A CYCLO-STATIONARY COMPLEX MULTICHANNEL WIENER FILTER FOR THE PREDICTION OF WIND SPEED AND DIRECTION

Jethro Dowell, Stephan Weiss, David Hill and David Infield

Department of Electronic & Electrical Engineering
University of Strathclyde, Glasgow G1 1XW, Scotland, UK
{jethro.dowell,stephan.weiss,david.c.hill,david.infield}@strath.ac.uk

ABSTRACT

This paper develops a linear predictor for application to wind speed and direction forecasting in time and across different sites. The wind speed and direction are modelled via the magnitude and phase of a complex-valued time-series. A multichannel adaptive filter is set to predict this signal, based on its past values and the spatio-temporal correlation between wind signals measured at numerous geographical locations. The time-varying nature of the underlying system and the annual cycle of seasons motivates the development of a cyclo-stationary Wiener filter, which is tested on hourly mean wind speed and direction data from 13 weather stations across the UK, and shown to provide an improvement over both stationary Wiener filtering and a recent auto-regressive approach.

Index Terms— Multichannel adaptive filtering; adaptive prediction; cyclo-stationary Wiener filter.

1. INTRODUCTION

There are numerous decision making problems which rely on short-term wind forecasts such as sailing, ship routing, air traffic control, etc. Wind forecasts are also used to produce predictions of wind farm power output, which are of significant value to power system operators, electricity generators and energy traders for look ahead times of up to 48 hours [1]. For forecast horizons of 6 to 72 hours numerical weather prediction (NWP) models are employed and achieve a significant improvement on persistence, the standard against which this type of forecast is measured [2, 3]. However, since NWPs are typically only run every 6 hours due to their computational expense, simpler techniques are used to produce short-term forecasts [4].

For the prediction of wind speed alone, a number of linear [5–7] and non-linear approaches [8–11] have been suggested that exploit the spatial correlation between geographically separated measurements. While the underlying system is generally considered non-linear, linear approaches have been justified by their reduced complexity and relatively straightforward operation. However, the omission of wind direction and the subsequent reliance on only speed is claimed to introduce a systematic error into forecasting [8].

Wind direction has been incorporated as a phase into complex-valued data models [21], with prediction performed e.g. by neural networks, such as in [12–16]. These approaches are non-linear and their inherent complexity makes them difficult to expand to a multichannel arrangement to capture spatial correlation. The idea of modelling wind speed and direction by complex-valued time series has been extended even to quaternion-based / hypercomplex techniques [17–19]. These three or four-dimensional methods have been constructed to predict single-channel three dimensional wind vectors, which however exceed the requirements of most applications of wind forecasting. Auto-regressive models to predict the wind vector have been developed but rely on NWP results as inputs [20]. All of these approaches are single-channel, i.e. only attempt to forecast at a single site and ignore spatial correlation, likely due to prohibitive cost and/or numerical difficulties of the multichannel case.

Therefore, this paper attempts to combine prediction of wind speed and direction by means of a complex-valued wind-data model with the exploitation of spatial correlation of measurements at different geographical sites. Driven, akin to [6], by a desire to keep the computational model simple and inexpensive, the analysis is restricted to a linear multichannel prediction approach. This results in the formulation of a cyclo-stationary Wiener filter to exploit the nature of the wind data.

This paper is organised as follows. In Sec. 2, the data and its statistical properties are reviewed, leading to the formulation of an optimum cyclo-stationary Wiener filter in Sec. 3. Sec. 4 demonstrates the advantages of assuming cyclo-stationarity, justifies the parameter setting of the cyclo-stationary Wiener predictor, and demonstrates and compares simulations and test results. Conclusions are drawn in Sec. 5.

The authors gratefully acknowledge the British Atmospheric Data Centre for their supply of meteorological data, and the support of the UK's Engineering and Physical Sciences Research Council via the University of Strathclyde's Wind Energy Systems Centre for Doctoral Training, grant number EP/G037728/1.

2. DATA MODEL AND COVARIANCE MATRIX

Wind speed and direction across M geographically separate sites are embedded in a vector-valued complex time series $\boldsymbol{x}[n] \in \mathbb{C}^M$, where the speed and direction of the wind form the magnitude and phase of the complex samples, and n is the discrete time index. Based on the expectation operator $\mathcal{E}\{\cdot\}$, we define the space-time covariance matrix $\mathbf{R}_{xx}[n, n-\tau] = \mathcal{E}\{\boldsymbol{x}[n]\boldsymbol{x}^*[n-\tau]\}$, which contains auto-correlation sequences of the M wind signals on its main diagonal, and the cross-correlation sequences between different site measurements on the off-diagonals. In the case of wide-sense stationary data, the space-time covariance matrix will only depend on the lag parameter τ and takes on the Hermitian form $\mathbf{R}_{xx}[\tau] = \mathbf{R}_{xx}^{\mathrm{H}}[-\tau]$, where $\{\cdot\}^{\mathrm{H}}$ indicates Hermitian transpose.

Akin to [6,8], for the sake of a somewhat simplified model we neglect non-linear effects. However, we assume a quasi-stationary model, where — for sufficiently short time windows — the space-time covariance matrix can be assumed to be stationary, and therefore only depends on the lag parameter τ. However, taking seasonal patterns into account, we will detail a cyclo-stationary model later in Sec. 3.2.

3. COMPLEX MULTI-CHANNEL PREDICTION

The above data model motivates a quasi-stationary linear predictor outlined in Sec. 3.1, with the estimation of required statistics based on cyclo-stationarity in Sec. 3.2.

3.1. Optimal MSE Predictor

We consider the problem of predicting Δ samples ahead, based on M spatial measurements in $\boldsymbol{x}[n]$ and a time window containing N past samples for each site. Therefore, the prediction error can be formulated as

$$\mathbf{e}_n = \boldsymbol{x}[n] - \sum_{\nu=0}^{N-1} \mathbf{W}^{\mathrm{H}}[n,\nu]\boldsymbol{x}[n-\Delta-\nu] = \boldsymbol{x}[n] - \mathbf{W}_n^{\mathrm{H}}\mathbf{x}_{n-\Delta} \qquad (1)$$

with

$$\mathbf{W}_n = \begin{bmatrix} \mathbf{W}[n,0] \\ \mathbf{W}[n,1] \\ \vdots \\ \mathbf{W}[n,N-1] \end{bmatrix}, \quad \mathbf{x}_n = \begin{bmatrix} \boldsymbol{x}[n] \\ \boldsymbol{x}[n-1] \\ \vdots \\ \boldsymbol{x}[n-N+1] \end{bmatrix} .$$

The matrices $\mathbf{W}[n,\nu] \in \mathbb{C}^{M\times M}$ describe the predictor's reliance on all spatial measurements taken $\nu+\Delta$ samples in the past, at time instance n. Specifically, $[\mathbf{W}[n,\nu]]_{m,\mu}$ addresses the influence of the measurement at site m onto the prediction at the μ^{th} location. In order to simply use the Hermitian transpose operator in (1), $\mathbf{W}[n,\nu]$ contains the complex conjugate prediction filter coefficients.

The error covariance matrix derived from (1), $\mathbf{R}_{ee}[n] = \mathcal{E}\{\boldsymbol{e}_n\boldsymbol{e}_n^{\mathrm{H}}\} \in \mathbb{C}^{M\times M}$, is obtained by taking expectations over the ensemble, and in itself may be varying with time n. Note that in case of stationarity, the dependency of both $\mathbf{W}_n$ and $\mathbf{R}_{ee}[n]$ on n vanishes.

Assume that $\boldsymbol{x}[n]$ is stationary over at least 2Δ samples, $\mathbf{R}_{\mathbf{xx}}[n]$ is Hermitian and therefore positive semi-definite, and, together with full rank of $\mathbf{R}_{\mathbf{xx}}[n]$, admits a unique solution that minimises the quadratic problem

$$\mathbf{W}_{n,\mathrm{opt}} = \arg\min_{\mathbf{W}_n} \mathrm{trace}\{\mathbf{R}_{ee}[n]\} \quad . \qquad (2)$$

The solution can be found by equating the gradient w.r.t. the unconjugated predictor coefficients in $\mathbf{W}_n^*$ to zero. Using matrix-valued calculus summarised in e.g. [23], for constant matrices $\mathbf{A}$ and $\mathbf{B}$, $\partial\mathrm{trace}\{\mathbf{A}\mathbf{W}_n^{\mathrm{H}}\mathbf{B}\}/(\partial\mathbf{W}_n^*) = \mathbf{BA}$ but $\partial\mathrm{trace}\{\mathbf{A}\mathbf{W}_n\mathbf{B}\}/(\partial\mathbf{W}_n^*) = \mathbf{0}$ hold for Wirtinger's complex differentiation. Applying this, and using the product rule for differentiation of the quadratic term in (7), yields

$$\frac{\partial}{\partial\mathbf{W}_n^*}\mathrm{trace}\{\mathbf{R}_{ee}[n]\} = -\mathbf{R}_{x\mathbf{x}}^{\mathrm{H}}[n] + \mathbf{R}_{\mathbf{xx}}[n]\mathbf{W}_{n,\mathrm{opt}} \stackrel{!}{=} \mathbf{0}$$

$$\longrightarrow \mathbf{W}_{n,\mathrm{opt}} = \mathbf{R}_{\mathbf{xx}}^{-1}[n]\mathbf{R}_{x\mathbf{x}}^{\mathrm{H}}[n] \quad , \qquad (3)$$

which is the well-known Wiener-Hopf solution. The time dependence leads to a Wiener solution that will rely on local stationarity, akin to recent results presented in [22].

3.2. Cyclo-Stationary Solution

In the estimation of $\mathbf{R}_{xx}[n,\tau]$ as required for (3) with (8) and (9), we assume both quasi-stationarity, i.e. over a window of L samples the signal $\boldsymbol{x}[n]$ is wide sense stationary, and cyclo-stationarity, i.e. $\mathbf{R}_{xx}[n,\tau] = \mathbf{R}_{xx}[n-kT,\tau]$, with $k \in \mathbb{N}$ and T the fundamental period. To capture the annual seasonal patterns, here T is selected as one year. It is noted that $L < T$ must be satisfied as part of the cyclo-stationary assumption. The estimation of $\hat{\mathbf{R}}_{xx}[n,\tau]$ for the stationary case is recovered when $L = T$.

On the basis of cyclo-stationarity and data available for K past years, the estimation of the covariance matrix for time n is performed as

$$\hat{\mathbf{R}}_{xx}[n,\tau] = \frac{1}{KL}\sum_{k=1}^{K}\sum_{\nu=1-\frac{L}{2}}^{\frac{L}{2}}\boldsymbol{x}[n-kT-\nu]\boldsymbol{x}^{\mathrm{H}}[n-kT-\nu-\tau] + \frac{2}{L}\sum_{\nu=1}^{\frac{L}{2}}\boldsymbol{x}[n-\nu]\boldsymbol{x}^{\mathrm{H}}[n-\nu-\tau] \quad . \qquad (4)$$

The optimal predictor for time n can be then be calculated as

$$\boldsymbol{W}_{n,\mathrm{opt}} = \hat{\mathbf{R}}_{\mathbf{xx}}^{-1}[n]\hat{\mathbf{R}}_{x\mathbf{x}}^{\mathrm{H}}[n] \quad , \qquad (5)$$

with the estimated r.h.s. quantities defined analogously to (8) and (9) based on (4). In determining the window length L, a trade-off is made between consistency of the estimation and the error caused using outdated statistics.

$$
\begin{aligned}
\mathbf{R}_{ee}[n] &= \mathcal{E}\left\{(\boldsymbol{x}[n] - \mathbf{W}_n^{\mathbf{H}}\mathbf{x}_{n-\Delta})(\boldsymbol{x}^{\mathrm{H}}[n] - \mathbf{x}_{n-\Delta}^{\mathbf{H}}\mathbf{W}_n)\right\} \\
&= \mathbf{R}_{xx}[n,0] - \mathcal{E}\{\boldsymbol{x}[n]\mathbf{x}_{n-\Delta}^{\mathbf{H}}\}\mathbf{W}_n - \mathbf{W}_n^{\mathrm{H}}\mathcal{E}\{\mathbf{x}_{n-\Delta}\boldsymbol{x}^{\mathbf{H}}[n]\} + \mathbf{W}_n^{\mathrm{H}}\mathcal{E}\{\mathbf{x}_{n-\Delta}\mathbf{x}_{n-\Delta}^{\mathbf{H}}\}\mathbf{W}_n \\
&= \mathbf{R}_{xx}[n,0] - \mathbf{R}_{x\mathbf{x}}[n]\mathbf{W}_n - \mathbf{W}_n^{\mathrm{H}}\mathbf{R}_{x\mathbf{x}}^{\mathrm{H}}[n] + \mathbf{W}_n^{\mathrm{H}}\mathbf{R}_{\mathbf{xx}}[n]\mathbf{W}_n
\end{aligned} \tag{7}
$$

$$
\mathbf{R}_{x\mathbf{x}}[n] = \begin{bmatrix} \mathbf{R}_{xx}[n,\Delta] \\ \mathbf{R}_{xx}[n,\Delta-1] \\ \vdots \\ \mathbf{R}_{xx}[n,\Delta-N+1] \end{bmatrix} \tag{8}
$$

$$
\mathbf{R}_{\mathbf{xx}}[n] = \begin{bmatrix} \mathbf{R}_{xx}[n-\Delta,0] & \mathbf{R}_{xx}[n-\Delta,1] & \dots & \mathbf{R}_{xx}[n-\Delta,N-1] \\ \mathbf{R}_{xx}[n-\Delta-1,-1] & \mathbf{R}_{xx}[n-\Delta-1,0] & & \mathbf{R}_{xx}[n-\Delta-1,N-2] \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{xx}[n-\Delta-N+1,-N+1] & \mathbf{R}_{xx}[n-\Delta-N+1,-N+2] & \dots & \mathbf{R}_{xx}[n-\Delta-N+1,0] \end{bmatrix} \tag{9}
$$

4. SIMULATIONS AND RESULTS

Based on test data characterised in Sec. 4.1, Sec. 4.2 justified the parameter selection for the proposed method, which is then tested and compared to a benchmark algorithm in Sec. 4.3.

4.1. Test Data

The proposed approach is tested on wind data provided by the British Atmospheric Data Centre, which comprises of recordings over 6 years — from 00:00h on 1/3/1992 to 23:00h on 28/2/1998 — obtained from 13 sites across the UK as detailed in Fig. 1. The data taken in open terrain at a hight of 10m [24], and is sampled at hourly intervals, providing hourly averages that are quantised to a 10° angular granularity and integer multiples of one knot ($0.515\mathrm{ms}^{-1}$).

Although the sites and time window were chosen to have a near-continuous record, the problem of missing data points had to be addressed. Firstly, for the estimation of the covariance matrices, missing data points were zero-padded, and the normalisation in calculating correlation coefficients was adjusted accordingly to provide unbiased estimates. Secondly for the prediction filtering, any errors and their resulting transients were discarded from the prediction output when assessing performance.

4.2. Estimation of Cyclo-Stationary Statistics

Quasi-Stationarity. To investigate the assumption of quasi-stationarity, the total squared error, $\sum_n \mathbf{e}_n^{\mathrm{H}}\mathbf{e}_n$, is compared for different window lengths L at look-ahead intervals $\Delta = 1\dots 6$, as depicted in Fig. 2. The total MSE is normalised w.r.t. a maximum window length L equivalent to one year, where all data is used to calculate a stationary estimate. From

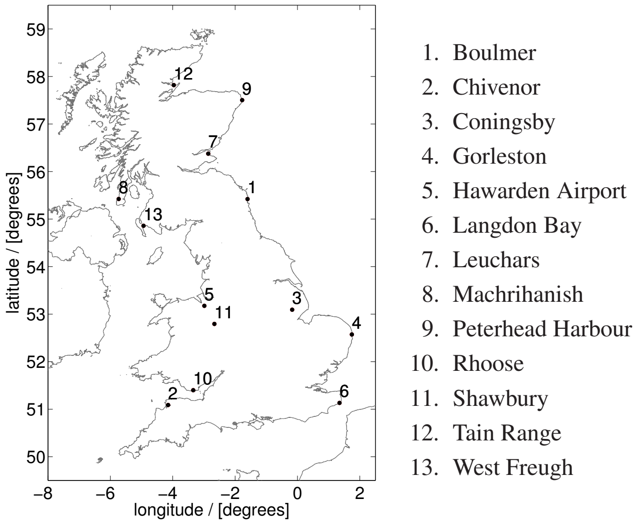

Fig. 1. Geographical distribution of 13 Met. Office stations supplying the test data.

Fig. 2, it is evident that the optimum window length L is approximately 15 weeks, which provides the best trade-off between inconsistent estimation and the use of outdated inputs. It is noted that individual errors for the 13 sites in Fig. 1 fluctuate and that quasi-stationarity across locations varies due to differing local geography, which will also manifest itself in the quality of prediction of wind speed and direction at individual sites later.

Cyclo-Stationarity. To underline the validity of cyclo-stationarity, Fig. 3 compares the total squared error of the Wiener solution under stationary, quasi-stationary, and cyclo-stationary assumptions against the error in persistence [2], $\mathbf{e}_{\mathrm{p}}[n] = \mathbf{x}[n-\Delta] - \mathbf{x}[n]$, which takes the current sample

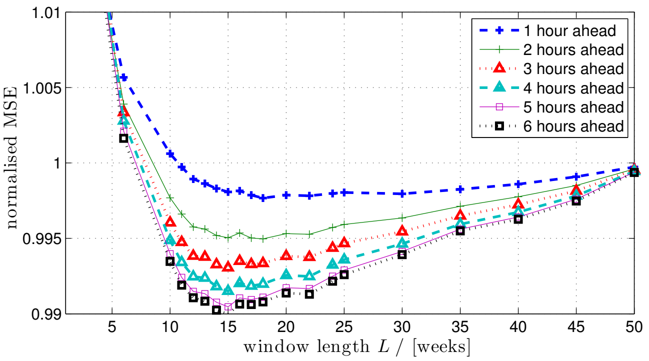

Fig. 2. Cyclo-stationary filter performance depending on the window length L in terms of total MSE normalised w.r.t. stationary Wiener filter for all look-ahead times Δ.

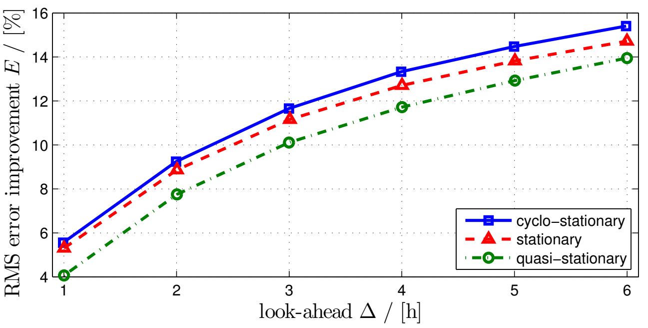

Fig. 3. Mean improvement over persistence across all 13 sites for the cyclo-stationary, quasi-stationary and stationary Wiener filter.

$\mathbf{x}[n]$ as an estimate for a look-ahead, i.e. expecting no change in wind speed or direction over the next Δ samples. The root mean square (RMS) error improvement E shown in Fig. 3 is therefore defined as

$$E = 1 - \left(\frac{\sum_n \mathbf{e}^{\mathrm{H}}[n]\mathbf{e}[n]}{\sum_n \mathbf{e}_p^{\mathrm{H}}[n]\mathbf{e}_p[n]} \right)^{\frac{1}{2}} \quad . \tag{6}$$

The relative improvement over persistence E in Fig. 3 depends on the look-ahead $\Delta = 1 \ldots 6$. The largest improvements are seen at greater look-ahead times where the performance of the persistence method worsens. The relatively poor performance of the quasi-stationary filter illustrates the need for multiple years of training data to smooth the filter coefficients and counter the likely effect of inconsistent estimation. In contrast, the cyclo-stationary solution with an estimate based on a window L of 15 weeks outperforms the two other models.

While the notation in (5) suggests to re-calculate the Wiener filter coefficients at every step in time, for the sake of computational complexity, the coefficient set was only updated once a day, which is sufficiently short compared to the much longer data window L and incurred no penalty in terms of performance.

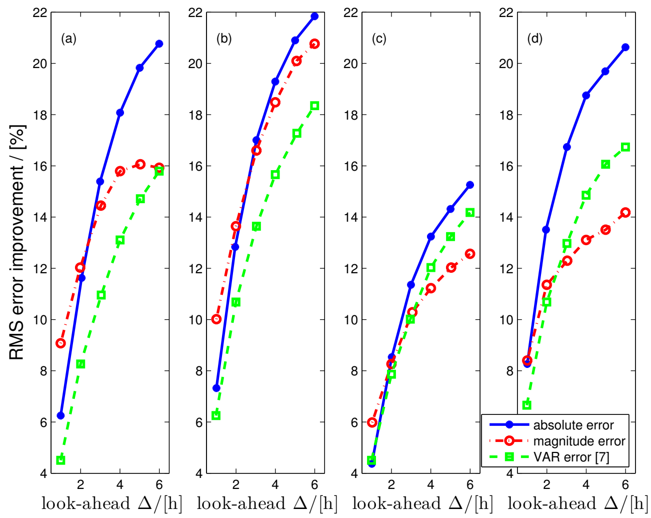

Fig. 4. Improvement over persistence by cyclo-stationary Wiener filter and VAR [7] for (a) Boulmer, (b) Coningsby, (c) Leuchars, and (d) Shawbury; the absolute error refers to $e_l[n]$, while the magnitude error $e_{s,l}[n]$ directly compares to the error in VAR.

4.3. Prediction Results

To compare the proposed approach, it is noted that for spatial multi-channel predictions to date only wind speed is considered. Compared to the prediction error $e_l[n]$ of the predicted estimate $\hat{x}_l[n]$ at a site l, i.e. the lth component in (1), an error for the speed-only component can be defined as

$$e_{s,l}[n] = |x_l[n]| - |\hat{x}_l[n]| \quad . \tag{7}$$

However, note that due to Schwarz' inequality, $|e_{s,l}[n]| \leq |e_l[n]|$, such a comparison is difficult.

In Fig. 4, the proposed approach is compared to a vector auto-regressive (VAR) method [7], a linear multichannel prediction approach for wind speed with a computational complexity that is comparable to the Wiener flter. The VAR method tackles non-stationarity through a detrending procedure applied to the wind speed time series on a site-by-site basis. For the error improvement over persistence, we see that the speed part of the Wiener filter's prediction is comparable to VAR though the performance of both approaches varies across the 4 depicted sites. The directional Wiener filter outperforms both other methods w.r.t. E, since the accuracy of persistence suffers significantly when direction is considered. The accuracy of the speed part of the prediction is almost identical to that of the VAR method applied to data from the same sites in [7], but here valuable directional information is provided as part of the forecast.

5. CONCLUSIONS

For the prediction of both wind speed and direction, modelled as a complex-valued time series, we have proposed a multichannel Wiener filter whose coefficients depend on statistics that have been approximated as cyclo-stationary. This assumption has been justified in simulations, where performance exceeds that of stationary and quasi-stationary solutions. In the prediction of wind data from 13 UK sites, we have demonstrated that the wind speed prediction can match a benchmark state-of-the-art algorithm, but additionally provide a valuable directional forecast.

6. REFERENCES

[1] S.J. Watson, L. Landberg, and J.A. Halliday, "Application of wind speed forecasting to the integration of wind energy into a large scale power system," *IEE Proceedings — Generation, Transmission and Distribution*, **141**(4):357–362, 1994.

[2] T.S. Nielsen, A. Joensen, H. Madsen, L. Landberg, and G. Giebel, "A new reference model for wind power forecasting," *Wind Energy*, **1**(1):2934, Sept. 1998.

[3] H. Madsen, P. Pinson, G. Kariniotakis, H.A. Nielsen, and T. S. Nielsen, "Standardizing the performance evaluation of short-term wind power prediction models," *Wind Engineering*, **29**(6):475–489, 2005.

[4] G. Giebel, "The state-of-the-art in short-term prediction of wind power - a literature overview," Tech. Rep., Risø National Laboratory, 2011.

[5] B. Ernst, B. Oakleaf, M.L. Ahlstrom, M. Lange, and C. Moehrlen, "Predicting the wind," in *IEEE Power & Energy Magazine*, **5**(6):78–89, Nov.-Dec. 2007.

[6] T.H.M. El-Fouly, E.F. El-Saadany, and M.A. Salama, "One day ahead prediction of wind speed and direction," *IEEE Trans. Energy Conversion*, **23**(1):191–201, Jan. 2008.

[7] D. C Hill, D. McMillan, K. R. W. Bell, and D. Infield, "Application of auto-regressive models to UK wind speed data for power system impact studies," *IEEE Trans. Sustainable Energy*, **3**(1):134–141, Jan. 2012.

[8] M.C. Alexiadis, P.S. Dokopoulos, and H.S. Sahsamanoglou, "Wind speed and power forecasting based on spatial correlation models," *IEEE Trans. Energy Conversion*, **14**(3):836–842, Sep. 1999.

[9] I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, and P.S. Dokopoulos, "A fuzzy model forwind speed prediction and power generation in wind parks using spatial correlation," *IEEE Trans. Energy Conversion*, **19**(2):352–361, June 2004.

[10] R. Karki, S. Thapa, and R. B. Billinton, "A simplified risk-based method for short-term wind power commitment," *IEEE Trans. Sustainable Energy*, **3**(3):498–505, Jul. 2012.

[11] G. Zhang, H. Li, and M. Gan, "Design a wind speed prediction model using probabilistic fuzzy system," *IEEE Trans. Industrial Informatics*, **8**(4):819–827, Nov. 2012.

[12] D.P. Mandic, S.L. Goh, and K. Aihara, "Sequential data fusion via vector spaces: Complex modular neural network approach," IEEE Workshop on Machine Learning for Signal Processing, Mystic, Connecticut, pp.147–151, Sept. 2005.

[13] S.L. Goh, M. Chen, D.H. Popović, K. Aihara, D. Obradovic, and D.P. Mandic, "Complex-valued forecasting of wind profile," *Renewable Energy*, **31**(11):1733–1750, Nov. 2006.

[14] S. Li, D.C. Wunsch, E.A. O'Hair, and M.G. Giesselmann, "Using neural networks to estimate wind turbine power generation," *IEEE Trans. Energy Conversion*, **16**(3):276–282, Sept. 2001.

[15] T. Kitajima and T. Yasuno, "Output prediction of wind power generation system using complex-valued neural network," *SICE Annual Conference*, Taipei, Taiwan, pp. 3610–3612, Aug. 2010.

[16] K. Subramanian, R. Savitha, and S. Suresh, "Complex-valued neuro-fuzzy inference system for wind prediction," IEEE World Congress on Computational Intelligence, Brisbane, AUstralia, pp. 1-7, June 2012.

[17] C. Jahanchahi, C.C. Took, and D.P. Mandic, "On HR calculus, quaternion valued stochastic gradient, and adaptive three dimensional wind forecasting," IEEE International Joint Conference on Neural Networks, Barcelona, Spain, pp. 1–5, July 2010.

[18] F.A. Tobar, A. Kuh, and D.P. Mandic, "A novel augmented complex valued kernel LMS," 7th Sensor Array and Multichannel Signal Processing Workshop, Hoboken, NJ, June 2012.

[19] C.C. Took and D.P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. Signal Processing*, **57**(4):1316–1327, April 2009.

[20] M. Khalid and A.V. Savkin, "A method for short-term wind power prediction with multiple observation points," *IEEE Trans. Power Systems*, **27**(2):579–586, May 2012.

[21] T. Gautama, D.P. Mandic and M.M. Van Hulle, "A non-parametric test for detecting the complex-valued nature of time series," *Int. J. Knowledge-based Intell. Eng. Systems*, **8**(2):99–106, Feb. 2004.

[22] A. Ispas, M. Dörpinghaus, G. Ascheid, and T. Zemen, "Characterization of non-stationary channels using mismatched wiener filtering," *IEEE Trans. Signal Processing*, **61**(2):274–288, Feb. 2013.

[23] A. Hjorungnes, and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Processing*, **55**(6):2740–2746, June 2007.

[24] http://badc.nerc.ac.uk