A WIDELY LINEAR MULTICHANNEL WIENER FILTER FOR WIND PREDICTION

Jethro Dowell, Stephan Weiss, David Infield

Dept. of Electronic and Electrical Engineering
University of Strathclyde, Glasgow, UK
{jethro.dowell,stephan.weiss,david.infield}@strath.ac.uk

Swati Chandna

Dept. of Electronic Engineering
University of Surrey, Guildford, UK
s.chandna@surrey.ac.uk

ABSTRACT

The desire to improve short-term predictions of wind speed and direction has motivated the development of a spatial covariance-based predictor in a complex valued multichannel structure. Wind speed and direction are modelled as the magnitude and phase of complex time series and measurements from multiple geographic locations are embedded in a complex vector which is then used as input to a multichannel Wiener prediction filter. Building on a $\mathbb{C}$-linear cyclo-stationary predictor, a new widely linear filter is developed and tested on hourly mean wind speed and direction measurements made at 13 locations in the UK over 6 years. The new predictor shows a reduction in mean squared error at all locations. Furthermore it is found that the scale of that reduction strongly depends on conditions local to the measurement site.

Index Terms— Widely linear processing, prediction, complex data, Wiener filter

1. INTRODUCTION

The short-term prediction of wind power generation is essential for reliable and economic power system operation and for that reason is the subject of much current research, see [1, 2] and references therein. A number of methods have been developed to model the spatio-temporal relationship between multiple measurement locations to produce accurate predictions [3–5]. In addition, wind direction is frequently modelled alongside wind speed in a complex-valued time series as in [6], among others. The spatio-temporal problem lends itself to a multichannel treatment, and complex valued filtering can be employed to predict both wind speed and direction as in [5].

Linear operations can be applied to complex quantities (then termed strictly linear or $\mathbb{C}$-linear) in exactly the same way as to real ones, but with some significant limitations that must be appreciated. Consider the linear transformation

$$y = kx \qquad (1)$$

with $x = x_r + jx_i$ and $y = y_r + jy_i$. Writing the product in terms of its real and imaginary parts

$$\begin{bmatrix} y_r \\ y_i \end{bmatrix} = \begin{bmatrix} \text{Re } k & -\text{Im } k \\ \text{Im } k & \text{Re } k \end{bmatrix} \begin{bmatrix} x_r \\ x_i \end{bmatrix} \qquad (2)$$

and comparing that to the more general $\mathbb{R}^2$ transformation

$$\begin{bmatrix} y_r \\ y_i \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} x_r \\ x_i \end{bmatrix} \qquad (3)$$

illustrates the limitation: the $\mathbb{R}^2$ transformation is only $\mathbb{C}$-linear iff $M_{11} = M_{22}$ and $M_{12} = -M_{21}$. The complex equivalent of (3) is the *widely linear* transformation

$$y = k_1 x + k_2 x^* \qquad . \qquad (4)$$

Detailed discussions on widely linear processing can be found in [7, 8].

Wind measurements have been modelled as a $\mathbb{C}$-linear cyclo-stationary time series in [5] with computational efficiency in mind. However, in order to quantify the potential gains that could achieved, in this paper we will explore a widely linear model even though this comes at the expense of doubling the filter order.

The data model is described in Section 2 and the minimum mean squared error predictor is derived in 2.1, with the cyclo-stationary estimation of the covariance matrices outlined in 2.2. The data used for testing and test results are presented in Section 3 and conclusions are drawn in Section 4.

2. DATA MODEL AND PREDICTION

At discrete time n, the wind speed and direction at M locations are embedded as the magnitude and phase of a complex valued vector $\boldsymbol{x}[n] \in \mathbb{C}^M$. The spatial covariance matrix is defined based on the expectation operator, $\mathcal{E}\{\cdot\}$, as $\mathbf{R}_{xx}[n, \tau] = \mathcal{E}\{\boldsymbol{x}[n]\boldsymbol{x}^H[n - \tau]\}$, where $\boldsymbol{x}^H[n]$ denotes the Hermitian transpose of $\boldsymbol{x}[n]$ and τ is a general lag parameter.

Since we are pursuing widely linear processing, we also define the complementary covariance matrix based on the expectation operator as $\tilde{\mathbf{R}}_{xx}[n, \tau] = \mathcal{E}\{\boldsymbol{x}[n]\boldsymbol{x}^T[n - \tau]\}$. In addition, by considering the *augmented* vector $\underline{\boldsymbol{x}}[n]$, which is the concatenation of $\boldsymbol{x}[n]$ and its conjugate, we can define the augmented covariance matrix $\underline{\mathbf{R}}_{xx}[n, \tau] = \mathcal{E}\{\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}^H[n - \tau]\}$,

$$\begin{aligned} \underline{\mathbf{R}}_{xx}[n, \tau] &= \mathcal{E}\left\{ \begin{bmatrix} \boldsymbol{x}[n] \\ \boldsymbol{x}^*[n] \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^H[n - \tau] & \boldsymbol{x}^T[n - \tau] \end{bmatrix} \right\} \\ &= \begin{bmatrix} \mathbf{R}_{xx}[n, \tau] & \tilde{\mathbf{R}}_{xx}[n, \tau] \\ \tilde{\mathbf{R}}_{xx}^*[n, \tau] & \mathbf{R}_{xx}^*[n, \tau] \end{bmatrix} \qquad . \end{aligned} \qquad (5)$$

Notice that since $\underline{\mathbf{R}}_{xx}$ is positive semi-definite, and therefore has a positive determinant, the limit $|\mathbf{R}_{xx}|^2 \geq |\tilde{\mathbf{R}}_{xx}|^2$ sets an upper bound for the magnitude of $\tilde{\mathbf{R}}_{xx}$.

It is well known that wind speed and wind direction are likely non-stationary and non-linear, both can be volatile and, direction particularly, can 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. In the succeeding text, we ignore the potential non-linear nature of the

$$\mathbf{R}_{ee}[n] = E\left\{(\boldsymbol{x}[n] - \boldsymbol{W}_n^{\mathrm{H}}\underline{\mathbf{x}}_{n-\Delta})(\boldsymbol{x}^{\mathrm{H}}[n] - \underline{\mathbf{x}}_{n-\Delta}^{\mathrm{H}}\boldsymbol{W}_n)\right\} \quad,$$

$$= \mathbf{R}_{xx}[n,0] - \mathcal{E}\{\boldsymbol{x}[n]\underline{\mathbf{x}}_{n-\Delta}^{\mathrm{H}}\}\boldsymbol{W}_n - \boldsymbol{W}_n^{\mathrm{H}}\mathcal{E}\{\underline{\mathbf{x}}_{n-\Delta}\boldsymbol{x}^{\mathrm{H}}[n]\} + \boldsymbol{W}_n^{\mathrm{H}}\mathcal{E}\{\underline{\mathbf{x}}_{n-\Delta}\underline{\mathbf{x}}_{n-\Delta}^{\mathrm{H}}\}\boldsymbol{W}_n \quad,$$

$$= \mathbf{R}_{xx}[n,0] - \underline{\mathbf{R}}_{x\underline{\mathbf{x}}}[n]\boldsymbol{W}_n - \boldsymbol{W}_n^{\mathrm{H}}\underline{\mathbf{R}}_{x\underline{\mathbf{x}}}^{\mathrm{H}}[n] + \boldsymbol{W}_n^{\mathrm{H}}\underline{\mathbf{R}}_{\mathbf{xx}}[n]\boldsymbol{W}_n \quad, \tag{9}$$

where

$$\underline{\mathbf{R}}_{x\mathbf{x}}[n] = [\ \mathbf{R}_{xx}[n,\Delta]\ ,\ \mathbf{R}_{xx}[n,\Delta+1]\ ,\ \ldots\ ,\ \mathbf{R}_{xx}[n,\Delta+N-1]\ ,$$
$$\tilde{\mathbf{R}}_{xx}[n,\Delta]\ ,\ \tilde{\mathbf{R}}_{xx}[n,\Delta+1]\ ,\ \ldots\ ,\ \tilde{\mathbf{R}}_{xx}[n,\Delta+N-1]\ ]\quad, \tag{10}$$

$$\underline{\mathbf{R}}_{\mathbf{xx}}[n] = \begin{bmatrix} \mathbf{R}_{\mathbf{xx}}[n] & \tilde{\mathbf{R}}_{\mathbf{xx}}[n] \\ \tilde{\mathbf{R}}_{\mathbf{xx}}^*[n] & \mathbf{R}_{\mathbf{xx}}^*[n] \end{bmatrix}. \tag{11}$$

$$\mathbf{R}_{\mathbf{xx}}[n] = \begin{bmatrix} \mathbf{R}_{xx}[n-\Delta,0] & \mathbf{R}_{xx}[n-\Delta,1] & \cdots & \mathbf{R}_{xx}[n-\Delta,N-1] \\ \mathbf{R}_{xx}[n-\Delta-1,-1] & \mathbf{R}_{xx}[n-\Delta-1,0] & & \mathbf{R}_{xx}[n-\Delta-1,N-2] \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{xx}[n-\Delta-N+1,-N+1] & \mathbf{R}_{xx}[n-\Delta-N+1,-N+2] & \cdots & \mathbf{R}_{xx}[n-\Delta-N+1,0] \end{bmatrix}. \tag{12}$$

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 [9].

2.1. MMSE 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, plus the complex conjugates of the same. Therefore, the prediction error can be formulated as

$$\boldsymbol{e}[n] = \boldsymbol{x}[n] - \sum_{\nu=0}^{N-1}\left(\boldsymbol{P}^{\mathrm{H}}[n,\nu]\boldsymbol{x}[n-\Delta-\nu]\right.$$
$$\left. + \boldsymbol{Q}^{\mathrm{H}}[n,\nu]\boldsymbol{x}^*[n-\Delta-\nu]\right) \tag{6}$$

$$= \boldsymbol{x}[n] - \boldsymbol{W}_n^{\mathrm{H}}\underline{\mathbf{x}}_{n-\Delta} \quad, \tag{7}$$

with

$$\boldsymbol{W}_n = \begin{bmatrix} \boldsymbol{P}[n,0] \\ \boldsymbol{P}[n,1] \\ \vdots \\ \boldsymbol{P}[n,N-1] \\ \boldsymbol{Q}[n,0] \\ \boldsymbol{Q}[n,1] \\ \vdots \\ \boldsymbol{Q}[n,N-1] \end{bmatrix}, \quad \underline{\mathbf{x}}_n = \begin{bmatrix} \boldsymbol{x}[n] \\ \boldsymbol{x}[n-1] \\ \vdots \\ \boldsymbol{x}[n-N+1] \\ \boldsymbol{x}^*[n] \\ \boldsymbol{x}^*[n-1] \\ \vdots \\ \boldsymbol{x}^*[n-N+1] \end{bmatrix} \tag{8}$$

The matrices $\boldsymbol{P}[n,\nu], \boldsymbol{Q}[n,\nu] \in \mathbb{C}^{M\times M}$ describe the predictor's reliance on all spatial measurements and their conjugates, respectively, taken $\nu + \Delta$ samples in the past, at time instance n.

The error covariance matrix derived from (7), $\mathbf{R}_{ee}[n] = \mathcal{E}\{\boldsymbol{e}[n]\boldsymbol{e}^{\mathrm{H}}[n]\} \in \mathbb{C}^{M\times M}$, 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 an approximately stationary solution in Section 2.2. Calculating $\mathbf{R}_{ee}[n]$ using (7) yields a quadratic expression in $\boldsymbol{W}_n$, Equation (9).

We assume that $\boldsymbol{x}[n]$ is stationary over at least 2Δ samples. As a result, $\underline{\mathbf{R}}_{\mathbf{xx}}[n]$ is Hermitian and therefore positive semi-definite [10]. This property together with full rank of $\underline{\mathbf{R}}_{\mathbf{xx}}[n]$ admits a unique solution to minimises the mean square error,

$$\boldsymbol{W}_{n,\mathrm{opt}} = \arg\min_{\boldsymbol{W}_n}\mathrm{trace}\{\mathbf{R}_{ee}[n]\} \quad. \tag{13}$$

It can be shown that $\mathrm{trace}\{\mathbf{R}_{ee}[n]\}$ is quadratic in $\boldsymbol{W}_n$, such that the solution to (13) can be found by matrix- and complex-valued calculus [11]. Finding the minimum requires equating the gradient with respect to the unconjugated predictor coefficients in $\boldsymbol{W}_n^*$ to zero. We utilise results from [11] which show that for constant matrices $\mathbf{A}$ and $\mathbf{B}$ the expressions $\partial\mathrm{trace}\{\mathbf{A}\boldsymbol{W}_n^{\mathrm{H}}\mathbf{B}\}/(\partial\boldsymbol{W}_n^*) = \mathbf{B}\mathbf{A}$ 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 (9), yields

$$\frac{\partial}{\partial\boldsymbol{W}_n^*}\mathrm{trace}\{\mathbf{R}_{ee}[n]\} = -\underline{\mathbf{R}}_{x\mathbf{x}}^{\mathrm{H}}[n] + \underline{\mathbf{R}}_{\mathbf{xx}}[n]\boldsymbol{W}_n \quad. \tag{14}$$

Finally, setting the gradient on the right-hand side of (14) equal to zero yields the optimum predictor coefficients that minimise $\mathrm{trace}\{\mathbf{R}_{ee}[n]\}$,

$$\boldsymbol{W}_{n,\mathrm{opt}} = \underline{\mathbf{R}}_{\mathbf{xx}}^{-1}[n]\underline{\mathbf{R}}_{x\mathbf{x}}^{\mathrm{H}}[n] \quad, \tag{15}$$

which is the well-known Wiener-Hopf solution [12, 13].

If the process $\boldsymbol{x}[n]$ is uncorrelated with its conjugate, i.e. $\tilde{\mathbf{R}}_{xx} = \mathbf{0}$, all the matrices $\boldsymbol{Q}[n,\nu] = \mathbf{0}$ and the prediction problem reduces to the $\mathbb{C}$-linear case.

2.2. Cyclo-Stationary Covariance Matrix

The cyclo-stationary covariance matrix (and its associated complementary covariance matrix) is formulated 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

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

$$+ \frac{2}{L} \sum_{\nu=1}^{\frac{L}{2}} \boldsymbol{x}[n-\nu]\boldsymbol{x}^{\mathrm{H}}[n-\nu-\tau] \quad , \quad (16)$$

and the complementary covariance matrix is calculated in the same way but with the Hermitian transpositions replaced by standard transpositions. The widely linear optimal prediction filter for time n can then be calculated by replacing the quantities in the Wiener solution (15) by their estimates derived from (16) inserted into (10)–(12).

3. TESTING AND RESULTS

3.1. 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. The measurements are taken in open terrain at a height of 10m and sampled at hourly intervals, comprise hourly averages that are quantised to a $10°$ angular granularity and integer multiples of one knot ($0.515\mathrm{ms}^{-1}$) [14].

Widely linear processing is advantageous for improper signals, or cross-improper in the multichannel case, i.e. if $\tilde{\mathbf{R}}_{xx} \neq \mathbf{0}$. The statistical hypothesis test for the impropriety of complex vectors described in [15] has been applied to the test data. The test unambiguously rejected the hypothesis $H_0 : \tilde{\mathbf{R}}_{xx} = \mathbf{0}$ in favour of $H_1 : \tilde{\mathbf{R}}_{xx} \neq \mathbf{0}$ indicating that the data is improper and therefore that widely linear processing is appropriate.

3.2. Cyclo-stationary Estimation

In the estimation of the cyclo-stationary covariance matrix, (16), $K = 5$ to make use of all available training data and the optimal window length L is chosen heuristically to be 15 weeks. The filter length is chosen to be $2N = 6$ since the gains from increasing it further are negligible.

3.3. Results

The widely linear predictor yields improved prediction performance in terms of root mean squared error for all 13 channels and at all look-ahead times. As one would expect the new predictor yields greater improvement over its $\mathbb{C}$-linear equivalent at sites with larger complementary correlation and lower directional variance.
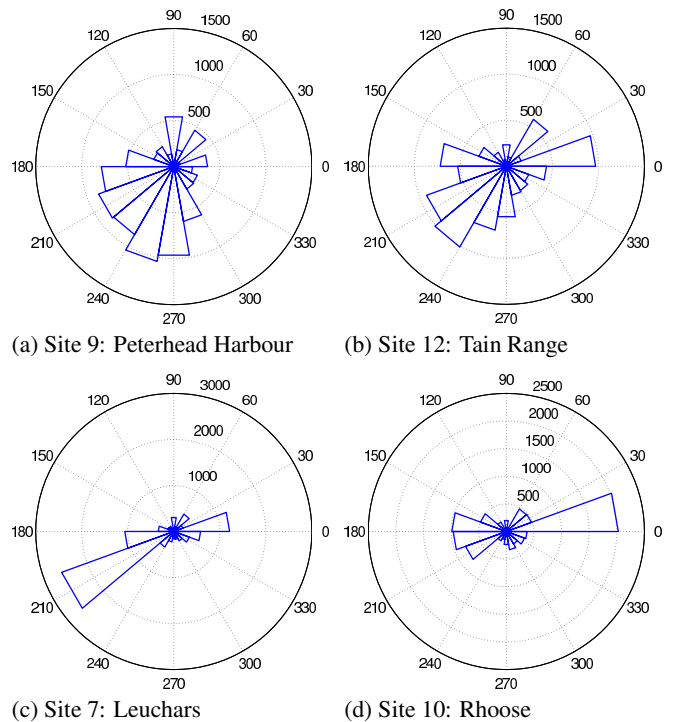

(a) Site 9: Peterhead Harbour   (b) Site 12: Tain Range

(c) Site 7: Leuchars   (d) Site 10: Rhoose

Fig. 1: Circular histograms of hourly-mean wind direction measurements at 4 selected sites from the 1 year of data used for testing.

Results from two channels that showed the least improvement (9 & 12) and the two that showed the most (7 & 10) are detailed in Table 1. The distribution of the arguments of these four channels are illustrated by the histograms in Figure 1. The sites in Figures 12a & 12b show little improvement and have arguments, or wind directions, spread evenly over a wide range of angles, whereas the sites in Figures 12c & 12d demonstrate large improvement and have very narrow distributions, corresponding to low directional variance and high complementary correlation.

The complementary autocorrelation coefficients for all 13 channels are plotted in Figure 2. The two channels showing the least improvement over the $\mathbb{C}$-linear predictor are not simply the two with the smallest complementary autocorrelation since the improvement is due to both complementary auto- and cross-correlation.

4. CONCLUSION

In this paper a multichannel widely linear cyclo-stationary Wiener filter for the prediction of hourly mean wind speed and direction from 1 to 6 hours ahead has been derived and tested. The performance of the proposed filter is compared to that of its $\mathbb{C}$-linear equivalent to quantify the benefits of increasing computational complexity to accommodate the widely linear model.

The widely linear model captures information contained in the complementary auto- and cross-covariance which is inaccessible in a strictly linear formulation. In addition, the cyclo-stationary estimation of the covariance matrices captures the seasonal behaviour of the wind which would otherwise lead to the inclusion of mismatched data in the estimation of the covariance matrices.

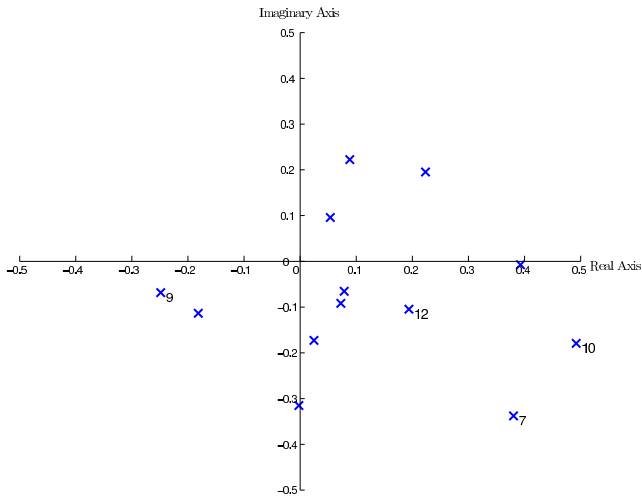The predictors are tested on wind measurements made at

Fig. 2: Scatter plot of complementary auto-correlation coefficients at zero lag for the 13 measurement locations. The four examples used in Fig. 1 and Table 1 are labelled by site number.

13 locations distributed geographically around the UK over a period of 6 years. The widely linear predictor shows improved prediction versus its $\mathbb{C}$-linear equivalent at all 13 locations. The locations which exhibit greater improvement are those with the least directional variation and associated high complementary auto-correlation.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, *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.

[2] X. Zhu and M.G. Genton, "Short-term wind speed forecasting for power system operations," *International Statistical Review*, vol. 80, no. 1, pp. 2–23, 2012.

[3] A.S. Hering and M.G. Genton, "Powering up with space-time wind forecasting," *Journal of American Statistical Association*, vol. 105, pp. 92–104, 2010.

[4] M. Khalid and A.V. Savkin, "A method for short-term wind power prediction with multiple observation points," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 579–586, May 2012.

[5] J. Dowell, S. Weiss, D. Hill, and D. Infield, "Short-term spatio-temporal prediction of wind speed and direction," *Wind Energy*, 2014.

Table 1: Root Mean Squared Errors (RMSE) for the cyclostationary (CSWF) and widely linear (WLCSWF) Wiener filters at look-ahead times (Δ) from 1–6 hours.

Site	Δ	RMSE CSWF	WLCSWF	% Improvement
9: Peterhead Harbour	1	1.73	1.72	0.5
	2	2.35	2.33	0.8
	3	2.77	2.74	1.0
	4	3.11	3.08	1.0
	5	3.40	3.37	1.1
	6	3.66	3.62	1.0
12: Tain Range	1	1.97	1.96	0.4
	2	2.47	2.45	0.6
	3	2.78	2.75	0.8
	4	3.03	3.00	1.1
	5	3.23	3.19	1.2
	6	3.41	3.36	1.3
7: Leuchars	1	1.52	1.50	1.1
	2	2.07	2.02	2.0
	3	2.43	2.36	2.9
	4	2.72	2.63	3.4
	5	2.98	2.87	3.7
	6	3.20	3.07	3.9
10: Rhoose	1	1.67	1.66	0.4
	2	2.18	2.15	1.3
	3	2.59	2.54	1.8
	4	2.98	2.91	2.3
	5	3.31	3.21	2.7
	6	3.59	3.48	3.0

[6] S.L. Goh, M. Chen, D.H. Popović, K. Aihara, D. Obradovic, and D.P. Mandic, "Complex-valued forecasting of wind profile," *Renewable Energy*, vol. 31, no. 11, pp. 1733–1750, 2006.

[7] D. Mandic and S.L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Modes*, John Wiley & Sons, 2009.

[8] P.J. Schreier and L.L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge University Press, 2010.

[9] A. Ispas, M. Dörpinghaus, G. Ascheid, and T. Zemen, "Characterization of non-stationary channels using mismatched Wiener filtering," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 274–288, 2013.

[10] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, third edition edition, 1996.

[11] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, part I, pp. 2740–2746, June 2007.

[12] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, New York, 1985.

[13] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 4th edition, 2002.

[14] UK Meteorological Office, "MIDAS Land Surface Stations data (1853–current), NCAS British Atmospheric Data Centre, accessed December 2008," .

[15] A.T. Walden and P. Rubin-Delanchy, "On testing for impropriety of complex-valued Gaussian vectors," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 825–834, March 2009.