmetry between the regulation unit cost for positive energy imbalances $\pi_{T_i}^{*,+}$ and the one for negative energy imbalances $\pi_{T_i}^{*,-}$. That makes the penalization function asymmetric. For the considered period, we check that $\pi_{T_i}^{*,+}$ tends to be higher than $\pi_{T_i}^{*,-}$:

$$mean(\pi_{T_i}^{*,+}) = 4.421€/MWh$$

$$mean(\pi_{T_i}^{*,-}) = 2.917€/MWh$$

where € refer to the Euro currency unit.

All the wind farms achieve a level of income of at least 94.27% of the perfect revenue, which is the income that producers would rake in if there were no imbalance penalty, i.e. if they would make a perfect prediction about the level of actual power generation. None of them reaches more than 96.65% of the perfect revenue. We now consider the relation between predictability and imbalance penalty by plotting the NMAE criterion versus the annual imbalance penalty.

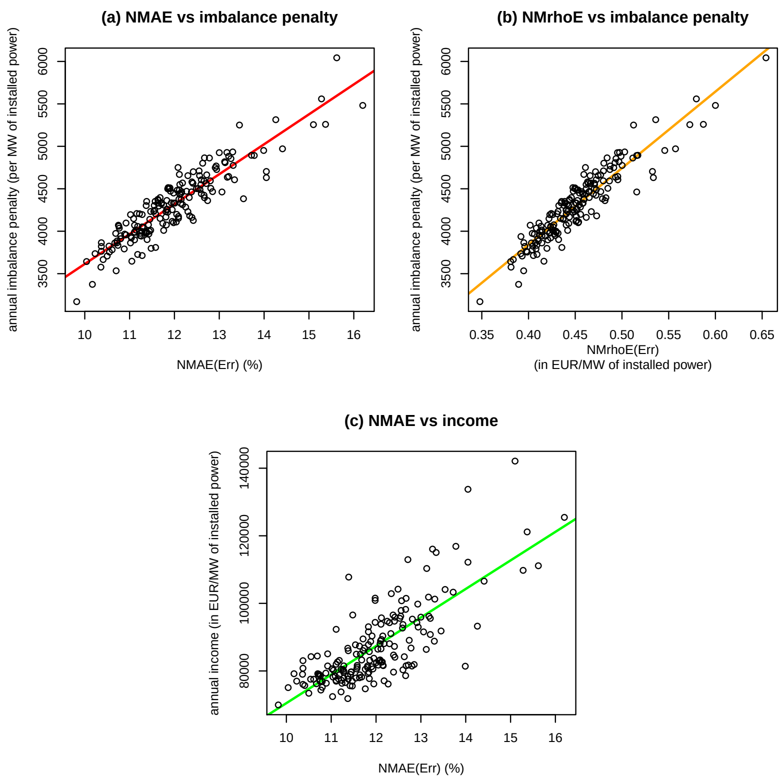

Figure 2.1: Plot of predictability versus imbalance penalty and income for every wind farm, market simulation from 18/11/2007 to 29/09/2009.

Considering Figure 2.1 (a), it is obvious that predictability is strongly and linearly related to the amount of imbalance penalty (Pearson's correlation coefficient[3] is equal to 0.912 if predictability is measured by NMAE). More precisely, a 1 point decrease in NMAE results in a 352.5€ reduction of the imbalance penalty per year and per MW of installed power. Indeed, imbalance penalty is proportional to the absolute deviation from contracted energy but not by the same proportion whether imbalance is positive or negative; otherwise it would be perfectly correlated.

$$\delta(d) \propto NMAE(Err), \quad \text{if } \pi_{T_i}^{*,+} = \pi_{T_i}^{*,-}$$

Note that we could avoid the market simulation[4] by constructing a kind of intertemporal penalization function, using the average unit regulation costs for both positive and negative imbalances, and associating weights to these imbalances to take account for the dual pricing mechanism. We thus create a new indicator, the NMrhoE, which applies the ρ intertemporal penalization function to the energy imbalance. It is an approximation of the mean hourly imbalance penalty. Unlike δ, ρ no more depends on the subscript T_i. NMrhoE is even more correlated with the total amount of imbalance penalties than NMAE (Pearson's correlation coefficient is equal to 0.931), see Figure 2.1 (b). NMrhoE can be interpreted as an approximation of the mean hourly imbalance penalty. If multiplied by the length of the period of market participation, it gives an approximation of the producer's total amount of imbalance penalties.

$$NMrhoE(Err^{\omega}) = \frac{1}{24N}\sum_{i=1}^{N}\sum_{h=12}^{H=35}\rho(\frac{Err^{\omega}_{i,h}}{P^{\omega}_{nom}}), \quad \forall i = 1,\ldots,N$$

where $\rho(x)$ is as follows :

$$\rho(x) = \begin{cases} mean(\pi_{T_i}^{*,+}).x, & \text{x} \geq 0 \\ -mean(\pi_{T_i}^{*,-}).x, & \text{x} < 0 \end{cases}$$

What is more, the NMAE appears to be highly correlated with the producer's income, as seen on Figure 2.1 (c). However it does not imply any direct causal relation. As we will see, low predictability is associated to high yield, which in its turn implies high income.

When mapping the results for the income against predictability, it confirms that the wind farms that have less predictability get higher amounts of imbalance penalties. On Figure 2.2, the points represents, respectively, (a) the income and (b) the amount of imbalance penalties achieved by 182 Danish wind farms and the value for entire Denmark is interpolated on a regular spatial grid by kriging (*fields* R package was used for the implementation). Wind farms are more likely to reach high imbalance penalty levels (i.e. to reach low predictability levels) in the southwestern and northwestern parts of Denmark. The rest is quite homogeneous. Wind farms are more likely to reach high income levels in the northern and northwestern parts of Denmark, as well as in the south west and in some eastern islands.

However, Figure 2.3 (a) shows that an even stronger relation exists between capacity factor and the producer's income. Pearson's correlation coefficient is

[3]The formula for Pearson's correlation coefficient is detailed in Appendix B.

[4]Even if it is very fast from a computational point of view.

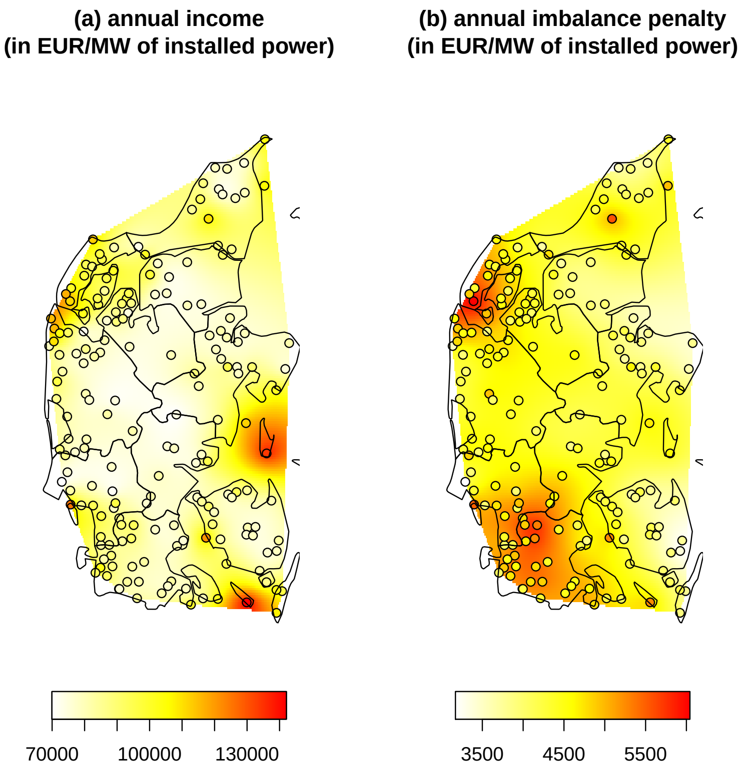

Figure 2.2: Maps of annual income and imbalance penalties for every wind farm, market simulation from the 18/11/2007 to the 29/09/2009.

equal to 0.998. More precisely, a 1 point increase in capacity factor results in a 3,907€increase in the producer's income per year and per MW of installed power. The capacity factor potentially states the wind energy yield. Making some assumptions (e.g. an homogeneous availability rate), the energy yield is not more that the capacity factor multiplied by the time step $\Delta\tau$, the length of the period when the wind farm was operational and the nominal power of the wind farm. In fact, this criterion appears to be sufficient if we are to explain the market revenue and compare wind farms by means of this variable. When considering only the wind farms whose capacity factor lies between 20% and 25% of the nominal power, it appears (Figure 2.3 (b)) that the income of some wind farms differ from their expected income. This could be the result of a relatively lower or greater predictability, given the capacity factor.

Looking at Figure 2.3 (c), it is not obvious that the NMAE impacts the level of producers' income for a given level of capacity factor. In other words, predictability does not seem to bring additional significant information on income, when having previously accounted for the capacity factor, which represents the energy yield.

In fact, Figure 2.3 (d) shows that the residuals of the linear fitting are negatively related with NMAE. This is quite logical, as negative deviation from

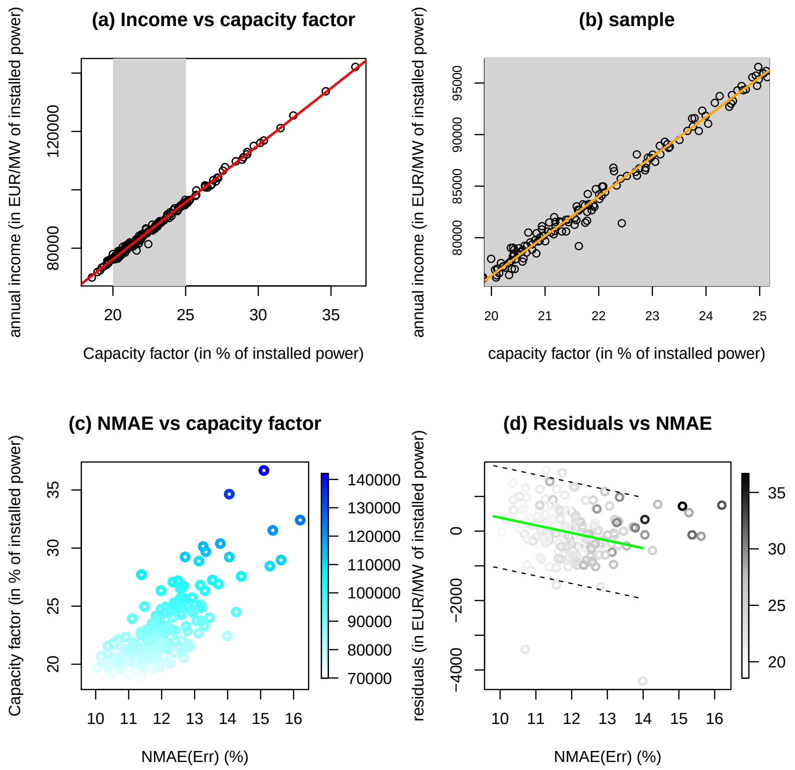

Figure 2.3: Plot of annual income versus capacity factor. (c) Points are colored by income, refer to the right scale (in € per MW of installed power) (d) Points are colored by capacity, refer to the right scale (in %).

the expected income are more likely to occur when predictability is low (i.e. NMAE is high), and positive ones when predictability is high (i.e. NMAE is low). Wind farms n°84 and n°183 are exceptionally underperforming with regard to their capacity factor. The linear fit (green curve) is rather coarse but the coefficient associated to NMAE is significant (p-value is equals to 1e-03). The dashed black line represents the 95% confidence interval around the fitted value. The wind farms that reach for NMAE lower than 12% tend to overperform in terms of income and the others tend to underperform. Here is the table giving the value for a range of NMAE values and the associated prediction interval bounds:

When predictability is very high, e.g. with NMAE equal to 0.1, the total income is *expected* to be greater by 387.63€ per year and per MW of installed power, which represents a relative gain of about 0.5% (the few wind farms with NMAE close to 0.1 have a capacity factor close to 20% and an annual income close to 80k€). When predictability is low, e.g. with NMAE equal to 0.14, the total income is *expected* to be lower by 489.94€ per year and per MW of installed power, which represents a relative loss of about 0.5% (when assuming a 100k€ annual income). In fact, these cases are limit cases and, most of the

NMAE(Err) (%)	lower bound	fitted	upper bound
10	-1071.00	387.63	1846.26
11	-1290.39	168.24	1626.87
12	-1509.78	-51.15	1407.47
13	-1729.17	-270.55	1188.08
14	-1948.57	-489.94	968.69

Table 2.2: Predictability versus the deviation from the expected income when having considered the capacity factor. Fitted values and associated prediction interval bounds for a range of predictability values.

time, the explained deviation from expected revenue is even more negligible. Anyway, the confidence intervals are very large so it is difficult to draw robust conclusions.

The low contribution of predictability is highlighted when carrying an analysis of variance on the income. NMAE explains 0.02% of its variance, whereas the capacity factor share is 99,59%. In other words, NMAE explains only 4.5% of the variance of the revenue that is not yet explained by the capacity factor.

Finally we can take a look at Figure 2.4. It tells us about the variability of the hourly normalized income achieved by each wind farm for the period from the 18/11/2007 to the 29/09/2009. The mean deviation from the average normalized income lies between 8.92 and 16.95€. Variability is higher in coastal regions, whether we consider western or eastern Denmark.

2.2 Investigating the influence of hedging solutions on the value of predictability

Hedging solutions refer to either physical or financial ways to cope with the uncertainty that arises from market participation. Three of them were introduced in Chapter 1: adding storage facilities, combining with conventional generation and intraday market participation. The impact of each method on both predictability and determination of imbalance penalty will be discussed in this section. The additional costs of hedging solution are taken into account.

2.2.1 Formulation of the imbalance penalty resulting from hedging solutions

It was shown in [6] that hedging solutions can be introduced in the market model by a modification of the reference penalization function. According to the author, the imbalance penalty related to the participation in intraday market can be written as (from now on, we neglect the subscript T_i because the total of either income or imbalance penalties is no more than their sum over each time unit):

$$\delta_{ID}^{C}(\tilde{E}, E^C) = X + \delta^{ref}(\tilde{E}, E^C + x)$$

where X and x are the additional cost and energy contract volume relative to the intraday solution. When developing the formula:

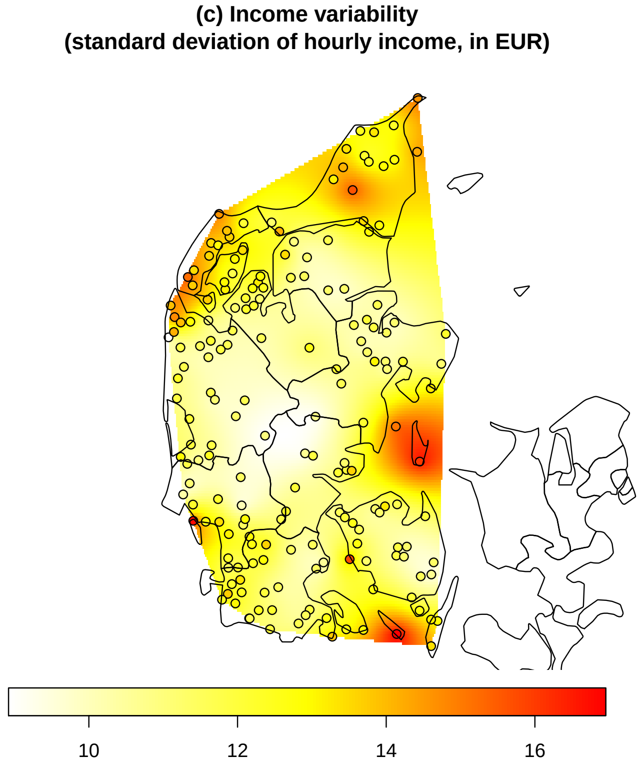

Figure 2.4: Standard deviation of hourly income

$$\delta^C_{ID}(\tilde{E}, E^C) = E^{ID} \times (\pi^C - \pi^{ID}) + \delta^{ref}(\tilde{E}, E^C + E^{ID})$$

In a very similar way, the imbalance penalty related to physical solutions can be written as:

$$\delta^C_{ST,CG}(\tilde{E}, E^C) = Y + \delta^{ref}(\tilde{E} + y, E^C)$$

where Y and y are the additional cost and the energy volume associated with the physical solution (ST for storage and CG for conventional generation). When developing the formula:

$$\delta^C_{ST}(\tilde{E}, E^C) = |\tilde{E}_{ST}| \times \pi^C \times \Gamma_{ST} + \delta^{ref}(\tilde{E} + \tilde{E}_{ST}, E^C)$$

where Γ_{ST} is a dimensionless quantity that represent the cost structure of storage.

$$\delta^C_{CG}(\tilde{E}, E^C) = e_{CG} \times (\pi^C - \pi_{CG}) + \delta^{ref}(\tilde{E} + e_{CG}, E^C)$$

where e_{CG} is the adjustment energy volume delivered by the conventional unit.

Basically, hedging solutions have two effects on the shape of the penalization function. They relocate the curve upward by adding additional cost (X and Y) and from left to right by permitting the adjustment of $\tilde{E}$ (physical solutions) or the revision of contracted energy E^C (financial solutions).

2.2.2 Results and analysis

We propose a measure $Imp_{ref,HS}$ to quantify the relative improvement in revenue and predictability we can obtain with hedging solutions with regard to the reference case :

$$Imp^{\omega}_{R,HS} = \frac{R^{\omega}_{ref} - R^{\omega}_{HS}}{R^{\omega}_{ref}}$$

$$Imp^{\omega}_{NMAE,HS} = \frac{NMAE^{\omega}_{ref} - NMAE^{\omega}_{HS}}{NMAE^{\omega}_{ref}}$$

where $Imp^{\omega}_{R,HS}$ refer to the improvement in producer's income for wind farm ω thanks to the hedging solution HS. $Imp^{\omega}_{NMAE,HS}$ refer to the improvement in predictability (measured by the NMAE criterion for wind farm ω thanks to the hedging solution HS).

What is more, we assume a storage capacity equal to the energy that the wind power plants that are represented by each transformation station could produce in one hour when working at full capacity. It is then proportional to the installed capacity in order to allow comparison. Similarly, the installed power capacity of the convention generation unit is set to 5% of the transformation station represented installed power capacity and the marginal cost is set to 45€/MWh, close to the average day-ahead market price. It has to be noticed that conventional generation is used only for the purpose of reducing the energy imbalance of the reference unit.

The result for each hedging solution is computed and the improvement in market revenue and predictability is resumed in Table 2.2.

	min	q1	q2	q3	max
Storage $Imp_{NMAE,ST}(d)$ (%)	15.7	27.7	29.1	30.4	34.6
Conventional generation $Imp_{NMAE,CG}(d)$ (%)	-1.1	9.1	9.7	10.4	14.8
Intraday $Imp_{NMAE,ID}(d)$ (%)	-0.2	0.5	0.9	1.2	2.1
Storage $Imp_{R,ST}(d)$ (%)	-8.3	-2.6	-0.6	1.1	8.4
Conventional generation $Imp_{R,CG}(d)$ (%)	-29.8	-20.1	-16.8	-14.3	-2.7
Intraday $Imp_{R,ID}(d)$ (%)	-2.6	-1.4	-1.0	-0.5	0.9

Table 2.3: Relative improvement for each strategy. Min, q1, q2, q3 and max respectively refer to the minimum, the first quartile, the median, the third quartile and the maximum that were found for each indicator.

It appears that combining with conventional generation is the less interesting solution in that case as it increases the amount imbalance penalty of more than 14.3% for 75% of the transformation stations, and at least of 2.7% for any of them. Storage is sometimes beneficial and sometimes not, but even if the

reduction of imbalance penalty is positive, it remains under 10% compared to the reference case. Intraday participation mostly brings negative results, as only a few transformation stations benefits from it. Maximum increase in imbalance penalty is equal to 2.1 %, which represents a quite negligible reduction of the total income.

The poor results for intraday participation can be given several explanations. First, minimum size for Elbas contract is 1MW/h, which automatically prevents the class of medium-size wind farms from participating most of the time. In fact, 25% of transformation stations represent less than 3 MW of installed power. In real world, only big energy producers participate in the intraday market in the DK-1 zone and no exchange takes place most of the time.

When looking at the relative penalty reduction and predictability improvement for each strategy, we can see that the improvement in NMAE is quite homogeneous among transformation stations, while the reduction of imbalance penalty is not. One reason for that is that the imbalance penalty is no more proportional to the energy imbalance due to the addition of the X and Y term in the expression of d_{T_i}. Those terms express the additional cost for using such strategies.

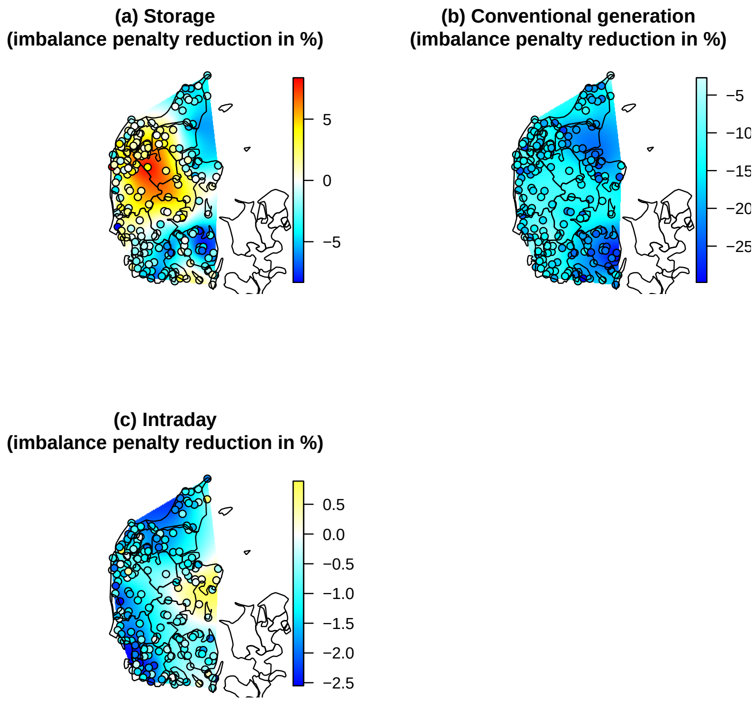

Figure 2.5: Reduction of imbalance penalty for each strategy, from the 18/11/2007 to the 29/09/2009.

The mapping of results (see Figure 2.5) leads to a few additional obser-

vations concerning this three hedging strategies. First of all, the imbalance penalty reduction $Imp^{\omega}_{R,HS}$ seems to have spatial patterns. In particular, the use of storage facilities has a positive impact in west central Denmark and is globally negative everywhere else. The combination of wind energy production with conventional generation is penalizing everywhere, but seems to bring even higher penalties in northern Denmark and Fyn Island. On the contrary, the participation in intraday markets is beneficial only in east central Denmark and particularly unfavorable in the south-eastern coastal region.

This result leads us to neglect those three solutions as for the rest of the study, as it doesn't bring any additional information while having a resource assessment perspective in mind. In fact, Spearman's rank-correlation coefficient[5] between reference income and the new income obtained by each hedging solution is almost equal to 1. That is to say that the hierarchy between wind farms has remained the same. By the light of those results, we can say that the possibility to resort neither to physical solutions nor financial solutions is not liable to influence the power producer's decision of installing a new wind farm because it does not significantly change wind farm's hierarchy in terms of revenue.

2.3 Conclusion

The study of the participation of 182 wind power plant (or more precisely of transformation stations which represent several wind power plants) in the Nordic power market during the period ranging from the 18/11/2007 to the 29/09/2009 brings us some meaningful results.

Predictability appears to be strongly and linearly correlated with the amount of imbalance penalties, when predictability is measured by the Normalized Mean Absolute Error. A 1 point decrease in NMAE results in a 352.5€ reduction of imbalance penalties per year and per MW of installed power, at best 0.5% of the total income.

However, the capacity factor, and thus the energy yield, is so strongly correlated with the producer's income that predictability does not seem to bring much additional information for the decision to make when installing a single wind farm. The part of the income that is not explained by capacity factor is partly explained by predictability. When predictability is low (resp. high), the deviation from the expected income is expected to be 0.5% lower (resp. higher). In addition to that, the uncertainty on the value of the capacity factor is lower than the one on the estimated predictability at the point of site resource assessment. According to the case-study presented in [3], the estimated capacity factor will be within 5.55 per cent of the long-term average band for a period of one year and a confidence interval of 90%.

Nevertheless, those results are to be put in perspective with the relatively small size of Denmark, which does not allow a large range of wind power output predictabilities and makes Danish wind farms no sot heterogeneous from this point of view. Secondly, market rules are designed so that imbalance penalties are not that high (more or less 5% of the income for wind power producers). Changes in the market design could lead us to revise those conclusions. At last, predictability might become a relevant criterion when we will consider

[5] The formula for Spearman's correlation coefficient is detailed in Appendix B

aggregation, which will allow to significantly reduce imbalances penalties by increasing predictability, while the average energy yield will decrease as the most productive sites will be occupied. In addition, the financial benefit of predictability, from the producer's point of view, can also be obtained by a reduction of operation and maintenance costs, especially offshore, where lack of predictability leads not only to market imbalance costs but also to loss of availability (reduction of AEP) due to downtime periods.

Finally, the possibility to resort to hedging solutions does not seem to have any influence on the decision-making problem, as it does not overturn the rank of wind farms in income's hierarchy, even if it sometimes permits to achieve better results in terms of predictability and income revenue. This should lead us to neglect those solutions when we will deal with wind farms aggregation. However, it has to be noticed that results are highly dependent on the assumptions we made about the technical characteristics of the devices involved in physical hedging solutions.

Chapter 3

The economical value of predictability in case of aggregated wind farms

This chapter presents the relations between aggregation, predictability and market revenue in case of aggregated wind farms. First, both spatial smoothing effect and portfolio effect are described. Then we investigate the role of the aggregation parameter on predictability and market revenue, especially spatial ones, and we define strategies for the aggregation of wind farms. Results are given for Denmark.

3.1 The spatial smoothing effect and the reduction of imbalance penalties

As previously said, wind is an intermittent energy source, which makes challenging the prediction of wind power production output. However, spatially distributed wind power plants can take advantage of a smoothing effect and thus reduce the uncertainty and the cost from it.

3.1.1 The spatial smoothing effect from aggregating wind farms

Spatially dispersed wind power plant can take advantage of the fact that, if the area is large enough, wind may be blowing at some locations although not in at a specific site. This makes the wind power production output less variable at a region size and thus more predictable. Consequently, prediction errors at one specific site are likely to be compensated by opposite prediction errors from other sites, which may reduce the energy imbalances when participating in the electricity market, and thus reduce the imbalance penalties.

This phenomenon has been widely studied in the literature, for limited to wide regions. For example, see [20] for getting european scale results. In turn, we calculated the Normalized Mean Absolute Deviation of wind power production for every possible combination Ω among a set of 16 Danish wind farms, where:

$$NMAD(P^{\Omega}) = \frac{1}{24N}\sum_{i=1}^{N}\sum_{h=12}^{H=35}(|\frac{P^{\Omega}_{i,h}}{P^{\Omega}_{nom}} - \frac{1}{24N}\sum_{i=1}^{N}\sum_{h=12}^{H=35}(\frac{P^{\Omega}_{i,h}}{P^{\Omega}_{nom}})|)$$

$$\forall i = 1, \ldots, N$$

$$P^{\Omega} = \sum_{\omega \in \Omega} P^{\omega}$$

Wind power production variability asymptotically decreases as the region size increases. This is shown in Figure 3.1. The points are colored by aggregation cardinality, $card(\Omega)$.

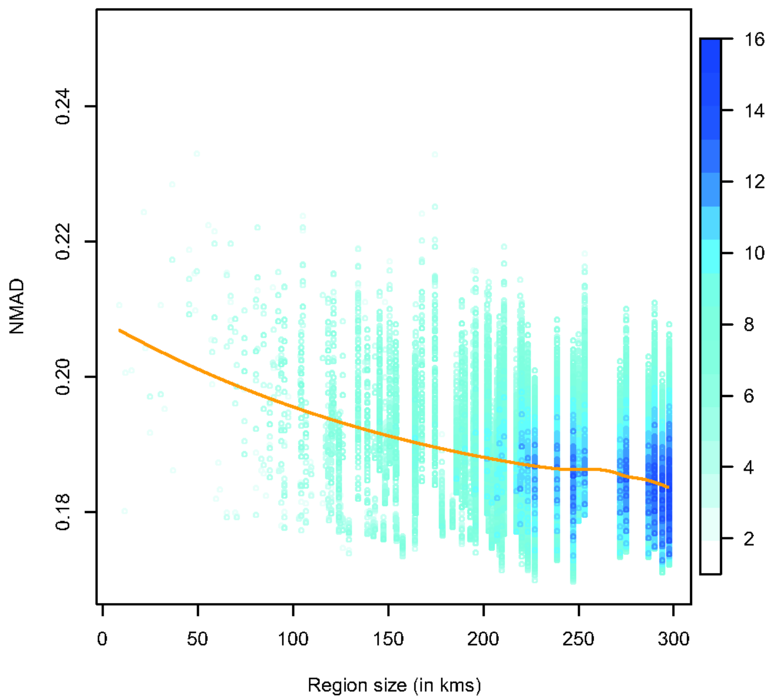

Figure 3.1: NMAD and region size. Points are colored by aggregation cardinality, refer to the right scale. The orange curve represents a local polynomial fit.

In [31], the limits of wind farm grouping for prediction error compensation are explored. It appears that the maximum reduction in the Error Index $EI(Err^{\Omega}) = \sum_{i=1}^{N} \frac{|\hat{P}^{\Omega}_i - P^{\Omega}_i|}{\hat{P}^{\Omega}_i}$ occurs when aggregating the third wind farm. The aggregation of more wind farms brings quite negligible improvement as regards to this indicator. However, the EI is not a suitable criterion in the context of

market participation as it gives different weights to the absolute deviation from the power output measure.

3.1.2 The portfolio effect : benefits from the market participation of aggregated wind farms

The smoothing of benefits due to lower variability is known as the portfolio effect. This is basic concept in finance and can be extended to the management of wind power plants portfolios. In our case, the portfolio effect refers to the mitigation of imbalance penalty within a portfolio of operating wind farms with regard to the case of those wind farms taken individually.

There exist several quantifications of the portfolio effect in the literature, when portfolio effect is studied for the power production output. Yet few studies have dealt with the portfolio effect understood as a mitigation of imbalance penalty.

For example, [14] simulates the participation of all the possible combinations among 20 Danish wind farm and find that the averaged imbalance penalty reduction is 7.58% for cardinality 2, 15.05% for cardinality 5 and reaches a 20.90% reduction level for cardinality 20. None of the combination allows a reduction greater than 25%.

In [41], a method is proposed to operate an optimal distribution of a given number of wind turbines over a given number of pre-selected sites, in terms of lowest wind power output variability.

3.2 The economical value of predictability when aggregating several wind farms

This section examines the value of predictability when wind farms are aggregated on the basis of highest yield. In this case, we seek to quantify the benefit for integrating predictability in this decision-making process. The idea is that the capacity factor decreases as wind farm with lower yield are aggregated while predictability tends to increase due to the smoothing effect.

3.2.1 Identifying the most relevant parameters of aggregation

The problem of aggregation characterization rapidly increases as the size of aggregation increases. Consequently, we are interested in selecting some particularly relevant variables that could characterize the inertia of aggregation and that appear to be key variables that perform well in explaining the income of a particular wind farms pool and that we could make use of it in an aggregation strategy. We distinguish three types of those parameters: technical parameters, spatial parameters and production-related parameters. Aggregation cardinality, total installed power and dispersion of nominal rated power belong to the first type. The distance between wind farms (whatever the metrics) belong to the second type. Wind power variability, wind power predictability, capacity factor and the correlation of prediction errors belong to the third type.

Here are some variable that could be considered to characterize the aggregation of wind farms.

- Technical parameters:
  - Aggregation cardinality is the number of wind farms within the wind farm pool, i.e. $card(\Omega)$.
  - Total installed power is the sum of installed powers within the wind farm pool, i.e. $\sum_{\omega\in\Omega} P^{\omega}_{nom}$.
  - Dispersion of nominal rated power is the standard deviation of individual nominal rated powers within the wind farm pool, i.e. $std(P^{\omega}_{nom}), \quad \omega \in \Omega$.
- Spatial parameters:
  - Region size is the maximum of pairwise distances within the wind farm pool, i.e. $max(distance(\omega_i, \omega_j)), \omega_i \neq \omega_j and \omega \in \Omega$.
  - Mean distance is the mean of pairwise distances within the wind farm pool, i.e. $mean(distance(\omega_i, \omega_j)), \omega_i \neq \omega_j and \omega \in \Omega$.
  - Standard deviation of distance is the standard deviation of pairwise distances within the wind farm pool, i.e. $std(distance(\omega_i, \omega_j)), \omega_i \neq \omega_j and \omega \in \Omega$.
- Production-related parameters:
  - Minimum of correlations is the minimum of pairwise prediction errors correlation within the pool, i.e. $min(corr(\omega_i, \omega_j)), \omega_i \neq \omega_j$, where $corr(\omega_i, \omega_j)$ is the correlation of prediction errors between wind farm ω_i and wind farm ω_j.
  - Mean of correlations is the mean of pairwise prediction errors correlation within the pool, i.e. $mean(corr(\omega_i, \omega_j)), \omega_i \neq \omega_j$.
  - Wind power variability, when measured by NMAD, expresses the dispersion of normalized aggregated wind power output from its median. The formula is given in 3.1.1.
  - Wind power predictability when measured by NMAE, expresses the normalized mean absolute wind power prediction error. The formula is given in 2.1.3.
  - Capacity factor is the mean rated power over a period, i.e. $mean(\frac{P^{\Omega}}{P^{\Omega}_{nom}})$.

As a first approach, we compute a market participation simulation for each of the possible combinations among a sample of 16 wind farms. The total number of simulations, N_{simu} is then the sum for each cardinality i of the number of permutation of i among 16 :

$$N_{simu} = \sum_{i=1}^{16} \binom{16}{i} = 2^{16} - 1 = 65535$$

Then, as suggested in [14], we conduct a normed PCA (Principal Component Analysis) of this aggregation data in order to identify the variables that most characterize data's dispersion. Excluding normalized imbalance penalties and producer's income from the set of data, and keeping them as illustrative variables, we can first compute the plot of the eigenvalue of each PCA dimension.

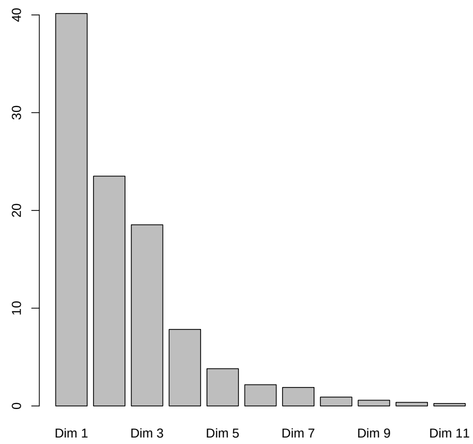

Figure 3.2: Plot of eigenvalues, for each PCA dimension

It appears that the first four axes explain 90% of the total inertia, which is quite satisfactory. To interpret the signification of the first four dimensions, we take a look at the dimension versus variables correlation graph.

	eigenvalue	inertia	cumulative inertia
comp 1	4.42	40.14	40.14
comp 2	2.59	23.51	63.65
comp 3	2.04	18.53	82.18
comp 4	0.86	7.83	90.01

Table 3.1: Table of eigenvalues, for each PCA dimension

It appears that the first component of PCA have a 40% part in the explanation of total inertia. It seems to characterize the magnitude of the smoothing effect as it is highly positively correlated with region size (i.e. maximum of distance between the aggregated wind farms) and negatively correlated with both minimum prediction error correlation among wind farms and predictability. Minimum prediction error is exactly opposite to region size so both variables bring the same information. With a wind resource assessment perspective, we should prefer to deal the region size because it is controllable, while correlation of prediction error is not known before wind farms are constructed. The second axe also contributes to explain the total inertia by almost a quarter. Wind farms groups which have positive coordinates along the second axes tend to have higher yield, higher power output variability and lower predictability. Not surprisingly, the normalized imbalance penalty, taken as illustrative variable, is well resumed by predictability. Similarly, the normalized income is well resumed

by the yield, measured by the capacity factor. Logically, the sum of nominal power and the cardinality of aggregation have opposite coordinates on this axe, which shows that the aggregation of more and more wind farms is associated to a lower average yield.

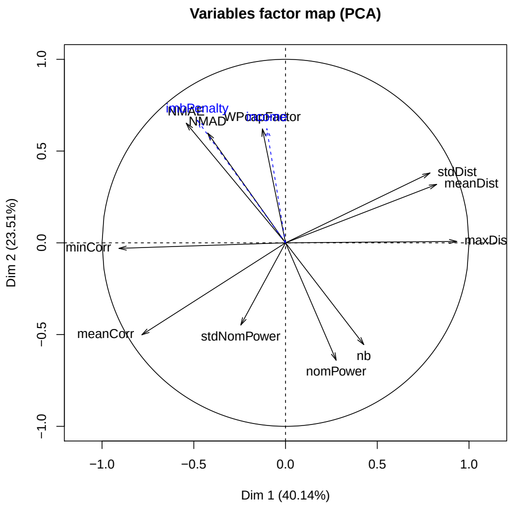

Figure 3.3: Correlation of variables with PCA components axes 1 and 2. Here are the signification of the variable labels: $NMAE$ is for NMAE criterion, $NMAD$ for NMAD criterion, *minCorr* and *meanCorr*, respectively for the minimum and the mean correlation of prediction errors between pairs of wind farms, *nomPower* and *stdNomPower*, respectively for the nominal power and the standard deviation of nominal powers among every combination of wind farms, *nb* for aggregation cardinality, and *stdDist*, *meanDist* and *maxDist* respectively for standard deviation, mean and maximum of pairwise distances between wind farms. Illustrative variable are reported in blue.

Axes 3 and 4 are less informative. Wind farms groups with positive coordinates along axe 3 have high power output variability but high capacity factor and thus high income and rather high cardinality. In axe 4, only the standard deviation of nominal power has a significant positive coordinate in the unit circle.

To sum up, and according to the PCA analysis we conducted, the main parameters of aggregation are:

- The distance between wind farms, understood as the maximum of distance between pairs of wind farms (i.e. region size).
- The aggregation cardinality.
- The dispersion of wind farms' nominal power.
- The sum of nominal power.
- Predictability and wind power output variability.

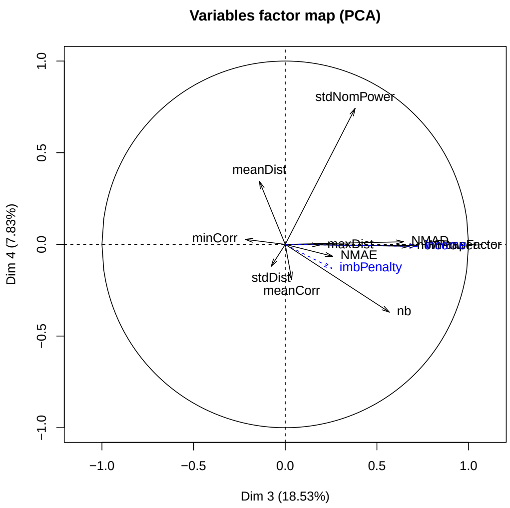

Figure 3.4: Correlation of variables with PCA components, axes 3 and 4.

From those parameters, we can first examine the influence of aggregation cardinality on predictability, imbalance penalties and market revenue. As previously been stated, the benefits from aggregation asymptotically decrease when cardinality increases and a large share of the potential benefit in terms of increased predictability and reduced imbalance penalty is reached at a relatively low cardinality. Exactly, the best combination of wind farms with respect to predictability (NMAE = 0.088) has a cardinality of 4 and the expected predictability show no significant gain after the aggregation of 10 wind farms. It seems that we have to focus on the results for low dimension aggregation of wind farms.

Expected income and capacity factor are not dependent on aggregation cardinality. This highlights a kind of decoupling in the relation we had stated between capacity factor and NMAE. The reason is the NMAE is now the sum of two terms. The first term is the average predictability of aggregated wind farms and the second term correspond to the smoothing effect.

3.2.2 Spatial analysis of the effect of aggregation on predictability and market revenue

We have seen from the previous analysis that spatial parameters are key, especially the region size of aggregation.

As previously been stated in [40], the main parameter of aggregation are the spatial extension and the number of wind farms. Figure 3.2.4 shows pairwise hourly prediction error correlations with respect to the distance between wind farms. The red curve is an exponential fit of shape $exp(\frac{-distance}{\delta})$ and the orange curve is a local polynomial fit. The decay parameter δ is equal to 210 and expresses the size of the region where the maximum smoothing effect has already occurred.

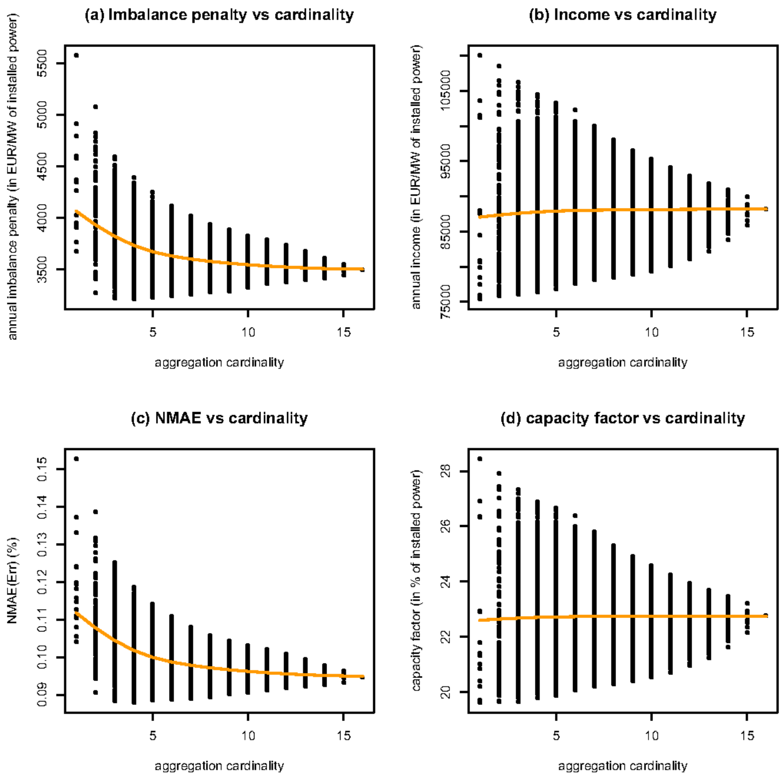

Figure 3.5: Annual income, annual imbalance penalty and NMAE with respect to aggregation cardinality.

We see that nearness not necessarily implies high correlation. Many pairs wind farms that are not farther than 50 kms have a correlation coefficient below 0.6, which is not that high. This highlights the possibility for smart aggregation strategies based on spatial criteria.

What is more, one should take care, not only on the distance between the initial pool and the wind farm to aggregate, but also the aggregation direction. When distance is fixed, some directions may be more favorable to the compensation of prediction errors, and thus to the reduction of imbalance penalties. In fact, there is prevailing wind speed directions. The compensation of errors will not be homogeneous at each direction. It is said to be *anisotropic*. This means that adding a wind farm to the pool is likely to be more or less interesting in terms of income depending on its relative position from the pool. To illustrate this, let's consider the case of one wind farm when wanting to aggregate another one. We can clearly see that direction matters. Correlations between actual prediction errors of wind farm n°59 and prediction errors of the others one hour ago are very small because it takes less than an hour for weather patterns to cross Denmark.

One can also consider the direction of aggregation. The rationale behind that

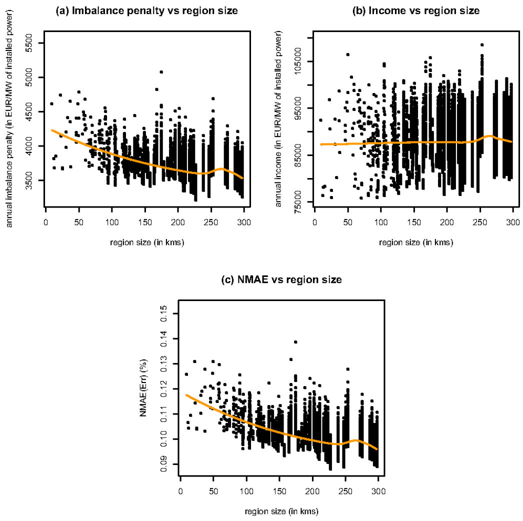

Figure 3.6: Annual income, annual imbalance penalty and NMAE with respect to region size.

is that the propagation of wind power prediction error is anisotropic because of the existence of prevailing wind direction. This means that adding a wind farm to the pool is likely to be more or less interesting in terms of income depending on its relative position from the pool.

According to Figure 3.7, there exists a prevailing west to east wind speed direction which makes the power output prediction errors between wind farms more correlated along this particular direction. So aggregation is likely to be more beneficial if the second wind farm's location is somewhere in the south of Denmark, while wind farm n°59 is located in the northeastern part of the country, in order to take advantage from the decorrelation of prediction errors.

3.2.3 Results and analysis

We can study the link between annual income, predictability and capacity factor, for the *aggregation* case (see Figure 3.10).

Not surprisingly, (a) income and capacity factor are still highly correlated (Pearson's correlation coefficient equals 0.998) ; (b) When considering medium-range capacity factor, some combination of wind farms appear to differ slightly from the average behavior. However, we can't see state clearly on graph (c) if there is gain to get from higher predictability when comparing combination of similar capacity factor. Graph (d) shows a negative relation between the residuals of the linear fit (red curve on graph (a)) and the NMAE criterion. There is no gain from predictability being exceptionally high (NMAE = 10%) but very low predictability (NMAE = 13%) is associated to an annual expected loss of 605€by MW of installed power, more or less 0.6% of the total income. When carrying an analysis of variance on the income, and considering both NMAE and capacity factor in the model, the NMAE appears to explain only

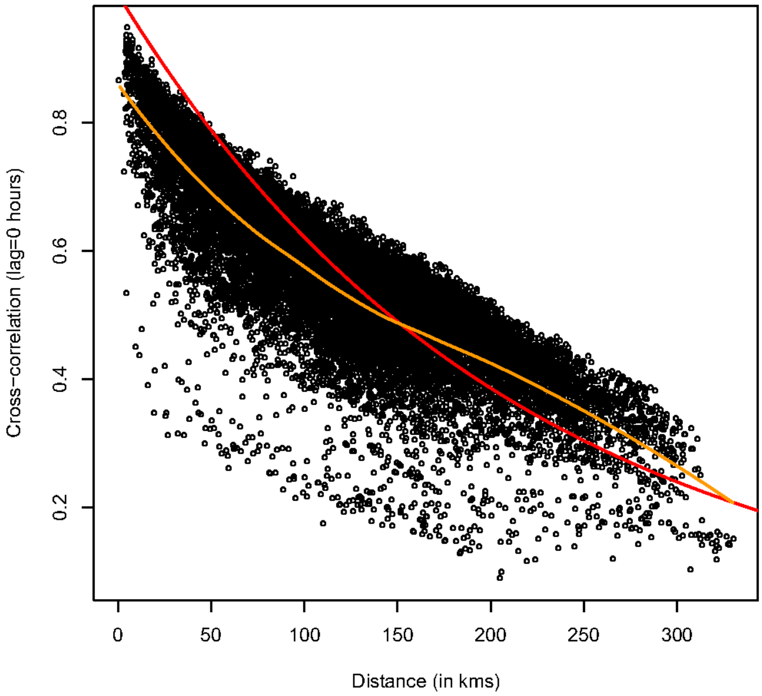

Figure 3.7: Cross-correlation for every pair of wind farms with respect to distance (lag=0hours).

0.04% of the overall variance.

3.3 Strategies for the aggregation of wind farms

3.3.1 Definition of different aggregation strategies

We need to investigate the ways that could provide optimal aggregation of wind farms. Many studies have dealt with this decision problem, for the optimization targets differ with one another: smoothest aggregate output, maximum benefit, maximum predictability, maximum reliability, maximum output.

Our aim is to find optimal solutions in terms of *maximum income* without explicitly evaluating every possible combination, while computational capabilities are limited. In fact, if considering the whole set of 182 wind farms, the total number of combination is the sum for each cardinality i of the number of permutation of i wind farms among 182:

$$N_{simu} = 2^{182} - 1$$

Using a greedy forward algorithm which aggregates farms on the basis of

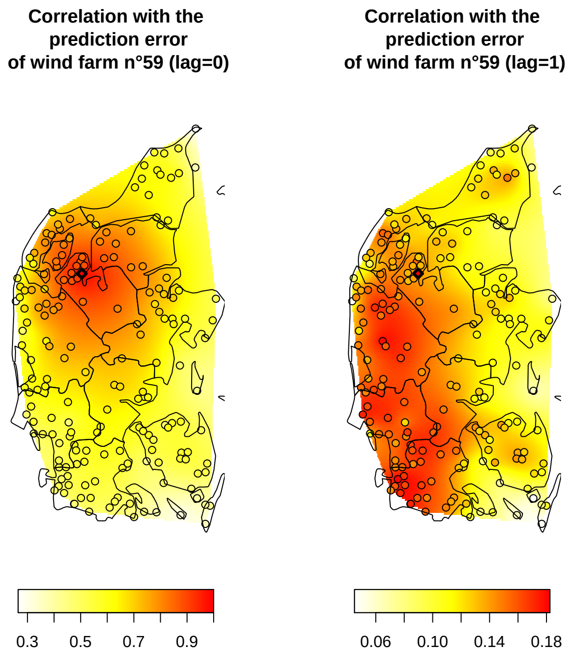

Figure 3.8: (a) Correlation between prediction errors of wind farm n°59 and prediction errors of the others (b) Correlation between actual prediction errors of wind farm n°59 and prediction errors of the others one hour ago. The black lozenge represents wind farm n°59.

the highest predictability, over a quite wide region including 17 representative wind farms from Ireland, France and Spain, the authors of [26], found that predictability rapidly increases as cardinality goes up. Up to 10 wind farms, the gain is small and the highest predictability level is reached when 14 wind farms are grouped. Consequently, the imbalance penalties are divided by two and the ratio of producer's income over reference income is 97.59%. In other words, the value of the increase in predictability due to the smoothing effect is close to 2.5% of the reference income. On the Danish data, we expect to find a lower value because the magnitude of the smoothing effect should be lower as the region is narrower. However, the absolute value for economic assessment is not reported and, by the light of the results of chapter 2, one should not construct an aggregation strategy which relies only on predictability. In fact, aggregation on the basis of predictability means that incremental wind farms are likely to have smaller and smaller capacity factors and thus provide a smaller income. Imbalance penalties may decrease in percentage but not in absolute value.

We saw in chapter 2 that the level of income that the producer can expect to earn is mainly explained by the energy yield of the specific site, that is to say by the capacity factor. Consequently, every aggregation strategy, from a power producer's point of view, should rely on the maximization of this target variable. There exist simple algorithms to construct wind farm aggregation. A first approach is to consider greedy forward selection. To consider which wind farm is

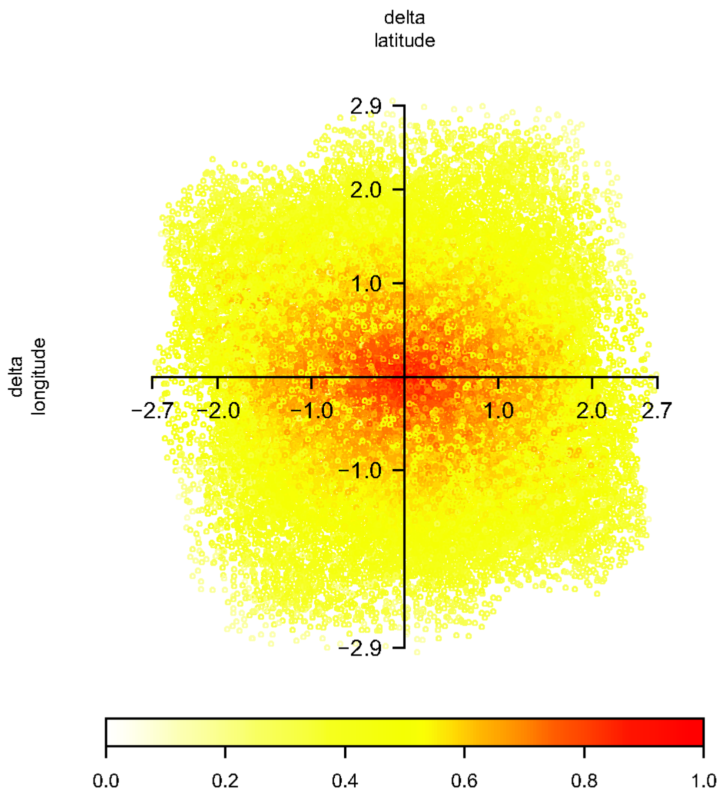

Figure 3.9: Correlation of prediction errors with respect to relative geographical coordinates. Each point represent the correlation for a pair of wind farms and delta longitude (latitude) refers to the difference in longitude (latitude) between those two wind farms.

to be aggregated to the pool, we consider alternatively three variable: capacity factor, mean of pairwise distances between wind farms and predictability. Here is the pseudo-code:

Algorithm 1 Greedy forward selection, by capacity factor

---

while $iter > maxIter$ **do**
  for i in pre-selected sites **do**
    aggregate i to the pool
    evaluate capacity factor
  end for
  select i such that the new combination has maximum capacity factor
  add wind farm i to aggregation
end while

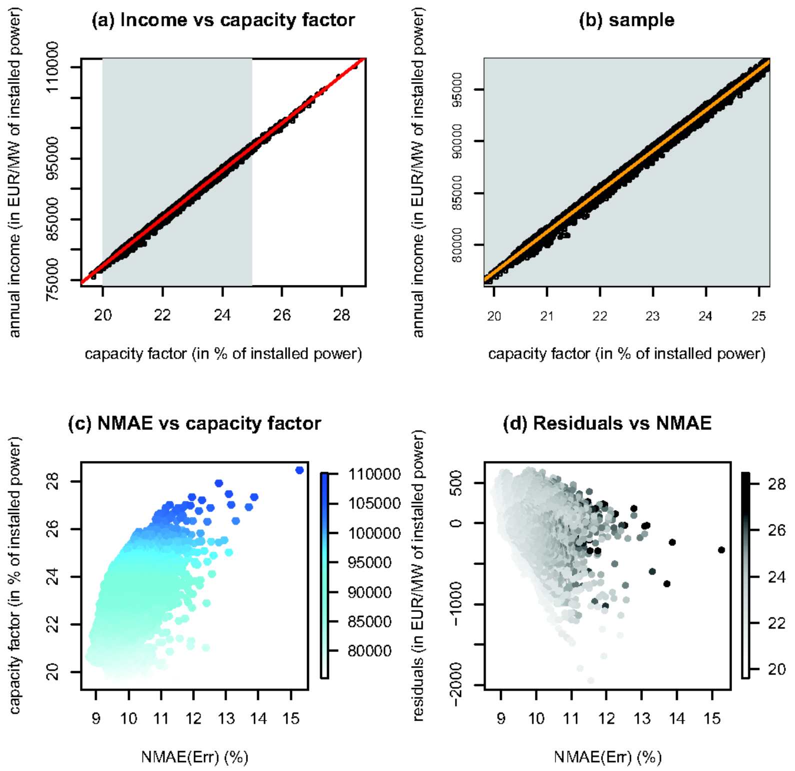

Figure 3.10: Plot of annual income versus capacity factor, *aggregation* case.

Algorithm 2 Greedy forward selection, by distance

while $iter > maxIter$ **do**
  for i in pre-selected sites **do**
    aggregate i to the pool
    evaluate distance
  end for
  select i such that the new combination has maximum mean pairwise distances
  add wind farm i to aggregation
end while

3.3.2 Results and analysis

Let us consider a single wind farm, n°59. We seek to evaluate each aggregation strategy in terms of improved revenue and reduced imbalance penalties. Here are the features of wind farm n°59 (Table 3.2):

When computing 1, 2 and 3, we obtain the following results (Table 3.3):

Algorithm 3 Greedy forward selection, by predictability

while $iter > maxIter$ **do**
  for i in pre-selected sites **do**
    aggregate i to the pool
    evaluate predictability
  end for
  select i such that the new combination has maximum predictability
  add wind farm i to aggregation
end while

Wind farm n°59	
Latitude	56°39'N
Longitude	9°02'E
Nominal power (MW)	8.355
Capacity factor	19.58
NMAE(Err) (%)	10.8
Revenue ($\times 1000€/MW$)	74.5
Imbalance penalty ($\times 1000€/MW$)	4.0
Revenue (in % of perfect revenue)	94.9

Table 3.2: Characteristics of wind farm n°59.

The aggregation of wind farms by the capacity factor performs well and permits to reach a normalized income that is 66% higher than the initial income at cardinality 10. Imbalance penalties decrease slightly. Between the two other strategies, the one using mean distance performs best.

3.4 Conclusion

When considering the installation of several wind farms, some parameters were shown to characterize aggregation. Among them, distance between wind farms and cardinality are key. Low cardinality, below 10, is sufficient to cause increased revenue and reduced imbalance penalties. The producer's income depends so highly on the capacity factor that no other variable was found to bring big additional information. This is illustrated by the results obtained when aggregating several wind farms with regard to different criterion.

Cardinality	2	5	10	15	20
Algorithm 1					
Capacity factor (%)	31.68	32.32	31.76	30.97	30.31
Revenue ($\times 1000€/MW$)	122.7	125.6	123.6	120.6	118.1
Imbalance penalty ($\times 1000€/MW$)	4.2	4.0	3.9	3.7	3.6
Revenue (in % of perfect revenue)	96.7	96.9	97.0	97.0	97.0
Algorithm 2					
Mean of pairwise distances (in kms)	232.6	328.4	330.3	330.3	330.3
Revenue ($\times 1000€/MW$)	97.4	93.4	96.8	97.1	95.8
Imbalance penalty ($\times 1000€/MW$)	3.4	3.1	3.1	3.1	3.1
Revenue (in % of perfect revenue)	96.6	96.8	96.9	96.9	96.9
Algorithm 3					
NMAE(Err) (%)	8.9	7.6	7.3	7.3	7.4
Revenue ($\times 1000€/MW$)	78.3	81.3	82.0	82.0	81.6
Imbalance penalty ($\times 1000€/MW$)	3.1	2.7	2.6	2.7	2.8
Revenue (in % of perfect revenue)	96.1	96.8	96.9	96.9	96.8

Table 3.3: Value of predictability for some cardinalities.

Conclusion and Perspectives

From our study, it appears that predictability is not likely to be a relevant parameter to consider when having to choose the best-suited location among two pre-selected sites to build a wind farm in terms of producer's income. The gain from including predictability is very small compared to that explained by capacity factor but it is still significant. As stated in chapter 2, a particularly high (resp. low) level of predictability can only cause an additional gain (resp. loss) of 0.5% of the reference income. But the imbalance penalties represent a small share of the revenue and making the market rule more penalizing would cause high predictability to represent 5% of the reference income. That could be important when comparing with the producer's profit margin.

Wind farm capacity is the main parameter to consider.

When considering wind farm aggregation, differences in predictability between several combinations only explain a marginal part of the discrepancy in terms of producer's income. The additional gain from high predictability is low so that the arbitrage window between two different combinations of wind farms with similar capacity factor is very narrow. Highly beneficial wind farm pools are characterized by low cardinality, and maximum coverage of Denmark (i.e. region size). They also group the most performing wind farms in terms of capacity factor (i.e. energy yield). Simple aggregation strategies lying on other parameters than capacity factor perform poorly.

Considering the aggregation benefits on predictability it is clear that the Danish test case is not the most suitable to put it in evidence due to the correlation of the wind climate and predictability over the area of study. This is stated in the last chapter. More significant figures are obtained when considering a portfolio of wind farms across Europe or in countries with high wind climate spatial variability like Spain or France. In these areas a large wind farm owner can improve predictability substantially by aggregation, reducing imbalance costs quite appreciably. In line with previous comments, it is not part of the decision making process for the developer since AEP will always lead. However, it can be an important aspect when considering a EU-wide electricity market where aggregation from distant wind farms can be a common trading strategy.

If we were to draw some perspectives, two directions arise for further work:

- One might carry a study on the relevance of predictability as a wind resource estimation parameter, not with a power producer's point of view, but through the eyes of a utility which aims at balancing electricity supply

and demand. Our work could be adapt to account for the utility specific requirements which are different from market benefit.

- Our results can be used to set up market rules that create incentives for higher predictability, which is desirable from the electric system point of view. It would be easy to increase this gain by changing the market rules. If multiplied by 10, the penalty would raise predictability to an important level hence creating incentives for the high predictability. This type of penalty transformation would certainly have drawback and more sophisticated changes in the market rules will be the purposes of further studies during the end of my training period.

Appendix A

Estimating predictability

The aim of this appendix is to provide evidences that the predictability of wind power output can be estimated and to quantify the incertitude that goes with this estimation, in the light of a French case-study.

Predictability is measured by the Normalized Mean Absolute Error. Here is the formula for NMAE for a prediction run at noon ($h = 0$):

$$NMAE(Err^{\omega}) = \frac{1}{35} \sum_{h=1}^{H=35} \frac{|Err_h^{\omega}|}{P_{nom}^{\omega}}$$

$$Err_h^{\omega} = P_h^{\omega} - \hat{P}_h^{\omega}$$

Another predictability criterion was considered, i.e. the Normalized Root Mean Square Error, but this did not change the results. Here is the formula for NRMSE:

$$NRMSE(Err^{\omega}) = \frac{1}{35} \sqrt{\sum_{h=1}^{H=36} (\frac{Err_h^{\omega}}{P_{nom}^{\omega}})^2}$$

A.1 Methodology for estimating predictability

We propose a methodology in order to estimate predictability before the wind farm is installed. This requires historical wind speed data from the prospecting site, i.e. one or more anemometers have been recording enough wind speed data.

Then, wind speed data are converted to power output by passing through the power curve. The power curve is provided by the wind turbine manufacturer and it describes the relation between the wind speed (usually in m/s) and the rated power of the turbine. Once having this "extrapolated" wind power output measurements, we set it as input in a Random forest regression model. Wind speed and direction predictions are also incorporated in the model as predictors. When comparing this prediction with the real wind power output measurements, we obtain the prediction errors of the model and we are able to compute the NMAE criterion for each prediction run. We call this result the *measured vs estimated* predictability.

In order to compare with the predictability that is recorded when the wind farm is constructed and operational, we work with data from an operational French wind farms where two masts were installed, thus having both historical wind power output and historical wind speed and direction time series. This time, we run the Random forest regression model on this real-world wind power output data. We obtain prediction errors and we are able to compute the NMAE criterion for each run. We call this result the *measured vs predicted* predictability.

Figure A.1 describes this methodology. Green boxes refer to the provided data. Blue (resp. orange) boxes refer to the case where real power production data are used (resp. where only wind speed data are available, i.e. when no wind farm is constructed yet).

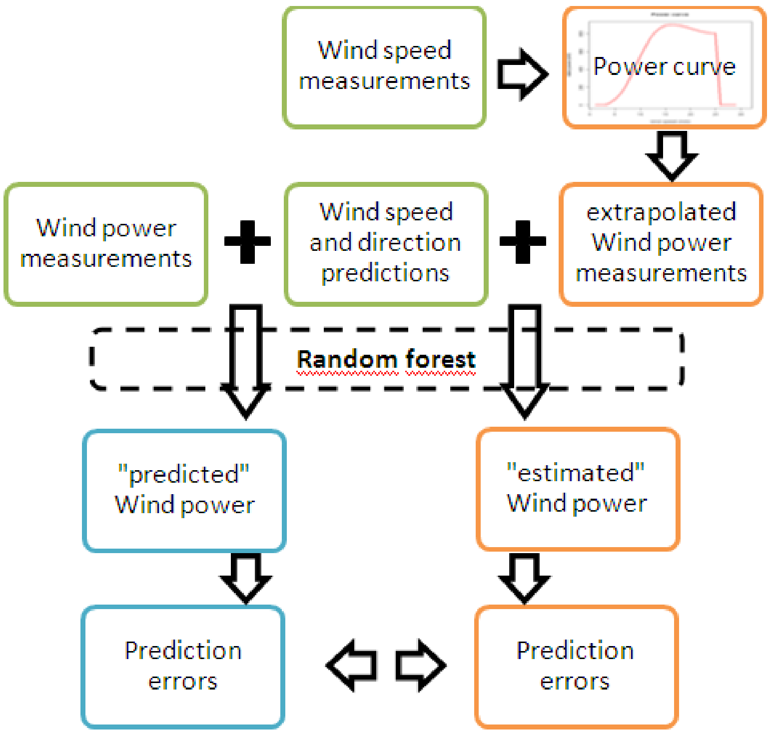

Figure A.1: Proposed methodology for estimating predictability.

A.2 Results for a French wind farm

We were supplied with the hourly power output data of a 9-turbines 9.720MW French wind farm, for a period ranging from February 2004 to September 2007. Two mast-based anemometers were recording the wind speed on the site during this period and we also got wind speed predictions. Wind speed and wind

power output times series were provided by the wind farm operator and the wind speed and direction predictions data were provided by Météo France, the French national meteorological service.

Consequently, we have two different *measured vs estimated* predictabilities, i.e. one per mast and nine different *measured vs predicted* predictabilities, i.e. one per turbine. We can first take a look at Figure A.2, which shows the two times series (for mast 1 and turbine 1). We can notice that the estimated predictability have difficulties in predicting relatively low or high values of real predictability. The discontinuity in both curves corresponds to missing data.

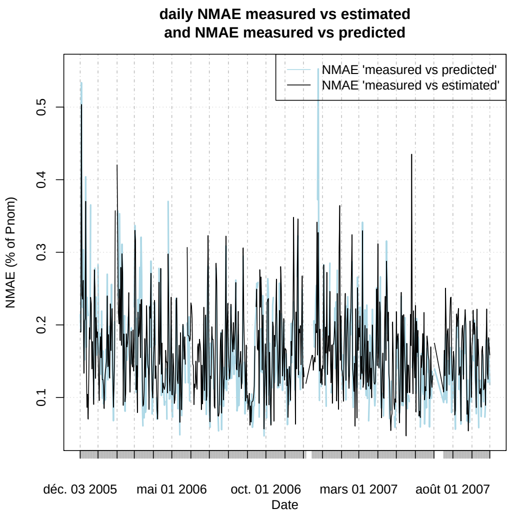

Figure A.2: Daily predictability, for mast 1 (black curve) and turbine 1 (blue curve).

However, the results are highly dependent on the turbine/mast that is considered. Figure A.3 (a) and A.3 (b) makes evident that not all turbines have the same predictability and that mast 1 performs better that mast 2 when estimating wind power output. Not surprisingly, predictability gets down when the look-ahead time increases, except between hours 18 and 27, approximately. It may be explained by some weather patterns.

If we focus on turbine 1 (turbine 1 has medium predictability), an interesting thing is to consider the two types of prediction errors, *measured vs predicted* and *measured vs estimated*. Figure A.4 shows for some look-ahead times. It is obvious that predictability is skewed to the right, whatever the horizon.

To make the discrepancy between the *measured vs predicted* prediction error distribution and the *measured vs estimated* prediction error distribution clearer, we can compare the quantiles of the first with those of the latter. This is done in Figure A.5 for horizons 12, 24 and 36. The empirical distribution estimation errors from mast 2 are skewed to the left compare with the empirical distribution of real errors, whatever the horizon that is considered. The distribution estimation errors from mast 1 better fits with the distribution of real errors. However, the estimations from both masts are less likely to produce high pos-

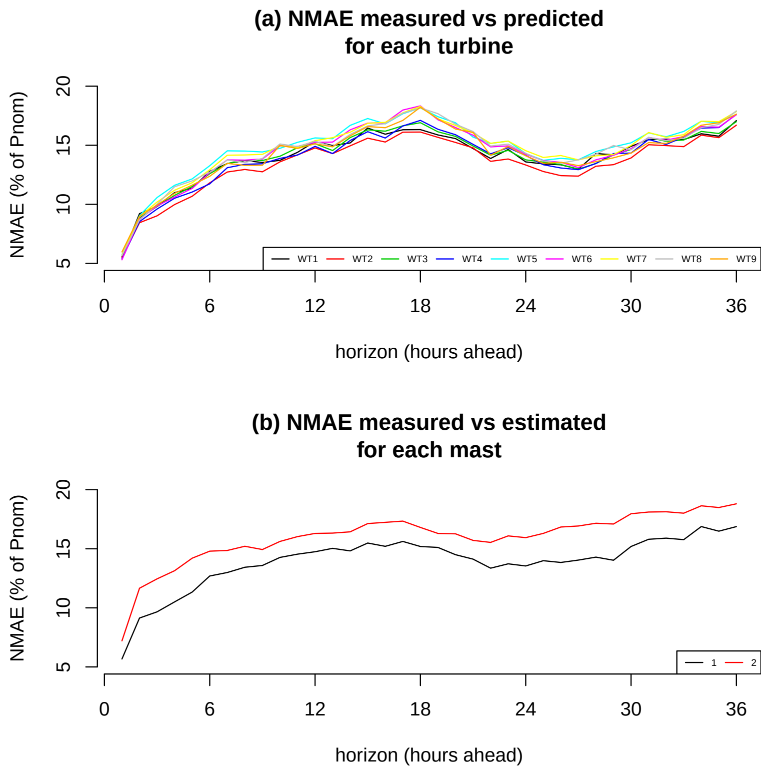

Figure A.3: Predictability h-hours ahead, (a) for each turbine (b) for each mast.

itive estimation errors than the model based on real power output data from turbine 1 does.

As regard to the quantification of NMAE uncertainty, we can compute the time series of NMAE errors, i.e. the difference between the real *measured vs predicted* NMAE and the *measured vs estimated* NMAE. For mast 2, this daily error is within $[-4.9, 4.2]$ points of percentage 80% of the time, within $[-6.3, 5.4]$ points of percentage 90% of the time and within $[-7.7, 8.0]$ points of percentage 95% of the time. The overall real NMAE (for all horizons and all prediction runs) is equal to 15.8%, and the estimated NMAE is slightly higher when considering mast 1 (16.3%) and much higher when considering mast 2 (18.3%).

An additional approach could have been tested. It consists in fitting a parametric distribution for wind data from a close wind farm/meteorological station (usually the Weibull distribution) and to generate hourly wind data according to this estimated distribution, transpose this data into power output data through the power curve, compute predictability and see how much the results differ from the both known and estimated predictability. The aim is to demonstrate whether it is possible to estimate predictability even if no on-site wind data is available.

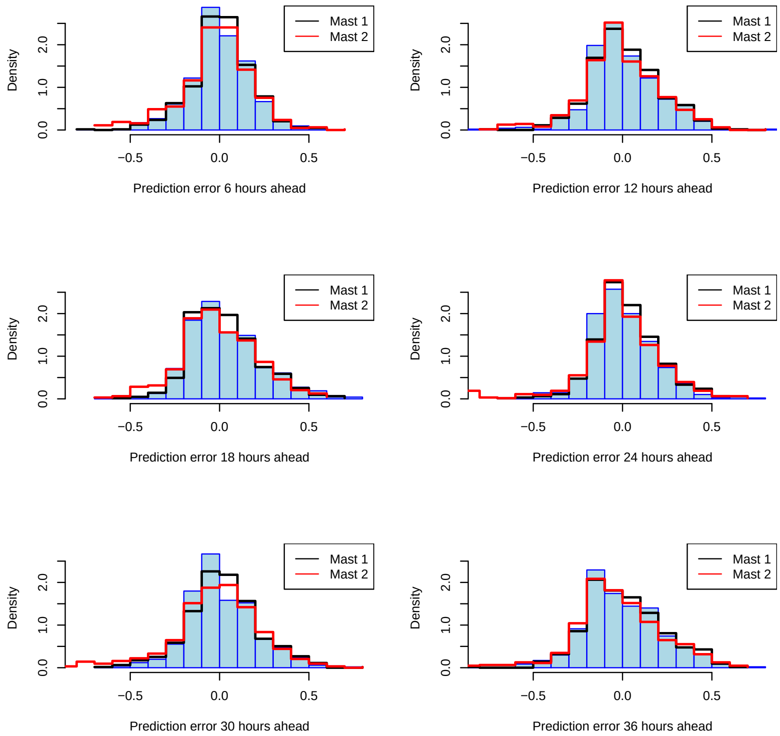

Figure A.4: Histogram of *measured vs estimated* prediction errors (black and red lines), compared to the histogram of *measured vs estimated* prediction errors.

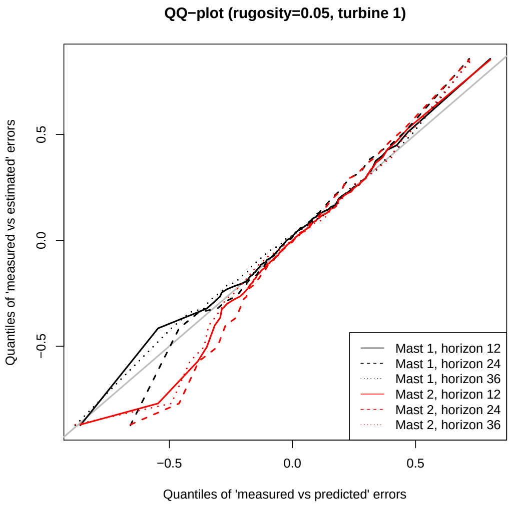

Figure A.5: Quantiles-Quantiles plot of *measured vs predicted* prediction errors and *measured vs estimated* prediction errors.

Appendix B

Pearson's and Spearman's correlation formulas

B.1 Linear correlation coefficient

Pearson's correlation coefficient between N realizations of two random variables $X(X_1, \dots, X_N)$ and $Y(Y_1, \dots, Y_N)$, is defined as follows:

$$\rho^P_{X,Y} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2}}$$

B.2 Rank coefficient correlation

Spearman's rank correlation coefficient is defined as the linear coefficient correlation between the ranked variables $x(x_1, \dots, x_N)$ and $y(y_1, \dots, y_N)$ associated to variables X, Y:

$$\rho^S_{x,y} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

Bibliography

[1] K.-O. Vogstad A. Coville, A. Siddiqui. The effect of missing data on wind resource estimation. *Energy*, 2011.

[2] W. Kempton A. Dhanju, P. Whitaker. Assessing offshore wind resources: An accessible methodology. *Renewable Energy*, 33:55–64, 2008.

[3] P. Armstrong A. Fontaine. Uncertainty analysis in energy yield assessment. In *Proc. of the European Wind Energy Conference*, Milan, Italy, May 2007.

[4] C. Johnson A. Schwarz A. Garrad G. Hassan A. Tindal, K. Harman. Validation of the gh energy and uncertainty predictions by comparison to actual production. In *In Proc. of the AWEA Wind Resource and Project Energy Assessment Workshop*, Portland, USA, September 2007.

[5] K. W. Ayotte. Computational modelling for wind energy assessment. *Journal of Wind Engineering and Industrial Aerodynamics*, 96:15711590, 2008.

[6] F. Bourry. *Management of uncertainties related to renewable generation participation in electricity market.* PhD thesis, Ecole des Mines de Paris, Center for Energy and Processes, Rue Claude Daunesse, Sophia Antipolis, 2009.

[7] J. Harding C. Bastide. Energy yield assessment - are you doing it right ? In *Proc. of the European Wind Energy Conference*, Milan, Italy, May 2007.

[8] G. Tina C. Brunetto. Wind generation imbalances penalties in day-ahead energy markets: The italian case. *Electric PowerSystemsResearch*, 2010.

[9] M. Saguan C. Hiroux. Large-scale wind power in european electricity markets : Time for revisiting support schemes and market designs? *Energy Policy*, 38:3135–3145, 2010.

[10] K. Burges C. Klessmann, C. Nabe. Pros and cons of exposing renewables to electricity market risks a comparison of the market integration approaches in germany, spain, and the uk. *Energy Policy*, 36:3646–3661, 2008.

[11] A. N. Celik. On the distributional parameters used in assessment of the suitability of wind speed probability density functions. *Energy Conversion and Management*, 45:1735–1747, 2004.

[12] G. Kariniotakis F. Bourry. Strategies for wind power trading in sequential short-term electricity markets. In *Proc. of the European Wind Energy Conference*, Marseille, France, March 2009.

[13] G. Kariniotakis F. Bourry, L. M. Costa. Risk-based strategies for wind/pumped-hydro coordination under electricity markets. In *In Proc. of the Bucarest Power Tech Conference*, Bucarest, Romania, July 2009.

[14] G. Kariniotakis F. Bourry, N. Siebert. Advantages of virtual power plants composed of aggregated wind farms for trading in short-term electricity markets. 2011.

[15] L. M. Costa G. Kariniotakis F. Bourry, J. Juban. Advanced strategies for wind power trading in short-term electricity markets. In *Proc. of the European Wind Energy Conference*, Brussels, Belgium, March/April 2008.

[16] L. Ferreira. Evaluation of short-term wind predictability. *IEEE TRANSACTIONS ON ENERGY CONVERSION*, 7, 1992.

[17] E. Sreevalsan S. Rajapandian G. M. J. Herbert, S. Iniyan. A review of wind energy technologies. *Renewable and Sustainable Energy Reviews*, 11:1117–1145, 2007.

[18] E. P. Martín G. M. Montes. Profitability of wind energy: short-term risk factors and possible improvements. *Renewable and Sustainable Energy Reviews*, 11:2191–2200, 2007.

[19] G. Strbac G. N. Bathurst, J. Weatherill. Trading wind generation in short term energy markets. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 17, 2002.

[20] G. Giebel. *On the benefits of distributed generation of wind energy in Europe*. PhD thesis, Carl von Ossietzky Universitt Oldenburg, 2000.

[21] J. Perron H. Ibrahim, A. Ilinca. Energy storage systemscharacteristics and comparisons. *Renewable and Sustainable Energy Reviews*, 12:12211250, 2008.

[22] H. Holttinen. Optimal electricity market for wind power. *Energy Policy*, 33:2052–2063, 2005.

[23] J. G. Usaola J. M. Angarita. Combining hydro-generation and wind energy biddings and operation on electricity spot markets. *Electric Power Systems Research*, 77:393–400, 2007.

[24] L. Söder J. Matevosyan. Minimization of imbalance cost tradingwind power on the short-term power market. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 21, 2006.

[25] S.-D. Kwon. Uncertainty analysis of wind energy potential assessment. *Applied Energy*, 87:856–865, 2010.

[26] E. Pascal E. Cantero Y. Loureiro P. M. Correia S. Lozano J. Sanz L. Fras, U. Irigoyen. Functional data analysis applied to the problem of wind farm aggregation. In *Proc. of the European Wind Energy Conference*, Brussels, Belgium, March 2011.

[27] O. Rathmann E. L. Petersen B. H. JJ. Bagder N. G. Mortensen L. Landberg, L. Myllerup. Wind resource estimation - an overview. *Wind Energy*, 6:261–271, 2003.

[28] J. Juban G. Kariniotakis L. M. Costa, F. Bourry. Management of energy storage coordinated with wind power under electricity market conditions. 2008.

[29] S. S. Fall M. Kasse E. Sylla S. Thioye L. Thiaw, G. Sow. A neural network based approach for wind resource and wind generators production assessment. *Applied Energy*, 87:17441748, 2010.

[30] J. F. Manwell M. A. Lackner, A. L. Rogers. Uncertainty analysis in wind resource assessment and wind energy production estimation. 2007.

[31] I. Mart M. Gastón, L. Fras. Exploring the limits of wind farm grouping for prediction error compensation. In *Proc. of the European Wind Energy Conference*, Athens, Greece, February/March 2006.

[32] W. L. Kling M. Gibescu, E. W. van Zwet. Optimal bidding strategy for mixed-portfolio producers in a dual imbalance pricing system. 2008.

[33] F. Chapman M. R. Milligan, A. H. Miller. Estimating the economic value of wind forecasting to utilities. In *Proc. of WindPower '95*, Washington, D.C., March 1995.

[34] P. Pinson M. Zugno, T. Jónsson. Trading wind energy in a liberalised market environment: a real-world test case in eastern denmark. 2010.

[35] P. E. Morthorst. Wind power and the conditions at a liberalized power market. *Wind Energy*, 6:297–308, 2003.

[36] J. S. Rodrigo. State-of-the-art of wind resource assessment. Technical report, WAUDIT Deliverable D1.1, 2010.

[37] A. Gastli S. Al-Yahyai, Y. Charabi. Review of the use of numerical weather prediction (nwp) models for wind energy assessment. *Renewable and Sustainable Energy Reviews*, 14:3192–3198, 2010.

[38] L. Söder T. Ackermann. An overview of wind energy-status 2002. *Renewable and Sustainable Energy Reviews*, 6:67–128, 2002.

[39] H.-P. Waldl U. Focken, M. Lange. Previento - a wind power prediction system with an innovative upscaling algorithm. 2001.

[40] K. Mnnich H.-P. Waldl H. G. Beyer A. Luig U. Focken, M. Lange. Short-term prediction of the aggregated power output of wind farms - a statistical analysis of the reduction of the prediction error by spatial smoothing effects. *Journal of Wind Engeneering and Industrial Aerodynamics*, 90:231–246, 2002.

[41] C. Singh Y. Degeilh. A quantitative approach to wind farm diversification and reliability. *Electrical Power and Energy Systems*, 33:303–314, 2011.