Influence of local wind speed and direction on wind power dynamics - Application to offshore very short-term forecasting

C. Gallego[a,*], P. Pinson[b], H. Madsen[b], A. Costa[a], A. Cuerva[c]

[a] *Wind Energy Unit, CIEMAT, Avd. Complutense 22, 28040. Madrid, Spain. Tel: +34 913466360*

[b] *DTU Informatics, Technical University of Denmark, Richard Petersens Plads 305, 2800 Kgs. Lyngby, Denmark.*

[c] *IDR/UPM, E.T.S.I.Aeronáuticos, Universidad Politécnica de Madrid, Pza. Cardenal Cisneros 3, 28040. Madrid, Spain*

Abstract

Wind power time series usually show complex dynamics mainly due to non-linearities related to the wind physics and the power transformation process in wind farms. This article provides an approach to the incorporation of observed local variables (wind speed and direction) to model some of these effects by means of statistical models. To this end, a benchmarking between two different families of varying-coefficient models (regime-switching and conditional parametric models) is carried out. The case of the offshore wind farm of Horns Rev in Denmark has been considered. The analysis is focused on one-step ahead forecasting and a time series resolution of 10 minutes. It has been found that the local wind direction contributes to model some features of the prevailing winds, such as the impact of the wind direction on the wind variability, whereas the non-linearities related to the power transfor-

*Corresponding author:
Email address: `cristobalj.gallego@ciemat.es` (C. Gallego)

mation process can be introduced by considering the local wind speed. In both cases, conditional parametric models showed a better performance than the one achieved by the regime-switching strategy. The results attained reinforce the idea that each explanatory variable allows the modelling of different underlying effects in the dynamics of wind power time series.
Keywords: Energy systems modelling, Forecasting, Wind power, Offshore, Varying-coefficient

1. Introduction

The explosive growth of installed wind power over the last 10 years combined with the progressive liberalization of electrical markets have given rise to some new challenges related to wind energy [1]. Special attention has turned towards wind power forecasting, concerning the activity of two agents: wind power producers need to provide accurate information about their energy production in order to take part in the electrical market and the Transmission System Operators (TSO's) need to keep the stability of the electrical system also facing fluctuations on the generation side. In fact, when a certain penetration of wind generation is attained, uncertainties about the evolution of the wind may force the TSO to switch-off a certain number of wind farms, even when the resource is available. These facts represent a clear limitation for wind power penetration, specially considering the ambitious development plans of the offshore industry for the next years [2]. However, accurate forecasts for horizons varying from few minutes to several days could help to mitigate the impact of the inherent uncertainty of the wind. As a result, the last decade has witnessed a rapid growth in the field of short-term wind power forecasting, for both statistical and physical approaches [3, 4, 5, 6, 7].

In this article we focus on the very-short term case, typically being based on a prediction horizon of some minutes to few hours. For such prediction horizons, it is generally accepted that statistical time series based models are more accurate than physical models, the latter ones being more appropriate for horizons beyond several hours [3, 5, 8]. The objective of statistical time series based models is to learn and replicate the dynamics shown by the temporal evolution of certain variables (such as the power output time series) under the hypothesis that these dynamics reflect different underlying effects of the wind power conversion process. Some of these effects would be atmospheric processes occurring at different scales [9], the electrical conversion carried out by the wind turbine, the wake effect generated by nearby wind turbines, etc. [10, 11].

The present work aims to disentangle some of the effects mentioned above by means of a set of available local measurements and an appropriate statistical model. Linear statistical models are characterized by their simplicity and reliability. Even though both wind speed and wind power time series show highly non-linear dynamics, several methodologies have been proposed based on a linear approach (see [12, 13, 14, 15, 16, 17, 18] among others). On the other hand, non-linear approaches are usually based on non parametric models such as Artificial Neural Networks [19], which does not permit a clear interpretation of the underlying processes being modelled. We focus on a non-linear approach based on varying-coefficient models [20] by generalising linear Autoregressive models (AR). The basic structure of an AR model considers the forecasted value as a linear combination of past values by employing fixed weights (see Eq. 6). The main idea is to replace these constant parameters by functions that take into account local observations such as wind speed and direction. This allows the modelling of dependencies in the time series dynamics based on other explanatory variables in a simple way.

Regime-switching autoregressive models are a particular case of varying-coefficient models that consider AR coefficients as constant piece-wise functions. In this case, the considered time series is supposed to evolve shifting between clearly differentiated dynamics (called regimes). These kind of models give rise to a new problem because regimes have to be identified and delimited in some sense [21]. If the shift between regimes is modelled as a function of lagged values of a time series, the process is called observable. This is the case of Threshold Autoregressive Open Loop (TARSO) models [22, 23, 24]. A different approach is considered by Markov Switching Autoregressive models (MSAR), where the current regime is a non-observable process following a first order Markov chain [25, 26, 27, 28, 29, 30].

On the other hand, Conditional Parametric Autoregressive models (CPARX) consider the AR coefficients as smooth functions of some explanatory variables [31, 32, 33]. There exist several approaches to estimate these coefficient-functions (see [34] and references therein). For example, the locally weighted linear regression introduced by Cleveland and Devlin [35] was applied in the design of the Danish Wind Power Prediction Tool WPPT4 [36]. In that case, the AR coefficients were modelled as a function of the forecasted wind speed and direction provided by physical Numerical Weather Prediction (NWP) model.

To the authors' knowledge, there is relatively little research concerning regime-switching models and conditional parametric models that take into account on-line available data such as local wind speed and direction. Thus, in this article we propose a benchmark between the two mentioned families of models (regime-switching and conditional parametric models) in order to clarify how this information can be added so as to model specific features of the wind power time series dynamics. Three reference models are also considered: Persistence, linear AR and MSAR models. Table 1 summarizes different regime-switching and conditional parametric models reviewed in the literature, as well as those considered in this study.

The paper is structured as follows: In Section 2 a theoretical description of the models considered in this article is presented. In Section 3 the database of the case study is described, the offshore wind farm of Horns Rev. The application of the models are detailed in Section 4, organized in four subsections:(i) Description of the reference models, (ii) Modelisation of the local wind direction influence, (iii) Modelisation of the local wind speed influence and (iv) Combining the effects of both local wind speed and direction. Results are presented and discussed in Section 5. Finally, the main findings of the article are summarized in Section 6.

2. Theoretical description of the models

From now, $\{y_t\}$, $t = 1, ..., N$ represents a discrete time series with N observations of averaged wind power production. $\{x_t\}$, $x_t \in \mathbb{R}$, $t = 1, ..., N$ is a discrete time series with N observations of a certain exogenous variable. Additionally, $\mathcal{Y}_T$ and $\mathcal{X}_T$ denote vectors gathering the first T values of the corresponding time series, e.g. $\mathcal{Y}_T = (y_1, ..., y_T)$. $\{y_t\}$ is supposed to follow a stochastic process like:

$$y_t = f(\mathcal{Y}_{t-k}, \mathcal{X}_{t-k}, \Theta) + \varepsilon_t \tag{1}$$

f provides the deterministic component of y_t as a function of a certain set of parameters Θ and the available observations $\mathcal{Y}_{t-k}$ and $\mathcal{X}_{t-k}$, k being the prediction horizon. $\{\varepsilon_t\}$ is a white noise process, that represents the noise of the stochastic process. The purpose of each model considered is to determine a certain function $\hat{f}$, this function being a proposal for the unknown deterministic component of the process. Nevertheless, there are some considerations that establish a common framework for the development of every model considered here. First, only the case of one-step ahead is considered, thus, $k = 1$. Moreover, the white noise is assumed to follow a centred Gaussian distribution with standard deviation σ, i.e., $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$. Hence, a certain model forecasts the value y_t, denoted with $\hat{y}$, as follows:

$$\hat{y}_t = E(y_t | \mathcal{Y}_{t-1}, \mathcal{X}_{t-1}, \Theta) = \hat{f}(\mathcal{Y}_{t-1}, \mathcal{X}_{t-1}, \Theta) \tag{2}$$

where $E(a|b)$ represents the expectation of the statistical variable a given b.

In order to estimate the set of parameters of a statistical model, Θ, the minimisation problem given by Eq. (3) has to be considered along with a score function. In this work we use the quadratic error function of Eq. (4) evaluated over a set of historical data (training-set) with N_{train} samples.

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}}\, \mathcal{E}(\Theta) \tag{3}$$

$$\mathcal{E}(\Theta) = \sum_{t=p+1}^{N_{train}} (y_t - \hat{y}_t)^2 \tag{4}$$

In the following subsections, the linear reference models are described first (Persistence and linear AR), then a non-linear reference model (the MSAR model, a regime-switching model without exogenous variables) and finally, TARSO and CPARX models, which comprise a set of varying-coefficient models that take into account the local wind direction and the local wind speed as explanatory variables.

2.1. Linear reference models: Persistence and autoregressive

Persistence is the most common reference forecasting method for prediction horizons up to 4-6 hours, due to the characteristic time of changes in the atmosphere [37]. A clear advantage of this model is that neither a parameter estimation nor exogenous variables are needed. Persistence states that the forecasted value at time t is the last available value:

$$\hat{y}_t = y_{t-1} \tag{5}$$

An AR(p) is an order-p linear model that considers $\hat{y}_t$ as a weighted sum of the previous p observed values:

$$\hat{y}_t = \theta_0 + \sum_{i=1}^{p} \theta_i \cdot y_{t-i} \tag{6}$$

In this case, given a certain order p, the set of parameters Θ gathers the $p+1$ AR coefficients. This set will be noted as $\Theta_{AR(p)}$

$$\Theta_{AR(p)} = \{\theta_0, \theta_1, ..., \theta_p\} \tag{7}$$

Since varying-coefficient models proposed in this article are obtained by generalising a linear AR model, comparison between them reveals the improvement obtained just related to the consideration of changing regimes or smooth dependencies.

2.2. Non-linear reference model: Markov-Switching Autoregressive Models

The first generalisation of linear AR models considered are the MSAR models. These models assume that a time series evolves switching between different autoregressive dynamics (called regimes). The shift between regimes is considered as a non observable process, which means that it cannot be determined by lagged values of the time series. Pinson et al. [29] demonstrated that MSAR models provided better results than other regime-switching models for two case studies of off-shore wind power forecasting, mainly because these models manage to capture more complex dynamics in regime-switching than when considering the regime as an observable process. Hence, MSAR models represent a suitable option to evaluate the improvement related to regime-switching hypothesis in the absence of exogenous variables. For this reason, MSAR models are here considered as the third reference model.

Let us consider that a time series evolves according to a certain number, r, of different regimes. The current regime at time t is given by the discrete state variable s_t, $t = 1, ..., N$, $s \in \{1, ..., r\}$. The shift between regimes is governed by a first order Markov chain, hence the probability $p(s_t|\mathcal{S}_{t-1}, \mathcal{Y}_{t-1}) = p(s_t|s_{t-1})$. These probabilities are collected in the so-called transition matrix P, where $P_{ij} = p(s_t = j|s_{t-1} = i)$. Since the process is considered unobservable, $\{s_t\}$ is hidden and has to be inferred from available data through the Hamilton filter introduced in Hamilton [38]. Each regime j, $j = 1, ..., r$, is supposed to follow an AR(p) process with coefficients $\Theta_{AR(p)}^{(j)} = \{\theta_0^{(j)}, ..., \theta_p^{(j)}\}$ and standard deviation $\sigma^{(j)}$. The set of parameters of the MSAR model, Θ_{MSAR}, gathers the transition matrix, the AR coefficients and the standard deviation for each regime:

$$\Theta_{MSAR} = \{P, \Theta_{AR(p)}^{(1)}, ..., \Theta_{AR(p)}^{(r)}, \sigma^{(1)}, ..., \sigma^{(r)}\} \tag{8}$$

As an example, Figure 1 illustrates the filtered probabilities of the current regime along with the power output time series for a short window time. It can be seen how the filtered probabilities balance depending on the level of fluctuations. During periods with missing-data, the transition matrix determines a smooth exponential convergence to the so-called ergodic probabilities (the probabilities of being in a certain regime at an arbitrary date).

MSAR models can be formulated in two different ways [39]: the Intercept-Form (MSAR-IF, Eq. 9) and the Mean Adjusted Form (MSAR-MAF, Eq. 10).

$$y_{t,IF}^{(s_t)} = \theta_0^{(s_t)} + \sum_{i=1}^{p} \theta_i^{(s_t)} \cdot y_{t-i} + \varepsilon_t^{(s_t)} \tag{9}$$

$$y_{t,MAF}^{(s_t)} - \mu_0^{(s_t)} = \sum_{i=1}^{p} \phi_i^{(s_t)} \cdot (y_{t-i} - \mu_0^{(s_{t-i})}) + \varepsilon_t^{(s_t)} \tag{10}$$

When no regimes are considered, both forms are equivalent by considering $\phi_i = \theta_i, \forall i > 0$ and $\mu_0 = \theta_0/(1 - \sum_{i=1}^{p} \theta_i)$. Nevertheless, MSAR-IF and MSAR-MAF model different underlying dynamics [39].

2.3. TARSO models

Open Loop Threshold Autoregressive models are a kind of regime-switching model where the current regime s_t is assessed by a predefined function of the available observations of exogenous variables, $s_t = s_t(\mathcal{X}_{t-1})$. Hence the process is called observable. Usually, only a certain lag of x_t is considered, $s_t = g(x_{t-lag})$. In that case, regimes are settled by a certain number of thresholds, $l_0, l_1, l_2, ..., l_r$, that divide the space spanned by $\{x_t\}$ in r subsets, called S_j, $j = 1, ..., r$ from now. Then, $x_{t-lag} \in S_j \Leftrightarrow l_{j-1} \leq x_{t-lag} < l_j$.

In this article only the previous lag of the exogenous variable is considered in assessing regimes. An AR process is assumed in each regime. For the sake of simplicity, all the AR processes will have the same order p. The model is given by:

$$y_t = \theta_0^{(s_t)} + \sum_{i=1}^{p} \theta_i^{(s_t)} \cdot y_{t-i} + \varepsilon_t^{(s_t)} \tag{11}$$

$$s_t = \begin{cases} 1, & x_{t-1} \in S_1 \\ 2, & x_{t-1} \in S_2 \\ ... \\ r, & x_{t-1} \in S_r \end{cases}$$

With the mentioned hypothesis, the implementation of a TARSO model gives rise to three questions: (i) what is the number, r, of regimes considered, (ii) what is the optimal value for the set of thresholds $\mathbf{l} = \{l_0, ..., l_r\}$ and (iii) what AR order p to choose.

Modelling a wind power time series with the described TARSO model implies that the wind farm output has clearly differentiated dynamics depending on the value of some observed variable. For example, in the case of the wind direction (wd), a different behaviour of the wind power time series would be expected depending on the local wind direction observed at the moment of making the forecasting, wd_{t-1}. If wd_{t-1} crosses one of the thresholds given by $\mathbf{l}$, then there is an abrupt change on the AR process that provides the forecast $\hat{y}_t$.

2.4. CPARX Models

Conditional parametric models are characterized by a smooth dependence of their coefficients with a certain variable. In particular, the CPARX models generalize an AR model by letting the coefficients depend on available observations of exogenous variables, $\theta_i = \theta_i(\mathcal{X}_{t-1})$. As in the preceding case, only the previous lag of the exogenous variable will be considered. The model is given by:

$$y_t = \theta_0(x_{t-1}) + \sum_{i=1}^{p} \theta_i(x_{t-1}) \cdot y_{t-i} + \varepsilon_t \tag{12}$$

A central point is how to define the coefficient-functions $\theta_i(x_{t-1})$. They can be estimated with non-parametric techniques from historical data or by means of a parametric function [40, 41]. In this work, the latter case will be considered.

Modelling a wind power time series with a CPARX model implies that the wind farm output dynamic is expected to change smoothly depending on the value of some observed variable x_{t-1}. For example, in the case of the wind speed, (ws), the observed local value ws_{t-1} fixes at each time step the

AR process (through the coefficient-functions $\theta_i(ws_{t-1})$) that provides the forecast $\hat{y}_t$.

3. Description of the data

The data considered originates from the offshore wind farm located at Horns Rev, off the west coast of Denmark. This wind farm has a rated power of 160 MW. Measurements of wind power output, wind speed and direction are available for each wind turbine, with a one-second sample rate. 10-minute resolution time series are derived by averaging raw data. At least 75% of the data within an interval has to be considered as valid in order to consider the averaged value also valid. The averaging process assures that the fast fluctuations related to the turbulent nature of the wind have been filtered. The period considered ranges from 16th February 2005 to 31st January 2006, consisting of 50,400 data points with 8,790 missing data. The data-base has been divided into the following 3 sets:

- Training-set, from 16th February to 31st May 2005: the parameters of the models are estimated considering this data set by solving the minimisation problem given by Eq. (3).

- Validation-set, from 1st June to 31st August 2005: the forecasts provided by the trained models are evaluated during this second period. By doing this, it is possible to assess the generalization capabilities of each model, which means that a certain model trained over a first period keeps its prediction performances over a different time period.

· Test-set, from 1st September 2005 to 31st January 2006: a benchmark analysis between validated models is carried out based on their forecasting performance in this period.

It should be notice that the division of the data-set does not permit models to capture seasonalities during the training process, which covers almost four months. This seasonalities are expected to be present in wind power time series considering the seasonal variability of wind at Horns Rev observed in Vincent et al. [42]. However, it does not necessarily imply that the optimal models would dramatically change from one month to another. In any case, the optimisation of the models taking into account seasonal variations would require several years of data (not available for this work) and the implementation of models with time-varying parameters being adaptively estimated. In this regard, the implementation of adaptive MSAR models was addressed in [30].

4. Application of the models

In this section, the implementation of the models considered in Section 2 in the case of data described in Section 3 is presented. The section is divided in four subsection on different alternatives about the explanatory variables considered. Each model is trained with different structures (concerning for example the AR order and the definition of regimes). The optimal parametrisation of each model was chosen regarding the generalisation capabilities across the validation-set. The performance of the models is evaluated in terms of the Normalized Root Mean Square Error ($NRMSE$) and the percentage of Improvement Over Persistence (IoP), defined as follows:

$$NRMSE = \frac{1}{P_N} \cdot \sqrt{\sum_{t=p+1}^{N} \frac{(y_t - \hat{y}_t)^2}{N-p}} \tag{13}$$

$$IoP(\%) = 100 \cdot \frac{NRMSE_0 - NRMSE}{NRMSE_0} \tag{14}$$

where P_N is the rated power of the wind farm and $NRMSE_0$ is the $NRMSE$ obtained with Persistence . Both criteria are suggested in Madsen et al. [37], which includes a broad overview of ways to evaluate wind power prediction methods.

4.1. Reference models

This subsection deals with the implementation of the reference models described in subsections 2.1 (Persistence and linear AR) and 2.2 (MSAR models). As previously mentioned, Persistence does not have free parameters to be estimated. Thus, the performance of this model is evaluated in a straightforward way. This is not the case for the linear AR models, since the appropriate AR order p and the set of parameters $\Theta_{AR(p)}$ need to be estimated. For a given value of p, $\Theta_{AR(p)}$ is estimated by means of the Yule-Walker equations (available in several works, e.g. [43]) over the training period. Then, the evaluation of the trained models over the validation-set allowed the optimal value of $p = 3$ to be identified.

Next, both MSAR-IF and MSAR-MAF architectures are employed to model the wind power time series of Horns Rev. In order to estimate Θ_{MSAR}, the Expectation-Maximization algorithm introduced in Dempster et al. [44] and further described in Hamilton [45] is applied (for further details, see [38, 46]). In the case of the MSAR-IF form, three regimes were identified with the following set of parameters:

Regime	θ_0	θ_1	θ_2	θ_3	σ
$s_t = 1$	0.01	1.24	-0.47	0.19	0.0573
$s_t = 2$	0.04	1.21	-0.24	0.00	0.0004
$s_t = 3$	0.00	1.45	-0.50	0.04	0.0075

$$P = \begin{bmatrix} 0.77 & 0.02 & 0.21 \\ 0.11 & 0.73 & 0.16 \\ 0.27 & 0.03 & 0.70 \end{bmatrix}$$

On the other hand, the MSAR-MAF model identified the two following regimes:

Regime	μ_0	ϕ_1	ϕ_2	ϕ_3	σ
$s_t = 1$	0.52	1.25	-0.46	0.18	0.0565
$s_t = 2$	0.53	1.38	-0.45	0.08	0.0121

$$P = \begin{bmatrix} 0.91 & 0.09 \\ 0.07 & 0.93 \end{bmatrix}$$

In both cases, the regimes were identified by sorting different levels of fluctuations, i.e., different values for $\sigma^{(i)}$, the standard deviation of the noise.

4.2. Modelling the influence of the local wind direction

In this subsection, the inclusion of the local wind direction into both TARSO and CPARX models is detailed. In order to get some clues about the dependence of wind power on wind direction, a preliminary analysis has been carried out. This would eventually suggest restrictions to the design of appropriate varying-coefficient models, e.g. the number of regimes and the shape of the parameter functions. Then, both the TARSO(wd) model and the CPARX(wd) model are implemented.

4.2.1. Preliminary analysis

The central idea is to train a linear AR model over a subset of the training data. The subset is given by the membership of the previous wind direction lag to a certain sector over the wind rose. The set of AR coefficients, Θ_{AR}, and the *NRMSE* obtained characterize the dynamic of the wind power output related to this particular sector. Then, by sliding smoothly the orientation of the sector and repeating the process, one observes the impact of wind direction on wind power dynamics.

Let us consider a main direction α_0 and a sector width h. The AR(p) model for this sector is given by:

$$\begin{cases} \hat{y}_t = \theta_0 + \sum_{i=1}^{p} \theta_i \cdot y_{t-i} \\ \forall t : wd_{t-1} \in \alpha_0 \pm h/2 \end{cases}$$

The estimation of this model provides specific values for $\Theta_{AR(p)}$ and *NRMSE*, related to α_0. Figure 2 illustrates the dependence of α_0 on $\theta_{AR(p)}$ and the *NRMSE*, when considering the case for $p = 2$ and $h = 90^{\text{o}}$. The following conclusions were derived from the previous analysis, where the considered values for p ranged from 1 to 5: (i) AR coefficients showed a certain dependence on α_0 for any value of p. This dependence is smooth sinus-shaped. (ii) The highest *NRMSE* (thus, the lowest predictability) is related to 270º-310º directions. (iii) The relationship between the *NRMSE* and α_0 shows a similar tendency in both the training-set and the validation-set. Hence, the influence of the wind direction learnt from historical data seems to be representative enough to model future behaviour.

4.2.2. TARSO models based on a wind direction criterion: TARSO(wd)

The previous analysis highlights different predictability levels, depending on the wind direction. Furthermore, there seems to be a high predictability orientation (E-SE), a low one (W-NW) and intermediate transitions. This fact suggests a low number of regimes to be considered a priori.

The TARSO model was introduced in Eq. (11). In this particular case, regime thresholds $\mathbf{l}$ will be related to wind direction sectors as follows: let us consider a main direction α_0 and a certain width sector h. For the sake of simplicity, the same h will be considered for every sector. The wind rose can be split in $r = 360^{\circ}/h$ sectors (the considered widths in the preliminary analysis assures that the number of sectors is a natural number between 2 and 8) by defining the following thresholds:

$$\begin{aligned} l_j = & \ \alpha_0 + \frac{2j-1}{2} \cdot h, \quad j = 1, ..., r \\ l_0 = & \ l_r \end{aligned}$$

This procedure provides the definition of $\mathbf{l}$ and r, given values of α_0 and h. Once the sectors have been defined, AR coefficients can be estimated for each regime once more by means of the Yule-Walker equations. Figure 3 shows the *NRMSE* obtained in the validation-set as a function of p and r, when considering the optimal orientation α_0 obtained. It can be noted that the model with the best generalization capability was obtained for the case of $p = 3$. In the same way, it does not seem to be worth increasing the number of regimes further than 3. In relation to the orientation sectors, Figure 4 illustrates the best ones for the six AR(3) models. It can be seen that the sectors are placed in such a way that the above mentioned low predictability orientation (W-NW) tends to form an independent regime, independently of the number of regimes considered.

The TARSO(wd) model that showed the best performance in the validation-set was:

$$\hat{y}_t = \begin{cases} 0.00 + 1.36 \cdot y_{t-1} - 0.51 \cdot y_{t-2} + 0.14 \cdot y_{t-3}, & s_t = 1 \\ 0.01 + 1.40 \cdot y_{t-1} - 0.54 \cdot y_{t-2} + 0.13 \cdot y_{t-3}, & s_t = 2 \\ 0.00 + 1.19 \cdot y_{t-1} - 0.43 \cdot y_{t-2} + 0.23 \cdot y_{t-3}, & s_t = 3 \end{cases}$$

The regimes were given by:

$$s_t = \begin{cases} 1, & wd_{t-1} \in [-41^{\mathrm{o}}, 79^{\mathrm{o}}) \\ 2, & wd_{t-1} \in [79^{\mathrm{o}}, 199^{\mathrm{o}}) \\ 3, & wd_{t-1} \in [199^{\mathrm{o}}, 319^{\mathrm{o}}) \end{cases}$$

4.2.3. CPARX models based on a wind direction criterion: CPARX(wd)

The description of CPARX models in Subsection 2.4 highlights that the crucial point is how to define the coefficients as a function of a certain exogenous variable. Considering the previous preliminary analysis, a sinus-shaped dependence is proposed:

$$\hat{y}_t = \theta_0(wd_{t-1}) + \sum_{i=1}^{p} \theta_i(wd_{t-1}) \cdot y_{t-i} \tag{15}$$

$$\theta_i(wd_{t-1}) = a_i + b_i \cdot \cos(wd_{t-1} - \phi_0), \quad i = 0, ..., p \tag{16}$$

a_i being the mean level of the i'th AR coefficient and b_i being the amplitude of the dependence of θ_i on the wind direction. Then, for a given value of p, the set of parameters is formed by:

$$\Theta_{CPARX} = \{a_0, ..., a_p, b_0, ..., b_p, \phi_0\} \tag{17}$$

Θ_{CPARX} is estimated in accordance with Eq. (3). As in the previous case, the best performance in the validation-set was achieved for the case of $p = 3$. Figure 5 collects the AR coefficients for the AR model, the TARSO(wd) model and the CPARX(wd) model.

4.3. Modelling the influence of the local wind speed

Following a similar methodology, this subsection focuses on how the local wind speed can be used to define regimes or smooth dependences in the wind power time series dynamics. A preliminary analysis between the predicted variable and the wind speed is firstly performed. Then, the TARSO(ws) model and the CPARX(ws) model are obtained.

4.3.1. Preliminary analysis

Let us consider the interval of wind speeds $I = [ws_0 - h/2, ws_0 + h/2)$. An AR($p$) model is trained taking into account only those data that satisfy at time t the condition $ws_{t-1} \in I$. For a certain h, the AR coefficients and the *NRMSE* obtained are related to the wind speed ws_0. Then, the interval

I slides over the spanned space of the wind speed in order to reveal how the time series dynamic and the predictability vary with ws_0. The following conclusions were obtained, where the considered values for p ranged from 1 to 5: (i) The AR coefficients show a certain dependence on the wind speed. This dependence is close to be linear in a substantial part of the wind speed range, as is shown in the Figure 6 (case $p = 2$, $h = 4\ m/s$). (ii) The *NRMSE* tends to be higher for high wind speeds, showing a maximum at a wind speed of around $10 - 12\ m/s$. However, a decrease in the *NRMSE* is observed for wind speeds beyond the nominal wind speed (at which the output power is constant up to the cut-off wind speed). (iii) A similar tendency of the relationship between *NRMSE* and wind speed has been found for both the training-set and the validation-set (see Figure 6). This fact suggests that the data sets are representative enough to consider this information valid for future time periods.

4.3.2. TARSO models based on a wind speed criterion: TARSO(ws)

The prior analysis reveals that a regime-switching model can be implemented in order to catch different predictability levels, though a low regimes number is suggested from Figure 6. In this case, the optimisation process considers the threshold values, $\mathbf{l}$, as parameters to be estimated. Then, for a certain number of regimes, r, and the AR order p, the set of parameters to estimate is given by:

$$\Theta_{TARSO} = \{\Theta_{AR^{(1)}}, ..., \Theta_{AR^{(r)}}, \mathbf{l}\} \tag{18}$$

Θ_{TARSO} is estimated by means of a numerical algorithm based on the criterion given by Eq. (3). Two and three regimes have been proposed with AR orders going from 1 to 5. In all the cases, the $AR(3)$ showed the best performance in the validation-set (see Figure 7). Furthermore, the two-regimes model was slightly better than the three-regimes one. Figure 8 illustrates the power curve depicted under the optimised regimes. In both cases, the thresholds obtained seems to be related to the shape of the power curve. First, considering two regimes lead to a threshold of around 10 m/s near the inflexion point. This value splits up the power curve in two regions: (i) the first one is characterized by a convex relationship between the wind speed and the output power. In an ideal case, this relationship is a cubic polynomial given by $P = \frac{1}{2}\rho C_p A v^3$, where ρ is the density of air, C_p is the power coefficient, A is the area swept by the rotor blades and v is the wind speed. (ii) The second part is characterized by a concave relationship, since the output power has to be limited by the rated power of the wind turbine. On the other hand, considering three regimes leads to a division clearly based on the slope of the power curve: two regimes for the two flat regions (for low and high wind speeds) and a third one for the steep part.

The TARSO(ws) model with best generalisation capabilities was:

$$\hat{y}_t = \begin{cases} 0.00 + 1.33 \cdot y_{t-1} - 0.50 \cdot y_{t-2} + 0.18 \cdot y_{t-3}, \ s_t = 1 \\ -0.02 + 1.22 \cdot y_{t-1} - 0.39 \cdot y_{t-2} + 0.18 \cdot y_{t-3}, \ s_t = 2 \end{cases}$$

The regimes were given by:

$$s_t = \begin{cases} 1, & ws_{t-1} < 10.08 \\ 2, & ws_{t-1} \geq 10.08 \end{cases}$$

4.3.3. CPARX models based on a wind speed criterion: CPARX(ws)

In this case, a linear dependence between AR coefficients and the last available data of wind speed ws_{t-1} is proposed (Eqs. (19) and (20)). This is partially supported by the preliminary analysis: even though this hypothesis does not seem to be accurate for low and high wind speeds, Figure 6 reveals that it is the case for a substantial part of the wind speed range.

$$\hat{y}_t = \theta_0(ws_{t-1}) + \sum_{i=1}^{p} \theta_i(ws_{t-1}) \cdot y_{t-i} \tag{19}$$

$$\theta_i(ws_{t-1}) = a_i + b_i \cdot (ws_{t-1}), \quad i = 0, ..., p \tag{20}$$

a_i being the i'th AR coefficient at null wind speed and b_i being the slope of the dependence of θ_i on the wind speed. The set of parameters is now given by $\Theta_{CPARX} = \{a_0, ..., a_p, b_0, ..., b_p\}$ and estimated in accordance with Eq. (3). The minimisation process has been carried out for several AR orders, $p = 1, 2, ..., 5$, giving $p = 3$ the optimal value in terms of generalisation capabilities. Figure 9 collects the AR coefficients obtained as a function of the wind speed for the AR model, the TARSO(ws) model and the CPARX(ws) model.

4.4. Combining both effects: CPARX(wd,ws)

Results concerning the incorporation of local wind direction and local wind speed in varying-coefficient models will be discussed in Section 5. However, at this point, it is worth noting that CPARX models showed a better performance than TARSO models when modelling the effect of the considered explanatory variable (see Figure 11). Additionally, each exogenous variable seems to provide information about different effects. In base of this hypothesis, the following CPARX model considering both wind speed and wind direction is proposed:

$$\hat{y}_t = \theta_0(wd_{t-1}, ws_{t-1}) + \sum_{i=1}^{p} \theta_i(wd_{t-1}, wd_{t-1}) \cdot y_{t-i} \tag{21}$$

$$\begin{aligned} \theta_i(wd_{t-1}, ws_{t-1}) &= a_i + b_i \cdot \cos(wd_{t-1} - \phi_0) \\ &+ c_i \cdot (ws_{t-1}), \quad i = 0, ..., p \end{aligned} \tag{22}$$

The set of parameters to be estimated is $\Theta_{CPARX} = \{a_0, ..., a_p, b_0, ..., b_p, c_0, ..., c_p, \phi_0\}$. In this case, the best model obtained was for an AR order of $p = 4$. The coefficient-functions $\theta_i(wd_{t-1}, ws_{t-1})$ are now surfaces that replicates the same trends found in the previous sections. As an example, the case of θ_1 is illustrated in Figure 10.

5. Results

This section gathers the results obtained over the test-set, when the optimal parametrisation of each model obtained in Section 4 is considered.

Globally, the improvements over Persistence ranged from almost 4% to more than 5.5% (see Fig. 11). This represents a good performance, since Persistence is traditionally difficult to improve on for a prediction horizon of 10 minutes. With regard to the reference models and in accordance with the previous studies [29, 30], improvements in very-short term point-forecasting can be attained when considering several regimes under the absence of other explanatory variables. In particular, MSAR models were able to capture shifts between non-observed meteorological states, delivering information about wind power fluctuations and providing a better performance than Persistence and linear AR models.

The models taking into account exogenous variables overcome the reference models. Regarding the influence of the local wind direction, a similar relationship between this variable and the AR parameters was identified by the TARSO(wd) and the CPARX(wd) models, as shown in Figure 5. In particular, given that Persistence can be considered as a particular case of AR model with $\theta_1 = 1$ and $\theta_i = 0$, $\forall i > 1$, both TARSO(wd) and CPARX(wd) models were likely to become globally closer to Persistence for wind directions related to the W-NW sector, characterized by a low predictability (the only exception being θ_3, which experiences a small increment for the mentioned wind directions). Additionally, a smooth dependence of the wind power dynamics on the local wind direction was found to be preferable to considering different regimes (though special attention was paid to track the optimal number of sectors and their orientation) given the IoP of 4.98% and 4.66% respectively. Similar conclusions were obtained when the local wind speed was considered as an exogenous variable: both models TARSO(ws) and CPARX(ws) became globally closer to Persistence (with the only exception of θ_3, which remains almost constant) for high wind speeds (Figure 9) characterized by a lower predictability, and a smooth dependence of the coefficient-functions on the wind speed provided a better result than the regime-switching strategy (an IoP of 4.82% compared to 4.58%).

In general, the models that took into account the wind direction attained slightly better results that those including the wind speed. This was also found when considering the results depicted monthly (Tables 2 and

3), the only exception being the month of January. However, in both a globally and a monthly basis, the best performance was clearly attained by the CPARX(wd,ws). This model attained a global IoP of 5.72%, which represents almost the addition of the single improvements obtained by the CPARX(wd) and the CPARX(ws) models with respect to the AR model. This finding is particularly significant as it supports the notion that each explanatory variable gives information about effects of a different nature.

5.1. Further discussion

It was found that the incorporation of the wind direction as an explanatory variable leads to an appreciable improvement of the prediction performance. It could be due to the fact that the proposed models managed to capture some influence of the local wind direction on the wind power time series dynamics. Vincent et al. [42] related the influence of the wind direction on the wind variability at Horns Rev to synoptic scale forcings combined with the location of the wind farm with respect to the shore. In particular, a high wind variability was observed for Westerly winds. According to Akhmatov [47], the implementation of the models of Subsection 4.2 evidences that these effects are propagated to the wind power time series. As mentioned above, it is interesting to note that a smooth dependence of the wind power dynamics on the local wind direction was preferable to a regime switching strategy. This could be explained by taking the following considerations: the present study is focused on an offshore wind farm, characterized by a flat topography with a uniform-clustered distribution of the wind turbines over a squared area. Hence, for this wind farm configuration no obstacle is introducing directional aerodynamic disturbances and, additionally, wind turbine wakes are likely to have a weaker impact on the dependence between the wind power and the local wind direction compared to the case of a single row wind farm configuration. Even though some works [10, 11] suggest a considerable influence of the wakes for very narrow sectors around the wind turbines line direction, this seems to be too specific to be relevant from a statistical point of view (at least with the models considered in this work). Our results suggest that the influence of the local wind direction on the wind power dynamics was likely to be related to synoptic conditions rather than microscale effects. However, microscale effect could become predominant in other study cases. Modelling the influence of the local wind direction in wind farms located in complex terrain, where topographic obstacles and non-homogeneity of the terrain introduce strong directional dependences on the power production, could require other AR coefficient-functions, instead of the sinus-shaped ones proposed here. Furthermore, wind farms with a non-squared distribution of wind turbines, for instance row-configured wind farms, could even require a regime switching strategy, since the wind turbine wakes would affect dramatically the performance of the wind farm for certain wind directions. In any case, further research on complex terrain and different configuration of wind farms would be required for confirmation.

On the other hand, when the local wind speed was considered as an exogenous variable, the optimisation of the models were likely to be related to the characteristics of the non-linear power transformation process. Considering that the power curve represents a non-linear transformation from wind speed to wind power, the slope of this curve provokes an amplification/reduction effect of the wind speed fluctuations. It has a direct impact on the output power dynamics, causing a dependence between the wind speed and the predictability of the output wind power. Hence, the improvement obtained could be due to the fact that the wind speed was employed as a signal about this non-linear effect. The regime-switching strategy provided thresholds of wind speed that divide the power curve into particular parts (convex-concave for the case of 2 regimes and low-high-low amplification level for the case of 3 regimes, see Figure 8). For the case of the conditional parametric model, a linear relationship between the AR coefficients and the wind speed seemed to be appropriate for a greater part of the wind speed range. However, the saturation effect of the output power related to extreme wind speeds (close to zero or above the nominal wind speed) has not been addressed. Future work could deal with this topic by considering the Generalized Logit transformation described in Pinson [48] or the so-called 'break-point models', a special subclass of varying-coefficient models that combine both CPARX and TARSO models (see the closing discussion in Hastie and Tibshirani [49]).

6. Conclusions

We have presented a study focused on modelling the influence of local wind speed and direction on the dynamics of a wind power time series. With this purpose, a benchmark between several varying-coefficient models for 10 minute-ahead forecasting was carried out. The models are built by generalising the conventional linear AR structure, following two approaches: regime switching models and conditional parametric models. By comparing the accuracy of the models, findings about the most suitable statistical approach were also obtained.

It was found that local measurements of both wind speed and direction provide useful information for a better comprehension of the wind power time series dynamics, at least when considering the case of the very-short term forecasting. In particular, the results suggest that different effects can be modelled depending on the considered explanatory variable: the local wind direction contributes to model some features of the prevailing winds, such as the impact of the wind direction on the wind variability, whereas the non-linearities related to the power transformation process can be introduced by considering the local wind speed. Additionally, for our particular case study, it was found that the conditional parametric models outperforms a regime-switching strategy.

It is interesting to note that the influence of both local wind speed and direction were modelled under the assumption of observable processes, and that only the last observation was taken into account. This study highlights two main lines for further research: the first one is to consider non-observable processes based on local observations, by incorporating exogenous variables whether in the transition matrix or in the definition of the AR coefficients of MSAR models. The second one is to include previous lags of the local observations in order to get a model sensitive to the evolution of the considered exogenous variable. By doing this, it would be possible to explore new effects that condition the dynamics of the output wind power time series (e.g. abrupt changes in local wind direction related to certain meteorological conditions).

Finally, the models here presented could be upgraded by letting the coefficients vary smoothly with time so as to capture seasonal variabilities of wind power dynamics due to climatological effects and the decrease of the wind turbine performance.

7. Acknowledgements

Acknowledgements are first due to CIEMAT who is founding the research of the first author through its PhD Scholarship Program. The work presented has also been partly supported by the Danish ForskEL programme through the project “Radar@Sea” (ForskEL 2009-1-0226) and the project “Mesoscale atmospheric variability and the variation of wind and production for offshore wind farms”, sponsored by the Danish Public Service Obligation (PSO) fund (PSO 7141), which are hereby acknowledged. We are thankful to Vattenfall Denmark for originally providing the wind and power measurements for the Horns Rev wind farm, and to Pierre-Julien Trombe for the data processing and quality checking.

References

[1] Purvins A, Zubaryeva A, Llorente M, Tzimas E, Mercier A. Challenges and options for a large wind power uptake by the European electricity system. Applied Energy 2011;88(5):1461–9.

[2] Snyder B, Kaiser MJ. A comparison of offshore wind power development in europe and the U.S.: Patterns and drivers of development. Applied Energy 2009;86(10):1845–56.

[3] Giebel G. The state of the art in short-term prediction of wind power - A literature overview. Tech. Rep.; ANEMOS EU project; 2003.

[4] Landberg L, Giebel G, Nielsen H, Nielsen T, Madsen H. Short-term prediction - An overview. Wind Energy 2003;6(3):273–80.

[5] Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. Renewable and Sustainable Energy Reviews 2008;12(6):1725–44.

[6] Pinson P, Nielsen H, Madsen H, Kariniotakis G. Skill forecasting from ensemble predictions of wind power. Applied Energy 2009;86(7-8):1326–34.

[7] Bouzgou H, Benoudjit N. Multiple architecture system for wind speed prediction. Applied Energy 2011;88(7):2463–71.

[8] Costa A. Mathematical/statistical and physical/meteorological models for short-term prediction of wind farms output. Ph.D. thesis; Escuela Técnica Superior de Ingenieros Industriales (Universidad Politécnica de Madrid); 2005.

[9] Orlanski I. A rational subdivision of scales for atmospheric processes. Bulletin of the American Meteorological Society 1975;56:527–30.

[10] Jensen L, Mørch C, Sørensen PB, Svendsen KH. Wake measurements from the Horns Rev off-shore wind farm. European Wind Energy Conference, London; 2004.

[11] Méchali M, Barthelmie R, Frandsen S, Jensen L, Réthoré PE. Wake effects at Horns Rev and their influence on energy production. European Wind Energy Conference, Greece; 2006.

[12] Brown B, Katz R, Murphy A. Time series models to simulate and forecast wind speed and wind power. Journal of Climate and Applied Meteorology 1984;23(8):1184–95.

[13] Huang Z, Chalabi ZS. Use of time-series analysis to model and forecast wind speed. Journal of Wind Engineering and Industrial Aerodynamics 1995;56(2-3):311–22.

[14] Kennedy S, Rogers P. A probabilistic model for simulating long-term wind-power output. Wind Engineering 2003;27(3):167–81.

[15] Torres J, Garcia A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Solar Energy 2005;79(1):65–77.

[16] De Giorgi MG, Ficarella A, Tarantino M. Error analysis of short term wind power prediction models. Applied Energy 2011;88(4):1298–311.

[17] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Applied Energy 2011;88(4):1405–14.

[18] Liu H, Erdem E, Shi J. Comprehensive evaluation of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed. Applied Energy 2011;88(3):724–32.

[19] Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Applied Energy 2010;87(7):2313–20.

[20] Cleveland WS, Grosse E, Shyu WM. Statistical models in S; chap.

Local regression models. Boca Raton, FL, USA: CRC Press, Inc. ISBN 0412052911; 1991, p. 309–76.

[21] Tong H. Threshold models in non-linear time series analysis. Springer-Verlag; 1983.

[22] Gneiting T, Larson K, Westrick K, Genton MG, Aldrich E. Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method. Journal of the American Statistical Association 2006;101(475):968–79.

[23] Tastu J, Pinson P, Kotwa E, Madsen H, Nielsen HA. Spatio-temporal modelling of short-term wind power prediction errors. Wind Energy 2011;14(1):43–60.

[24] Hering AS, Genton MG. Powering up with space-time wind forecasting. Journal of the American Statistical Association 2010;105(489):92–104.

[25] Castino F, Festa R, Ratto C. Stochastic modelling of wind velocities time series. Journal of Wind Engineering and Industrial Aerodynamics 1998;74-6:141–51. 2nd European and African Conference on Wind Engineering, Genoa, Italy, Jun 22-26, 1997.

[26] Hughes JP, Guttorp P, Charles SP. A non-homogeneous Hidden Markov Model for precipitation occurrence. Journal of The Royal Statistical Society Series C 1999;48(1):15–30.

[27] Ailliot P. Modèles autorégressifs à changements de régimes markoviens. Applications aux séries temporelles de vent. Ph.D. thesis; University of Rennes; 2004.

[28] Kosater P, Mosler K. Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices. Applied Energy 2006;83(9):943–58.

[29] Pinson P, Christensen LEA, Madsen H, Sørensen PE, Donovan MH, Jensen LE. Regime-switching modelling of the fluctuations of offshore wind generation. Journal of Wind Engineering and Industrial Aerodynamics 2008;96(12):2327–47.

[30] Pinson P, Madsen H. Adaptive modeling and forecasting of wind power fluctuations with Markov-switching autoregressive models. Journal of Forecasting 2010;in press.

[31] Chen R, Tsay RS. Functional-coefficient Autoregressive models. Journal of the American Statistical Association 1993;88(421):298–308.

[32] Nielsen HA, Nielsen TS, Joensen AK, Madsen H, Holst J. Tracking time-varying-coefficient functions. International Journal of Adaptive Control and Signal Processing 2000;14(8):813–28.

[33] Sánchez I. Short-term prediction of wind energy production. International Journal of Forecasting 2006;22(1):43–56.

[34] Fan J, Zhang W. Statistical methods with varying-coefficient models. Statistics and its Interface 2008;1:179–95.

[35] Cleveland WS, Devlin SJ. Locally-weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association 1988;83:596–610.

[36] Nielsen T, Madsen H, Nielsen H. Prediction of wind power using time-varying coefficient-functions. Proceedings of the 15th IFAC World Congress on Automatic Control, Barcelona (Spain); 2002.

[37] Madsen H, Pinson P, Kariniotakis G, Nielsen HA, S.Nielsen T. Standardizing the performance evaluation of short-term wind power prediction models. Wind Engineering 2005;29(6):475–89.

[38] Hamilton JD. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 1989;57(2):357–84.

[39] Krolzig HM. Markov-switching vector autoregressions. Springer; 1997.

[40] Nielsen HA, Nielsen TS, Madsen H. ARX-models with parameter variations estimated by local fitting. In: Sawaragi Y, Sagara S, editors. 11th IFAC Symposium on System Identification; vol. 2. 1997, p. 475–80.

[41] Madsen H, Holst J. Modelling non-linear and non-stationary time series. Technical University of Denmark, DTU Informatics; 2000.

[42] Vincent CL, Pinson P, Giebel G. Wind fluctuations over the North Sea. International Journal of Climatology 2010;in press.

[43] Peña D. Estadística. Modelos y métodos; vol. 2. Alianza Editorial; 2nd ed.; 1987.

[44] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 1977;39(1):1–38.

[45] Hamilton JD. Analysis of time series subject to changes in regime. Journal of Econometrics 1990;45(1-2):39–70.

[46] Patterson DM. Nonlinear time series analysis of economic and financial data; chap. A markov switching cookbook. Kluwer Academic; 1999, p. 33–43.

[47] Akhmatov V. Influence of wind direction on intense power fluctuations in large offshore windfarms in the North Sea. Wind Engineering 2007;31(1):59–64.

[48] Pinson P. On probabilistic forecasting of wind power time-series. Tech. Rep.; Technical University of Denmark, DTU Informatics; 2010.

[49] Hastie T, Tibshirani R. Varying-coefficient models. Journal of the Royal Statistical Society, Series B (Methodological) 1993;55(4):757–96.

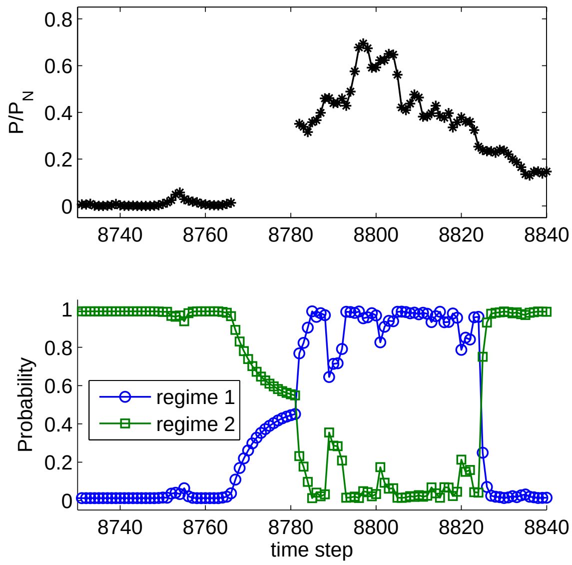

Figure 1: Filtered probabilities of the current regime provided by the MSAR model during periods with missing data. P/P_N represents the output power (P) normalized with the rated power of the wind farm (P_N).

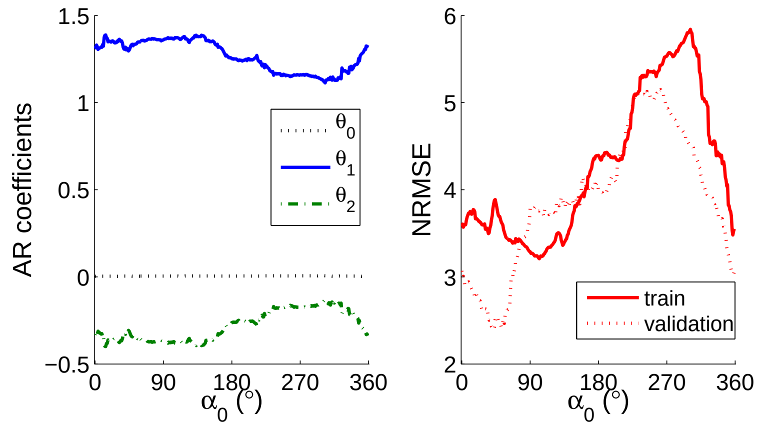

Figure 2: Dependence of the AR coefficients (left) and *NRMSE* in $\%P_N$ (right) with local wind direction. Case for AR order $p = 2$

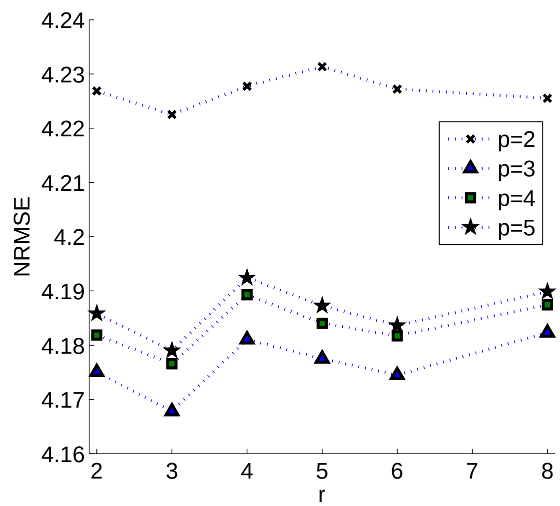

Figure 3: *NRMSE* (in $\%P_N$) of TARSO(wd) over the validation-set, as a function of the number of regimes, r, and the AR order, p. Results for $p = 1$, layout of the picture due to a higher *NRMSE*

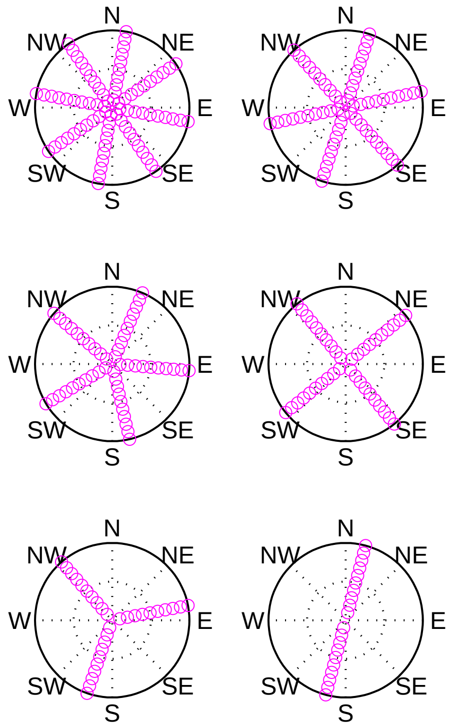

Figure 4: Optimal orientation of the sectors depending on the number of regimes considered in a TARSO(wd) model, case for AR order $p = 3$

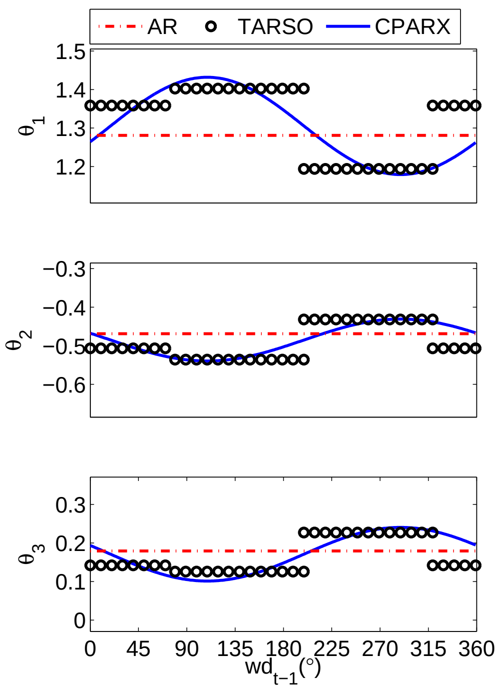

Figure 5: Dependence of the AR coefficients with local wind direction for AR, TARSO(wd) and CPARX(wd) models (Θ_0 is omitted, since it is very close to zero for every model)

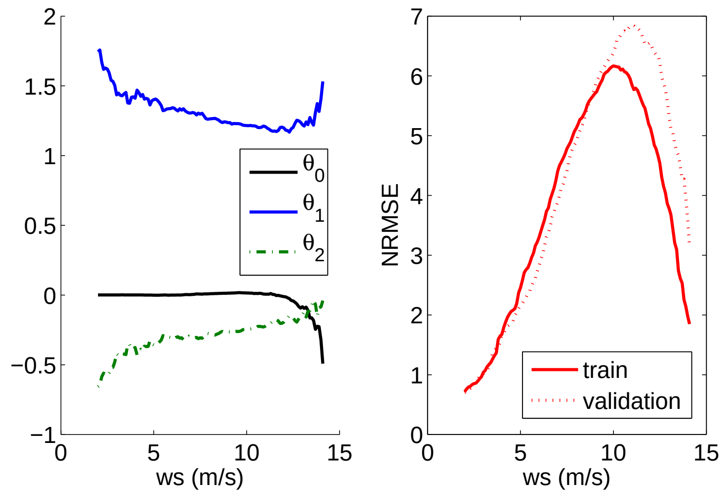

Figure 6: Dependence of the AR coefficients and $NRMSE$ (in $\%P_N$) with local wind speed. Case for AR order $p = 2$

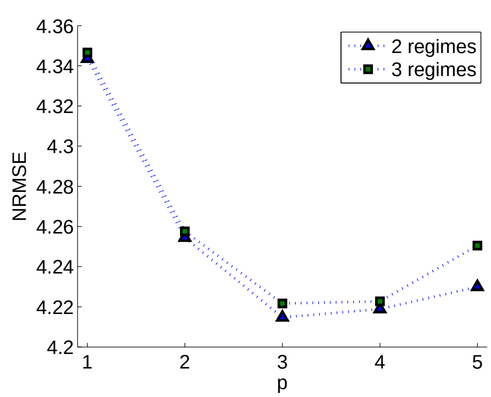

Figure 7: $NRMSE$ (in $\%P_N$) of TARSO(ws) over the validation-set, as a function of the number of regimes, r, and the AR order, p

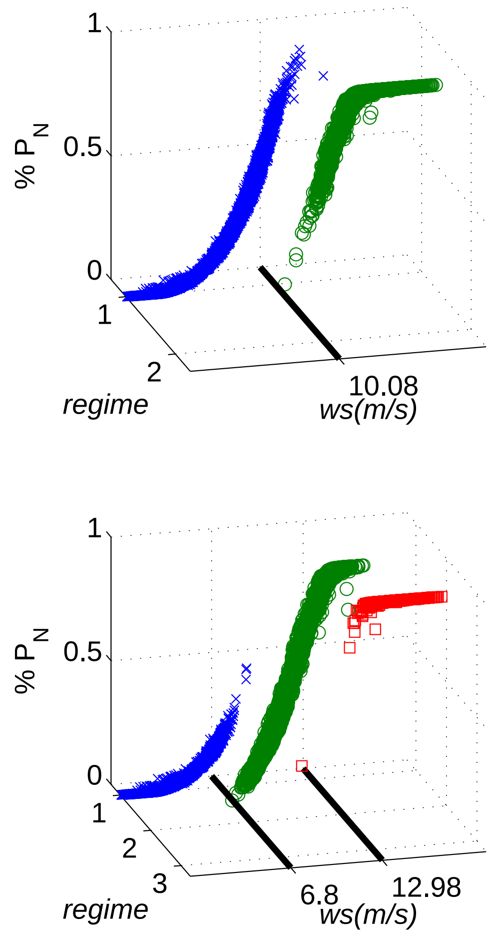

Figure 8: Optimal splitting of the power curve for TARSO(ws) models

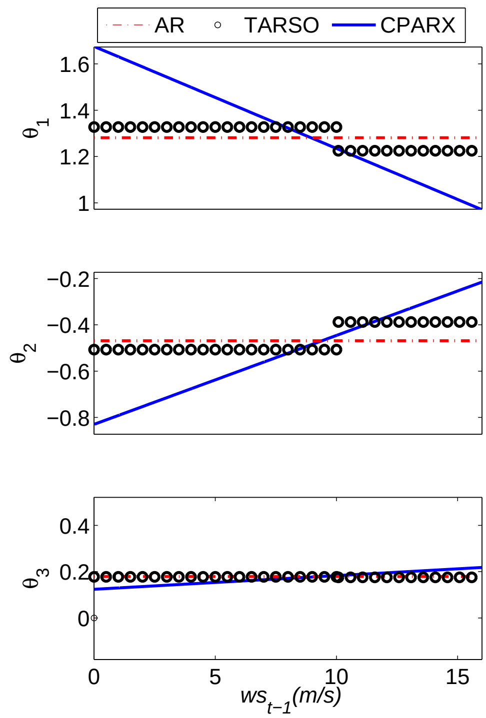

Figure 9: Dependence of the AR coefficients for AR, TARSO(ws) and CPARX(ws) models. (Θ_0 is omitted, since it is very close to zero for every model)

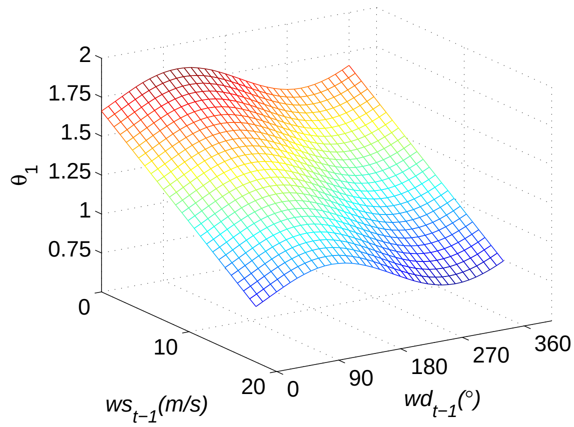

Figure 10: θ_1 as a function of local wind direction and local wind speed for the CPARX(wd,ws) model

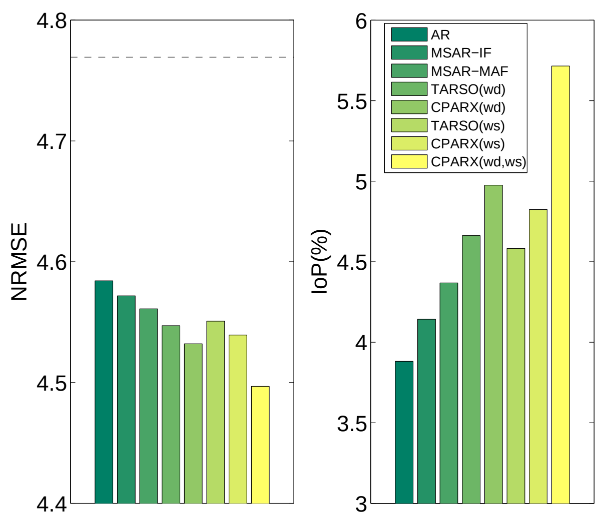

Figure 11: *NRMSE* (in $\%P_N$) and *IoP* for the test-set. Dashed line of the figure on the left refers to the *NRMSE* of Persistence

Constant coefficients	
	Persistence[1,2,3]**, AR**[2,3,4], ARMA[1,5,6]
Varying coefficients	
R-S (Obs)	STAR[1], SETAR[1], **TARSO**[7]
R-S (Non-Obs)	**MSAR**[1,2,8]
C-P	**CPARX**[3,9]

Table 1: Summary of models applied in some studies related to short-term wind and wind power forecasting. In bold, models considered in the present study. **R-S**: Regime-Switching, **C-P**: conditional parametric, **Obs**: Observable process. [1]Pinson et al. [29], [2]Pinson and Madsen [30], [3]Pinson [48], [4]Brown et al. [12], [5]De Giorgi et al. [16], [6]Erdem and Shi [17] [7]Tastu et al. [23], [8]Ailliot [27], [9]Nielsen et al. [36]

	September	October	November	December	January
Persistence	4.66	4.16	6.25	4.76	4.07
AR	4.44	3.96	6.03	4.42	3.98
MSAR-IF	4.43	3.95	5.97	4.47	3.97
MSAR-MAF	4.41	3.96	5.95	4.41	4.00
TARSO(wd)	4.42	**3.92**	5.94	4.37	3.99
CPARX(wd)	**4.41**	3.92	**5.91**	**4.37**	3.97
TARSO(ws)	4.44	3.93	5.94	4.39	3.97
CPARX(ws)	4.42	3.94	5.92	4.37	**3.94**
CPARX(wd,ws)	**4.41**	**3.91**	**5.82**	**4.35**	**3.93**

Table 2: *NRMSE* depicted monthly. The two lowest values in each column are given in bold fonts. The overall results are gathered in Figure 11

	September	October	November	December	January
AR	4.73	4.89	3.54	7.19	2.33
MSAR-IF	4.74	5.04	4.43	6.10	2.50
MSAR-MAF	5.30	4.81	4.80	7.46	1.80
TARSO(wd)	5.07	**5.72**	4.87	8.17	2.07
CPARX(wd)	**5.34**	5.71	**5.41**	**8.34**	2.54
TARSO(ws)	4.72	5.65	4.89	7.89	2.57
CPARX(ws)	5.07	5.39	5.16	8.14	**3.30**
CPARX(wd,ws)	**5.37**	**5.96**	**6.78**	**8.68**	**3.54**

Table 3: *IoP* depicted monthly. The two highest values in each column are given in bold fonts. The overall results are gathered in Figure 11