WIND ENERGY
Wind Energ. 2014; **00**:1–10
DOI: 10.1002/we

RESEARCH ARTICLE

This is the accepted version of the following article: J. Dowell, S. Weiss, D. Hill & D. Infield, ”Short-Term Spatio-Temporal Prediction of Wind Speed and Direction,” *Wind Energy*, 17, pp. 1945–1955, 2014, which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1002/we.1682/abstract.

Short-Term Spatio-Temporal Prediction of Wind Speed and Direction

Jethro Dowell, Stephan Weiss, David Hill and David Infield

Dept. Electronic & Electrical Engineering, University of Strathclyde, Glasgow.

ABSTRACT

This paper aims to produce a low-complexity predictor for the hourly mean wind speed and direction from 1 to 6 hours ahead at multiple sites distributed around the UK. 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 filter coefficients are determined by minimising the mean square prediction error. To account for the time-varying nature of the wind data and the underlying system, we propose a cyclo-stationary Wiener solution, which is shown to produce an accurate predictor. An iterative solution, which provides lower computational complexity, increased robustness towards ill-conditioning of the data covariance matrices, and the ability to track time-variations in the underlying system, is also presented. The approaches are tested on wind speed and direction data measured at various sites across the UK. Results show that the proposed techniques are able to predict wind speed as accurately as state-of-the-art wind speed forecasting benchmarks while simultaneously providing valuable directional information. Copyright © 2014 John Wiley & Sons, Ltd.

KEYWORDS

Spatio-temporal Prediction, Wind Forecasting, Multichannel Adaptive Filter, Wiener Filter.

Correspondence

Jethro Dowell, Wind Energy Systems CDT, R3.36 Royal College Building, University of Strathclyde, 204 George Street, Glasgow, G1 1XW, UK.

Email: jethro.dowell@strath.ac.uk

Received . . .

1. INTRODUCTION

As the penetration of wind on power systems across the UK, Europe and indeed much of the developed world increases, it is becoming increasingly important to accurately forecast wind power generation for reliable and economic power system operation since, among other reasons, wind generation must be backed-up by conventional generation. Most wind impact studies have focused on penetration below 20% (UK wind penetration is currently below 10%, however targets as high as 30% penetration by the year 2030 have been suggested [1]) and even they recognise that a substantial increase in operational reserve and storage is required to compensate for the variability of wind and other renewables. Accurate forecasting of wind power generation will reduce imbalance charges to generators making wind more competitive in the energy market [2], and allow for minimal back-up generation, helping to reduce carbon emissions [3].

Accurate predictions of hourly wind farm power output, which rely heavily on accurate wind speed predictions, are of significant value to power system operators, generators and energy traders for look-ahead times of up to 48 hours [4]. For forecasts horizons of 6 to 72 hours numerical weather prediction (NWP) models are employed and are a significant improvement on persistence, the standard against which this type of forecast is measured [5, 6]. However, since NWPs are typically run only every 6 hours due to their computational expense, simpler techniques are used to improve short-term forecasts [7, 8].

Many approaches to point wind speed prediction have been presented [9–17]. Others propose forecasting the probability distribution of possible future wind speeds [18, 19]. Spatial correlation is explicitly mentioned in [9–11, 14, 17]; these methods employ a multichannel-type approach, i.e. simultaneously processing data from multiple locations; the latter on a larger, more generally applicable scale, while the others examine only a few specific sites with particular climatology. Other authors have taken advantage of local climatology and identified wind *regimes* at specific locations, usually based on wind direction, and build separate wind speed forecast models for each regime [11, 14, 20].

Wind direction has been accounted for using complex-valued data models in artificial neural networks, see e.g. [21–25] or even quaternion-based / hypercomplex techniques [26–28]. The neural network approaches are non-linear and their inherent complexity makes them difficult to expand to a multichannel arrangement. The other methods are three dimensional and therefore exceed the requirements of most applications of wind forecasting. Autoregressive models to predict the wind vector have been developed; two bivariate approaches are suggested in [14] while others rely on additional NWP results as inputs such as [29]. In [21] it is claimed that according to [9], the omission of wind direction and reliance on speed only introduces a systematic error into forecasting, and in [22] the decision to use a complex-valued rather than a bivariate signal is justified based on discussions in [30], where the ease of accounting for correlation between parameters in the processing is used as a justification for a complex-valued representation. All of these approaches are single-channel, i.e. only attempt to forecast at a single site and ignore spatial correlation, perhaps due to prohibitive cost and/or numerical difficulties of the multichannel case.

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 [13], by a desire to keep the computational model simple and low-cost, the analysis is restricted to a linear multichannel prediction approach. The novelty of this work lies in the formulation of a cyclo-stationary Wiener filter which utilises the cyclic statistical properties of the wind signal over a number of years.

In Section 2 the data structure and approach to prediction are detailed: the optimal mean squared error predictor is derived in 2.2, the cyclo-stationary solution is described 2.3 and an iterative method in 2.4. Section 3.1 introduces the data used to test the predictors, Sections 3.2 and 3.3 explain the implementation of the Wiener filter and least mean squares algorithm, and the results of their implementation are summarised and compared in 3.4. Conclusions are presented in Section 4.

2. DATA MODEL AND ADAPTIVE PREDICTION

2.1. Wind Data

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. The measured time series from each spatial location forms a channel of the multichannel data model. Since the real and complex components of the wind signal are correlated, i.e. $\boldsymbol{x}[n]$ is an improper complex process [31], it is both sensible and simple to pursue complex processing. As well as the mathematical economies of complex processing, the complex representation offers geometrical insight without the need for a bivariate coordinate system.

Based on the expectation operator $\mathcal{E}\{\cdot\}$, we define the space-time covariance matrix $\mathbf{R}_{xx}[n,\tau] = \mathcal{E}\{\boldsymbol{x}[n]\boldsymbol{x}^{\mathrm{H}}[n-\tau]\}$, 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. The vector $\boldsymbol{x}^{\mathrm{H}}[n]$ denotes the conjugate transpose of $\boldsymbol{x}[n]$. 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 $\mathbf{R}_{xx}[\tau] = \mathbf{R}_{xx}^{\mathrm{H}}[-\tau]$.

With respect to wind speed and wind direction, the former is likely non-stationary and non-linear, while the latter can be volatile and depend heavily on the physical characteristics of the measurement site. Furthermore, the seasonal and diurnal trends that characterise our human experience of the wind are themselves variable. Below, we ignore the potential non-linear nature of the system and restrict ourselves to linear processing, but drop the assumption of stationarity for a quasi-stationary behaviour, whereby the space-time covariance matrix can be assumed to be stationary — and therefore only dependent on the lag parameter τ — for sufficiently short time windows [32].

2.2. Optimal Mean-Squared Error Predictor

We consider the problem of predicting Δ samples ahead while minimising the mean-squared prediction error (MSE), 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

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

 Wind Energ. 2014; **00**:1–10 © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/we
Prepared using weauth.cls

with

$$\boldsymbol{W}_n = \begin{bmatrix} \boldsymbol{W}[n,0] \\ \boldsymbol{W}[n,1] \\ \vdots \\ \boldsymbol{W}[n,N-1] \end{bmatrix} \in \mathbb{C}^{MN\times M} \quad , \quad \mathbf{x}_n = \begin{bmatrix} \boldsymbol{x}[n] \\ \boldsymbol{x}[n-1] \\ \vdots \\ \boldsymbol{x}[n-N+1] \end{bmatrix} \in \mathbb{C}^{MN} \quad . \tag{2}$$

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

The error covariance matrix derived from (1), $\mathbf{R}_{ee}[n] = \mathcal{E}\{\boldsymbol{e}[n]\boldsymbol{e}^{\mathrm{H}}[n]\} \in \mathbb{C}^{M\times M}$, 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 $\boldsymbol{W}_n$ and $\mathbf{R}_{ee}[n]$ on n vanishes. We will carry forward n since it is well known that the wind signal is non-stationary and develop a cyclo-stationary solution in Section 2.3.

Calculating $\mathbf{R}_{ee}[n]$ using (1) yields a quadratic expression in $\boldsymbol{W}_n$,

$$\begin{aligned} \mathbf{R}_{ee}[n] &= \mathcal{E}\left\{(\boldsymbol{x}[n] - \boldsymbol{W}_n^{\mathrm{H}}\mathbf{x}_{n-\Delta})(\boldsymbol{x}^{\mathrm{H}}[n] - \mathbf{x}_{n-\Delta}^{\mathrm{H}}\boldsymbol{W}_n)\right\} \quad , \\ &= \mathbf{R}_{xx}[n,0] - \mathcal{E}\{\boldsymbol{x}[n]\mathbf{x}_{n-\Delta}^{\mathrm{H}}\}\boldsymbol{W}_n - \boldsymbol{W}_n^{\mathrm{H}}\mathcal{E}\{\mathbf{x}_{n-\Delta}\boldsymbol{x}^{\mathrm{H}}[n]\} + \boldsymbol{W}_n^{\mathrm{H}}\mathcal{E}\{\mathbf{x}_{n-\Delta}\mathbf{x}_{n-\Delta}^{\mathrm{H}}\}\boldsymbol{W}_n \quad , \\ &= \mathbf{R}_{xx}[n,0] - \mathbf{R}_{x\mathbf{x}}[n]\boldsymbol{W}_n - \boldsymbol{W}_n^{\mathrm{H}}\mathbf{R}_{x\mathbf{x}}^{\mathrm{H}}[n] + \mathbf{W}_n^{\mathrm{H}}\mathbf{R}_{\mathbf{xx}}[n]\boldsymbol{W}_n \quad , \end{aligned} \tag{3}$$

where

$$\mathbf{R}_{x\mathbf{x}}[n] = \left[\ \mathbf{R}_{xx}[n,\Delta]\,, \quad \mathbf{R}_{xx}[n,\Delta{+}1]\,, \quad \ldots\,, \quad \mathbf{R}_{xx}[n,\Delta{+}N{-}1]\ \right] \quad , \tag{4}$$

$$\mathbf{R}_{\mathbf{xx}}[n] = \begin{bmatrix} \mathbf{R}_{xx}[n{-}\Delta,0] & \mathbf{R}_{xx}[n{-}\Delta,1] & \ldots & \mathbf{R}_{xx}[n{-}\Delta,N{-}1] \\ \mathbf{R}_{xx}[n{-}\Delta{-}1,-1] & \mathbf{R}_{xx}[n{-}\Delta{-}1,0] & & \mathbf{R}_{xx}[n{-}\Delta{-}1,N{-}2] \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{xx}[n{-}\Delta{-}N{+}1,-N{+}1] & \mathbf{R}_{xx}[n{-}\Delta{-}N{+}1,-N{+}2] & \ldots & \mathbf{R}_{xx}[n{-}\Delta{-}\,N{+}1,0] \end{bmatrix}. \tag{5}$$

We assume that $\boldsymbol{x}[n]$ is stationary over at least 2Δ samples. As a result, $\mathbf{R}_{\mathbf{xx}}[n]$ is Hermitian and therefore positive semi-definite [33]. This property together with full rank of $\mathbf{R}_{\mathbf{xx}}[n]$ admits a unique solution to minimises the mean square error,

$$\boldsymbol{W}_{n,\mathrm{opt}} = \arg\min_{\boldsymbol{W}_n} \mathrm{trace}\{\mathbf{R}_{ee}[n]\} \quad . \tag{6}$$

It can be shown that $\mathrm{trace}\{\mathbf{R}_{ee}[n]\}$ is quadratic in $\boldsymbol{W}_n$, such that the solution to (6) can be found by matrix- and complex-valued calculus [34]. Finding the minimum requires equating the gradient w.r.t. the unconjugated predictor coefficients in $\boldsymbol{W}_n^*$ to zero. We utilise results from [34] which show that for constant matrices $\mathbf{A}$ and $\mathbf{B}$ the expressions $\partial\mathrm{trace}\{\mathbf{A}\boldsymbol{W}_n^{\mathrm{H}}\mathbf{B}\}/(\partial\boldsymbol{W}_n^*) = \mathbf{BA}$ but $\partial\mathrm{trace}\{\mathbf{A}\boldsymbol{W}_n\mathbf{B}\}/(\partial\boldsymbol{W}_n^*) = \mathbf{0}$ hold. Applying this, and using the product rule for differentiation of the quadratic term in (3), yields

$$\frac{\partial}{\partial\boldsymbol{W}_n^*}\mathrm{trace}\{\mathbf{R}_{ee}[n]\} = -\mathbf{R}_{x\mathbf{x}}^{\mathrm{H}}[n] + \mathbf{R}_{\mathbf{xx}}[n]\boldsymbol{W}_n \quad . \tag{7}$$

Finally, setting the gradient on the r.h.s. of (7) equal to zero yields the optimum predictor coefficients that minimise $\mathbf{R}_{ee}[n]$ $\mathrm{trace}\{\mathbf{R}_{ee}[n]\}$,

$$\boldsymbol{W}_{n,\mathrm{opt}} = \mathbf{R}_{\mathbf{xx}}^{-1}[n]\mathbf{R}_{x\mathbf{x}}^{\mathrm{H}}[n] \quad , \tag{8}$$

which is the well-known Wiener-Hopf solution [35, 36].

2.3. Cyclo-stationary Solution

The cyclo-stationary solution is based on the assumption that windows of data of length L are approximately stationary, and furthermore, that the statistics of that period are the same during the equivalent window in all years. The covariance matrix $\mathbf{R}_{xx}[n,\tau]$ is estimated by calculating the expectation using only data in the quasi-stationary window centred on n from each year of available training data. In the estimation of $\mathbf{R}_{xx}[n,\tau]$, we assume 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, i.e. 1 year. 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

$$\hat{\mathbf{R}}_{xx}[n,\tau] = \frac{1}{K(L+1)} \sum_{k=1}^{K} \left(\sum_{\nu=\frac{L}{2}}^{\frac{L}{2}} \boldsymbol{x}[n-kT-\nu]\boldsymbol{x}^{\mathrm{H}}[n-kT-\nu-\tau] \right) + \frac{2}{L} \sum_{\nu=1}^{\frac{L}{2}} \boldsymbol{x}[n-\nu]\boldsymbol{x}^{\mathrm{H}}[n-\nu-\tau] \quad . \tag{9}$$

The optimal prediction filter for time n can then be calculated as

$$\boldsymbol{W}_{n,\mathrm{opt}} = \hat{\mathbf{R}}_{\mathbf{xx}}^{-1}[n]\hat{\mathbf{R}}_{x\mathbf{x}}^{\mathrm{H}}[n] \quad . \tag{10}$$

Determining the window length, L, is a trade-off between consistency of performance and excess error caused by the inclusion of mismatched statistics. The window must be short enough to capture the common properties of the season but also long enough to smooth the effects of extreme events from individual years.

2.4. Iterative Prediction Filter

As an alternative to the Wiener-Hopf solution defined by (8), the quadratic MSE cost function has motivated lower-cost iterative approaches such as the method of steepest descent where

$$\boldsymbol{W}_{n+1} = \boldsymbol{W}_n - \mu \frac{\partial}{\partial \boldsymbol{W}_n^*} \mathrm{trace}\{\mathbf{R}_{ee}[n]\} \quad , \tag{11}$$

i.e. the algorithm steps in the direction of the negative gradient of the cost function in proportion to the learning rate, μ. Amongst iterative schemes, Widrow's stochastic gradient technique called the least mean squares (LMS) algorithm [37] has proven simple and robust, whereby $\mathbf{R}_{ee}[n]$ is replaced by the poor instantaneous estimate $\hat{\mathbf{R}}_{ee}[n] = \boldsymbol{e}[n]\boldsymbol{e}^{\mathrm{H}}[n]$. The differentiation $\frac{\partial}{\partial \boldsymbol{W}_n^*} \mathrm{trace}\{\boldsymbol{e}[n]\boldsymbol{e}^{\mathrm{H}}[n]\} = -\mathbf{x}_n \boldsymbol{e}^{\mathrm{H}}[n]$ leads to the straightforward update equation

$$\boldsymbol{W}_{n+1} = \boldsymbol{W}_n + \mu \mathbf{x}_n \boldsymbol{e}^{\mathrm{H}}[n] \quad . \tag{12}$$

Assuming a sufficiently small value of μ, the iterative nature of (12) averages out the gradient noise that results from the poor estimation of $\mathbf{R}_{ee}[n]$.

For the stationary case, selecting μ presents a trade-off between convergence speed and excess MSE. For a large value of μ within the learning rate bounds [35,36], the filter coefficients will quickly converge towards the Wiener-Hopf solution, however the gradient noise contributes inaccuracy to $\boldsymbol{W}$ which negatively impacts the predictor's performance. Choosing a smaller μ reduces the effect of noise on the filter coefficients — reducing excess MSE — at the cost of convergence speed.

Under non-stationary conditions, the optimal filter coefficients are time dependent; the LMS algorithm will track this dependence, albeit with some lag [38]. Now the trade-off when choosing μ lies between accurate tracking and minimising lag. Convergence speed is still also a consideration. The tracking ability and relative simplicity of the LMS algorithm offer a computationally inexpensive predictor which can potentially provide better tracking abilities than algorithms with a significantly higher complexity [39, 40].

3. TESTING AND RESULTS

3.1. Data Used for Testing

The proposed approaches are 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 Figure 1. The measurements are taken in open terrain at a height of 10m and comprise hourly averages that are quantised to a 10° angular granularity and integer multiples of one knot (0.515ms^{-1}) [41]. For this study, only sites with near complete continuous data are used and any prediction errors affected by a missing or erroneous data points are discarded. For the purposes of calculating data covariance matrices, missing and erroneous data was again discarded and the normalisation factors in (9) adjusted accordingly. Predictions and errors affected by missing data are thus mitigated. The performance of each filter is assessed by measuring the MSE and improvement over the persistence method, which is a common benchmark for such forecasts [5].

3.2. Wiener Solution

The stationary Wiener filter for the complete data set was calculated on the 5 year training data and tested on the remaining year of data for comparison to the cyclo-stationary Wiener and LMS approaches. The filter length was set to be $N = 3$ after

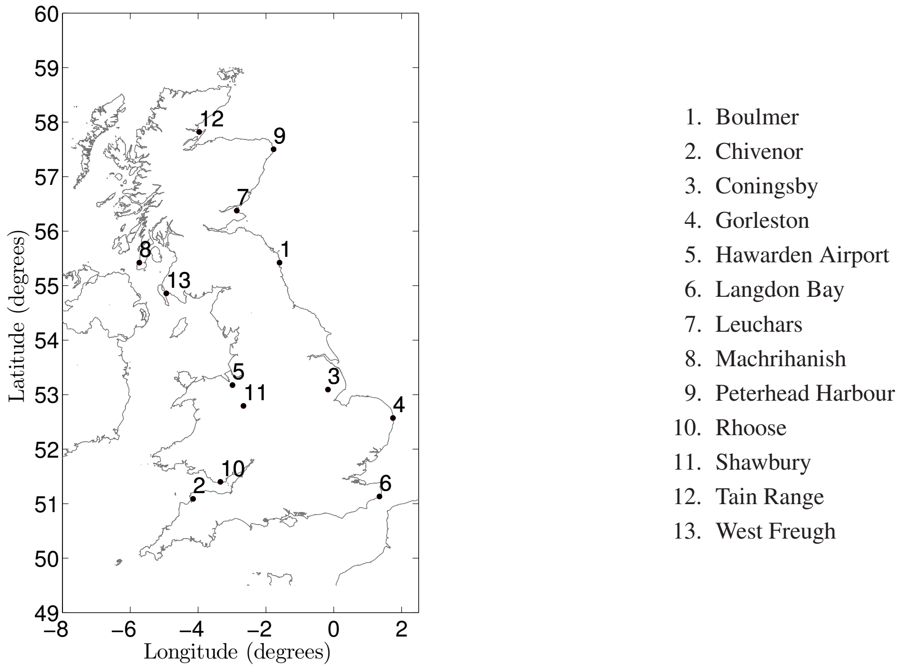

Figure 1. Geographical distribution of 13 Met Office stations supplying test data.

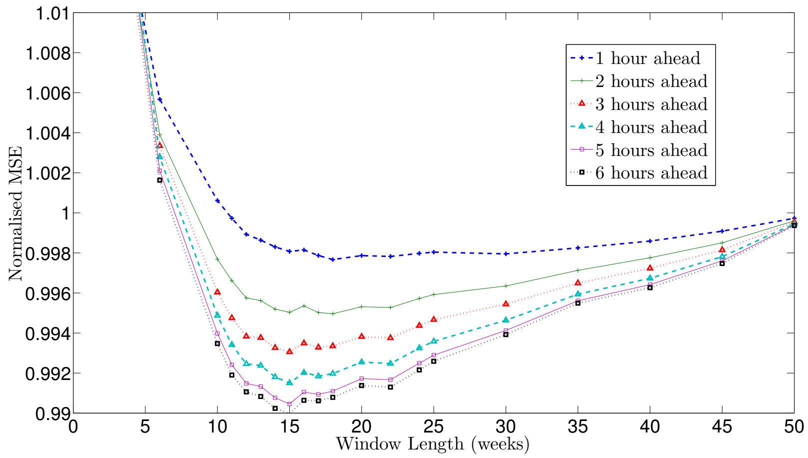

Figure 2. Cyclo-stationary filter performance depending on the window length L in terms of MSE averaged across all locations and normalised w.r.t. a stationary Wiener filter for all look-ahead times Δ.

exhaustive testing showed no benefit from increasing it further. In order to implement the cyclo-stationary approximation, the optimum window size L that best approximates stationarity with a sufficiently consistent estimation was found through numerical testing, shown in Figure 2, to be L equivalent to 15 weeks. Data windows from some sites are closer to being stationary than others and this is reflected in the final filter's performance.

While the notation in (9) suggests to re-calculate the Wiener filter coefficients at every time step, for the sake of computational complexity, the coefficient set was only updated once a day, i.e. every 24 time steps. In tests this proved to be sufficiently short compared to the much longer data window L, and incurred no penalty in terms of performance.

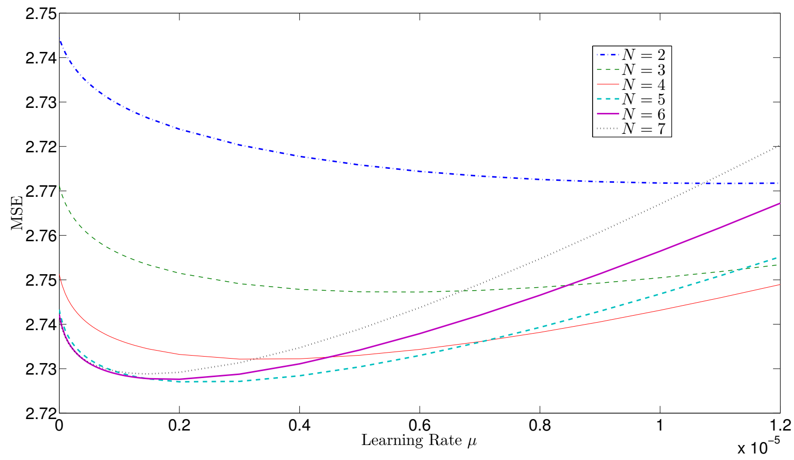

Figure 3. Tracking performance of LMS algorithm for different filter lengths when initialised with the stationary Wiener solution.

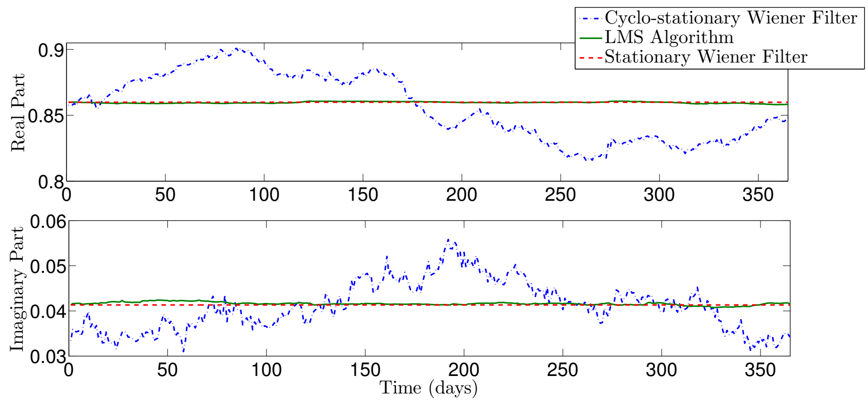

Figure 4. Variation of a typical Cyclo-stationary, LMS and stationary Wiener auto-correlative filter coefficient during prediction year.

3.3. Least Mean Squares Algorithm

As discussed in Section 2.4, there are trade-offs to be made when choosing the filter length, N, and the learning rate, μ, of the LMS algorithm. The filter length and learning rate are chosen to minimise excess MSE caused by poor tracking of the non-stationarity and the effects of noise.

In order to characterise the tracking ability of the LMS algorithm, we initialise the filter weights with the Wiener-Hopf solution of Section 2.2, using (8) and the training data. Based on the resulting tracking performance in Figure 3 for a combination of values for the filter length N and the learning rate μ, we have determined approximately optimal performance for $N = 5$ and $\mu = 2.5 \times 10^{-6}$, which have been employed for all further tests with the LMS predictor below.

Figure 4 shows the typical variation of the filter coefficients during the year of test data. Clearly the LMS algorithm fails to match the tracking ability of the cyclo-stationary Wiener filter. This should not come as a surprise since each LMS update relies on only very recent information where as the cyclo-stationary filter additionally relies on information from previous weeks and years.

3.4. Results

Prediction results from the cyclo-stationary Wiener solution and LMS algorithm in terms of root mean-squared error are shown in Table I. The mean improvement over persistence is shown in Figure 5 and compared to results from the stationary Wiener filter. The LMS algorithm's tracking ability provides a clear improvement on the stationary Wiener filter though not as much as the cyclo-stationary Wiener solution. The largest improvements are seen at greater look-ahead times where the performance of the persistence method worsens.

Wind Energ. 2014; **00**:1–10
DOI: 10.1002/we
Prepared using weauth.cls

		Speed RMSE				Velocity RMSE		
Site	Δ	CSWF	LMS	VAR(2)	Persistence	CSWF	LMS	Persistence
Boulmer	1	**1.09**	**1.09**	1.15	1.20	**1.54**	1.55	1.64
	2	1.50	**1.49**	1.57	1.71	**2.08**	2.09	2.35
	3	1.76	**1.74**	1.83	2.06	**2.41**	2.43	2.85
	4	1.97	**1.95**	2.03	2.34	**2.68**	2.70	3.27
	5	2.14	**2.12**	2.18	2.56	**2.91**	2.95	3.64
	6	2.30	**2.27**	2.30	2.73	**3.12**	3.17	3.95
Coningsby	1	**0.84**	0.85	0.87	0.93	**1.22**	**1.22**	1.32
	2	**1.17**	1.18	1.21	1.36	**1.67**	**1.67**	1.92
	3	**1.40**	1.41	1.45	1.68	**1.98**	1.99	2.38
	4	**1.58**	1.60	1.64	1.94	**2.23**	2.24	2.77
	5	**1.74**	1.76	1.80	2.17	**2.45**	2.46	3.09
	6	**1.87**	1.91	1.93	2.37	**2.63**	2.65	3.37
Leuchars	1	**1.08**	**1.08**	1.09	1.15	**1.52**	**1.52**	1.59
	2	**1.49**	**1.49**	1.50	1.63	**2.07**	**2.07**	2.25
	3	1.77	**1.75**	1.76	1.97	**2.43**	2.44	2.74
	4	2.00	**1.98**	**1.98**	2.25	**2.72**	2.74	3.14
	5	2.20	2.18	**2.17**	2.50	**2.98**	3.01	3.48
	6	2.37	2.36	**2.32**	2.71	**3.20**	3.24	3.77
Shawbury	1	1.01	**1.00**	1.03	1.10	1.43	**1.42**	1.56
	2	1.37	**1.33**	1.38	1.54	1.90	**1.89**	2.20
	3	1.62	**1.56**	1.60	1.85	**2.23**	**2.23**	2.67
	4	1.83	**1.77**	1.79	2.10	**2.50**	2.51	3.07
	5	2.01	**1.94**	1.95	2.32	**2.73**	2.76	3.41
	6	2.15	2.10	**2.09**	2.51	**2.93**	2.98	3.70
Mean Error	1	**1.11**	**1.11**	1.15	1.20	**1.64**	**1.64**	1.74
	2	1.51	**1.50**	1.52	1.66	**2.20**	**2.20**	2.42
	3	1.78	**1.77**	1.78	1.98	**2.59**	**2.59**	2.92
	4	2.01	2.00	**1.98**	2.24	**2.90**	2.91	3.34
	5	2.20	2.20	**2.14**	2.46	**3.16**	3.19	3.70
	6	2.36	2.37	**2.28**	2.64	**3.39**	2.43	4.01

Table I. Root mean-squared speed and vector prediction errors (RMSE) from four selected sites, and the mean taken across all thirteen sites in the model for the cyclo-stationary Wiener filter (CSWF), least mean squares algorithm (LMS), vector autoregressive (VAR(2)) and persistence.

To compare the proposed approaches with others, it is noted that for non-site-specific spatial multichannel prediction to date only wind speed has been considered. Compared to the complex 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, $e_{\mathrm{s},l}[n]$, can be defined as

$$e_{\mathrm{s},l}[n] = |x_l[n]| - |\hat{x}_l[n]| \quad . \tag{13}$$

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

The accuracy of the cyclo-stationary Wiener filter's wind speed prediction for specific look-ahead periods, Δ, has been calculated and is compared to the mathematically similar vector autoregressive (VAR(2)) method of [17] in Figure 6. The autoregressive coefficients of the VAR(2) model are static and calculated using the Yule-Walker approach on the detrended test data [42]. The annual and diurnal trends were determined by fitting Fourier series to the test data for individual sites: a single three term series for the annual trend and four two-term series for the diurnal trend, one for each season [17].

We see that the speed part of the cyclo-stationary Wiener filter's prediction is comparable to the VAR(2) method though the performance of both approaches varies from site to site. Unsurprisingly, the directional Wiener filter outperforms both other methods by this measure since the accuracy of persistence suffers significantly when direction is considered.

4. CONCLUSIONS

The aim of this paper was to propose a relatively low-cost spatio-temporal adaptive filter for predicting the hourly mean of both wind speed and direction, based on spatial information drawn from geographically separated sites. With prediction

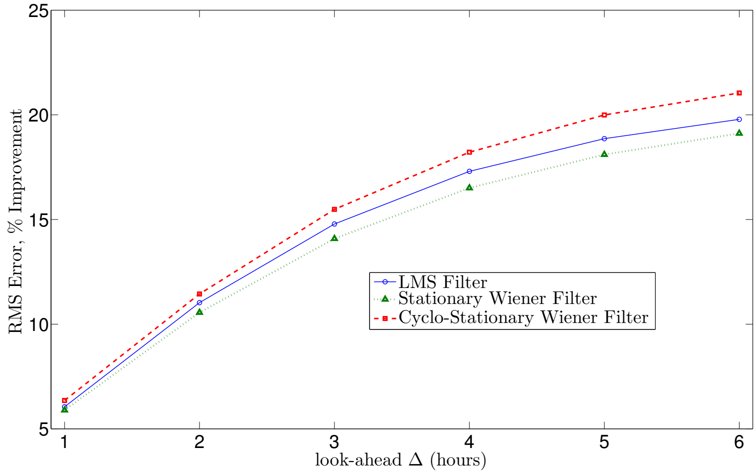

Figure 5. Improvement over persistence for the stationary Wiener filter, cyclo-stationary Wiener filter and LMS filter.

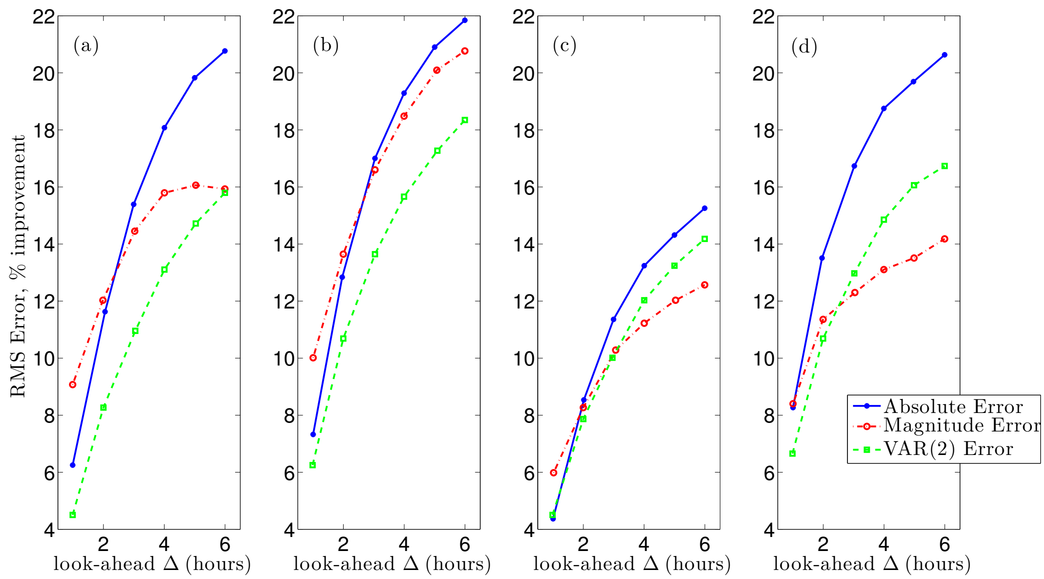

Figure 6. Improvement over persistence of VAR(2) and cyclo-stationary Wiener predictions for four sites. Both vector error and the error in magnitude of the Wiener predictions are shown for comparison to the speed only VAR(2) method. Site (a): Boulmer, Site (b): Coningsby, Site (c): Leuchars, Site (d): Shawbury.

methods of comparable complexity to date either restricted to single sites but with multiple parameters captured e.g. in complex valued time series, or restricted to speed prediction only when based on multiple measurement locations, the proposed method fills a gap in research.

We have developed a new cyclo-stationary Wiener filter which is motivated by the approximately annual cycles in the data, and leads to the estimation of a cyclo-stationary covariance matrix, which is assumed to be quasi-stationary over sufficiently small intervals. The calculation of this covariance matrix aims to keep the data window sufficiently short in order to discard out-dated samples from the estimation, while the cyclic inclusion of several years' of data enhances consistency of estimates. An iterative, stochastic gradient predictor has also been suggested, which utilises a multichannel least mean squares algorithm. The LMS is motivated by its significantly lower complexity compared to the Wiener solution, its enhanced numerical stability due to the avoidance of matrix inversions, and its favourable tracking performance when compared to other, much more numerically involved schemes like the recursive least squares algorithm.

DOI: 10.1002/we
Prepared using weauth.cls

The proposed methods have been tested on wind measurements obtained at 13 locations in the UK over a period of 6 years. The results have been assessed against persistence, and generally show superior performance of the cyclo-stationary Wiener filter over a stationary version and the LMS, which supports the assumption of cyclo-stationarity of the data. The cyclo-stationary Wiener filter and LMS provide speed and vector predictions with greater accuracy than persistence and the simplicity of the LMS algorithm is found to come at only a small cost in vector prediction accuracy and no cost in speed prediction accuracy. Finally, the performance of the proposed methods is compared to the speed-only spatial prediction VAR(2), as described in [17]. The proposed filters are found to produce wind speed predictions of comparable accuracy to VAR(2) while, significantly, also providing directional information. The wide applicability of the cyclo-stationary Wiener filter and the multichannel LMS prediction methods provide a valuable alternative to other statistical techniques which are often have to be tailored to local conditions or are computationally demanding.

ACKNOWLEDGEMENTS

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.

REFERENCES

1. Committee on Climate Change. Building a Low Carbon Economy — The UK's Innovation Challenge. UK Govenment Policy Report 2010.
2. Girard R, Laquaine K, Kariniotakis G. Assessment of wind power predictability as a decision factor in the investment phase of wind farms. *Applied Energy* 2013; **101**(0):609–617.
3. Denny E, O'Malley M. Wind generation, power system operation, and emissions reduction. *IEEE Transactions on Power Systems* 2006; **21**(1):341–347.
4. Watson SJ, Landberg L, Halliday JA. Application of wind speed forecasting to the integration of wind energy into a large scale power system. *Generation, Transmission and Distribution, IEE Proceedings-* 1994; **141**(4):357–362.
5. Nielsen T, Joensen A, Madsen H, Landberg L, Giebel G. A new reference model for wind power forecasting. *Wind Energy* 1998; **1**:29–36.
6. Madsen H, Pinson P, Kariniotakis G, Nielsen HA, Nielsen TS. Standardizing the performance evaluation of short-term wind power prediction models. *Wind Engineering* 2005; **29**(6):475–489.
7. Giebel G, Brownsword R, Kariniotakis G, Denhard M, Draxl C. *The State-Of-The-Art in Short-Term Prediction of Wind Power*. ANEMOS.plus, 2011. Project funded by the European Commission under the 6th Framework Program, Priority 6.1: Sustainable Energy Systems.
8. Zhu X, Genton MG. Short-term wind speed forecasting for power system operations. *International Statistical Review* 2012; **80**(1):2–23, doi:10.1111/j.1751-5823.2011.00168.x.
9. Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS. Wind speed and power forecasting based on spatial correlation models. *IEEE Transactions on Energy Conversion* 1999; **14**(3):836–842.
10. Damousis IG, Alexiadis MC, Theocharis JB, Dokopoulos PS. A fuzzy model forwind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion* 2004; **19**(2):352–361.
11. Gneiting T, Larson K, Westrick K, Genton MG, Aldrich E. Calibrated probabilistic forecasting at the stateline wind energy center. *Journal of the American Statistical Association* 2006; **101**(475):968–979.
12. Ernst B, Oakleaf B, Ahlstrom ML, Lange M, Moehrlen C. Predicting the wind. *78 IEEE power & energy magazine*, IEEE, 2007.
13. El-Fouly THM, El-Saadany EF, Salama MA. One day ahead prediction of wind speed and direction. *IEEE Transactions on Energy Conversion* 2008; **23**(1):191–201.
14. Hering AS, Genton MG. Powering up with space-time wind forecasting. *Journal of American Statistical Association* 2010; **105**:92–104.
15. Karki R, Thapa S, Billinton RB. A simplified risk-based method for short-term wind power commitment. *IEEE Transactions on Sustainable Energy* July 2012; **3**(3):498–505, doi:10.1109/TSTE.2012.2190999.
16. Zhang G, Li H, Gan M. Design a wind speed prediction model using probabilistic fuzzy system. *IEEE Transactions on Industrial Informatics* November 2012; **8**(4):819–827, doi:10.1109/TII.2012.2205392.
17. Hill DC, McMillan D, Bell KRW, Infield D. Application of auto-regressive models to UK wind speed data for power system impact studies. *IEEE Transactions on Sustainable Energy* 2012; **3**(1):134–141.
18. Sloughter JM, Gneiting T, Raftery AE. Probabilistic wind speed forecasting using ensembles and bayesian model averaging. *Journal of the American Statistical Association* 2010; **105**(489):25–35.

19. Jeon J, Taylor JW. Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association* 2012; **107**(497):66–79.
20. Pinson P, Madsen H. Adaptive modelling and forecasting of offshore wind power fluctuations with Markov-switching autoregressive models. *Journal of Forecasting* 2012; **31**(4):281–313, doi:10.1002/for.1194.
21. Mandic DP, Goh SL, Aihara K. Sequential data fusion via vector spaces: Complex modular neural network approach. IEEE, 2005.
22. Goh SL, Chen M, Popović DH, Aihara K, Obradovic D, Mandic DP. Complex-valued forecasting of wind profile. *Renewable Energy* 2006; **31**(11):1733–1750.
23. Li S, Wunsch DC, O'Hair EA, Giesselmann MG. Using neural networks to estimate wind turbine power generation. *IEEE Transactions on Energy Conversion* 2001; **16**(3).
24. Kitajima T, Yasuno T. Output prediction of wind power generation system using complex-valued neural network. *SICE Annual Conference 2010 August 18–21, 2010, The Grand Hotel, Taipei, Taiwan*, 2010.
25. Subramanian K, Savitha R, Suresh S. Complex-valued neuro-fuzzy inference system for wind prediction. IEEE World Congress on Computational Intelligence, 2012.
26. Took C, Mandic DP. The quaternion LMS algorithm for adaptive filtering of hypercomplex processes. *IEEE Transactions on Signal Processing* Apr 2009; **57**(4):1316–1327, doi:10.1109/TSP.2008.2010600.
27. Jahanchahi C, Took CC, Mandie DP. On HR calculus, quaternion valued stochastic gradient, and adaptive three dimensional wind forecasting. IEEE, 2010.
28. Took C, Strbac G, Aihara K, Mandic D. Quaternion-valued short-term joint forecasting of three-dimensional wind and atmospheric parameters. *Renewable Energy* 2011; **36**(6):1754 – 1760,
29. Khalid M, Savkin AV. A method for short-term wind power prediction with multiple observation points. *IEEE Transactions on Power Systems* May 2012; **27**(2):579–586, doi:10.1109/TPWRS.2011.2160295.
30. Gautama T, Mandic D, Hulle M. A non-parametric test for detecting the complex-valued nature of time series. *Int J Knowledge-based Intell Eng Systems* 2004; **8**(2):99–106.
31. Walden A, Rubin-Delanchy P. On testing for impropriety of complex-valued Gaussian vectors. *IEEE Transactions on Signal Processing* March 2009; **57**(3):825–834, doi:10.1109/TSP.2008.2008970.
32. Ispas A, Dörpinghaus M, Ascheid G, Zemen T. Characterization of non-stationary channels using mismatched Wiener filtering. *IEEE Transactions on Signal Processing* 2013; **61**(2):274–288.
33. Golub GH, Loan CFV. *Matrix Computations*. The John Hopkins University Press, 1996.
34. Hjørungnes A, Gesbert D. Complex-valued matrix differentiation: Techniques and key results. *IEEE Transactions on Signal Processing* June 2007; **55**(6, part I):2740–2746, doi:10.1109/TSP.2007.893762.
35. Widrow B, Stearns SD. *Adaptive Signal Processing*. Prentice Hall: Englewood Cliffs, New York, 1985.
36. Haykin S. *Adaptive Filter Theory*. 4th edn., Prentice Hall, 2002.
37. Widrow B, McCool J, Ball M. The complex LMS algorithm. *Proceedings of the IEEE* 1975; **63**(4):719–720, doi: 10.1109/PROC.1975.9807.
38. Widrow B, McCool J, Larimore M, Johnson C. Stationary and nonstationary learning characteristics of the LMS adaptive filter. *Proceedings of the IEEE* aug 1976; **64**(8):1151–1162, doi:10.1109/PROC.1976.10286.
39. Bershad NJ, Macchi OM. Adaptive Recovery of a Chirped Sinusoid in Noise, Part 2: Performance of the LMS Algorithm. *IEEE Transaction on Signal Processing* March 1991; **39**(3):595–602, doi:10.1109/78.80879.
40. Macchi OM, Bershad NJ. Adaptive Recovery of a Chirped Sinusoid in Noise, Part 1: Performance of the RLS Algorithm. *IEEE Transaction on Signal Processing* March 1991; **39**(3):583–594, doi:10.1109/78.80878.
41. URL `http://badc.nerc.ac.uk`.
42. Lütkepohl H. *New Introduction to Multiple Time Series Analysis*. Springer, 2005.

Wind Energ. 2014; **00**:1–10
DOI: 10.1002/we
Prepared using weauth.cls