

Wind prediction enhancement by exploiting data non-stationarity

A Malvaldi, J Dowell, S Weiss, D Infield

*Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, Scotland
{alice.malvaldi, jethro.dowell, stephan.weiss, david.infield}@strath.ac.uk*

Keywords: wind forecasting; multichannel prediction; non-stationary filtering; adaptive signal processing.

Abstract

The short term forecasting of wind speed and direction has previously been improved by adopting a cyclo-stationary multichannel linear prediction approach which incorporated seasonal cycles into the estimation of statistics. This paper expands previous analysis by also incorporating diurnal variation and time-dependent window lengths. Based on a large data set from the UK's Met Office, we demonstrate the impact of this proposed approach.

1. Introduction

Accurate forecasts of wind speed and direction are key to a number of applications. For example, the stability and reliability of power system operations are highly dependent on the predicted power outputs [13,12]. Wind farm operators are penalised if their predicted power in a bid on the electricity market is not met [7]. Accurate wind forecasting can also help to schedule off-shore maintenance access and down-time [2].

The literature distinguishes two main approaches to predicting wind speed and direction [6]. Firstly, numerical weather prediction (NWP) models are highly complex, typically run every six hours and are used for longer (greater than six hours ahead) forecasts due to their good accuracy [8,17]. Secondly, statistical models are preferred for short term (one to six hours ahead) forecasts, due to their lower computational complexity. Statistical approaches generally also offer a spatial resolution which NWP approaches cannot match.

Amongst the statistical approaches, linear [6,5,10] and non-linear methods [1,3,16] have been proposed, to exploit the spatio-temporal correlation between geographically separated measurements. Despite the non-linear nature of the system, linear methods are often preferred due to their relative computational simplicity when compared to non-linear approaches. This is particularly true when considering methods that exploit the spatial dimension of the data, which in their majority are linear.

Recently in [6], a multichannel linear predictor has been proposed, which models both wind speed and direction by utilising complex time series with magnitude as wind speed and phase as direction. Further, this method exploits the seasonal cyclic behaviour by means of a cyclo-stationary assumption of

the data, which impacts on the way the prediction filter coefficients are estimated and applied. Because the cyclo-stationary approach in [6] provides predictions with greater accuracy than persistence, a similar approach is chosen here.

In this paper, the cyclo-stationary model proposed in [5] is expanded to consider not only the seasonal variation of the wind signal but also its diurnal oscillation. It is investigated how this can impact on and potentially enhance the accuracy of forecasts. To demonstrate this, Sec. 2 describes the non-stationary prediction filter, and how the prediction filter coefficients are estimated, exploiting seasonal and diurnal cycles. Sec. 3 provides some comments on the computational complexity of this filter. The system is tested on real data from the UK Met. Office, which is characterised on Sec. 4 followed by simulations and results. Conclusions are presented in Sec. 5.

2. Methodology

This section, based on the data model in Sec. 2.1, details a multichannel linear predictor in Sec. 2.2. The assumed cyclo-stationarity of the data is exploited in Sec. 2.3 by appropriate windowing of the data.

2.1 Complex Multichannel Data

The proposed predictor is based on a complex-valued vector time series $\mathbf{x}_n \in \mathbb{C}^M$, with

$$\mathbf{x}_n = [x_1[n] \ x_2[n] \ \dots \ x_M[n]]^T \quad (1)$$

where individual time series $x_m[n] \in \mathbb{C}$, $m = 1 \dots M$ represent hourly mean wind speed and direction at a specific geographical site. Specifically, the wind speed is the magnitude of the complex variable, and the wind direction is its phase, with n being the discrete time index.

In terms of their statistics, all signals are pre-processed to have zero mean. Due to their potential non-stationarity, a time-dependent cross-covariance of the data, $r_{x_i x_j}[n, \tau] = \mathcal{E}\{x_i[n]x_j^*[n - \tau]\}$, $i, j = 1 \dots M$, includes the time instance n at which this measure is taken.

2.2 Minimum Mean Square Prediction Error

The general structure of the multichannel predictor is outlined in Fig. 1. Based on past measurements taken at M sites, the predictor tries to forecast the time series $x_m[n]$ at site m at time

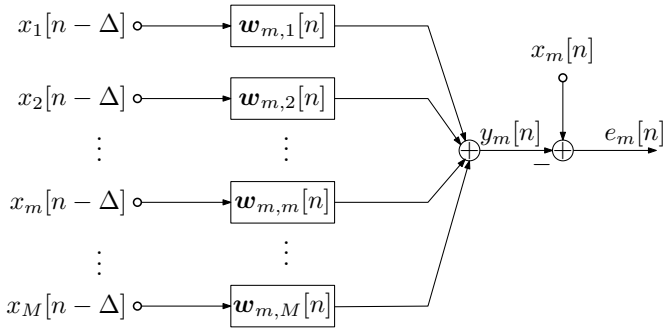


Figure 1. Multi-channel filter to predict the m th signal Δ samples ahead from a total of $m = 1 \dots M$ inputs.

index n , whereby $m = 1 \dots M$. The prediction filter coefficients are contained in finite impulse response filter $\mathbf{w}_{m,i}[n] \in \mathbb{C}^N$, where N is the temporal window over which prediction is performed, and $m, i = 1 \dots M$. A tap delay line vector

$$\mathbf{x}_m[n] = \begin{bmatrix} x_m[n] \\ x_m[n-1] \\ \vdots \\ x_m[n-N+1] \end{bmatrix} \quad (2)$$

holds this data window at the m^{th} site during iteration n .

The coefficients $\mathbf{w}_{m,i}[n] \in \mathbb{C}^N$, $i, m = 1 \dots M$ are adjusted such that the prediction error

$$e_m[n] = d_m[n] - \sum_{i=1}^M \mathbf{w}_{m,i}^H[n] \mathbf{x}_i[n] = d_m[n] - \mathbf{w}_m^H \mathbf{x}[n] \quad (3)$$

is minimised in the mean square error (MSE) sense with $d_m[n] = x_m[n + \Delta]$. The vectors $\mathbf{w}_m[n]$ and $\mathbf{x}[n]$ in (3) are formed from concatenations of $\mathbf{w}_{m,i}[n]$ and $\mathbf{x}_i[n]$, $i = 1 \dots M$,

$$\mathbf{x}[n] = \begin{bmatrix} \mathbf{x}_1[n] \\ \mathbf{x}_2[n] \\ \vdots \\ \mathbf{x}_M[n] \end{bmatrix}, \quad \mathbf{w}_m[n] = \begin{bmatrix} \mathbf{w}_{m,1}[n] \\ \mathbf{w}_{m,2}[n] \\ \vdots \\ \mathbf{w}_{m,M}[n] \end{bmatrix}, \quad (4)$$

thus containing all data points within the spatio-temporal window and all filter coefficients, respectively.

The MSE of the prediction error $e_m[n]$ is given by

$$\begin{aligned} \xi_m &= \mathcal{E}\{e_m[n]e_m^*[n]\} \\ &= \mathcal{E}\{(d_m[n] - \mathbf{w}_m^H[n]\mathbf{x}[n]) \cdot (d_m^*[n] - \mathbf{x}[n]^H\mathbf{w}_m[n])\} \quad (5) \end{aligned}$$

$$= \sigma_{x_m}^2 - \mathbf{w}_m^H[n]\mathbf{p}_m[n] - \mathbf{p}_m^H[n]\mathbf{w}_m[n] + \mathbf{w}_m^H[n]\mathbf{R}[n]\mathbf{w}_m[n] \quad (6)$$

In (5), $\mathbf{R}[n] = \mathcal{E}\{\mathbf{x}[n]\mathbf{x}^H[n]\}$ is the covariance matrix of the data, and $\mathbf{p}_m[n] = \mathcal{E}\{d_m[n]\mathbf{x}^*[n]\}$ the cross-covariance vector between the desired signal $d_m[n]$ for site m and the data vector.

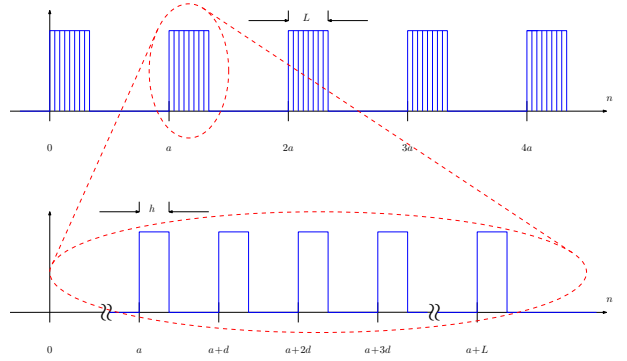


Figure 2. Data window $\lambda[n]$ for the cyclo-stationary model: time window of length L and number of hours per day h . The upper coarse part of the diagram captures seasonal cyclo-stationarity, with a sampled per annum, while lower part exploits potential diurnal cyclo-stationarity, with d samples per diem.

The MSE can be minimised by equating the first derivative of (5) with respect to the coefficients \mathbf{w}_m to zero, resulting in

$$\mathbf{w}_{m,\text{opt}}[n] = \mathbf{R}^{-1}[n]\mathbf{p}_m[n], \quad (7)$$

which is commonly known as the Wiener-Hopf solution [4,15,9]. The minimum MSE (MMSE) in (6) is obtained by inserting (7), such that

$$\xi_{m,\text{min}} = \sigma_{x_m}^2 - \mathbf{p}_m^H[n]\mathbf{R}^{-1}[n]\mathbf{p}_m[n] \quad (8)$$

Therefore, N previous values of the M time series that are weighted by the optimal coefficients $\mathbf{w}_{m,\text{opt}}[n]$ as given by in (7), such that the mean square value of the prediction error for the m th site is minimised, taking on the MMSE value in (8).

2.3 Cyclo-stationary Model

In previous work, a cyclo-stationary Wiener filter was proposed in [6], which exploits seasonal dependencies of the wind. There, it was found that the optimum window length over which cross- and auto-correlation values would be calculated was 15 weeks. While this model provided an advantage over the stationary Wiener filter, it did not provide variable window lengths for different seasons, nor did it account for the diurnal variations that are often apparent in the data.

Therefore, in this study the diurnal variation is taken into account in addition to seasonal dependencies. Our model assumes that data can be assumed to be stationary over h hours per day over a total window L hours per season. The details of this model are sketched in Fig. 2, whereby the higher level structure on top exploits the seasonal cyclo-stationarity as in [6]. The detailed binary ‘‘comb’’ structure, where h hours per day are selected, is to exploit the diurnal behaviour, and represents the innovation over [6].

The statistics are calculated as previously described in Sec. 2.2, once that data is windowed. Some of the computational aspects of evaluating auto- and cross-correlations on windowed data will be addressed further below. In order to still retain zero mean processes, the data-dependent mean is removed from individual segments of windowed data.

3. Implementation Aspects

This section highlights some of the considerations of the proposed approach with respect to computational cost. While the calculation of the Wiener solution and the application of the prediction filter to the actual data can also be costly, we here focus on computation aspects in calculating the statistics of the data required to obtaining the Wiener solution.

To estimate the statistics, the expectation operator is replaced by temporal averages. Therefore e.g. a cross-correlation term, assuming stationarity and ergodicity over the window of length T for which it is computed, is obtained as

$$r_{x_i x_j}[\tau] \approx \frac{1}{T} \sum_{n=0}^{T-1} x_i[n] x_j^*[n - \tau]. \quad (9)$$

The computation on the r.h.s. of (9) is intense, particularly since $\frac{1}{2}M(M-1)$ combinations need to be calculated. With one lag value requiring T multiply accumulates, and a total of $2L-1$ lags to be evaluated, the total cost to evaluate the auto- and cross-correlation terms is

$$C_1 = \frac{1}{2}M(M-1)T(2L-1), \quad (10)$$

such that the complexity is of order $\mathcal{O}\{M^2TL\}$.

If evaluated in the frequency domain, M FFTs of length T are required for the time series. In the DFT domain, a convolution-type operation turns into a multiplication, such that $\frac{1}{2}M(M-1)$ products each over T elements have to be computed. This is followed by $\frac{1}{2}M(M-1)$ inverse FFTs. Therefore the complexity of implementing the auto- and cross-correlation sequences in the DFT domain are

$$C_2 = MT \log_2 T + \frac{1}{2}M(M-1)T + \frac{1}{2}M(M-1)T \log_2 T \\ = \frac{1}{2}M(M+1)T \log_2 T + \frac{1}{2}M(M-1)T, \quad (11)$$

i.e. the complexity is of order $\mathcal{O}\{M^2T \log_2 T\}$, which is below the time domain implementation cost C_1 in (10) if $L > \log_2 T$. Hence, if only a limited number of lag values L need to be evaluated, the time domain approach is advantageous.

If the data window characterised in Fig. 2 is labelled as $\lambda[n]$, the calculation in (9) has to be performed on $\tilde{x}_m[n] = \lambda[n]x_m[n]$ instead of $x_m[n]$, $m = 1 \dots M$. The frequency domain implementation now becomes unattractive, as a multiplication in the time turns to a convolution-type operation in the DFT domain. In the time domain however, the evaluation of $\lambda[n]x_m[n]$ can be performed multiplied-free through clever indexing, since $\lambda[n]$ represents a binary mark. In this case, the cost remains unaltered compared to C_1 in (10). Tapered instead of binary windows for $\lambda[n]$ might be preferable in terms of estimation accuracy, but will also incur a much greater computational cost.

4. Data, Testing, and Results

4.1 Dataset

The data used in this study has been provided by the British Atmospheric Data Centre [14]. The data set of onshore weather

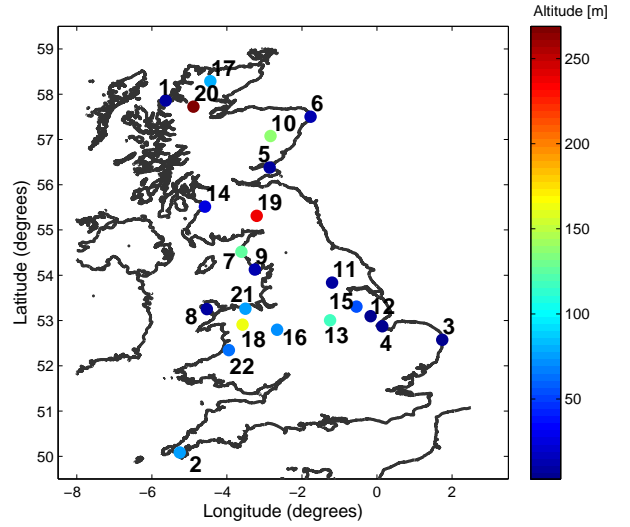


Figure 3. Geographical distribution of 22 Met. Office stations supplying the test data.

data, from the Met Office Integrated Data Archive System (MIDAS), comprises of hourly wind speed [knots] and direction [deg] observations taken at a height of 10m from ground. For this analysis, 22 weather stations have been selected considering those with less than 2% missing and invalid data between 00:00h on 1/01/2002 till 23:00h on 31/12/2007. Figure 3 shows the locations of the 22 Met Office stations.

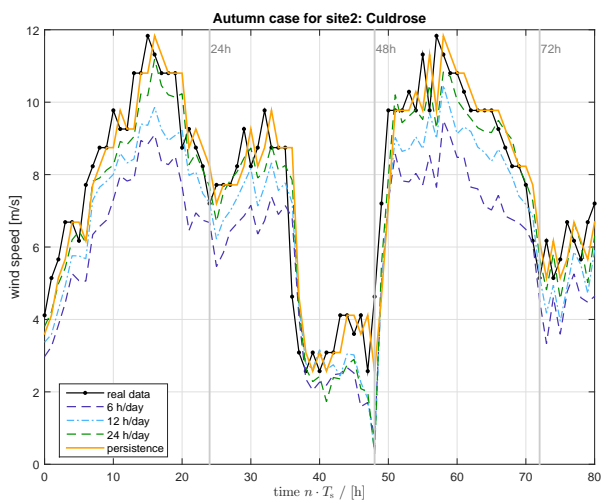
4.2 Testing and Results

Previously in [11], the cyclo-stationary Wiener filter has been tested on the same MIDAS dataset for summer and winter cases only, and using a fixed window length of $L = 15$ weeks. We here extend this analysis to spring and autumn, and perform these tests for different window lengths of 5, 10, 15, and 20 weeks. The data has been partitioned such that measurement from 2002 until 2006 have been used for training to calculate the filter coefficients, while the data from 2007 has been employed as test data to forecast wind speed and direction. The look-ahead time Δ in this paper is one hour.

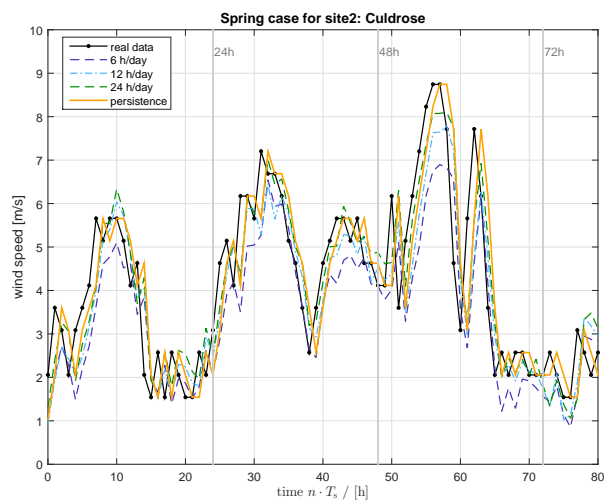
To evaluate the proposed algorithms, their performance is benchmarked against persistence, which assumes that the wind will not change compared to the current measurement, i.e. $x_m[n] = \hat{x}_{m,\text{measured}}[n-1]$. The cyclo-stationary Wiener filter has been tested for four different window lengths (5, 10, 15, and 20 weeks), with three different cases considered:

- (1) 6 hours per day;
- (2) 12 hours per day;
- (3) 24 hours per day (which corresponds to the cyclo-stationary case previously introduced by [5]).

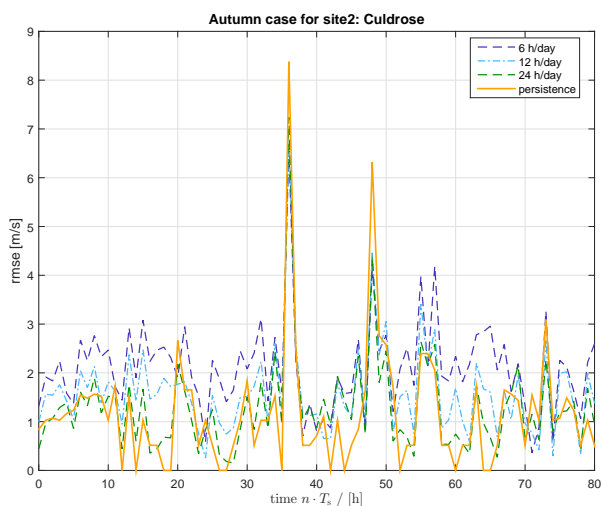
Prediction results from the cyclo-stationary algorithms and the stationary Wiener filter have been compared with persistence in terms of root mean-squared error (RMSE). Here, results from spring and autumn are reported for 2 different sites, similar results are found for all the other MIDAS sites and seasons. In order to perform the forecasts, it is necessary



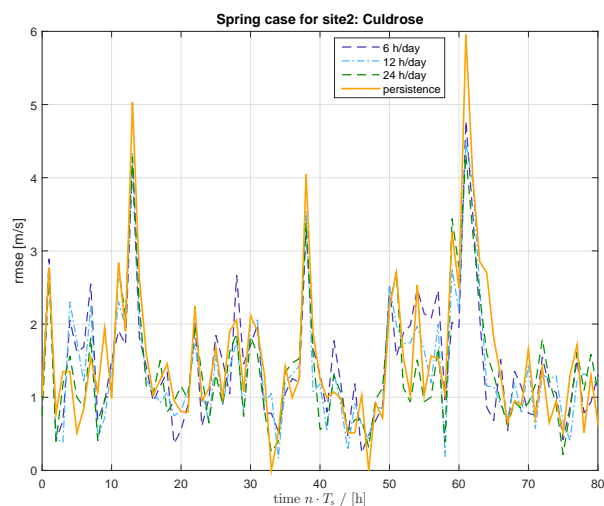
(a) Predicted wind speed, autumn case.



(a) Predicted wind speed, spring case.



(b) Root mean squared error, autumn case.



(b) Root mean squared error, spring case.

Figure 4. Comparison between the real wind speed (black solid line) and the predicted wind speeds for site 2 (Culdrose) using: persistence method (red solid line), the cyclo-stationary solution with 6 hours per day (orange dash line), 12 hours per day (green dash-dot line), and 24 hours per day (purple dot line).

Figure 5. Comparison between the real wind speed (black solid line) and the predicted wind speeds for site 2 (Culdrose) using: persistence method (red solid line), the cyclo-stationary solution with 6 hours per day (orange dash line), 12 hours per day (green dash-dot line), and 24 hours per day (purple dot line).

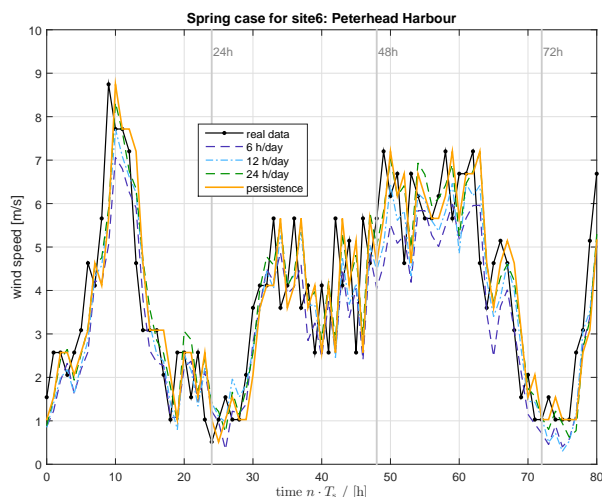
that the time series do not have missing data in order to avoid transient behaviours and problems in the prediction. For this purpose, shorter time windows have been selected for each season:

- Spring: from 29/03/2007 to 04/04/2007;
- Summer: from 23/07/2007 to 25/07/2007;
- Autumn: from 22/10/2007 to 26/10/2007;
- Winter: from 15/01/2007 to 18/01/2007.

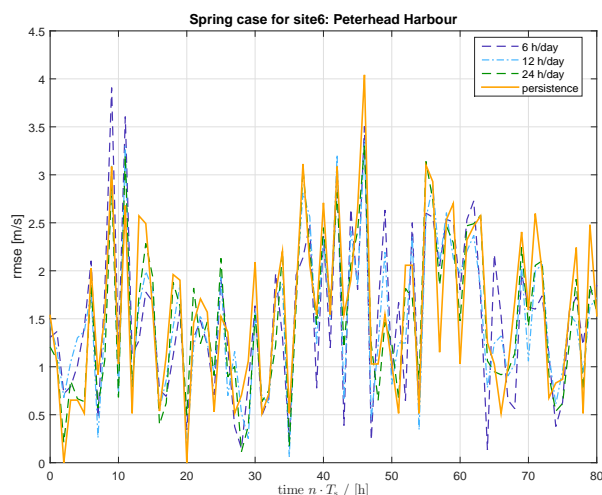
Figs. 4 and 5 show the RMSE and wind speed forecast as a function of time for site 2 (Culdrose, in Cornwall) in spring and autumn. Fig. 6 shows the same results but for a different location: site 6, Peterhead Harbour, in the north east of Scotland. The three cases of the cyclo-stationary model are compared with persistence. From the graphs it is evident that the performance of each model varies from site to site, with the season and time of day.

RMSEs are generally higher in winter and autumn compared with summer and spring. Moreover, some sites present a more evident diurnal variation than others, which depends on the season as well. For instance, in summer the diurnal variation is generally stronger than in winter.

The performance of the models are site specific, but it is evident from the RMSEs graphs reported in Figs. 4b, 5b, and 6b that the 6 h per day cyclo-stationary model has a poorer performance when the wind is less variable but performs better when there are abrupt changes in the wind speed. In those cases it has a lower error than persistence. This is very important in wind power forecasting as abrupt changes in wind can cause significant problems for power system operations.



(a) Predicted wind speed, spring case.



(b) Root mean squared error, spring case.

Figure 6. Comparison between the real wind speed (black solid line) and the predicted wind speeds for site 6 (Peterhead Harbour) using: persistence method (red solid line), the cyclo-stationary solution with 6 hours per day (orange dash line), 12 hours per day (green dash-dot line), and 24 hours per day (purple dot line).

5. Conclusions

In this article data non stationarity has been considered focusing on the diurnal patterns of the wind speed. This work is part of a previous paper [11] where the model has been introduced. The suggested model is a complex-valued multichannel Wiener filter where it is assumed that the data are stationary on windows of length L during h hours per day. Therefore, the coefficients depend on statistics that are the same during that time period in all years.

The testing of the algorithm has been performed on different window lengths $L = 5, 10, 15$ and 20 weeks, and 3 different cases: selection of 6, 12 and 24 hours per day. The persistence method has been used as a benchmark to evaluate the performance of the models. Results show that the 6h per day model has a better performance when the wind has abrupt changes (either speed up or slow down in a very short period of time).

This result is valid for all sites and all the cases considered. Future work will carry out a statistical analysis considering forecasts horizons of more than 1 hour as it has been shown in [6,5] that improvements of the prediction error of the cyclo-stationary Wiener filter over persistence depends on the look-ahead time.

References

- [1] M. Alexiadis, P. Dokopoulos, and H. Sahsamanoglou. Wind speed and power forecasting based on spatial correlation models. *Energy Conversion, IEEE Transactions on*, 14(3):836–842, Sep 1999.
- [2] A. Botterud, Z. Zhou, J. Wang, R. Bessa, H. Keko, J. Mendes, J. Sumaili, and V. Miranda. *Use of Wind Power Forecasting in Operational Decisions*. Argonne National Laboratory, 2011.
- [3] I. Damousis, M. Alexiadis, J. Theocharis, and P. Dokopoulos. A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *Energy Conversion, IEEE Transactions on*, 19(2):352–361, June 2004.
- [4] J. Dowell, S. Weiss, D. Hill, and D. Infield. A cyclo-stationary complex multichannel Wiener filter for the prediction of wind speed and direction. In *21st European Signal Processing Conference*, Marrakech, Morocco, September 2013.
- [5] J. Dowell, S. Weiss, D. Hill, and D. Infield. A cyclo-stationary complex multichannel Wiener filter for the prediction of wind speed and direction. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pages 1–5, Sept 2013.
- [6] J. Dowell, S. Weiss, D. Hill, and D. Infield. Improved spatial modelling of wind fields. In *European Wind Energy Association Conference*, Vienna, Austria, February 2013.
- [7] E. M. Garrigle and P. Leahy. Quantifying the value of improved wind energy forecasts in a pool-based electricity market. *Renewable Energy*, 80(0):517 – 524, 2015.
- [8] G. Giebel, G. Kariniotakis, and R. Brownsword. The state-of-the-art in short-term prediction of wind power. a literature overview. *Deliverable Report D1.1 of the Anemos project*, 2011. available online at <http://anemos.cma.fr>.
- [9] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, 4th edition, 2002.
- [10] D. Hill, K. R. W. Bell, D. McMillan, and D. Infield. A vector auto-regressive model for onshore and offshore wind synthesis incorporating meteorological model information. *Advances in Science and Research*, 11:35–39, 2014.
- [11] A. Malvaldi, J. Dowell, S. Weiss, and D. Infield. Short-term forecasting of wind speed and direction exploiting data non-stationarity. In *International Work-Conference on Time Series*, Granada, Spain, July 2015.
- [12] J. Morales, A. Conejo, H. Madsen, P. Pinson, and M. Zugno. *Integrating Renewables in Electricity Markets - Operational Problems*. Springer, 2014.
- [13] L. Söder and T. Ackermann. *Wind Power in Power Systems: An Introduction*, pages 47–72. John Wiley & Sons, Ltd, 2012.
- [14] U.K. Meteorological Office. Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current), NCAS British Atmospheric Data Centre, 2012.
- [15] B. Widrow, J. McCool, and M. Ball. The complex lms algorithm. *Proceedings of the IEEE*, 63(4):719–720, April 1975.
- [16] G. Zhang, H.-X. Li, and M. Gan. Design a wind speed prediction model using probabilistic fuzzy system. *Industrial Informatics, IEEE Transactions on*, 8(4):819–827, Nov 2012.
- [17] X. Zhu and M. G. Genton. Short-term wind speed forecasting for power system operations. *International Statistical Review*, 80(1):2–23, 2012.