Available online at www.sciencedirect.com
SCIENCE DIRECT®
Renewable Energy 31 (2006) 1733–1750
RENEWABLE ENERGY
www.elsevier.com/locate/renene

Complex-valued forecasting of wind profile

S.L. Goh[a,*], M. Chen[a], D.H. Popović[a], K. Aihara[b], D. Obradovic[c], D.P. Mandic[a]

[a]*Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, UK*
[b]*Department of Complexity Science and Engineering, Graduate School of Frontier Sciences, University of Tokyo, Tokyo 113-8656, Japan*
[c]*Siemens AG, Corporate Technology IC 4, Otto-Hahn-Ring 6, D-81739 Munich, Germany*

Received 3 February 2005; accepted 27 July 2005
Available online 29 September 2005

Abstract

This paper presents a novel approach for the simultaneous modelling and forecasting of wind signal components. This is achieved in the complex domain by using novel neural network algorithms and architectures. We first perform a signal nonlinearity and component-dependent analyses, which suggest the use of modular complex-valued recurrent neural networks (RNNs). This RNN-based modelling rests upon a combination of nonlinearity, complexity and internal memory and allows for the multiple step ahead forecasting of the wind signal in its complex form (speed and direction). The approach is first verified on benchmark Data Set A (NH_3 laser data) of the Santa Fe Time Series Prediction Competition together with artificial data generated by chaotic Mackey–Glass equations, and then applied to the real-world wind measurements. Simulations support the proposed architecture and algorithms.

Keywords: Wind forecasting; Complex-valued representation; Recurrent neural networks; Pipelined recurrent neural network; Delay vector variance

*Corresponding author.
E-mail addresses: su.goh@imperial.ac.uk (S.L. Goh), mo.chen@imperial.ac.uk (M. Chen), dp@imperial.ac.uk (D.H. Popović), aihara@sat.t.u-tokyo.ac.jp (K. Aihara), Dragan.Obradovic@siemens.com (D. Obradovic), d.mandic@imperial.ac.uk (D.P. Mandic).

0960-1481/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.renene.2005.07.006

1. Introduction

There are many potential applications which require some kind of wind forecasting, such as those in the power system operation, air pollution modelling and aviation safety [1]. Wind power is envisaged as becoming one of the dominant sources of energy as soon as the year 2020. Accurate estimation of the wind turbines (WT) power output is required in order to incorporate wind-generated electric power into the grid [2,3]. Nowadays, it is widely accepted that WT power forecasts should be based on the actual wind signal forecasts rather than on the output power of the WT [4], which highlights the need for precise and reliable modelling of not only the actual wind data, but also the underlying wind dynamics.

The power generated by WTs is difficult to forecast, due to the continuous fluctuation of both the wind speed and direction. Various field measurements have shown that the direction of wind as compared with wind speed has less influence on WT power output because each turbine is usually built to face into the wind when operating. Consequently, and especially at stronger winds, there is no significant difference in the power generated for different wind directions. However, the impact of wind direction on output power is more prominent at milder winds since they usually come from much wider directions [4]. The importance of wind direction is of further significance in spatial correlation studies which aim to assess the influence of WT position in a wind park [4]. Studies on multivariate wind forecasts do not simultaneously model all of the wind parameters. They use several relevant variables as model inputs in order to predict a single output variable. While wind speed and direction are shown to influence turbine power simultaneously, they are separate forecasts and introduce an error in both the wind dynamics and wind power forecasts. All this emphasizes the need to process wind signal as a vector field defined by wind speed and direction, amongst other factors. This work therefore focuses on the development of a new methodology and design of a neural network-based forecasting system to be used for the estimation of WT power output. This is achieved based on a complex-valued vector-field wind signal representation rather than on wind speed and wind direction being modelled separately.

Recent results have shown that neural networks are powerful tools for forecasting real-world data [5–7]. The forecasting methods based on a neural network approach have been shown to be most promising in terms of forecasting accuracy and efficient computation, due to their function approximation and generalization ability in a non-parametric fashion [8–13]. In particular, recurrent neural networks (RNNs) possess rich internal nonlinear dynamics, which make them capable of modelling a wider class of dynamical processes (such as those with temporal dependencies within the signal) than their feedforward counterparts [5,14,15].[1] Fully connected recurrent neural networks (FCRNNs) possess both short- and long-term memory (due to their feedback) and exhibit attractor dynamics shown to be particularly suitable for forecasting of nonlinear and non-stationary signals [16]. For real-time applications, the Real-Time Recurrent Learning (RTRL) algorithm [17] has been widely used to train FCRNNs.

We set out to investigate the possibility of forecasting the wind speed and direction simultaneously by making use of a complex-valued representation of the wind signal. A

[1]Nonlinear autoregressive (NAR) processes can be modelled using feedforward networks, whereas nonlinear autoregressive moving average (NARMA) processes can be represented using RNNs

recently proposed complex-valued real-time recurrent learning (CRTRL) algorithm for RNNs [18] has shown the possibility of using nonlinear feedback architectures for the forecasting of complex-valued wind field. Initial results on using single complex RNNs for wind forecasting are given in [19].

Here, we first provide the theoretical justification for the use of complex-valued representation of wind. This is achieved within the framework of surrogate data testing and component correlation. The proposed RNN architecture was a modular complex-valued pipelined recurrent neural network (CPRNN), which accounts for both temporal and spatial correlations within a signal. Simulations on both the benchmark and real-world complex wind measurements support the analysis.

The paper is organized in the following manner: in Sections 2 and 3, we present theoretical background on wind power and wind signal characteristics. In Section 4, an analysis of predictability is introduced. In Section 5, the network architecture and prediction configuration are given. This is followed by comprehensive simulations in Section 6. The paper concludes in Section 7.

2. Wind and wind power characteristics

The power generated by a WT is nonlinearly dependent on the wind speed. More specifically, the power which can be extracted from the airflow is given by [1]

$$P_w = \frac{\rho}{2} C_{\mathrm{p}}(\alpha, \theta) A\, w^3 (\mathrm{W}), \tag{1}$$

where ρ is the air density (kg/m^3), C_{p} is the performance or power coefficient, α is tip speed ratio, that is, the ratio between the blade tip speed w_{t} and the wind speed upstream the rotor w (m/s), θ is the blade pitch angle and A is the area swept by the rotor (m^2). The relation between wind speed and generated power is usually given by the manufacturer in the power curve of the WT. A power curve, however, is derived under a set of assumptions regarding the wind speed and air density. These so-called 'ideal data' are often impractical in estimating the actual power output of each WT due to the distances and relative positions of turbines in the wind park with respect to the meteorological tower(s). Fig. 1a shows the wind rose plot of the direction versus magnitude (speed) of the wind at a site in Iowa (USA) over 5000 min. As a result, the estimation of actual WT power should be based on processing the wind-vector (speed and direction) measurements (as two strongly correlated inputs which define a wind signal (see Fig. 1b)). Since wind signal has a spatial and temporal dimension, it can be treated as a time series. Time series prediction problems are traditionally approached from a stochastic modelling perspective or more recently from a nonparametric neural network perspective [6]. Either approach has advantages and disadvantages: stochastic methods are usually fast, but limited in their applicability since they are parametric and based on linear models [6,20,21]. Since wind signal is highly nonlinear and non-stationary, the identification of parameters and contributing factors to describe the power supplied by this non-controllable and intermittent source is not trivial. Thus, classical parametric techniques such as time-series (AR, ARMA) methods used widely in short- and long-term forecasting have limitations in dealing with the nonlinear and nonstationary nature of wind signals and are prone to numerical instability and inaccuracy. Neural network methods, on the other hand, are powerful enough, but the

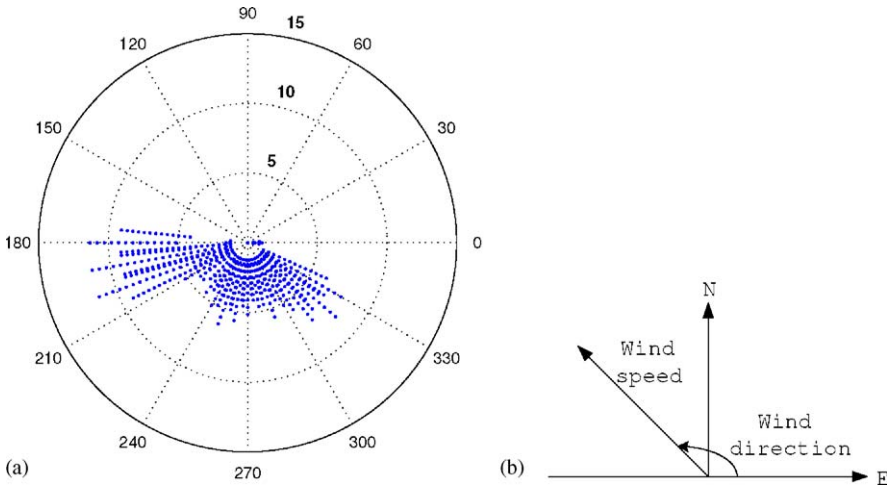

Fig. 1. Wind as a complex (speed,direction) quantity, $v = e^{j\theta} = v_N + jv_E$. (a) Wind rose representation, (b) wind vector representation.

Table 1
Statistical properties of the wind data sets

Set	1 h	3 h	6 h
Cumulative samples	1200	1200	1200
Minimum speed (m/s)	0	0	0
Maximum speed (m/s)	13.0582	12.2865	11.4663
Mean speed (m/s)	3.2905	3.2905	3.2905
Standard deviation (m/s)	2.3387	2.2653	2.1520

selection of an appropriate architecture and its parameters is usually a time-consuming trial and error procedure [21,22].

3. Analysis of wind characteristics

The data used in the simulations were obtained from the Iowa (USA) Department of Transport.[2] The data were sampled at 1-, 3- and 6-h intervals. Table 1 shows the statistical properties of the wind data sets considered. To provide insight into the variation of the complex-valued wind signal, three tests were performed. Firstly, the autocovariance of the wind data with lags up to 25 was calculated. The autocorrelation coefficients for average wind data over 1-, 3- and 6-h intervals are shown, respectively, in Figs. 2a, 3a and 4a. A correlogram can be used to achieve a general understanding of the behaviour of the wind time series with respect to averaging. If the autocorrelation coefficients decay slowly, this indicates short-term correlation within the data. For a time series that contains a trend, the

[2]Real-life wind measurements are publicly available from "http://mesonet.agron.iastate.edu/request/awos/1min.php".

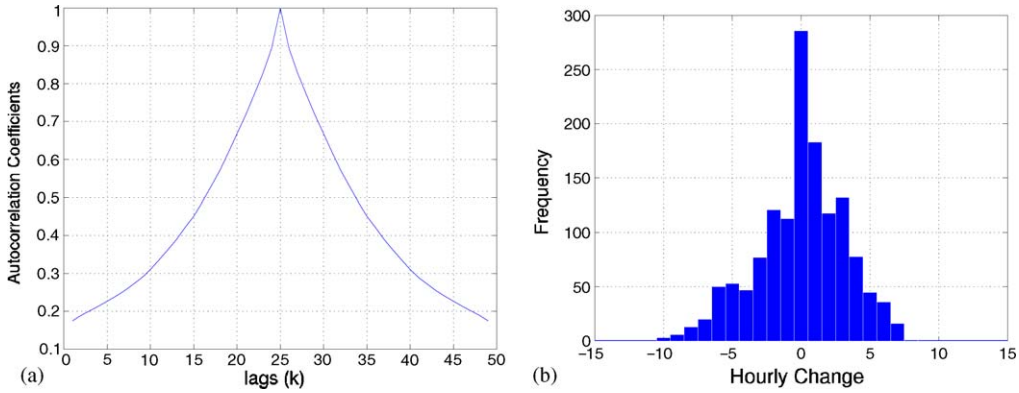

Fig. 2. Wind signal sampled at a 1 h interval. (a) Autocorrelation coefficients, (b) histogram of wind data points.

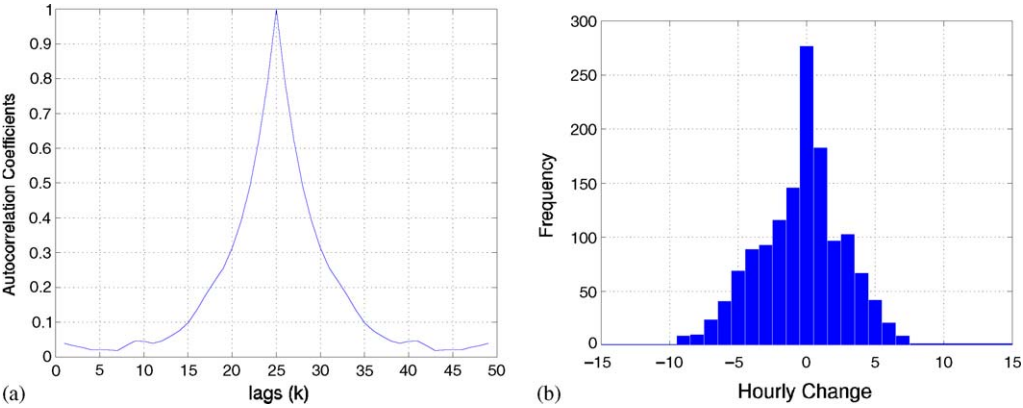

Fig. 3. Wind signal sampled at 3 h interval. (a) Autocorrelation coefficients, (b) histogram of wind data points.

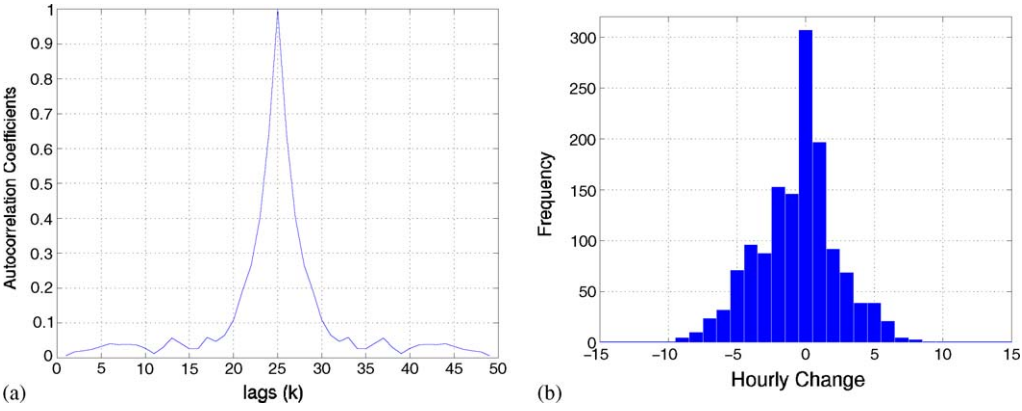

Fig. 4. Wind signal sampled at 6 h interval. (a) Autocorrelation coefficients, (b) histogram of wind data points.

values of the autocorrelation coefficients will not approach zero except for very large values of the lag [23]. From Figs. 2–4, the short-term correlation in the time series decreases as the sample average of the wind signal increases.

In the second test, the amplitude distributions of wind were plotted in the form of a histogram. This is shown in Figs. 2b, 3b and 4b. The distribution of the histogram is usually represented by the Weibull distribution [24]. The amplitude distribution curve becomes slightly narrower with the increase in the sample interval.

3.1. Insight into complex-valued nature of the (speed, direction) representation of wind data

Some recent work showed that wind signal components are only locally predictable and correlated [25], which is a strong indication that wind measurements could be treated as a complex-valued compact signal rather than two separate univariate variables [26].

To verify this, we perform an experiment in which one step ahead predictions were undertaken using three approaches:

(1) wind speed and direction are univariate[3] signals and should be modelled independently;
(2) wind speed and direction can be modelled by a complex number, but the representation is split[4] complex;
(3) 'fully' complex representation.

The used neural network configuration was the FCRNN architecture trained by the CRTRL algorithm. The simulation results on the prediction performance applied to the complex-valued real-world (speed and direction components) wind signal for the univariate, split and 'fully' complex case are shown on Figs. 5–7. Observe that the 'fully' CRTRL algorithm was more stable and has exhibited better and more consistent performance than the split and univariate approaches as indicated by the solid line (predicted signal) being much more in accordance with the dotted line (actual signal). Local predictability and associated short-tailed component correlations imply that the complex representation of wind is likely to have advantage over the real-valued univariate ones. This empirical result will be further theoretically justified in Section 4.

4. Predictability analysis

There are no general guidelines in the literature as how to select the parameters of a neural network predictor architecture. This leads to time consuming trial-and-error procedures. Some recent results from signal modality analysis [27], which is based on predictability of a signal, provide a method to determine the embedding parameters of a signal. This method is based on surrogate data representation and differential entropy, which allows for a convenient graphical representation [22,28]. The Takens' theorem [29]

[3]Here, we mean that each component (speed and direction) is being predicted individually and then put back together to form a complex vector.

[4]In a split-complex AF, the real and imaginary component of the complex-valued; input signal x are split and fed through the real-valued activation function $f_{\mathrm{R}}(x) = f_{\mathrm{I}}(x), x \in \mathbb{R}$. The functional expression of the split-complex activation function is given by $f(x) = f_{\mathrm{R}}(\mathtt{Re}(x)) + \jmath f_{\mathrm{I}}(\mathtt{Im}(x))$. We can see that this approach does not account for a 'fully' complex signal where the signal components are not statistically independent.

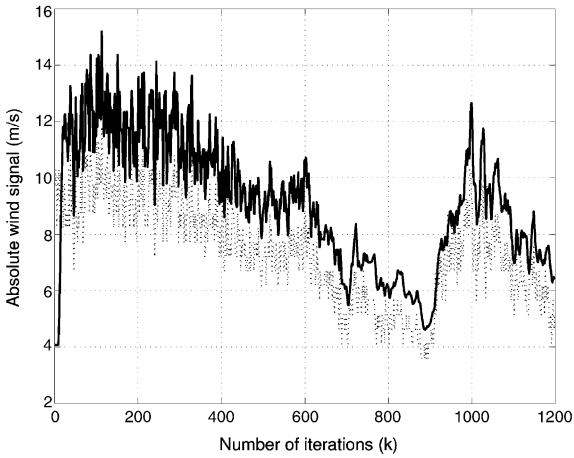

Fig. 5. Prediction of univariate wind signal using two real-valued RTRL algorithms (FCRNN). Solid curve: nonlinear prediction of wind signal. Dashed curve: actual wind signal.

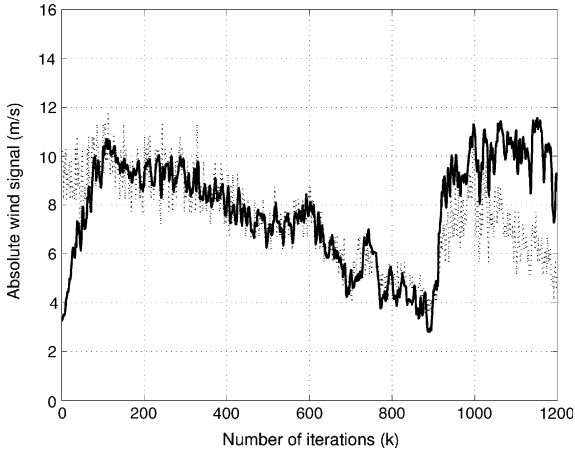

Fig. 6. Prediction of complex wind signal using CRTRL algorithm employing split activation function (FCRNN). Solid curve: nonlinear prediction of wind signal. Dashed curve: actual wind signal.

implies that for a wide class of deterministic systems, there exists a multiple-input single-output mapping (delay coordinate embedding) so that

$$x(k) = \Phi(x(k-1), x(k-2), \ldots, x(k-f)), \tag{2}$$

where Φ is any nonlinear function and f is the number of past values taken into consideration. The process of representing a system by one variable and its lagged versions is called embedding [22]. According to Takens' embedding theorem [30], the delay coordinate map from a τ-dimensional compact manifold to an m-dimensional Euclidean space is embedding provided that

$$m \geqslant 2\tau + 1. \tag{3}$$

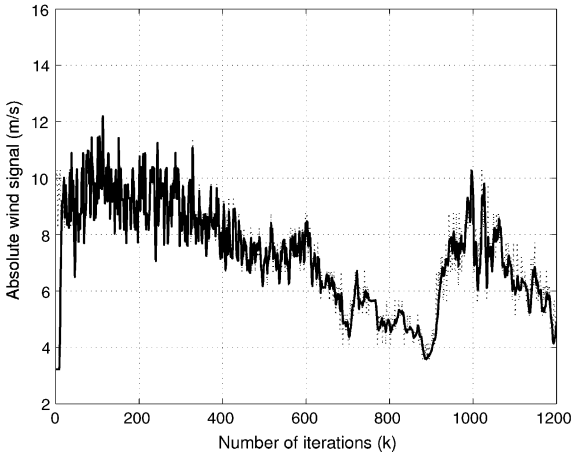

Fig. 7. Prediction of complex wind signal using 'fully' CRTRL algorithm (FCRNN). Solid curve: nonlinear prediction of wind signal. Dashed curve: actual wind signal.

Moreover, Sauer et al. [30] showed an extension of the theorem with the following condition

$$m > 2\,\mathrm{boxdim}(A), \tag{4}$$

where boxdim(A) is the box-counting dimension[5] of A, and A is a compact subset of the n-dimensional Euclidean space. This underpins the theory of low dimensional chaos, which has found many applications in engineering [31].

The choice of parameters of neural network and the phase-space characteristics of a time series are strongly related. For instance, calculating the embedding parameters of a series using the method by Gautama et al. [27] gives us the estimated number of degrees of freedom in the system. These results, shown in Fig. 8, are inconclusive, but suggest the use of a robust feedback architecture for forecasting. The resulting embedding parameters obtained are shown in Table 2.

To support the complex-valued representation, following the approach from [26], we perform a component dependence test for the complex-valued wind representation. The test is based on the complex-valued surrogate data analysis (see Fig. 9). From the figure, there is a significant component dependence within the complex-valued wind signal presentation as indicated by the rejection ratio of the null hypothesis of fully complex wind signal nature being significantly greater than zero.[6] Overall, there are stronger indications of a complex-valued nature when the wind is averaged at a 1 h interval than when it is averaged at a 6 h interval as represented by the respective percentage values of the rejection ratio. Complex-valued wind signal averaged at a longer interval will become more univariate and linear, and as indicated by the test, any linear complex-valued system has a

[5]Given that $N(\varepsilon)$ is the number of boxes of side length ε required to cover the fractal dimension of a set S in a Euclidean space $\mathbb{R}^n$, then the box-counting dimension is defined as $\dim_{\mathrm{box}}(S) = \lim_{\varepsilon\to 0} \log N(\varepsilon)/\log(1/\varepsilon)$.

[6]We adopt the statistical testing methodology proposed by Gautama et al. [28,32], where the null hypothesis is the original signal is complex-valued and signals are characterised by delay vector variance (DVV) method. Then we applied the proposed test on three wind signals with different intervals, each for 100 times and compute the number of times when null hypothesis is rejected out of 100.

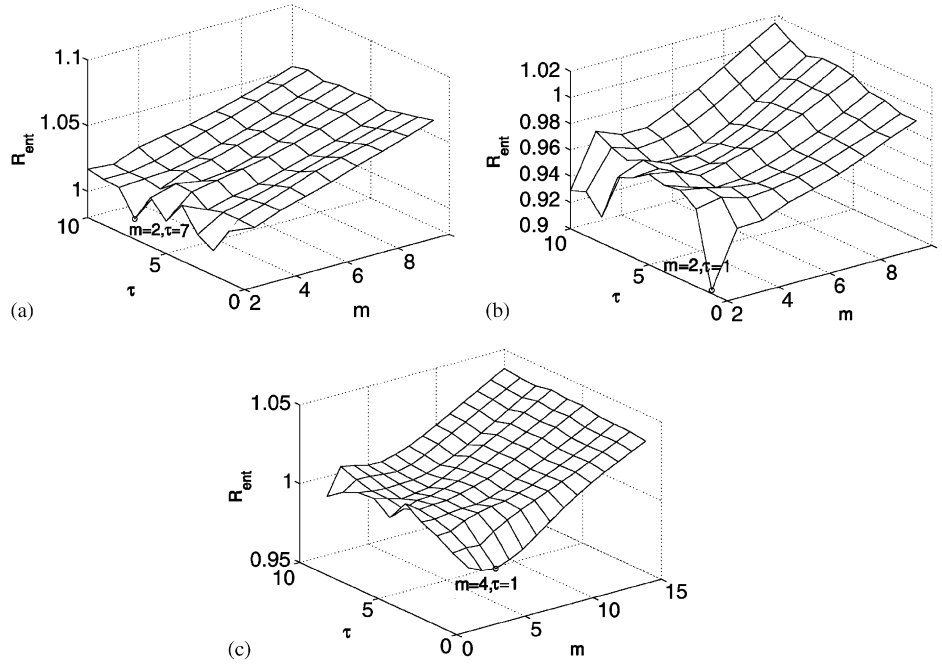

Fig. 8. Plots of the entropy ratio for: (a) wind speed, (b) wind direction and (c) complex-valued wind. The minima of the entropy ratio plots are indicated by open circles.

Table 2
Optimal embedding parameters using entropy ratio method

Wind component	m	τ
Wind speed	2	7
Wind direction	2	1
Complex-valued wind	4	1

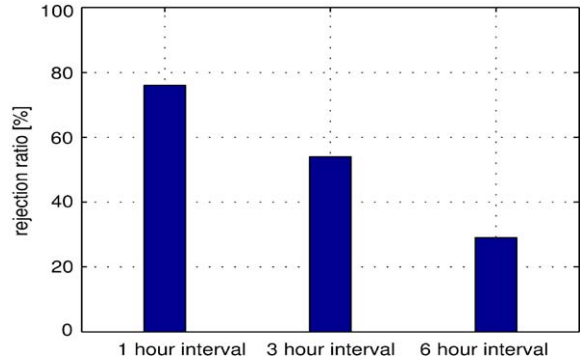

Fig. 9. The complex-valued surrogate data test for the complex nature of wind signal (speed and direction).

bivariate equivalent, though not vice versa. This provides an additional theoretical justification for the use of complex-valued nonlinear neural networks with feedback, CRNNs [7].

5. Network architecture and prediction configuration

5.1. Forecasting configuration

Forecasting of more than one step ahead can be achieved either in a direct or a recursive manner. The approach used in this paper is the recursive method [33]. In the recursive approach, once the network is trained, it predicts all the intermediate values up to T steps ahead by using the previously predicted values as inputs when predicting the next value [20,6].

5.2. The complex RNN

Fig. 10 shows an FCRNN, which consists of N neurons with p external inputs and N feedback connections. The network has two distinct layers, namely the external

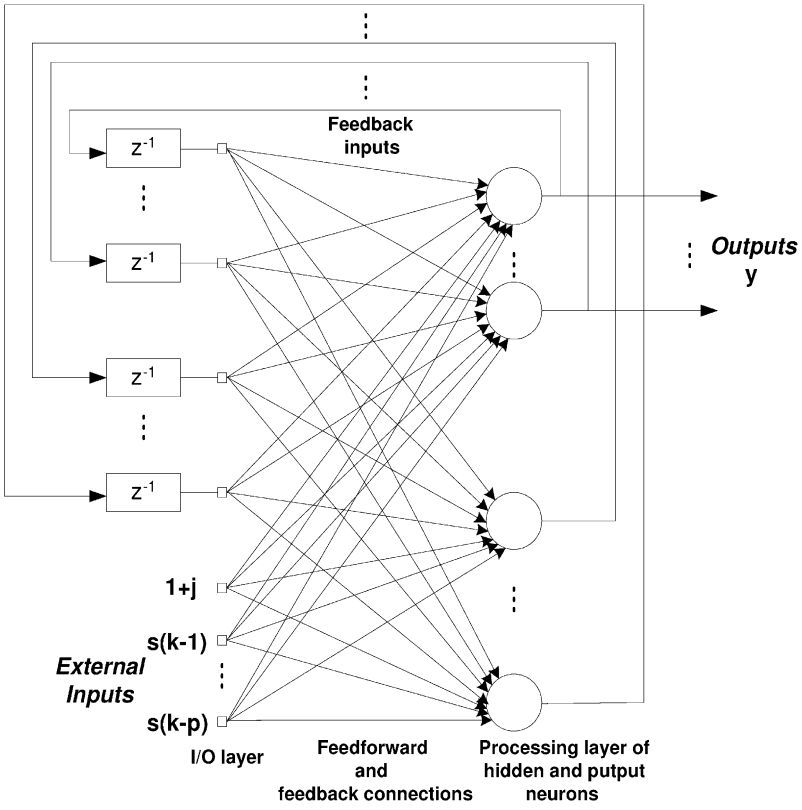

Fig. 10. A fully connected recurrent neural network for prediction.

input-feedback layer and a layer of processing elements. Let $y_l(k)$ denote the complex-valued output of a neuron, $l = 1, \ldots, N$ at time index k and $\mathbf{s}(k)$ the $(1 \times p)$ external complex-valued input vector. The overall input to the network $\mathbf{P}(k)$ then represents a concatenation of vectors $\mathbf{y}(k)$, $\mathbf{s}(k)$ and the bias input $(1 + \mathrm{j})$, and is given by

$$\begin{aligned} &\mathbf{P}(k) = [s(k-1), \ldots, s(k-p), 1+\mathrm{j}, y_1(k-1), \ldots, y_N(k-1)]^{\mathrm{T}}, \\ &P_n(k) \in \mathbf{P}(k) = P_n^{\mathrm{r}}(k) + \mathrm{j}P_n^{\mathrm{i}}(k), \quad n = 1, \ldots, p+N+1, \end{aligned} \tag{5}$$

where $\mathrm{j} = \sqrt{-1}$, $(\cdot)^{\mathrm{T}}$ denotes the vector transpose operator, and superscripts $(\cdot)^{\mathrm{r}}$ and $(\cdot)^{\mathrm{i}}$ denote, respectively, the real and imaginary parts of a complex number, or complex vector.

For the lth neuron, its weights form a $(p + N + 1) \times 1$-dimensional weight vector $\mathbf{w}_l^{\mathrm{T}} = [w_{l,1}, \ldots, w_{l,p+N+1}]$, $l = 1, \ldots, N$, which are encompassed in the complex-valued weight matrix of the network $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_N]$.

The output of every neuron can be expressed as

$$y_l(k) = \Phi(\mathrm{net}_l(k)), \quad l = 1, \ldots, N, \tag{6}$$

where Φ is a complex nonlinear activation function of a neuron and

$$\mathrm{net}_l(k) = \sum_{n=1}^{p+N+1} w_{l,n}(k) P_n(k) \tag{7}$$

is the net input to lth node at time index k.

5.3. The complex-valued pipelined recurrent neural network (CPRNN)

The nonlinear adaptive filtering architecture as proposed by Haykin and Li [16] consists of two sections, namely the nonlinear and linear ones. The nonlinear section, called the pipelined recurrent neural network (PRNN), is essentially a modular RNN, and performs nonlinear filtering, while the linear section represented by an FIR filter performs linear filtering of the signal. This cascaded combination of the PRNN and an FIR filter has been shown to be suitable for nonlinear prediction of real-valued non-stationary signals. The CPRNN architecture contains M modules of FCRNNs connected in a nested manner as shown in Fig. 11. The $(p \times 1)$-dimensional external complex-valued signal vector $\mathbf{s}^{\mathrm{T}}(k) = [s(k-1), \ldots, s(k-p)]$ is delayed by m time steps ($z^{-m}\mathbf{I}$) before feeding the module m, where z^{-m}, $m = 1, \ldots, M$ denotes the m-step time delay operator, and $\mathbf{I}$ is the $(p \times p)$-dimensional identity matrix. The complex-valued weight vectors $\mathbf{w}_l$ are embodied in an $(p + N + 1) \times N$-dimensional weight matrix $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_N]$. All the modules operate using the same weight matrix $\mathbf{W}$ (a full mathematical description of the PRNN is given [16,34]). The following equations provide a mathematical description of the CPRNN:

$$y_{t,l}(k) = \Phi(\mathrm{net}_{t,l}(k)), \quad t = 1, 2, \ldots, M, \tag{8}$$

$$\mathrm{net}_{t,l}(k) = \sum_{n=1}^{p+N+1} w_{l,n}(k) P_{t,n}(k), \quad \begin{matrix} l = 1, \ldots, N, \\ n = 1, \ldots, p+N+1, \end{matrix} \tag{9}$$

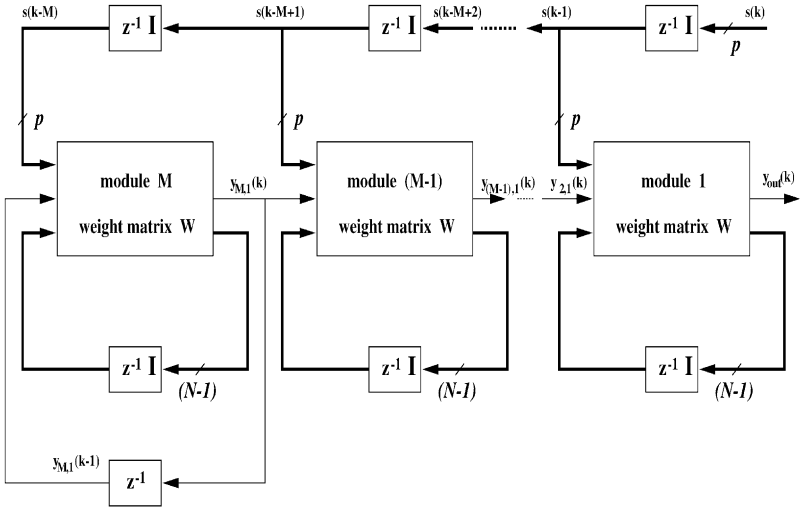

Fig. 11. Pipelined recurrent neural network (PRNN).

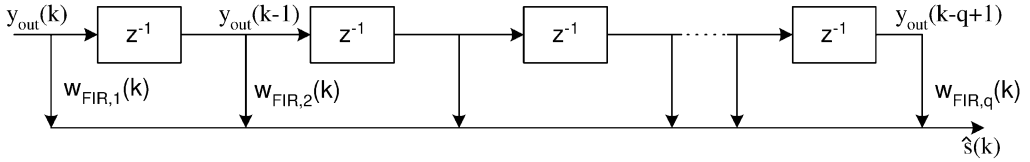

Fig. 12. An FIR filter.

$$\mathbf{P}_t^{\mathrm{T}}(k) = [s(k-t), \ldots, s(k-t-p+1), 1+\mathrm{j}, \\ y_{t+1,1}(k-1), y_{t,2}(k-1), \ldots, y_{t,N}(k-1)] \\ \text{for} \quad 1 \leqslant t \leqslant M-1, \tag{10}$$

$$\mathbf{P}_M^{\mathrm{T}}(k) = [s(k-M), \ldots, s(k-M-p+1), 1+\mathrm{j}, \\ y_{M,1}(k-1), y_{M,2}(k-1), \ldots, y_{M,N}(k-1)] \\ \text{for} \quad t = M. \tag{11}$$

5.4. Linear subsection

The linear subsection of the CPRNN consists of an FIR filter, shown in Fig. 12. The complex-valued least mean squares (CLMS) algorithm is used to update the tap weights of this filter, for which the output is given by

$$\hat{s}(k) = \mathbf{w}_{\mathrm{FIR}}^{\mathrm{T}}(k)\mathbf{y}_{\mathrm{out}}(k), \tag{12}$$

where $\mathbf{y}_{\mathrm{out}}(k) \triangleq [y_{\mathrm{out},1}(k), \ldots, y_{\mathrm{out},q}(k)]^{\mathrm{T}}$ is the output from the first CPRNN module ($y_{1,1}(k)$), $\mathbf{w}_{\mathrm{FIR}}(k) \triangleq [w_{\mathrm{FIR},1}(k), \ldots, w_{\mathrm{FIR},q}(k)]^{\mathrm{T}}$ the complex weight vector and q the number of tap inputs.

Table 3
Parameter selection details

Time series	M	N	p	q	η	μ
NH_3	8	2	35	8	0.01	0.1
Mackey–Glass	8	2	50	25	0.01	0.1
W-1 (1 h ave.)	5	3	6	5	0.01	0.1
W-3 (3 h ave.)	5	3	6	5	0.01	0.1
W-6 (6 h ave.)	5	3	6	5	0.01	0.1

6. Simulation results

The presented methodology was firstly applied to two different time series which were made artificially complex-valued by using the real part as the imaginary part: on choatic laser series NH_3[7] and nonlinear deterministic Mackey–Glass (MG) series.[8] Next, average wind data for 1-, 3- and 6-h intervals denoted, respectively, by W-1, W-3 and W-6 were used as inputs to the network.

Selection of network dimension and parameters: For training purposes, all the data were scaled to zero mean and unit variance. Initial weight values of the network were chosen randomly. The network predictor was trained with 1200 data points from the complex wind measurements. For the experiments, the nonlinearity at the neuron was chosen to be the complex tanh function

$$\Phi(x) = \frac{e^{\beta x} - e^{-\beta x}}{e^{\beta x} + e^{-\beta x}}, \tag{13}$$

where $x \in \mathbb{C}$. The value of the slope of $\Phi(x)$ was $\beta = 1$. The value of the learning rate for the CPRNN architecture was $\eta = 0.01$, while the learning rate for the FIR filter was $\mu = 0.1$. The forgetting factor for the CPRNN architecture was $\lambda = 0.995$. Table 3 shows the parameters selected for each time series.

Performance measure: The measurement used to assess the performance was the prediction gain R_p given by [16]

$$R_p(k) \triangleq 10 \log_{10}\left(\frac{\sigma_x^2}{\hat{\sigma}_e^2}\right)(\text{dB}), \tag{14}$$

where σ_x^2 denotes the variance of the input signal $x(k)$, whereas $\hat{\sigma}_e^2$ denotes the estimated variance of the forward prediction error $e(k)$. One more performance index was used to measure the forecasting performance of the neural network, and the error mean and the coefficient of multiple determination were used and given as [6]

$$B = \frac{1}{T}\sum_{\alpha=1}^{T} |x_\alpha - \hat{x}_\alpha|, \tag{15}$$

[7]Data Set A (NH_3 laser data) of the Santa Fe Time Series Prediction. The data consists of the fluctuation points of a far-infrared laser, approximately described by three coupled nonlinear ordinary differential equations.

[8]The Mackey–Glass benchmarks are generated by a nonlinear deterministic differential equation and are well known for the evaluation of forecasting methods.

Table 4
Analysis of configuration of architecture

Architecture	R_p (dB)
FIR	6.221
FCRNN ($M = 1$)	8.774
CPRNN + FIR	13.578

Table 5
Performance measures for one step ahead prediction

Measurements	NH_3	MG	W-1	W-3	W-6
B	0.0173	0.0233	0.0889	0.0813	0.0917
r^2	0.9136	0.9081	0.8217	0.8364	0.8667

$$r^2 = 1 - \frac{\sum_{\alpha=1}^{T} |x_\alpha - \hat{x}_\alpha|^2}{\sum_{\alpha=1}^{T} |x_\alpha - \bar{x}|^2}, \tag{16}$$

where T is the number of samples forecasted, x_α is the actual value, $\hat{x}_\alpha$ is the forecasted value, and $\bar{x}$ is the mean of the actual data. The error mean B is used to measure whether the predictor is biased or not. A predictor with the error mean close to zero is called unbiased. The coefficient of multiple determination can have several possible value ranges

$$r^2 = \begin{cases} 1 & \text{if } \forall_\alpha\ \hat{x}_\alpha = x_\alpha, \\ 0 > r^2 > 1 & \text{if } \hat{x}_\alpha \text{ is a better forecast than } \bar{x}, \\ 0 & \text{if generally } \hat{x}_\alpha = \bar{x}, \\ r^2 < 0 & \text{if } \hat{x}_\alpha \text{ is a worse forecast than } \bar{x}. \end{cases}$$

A coefficient of multiple determination close to one is preferable.

To verify the advantage of using the architecture proposed (CPRNN + FIR) over the FIR and single FCRNN architectures, we compared the performances of these architectures on complex wind data sampled at a 1-h interval. Table 4 shows the comparison of average prediction gains between the FIR, a single FCRNN and CPRNN + FIR configurations. There was a significant improvement in the prediction gain when the CPRNN + FIR architecture was employed over the performance of an FIR filter and a single module FCRNN.

Table 5 shows the results on error mean B and the coefficient of multiple determination r^2 for one step ahead forecast on various signals. The results obtained for the error mean B were close to zero for all cases, meaning the network is unbiased. The values of r^2 obtained for all the time series in Table 5 were relatively close to one, meaning that these forecasts were indeed better forecasts than the persistent one. Figs. 13 and 14 show the best prediction results for one step ahead prediction on NH_3, Mackey–Glass and wind (a 1 h average) time series. Time series prediction over longer horizons is highly desirable. To access how far in the future prediction is feasible, experiments are performed for 10, 20 and

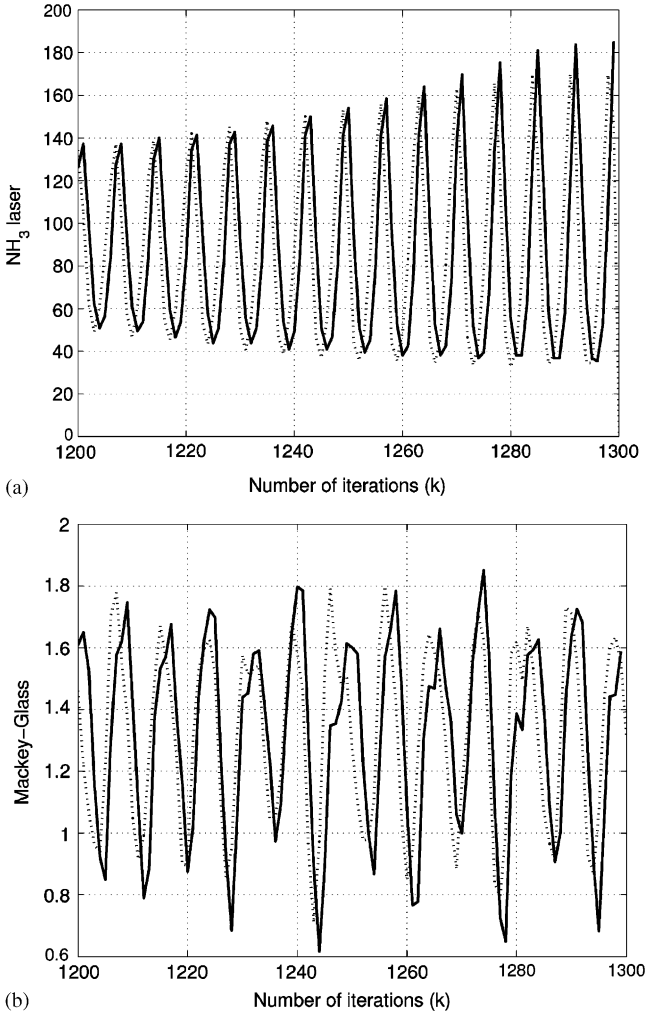

Fig. 13. One step ahead prediction of time series: (a) NH_3 laser and (b) Mackey–Glass. Solid curve: nonlinear prediction. Dashed curve: actual wind signal.

30 steps ahead. Table 6 shows the results obtained on r^2 for NH_3, Mackey–Glass and W-1 input signals. For both NH_3 and Mackey–Glass time series, longer prediction horizons were achievable. However, for the case of wind signal, the forecasts for more than 10 steps ahead were unreliable.

Table 7 shows the results of six steps ahead prediction for wind data average at 1-, 3- and 6-h intervals. The performance measures (r^2) shown in Table 7 indicate small improvement in the performance when the average interval of wind measurements increases. These results corresponded to the analysis done in Section 4. Wind averaged at longer intervals exhibits shorter correlation in the time series, which increases the predictability.

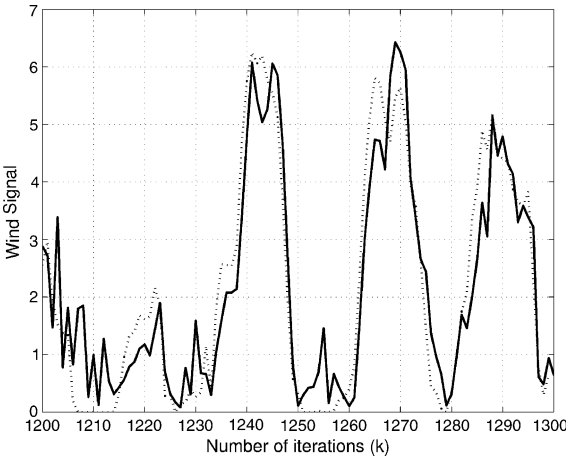

Fig. 14. One step ahead prediction of wind time series sampled at a 1 h interval. Solid curve: nonlinear prediction. Dashed curve: actual wind signal.

Table 6
Performance measures (r^2) for 10, 20 and 30 steps ahead prediction

Measurements (r^2)	NH_3	Mackey–Glass	W-1 (1 h ave.)
10 steps	0.9112	0.8846	0.6523
20 steps	0.8752	0.8124	0.2144
30 steps	0.8014	0.8093	0.1258

Table 7
Performance measures for 1-, 3- and 6-h average of wind data for six steps ahead prediction

Wind	Measurements (r^2)
W-1 (1 h ave.)	0.7003
W-3 (3 h ave.)	0.7578
W-6 (6 h ave.)	0.7671

7. Conclusions

This paper has introduced a novel approach for estimation of the wind signal in the complex domain, taking into account the strong correlation between the two wind components, speed and direction. A cascaded complex-valued pipelined recurrent neural network (CPRNN) architecture for prediction of nonlinear and non-stationary signals has been used as a forecasting model. The complex-valued real time recurrent learning (CRTRL) algorithm has been used for nonlinear adaptive filtering performed by fully connected recurrent neural networks (FCRNNs) in the complex domains. Unlike the previous algorithms of this kind, the proposed CRTRL algorithm is generic and applicable for a variety of complex signals including those with strong component correlations. It has

been demonstrated that the complex-valued representation of the wind signal provides a better model than a univariate real-valued one. The performance of the architecture has been evaluated on forecasting real-life wind signals and has been shown to give a reasonable prediction up to six steps ahead.

References

[1] Manwell JF, McGowan JG, Rogers AL. Wind energy explained: theory, design and application. New York: Wiley; 2002.

[2] Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS, Manousaridis IM. Short-term forecasting of wind speed and related electrical power. Solar Energy 1998;63(1):61–8.

[3] Kelouwani S, Agbossou K. Nonlinear model identification of wind turbine with a neural network. IEEE Trans Energy Conversion 2004;19(3):607–12.

[4] Li S, Wunsch DC, O'Hair EA, Giesselmann M. Using NNs to estimate wind turbine power generation. IEEE Trans Energy Conversion 2001;16(3):276–82.

[5] Connor JT, Martin RD, Atlas LE. Recurrent neural networks and robust time series prediction. IEEE Trans Neural Networks 1994;5(2):240–53.

[6] Drossu R, Obradovic Z. Rapid design of neural networks for time series prediction. IEEE Comput Sci Eng 1996;3(2):78–89 [see also Comput Sci Eng].

[7] Mandic DP, Chambers JA. Recurrent neural networks for prediction: learning algorithms, architectures and stability. New York: Wiley; 2001.

[8] Cybenko G. Approximation by superpositions of a sigmoidal function. Math Control Signals Systems 1989;2:303–14.

[9] Funahashi KI. On the approximate realization of continuous mappings by neural networks. Neural Networks 1989;2:183–92.

[10] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximations. Neural Networks 1989;2:359–66.

[11] Kim T, Adali T. Approximation by fully complex multilayer perceptrons. Neural Comput 2003;15(7): 1641–66.

[12] Kalogirou SA. Artificial neural networks in renewable energy systems applications: a review. Renewable Sustainable Energy Rev 2001;5(4):373–401.

[13] Kalogirou SA, Bojic M. Artificial neural networks for the prediction of the energy consumption of a passive solar building. Energy 2000;25(5):479–91.

[14] Medsker LR, Jain L. Recurrent neural networks: design and applications, The CRC Press International Series on Computational Intelligence, 2000.

[15] Narendra KS, Parthasarathy K. Identification and control of dynamical systems using neural networks. IEEE Trans Neural Networks 1990;1(1):4–27.

[16] Haykin S, Li L. Nonlinear adaptive prediction of nonstationary signals. IEEE Trans Signal Process 1995;43(2):526–35.

[17] Williams RJ, Zipser DA. A learning algorithm for continually running fully recurrent neural networks. Neural Comput 1989;1(2):270–80.

[18] Goh SL, Mandic DP. A complex-valued RTRL algorithm for recurrent neural networks. Neural Comput 2004;16(12):2699–713.

[19] Goh SL, Popovic DH, Mandic DP. Complex-valued estimation of wind profile and wind power (Best student paper award). Proceedings of the 12th IEEE mediterranean electrotechnical conference, MELECON, vol. 3. 2004. p. 1037–40.

[20] Atiya AF, El-Shoura SM, Shaheen SI, El-Sherif MS. A comparison between neural-network forecasting techniques-case study: river flow forecasting. IEEE Trans Neural Networks 1999;10(2):402–9.

[21] Parlos AG, Rais OT, Atiya AF. Multi-step-ahead prediction using dynamic recurrent neural networks. Neural Networks 2000;13:765–86.

[22] Saad EW, Prokhorov DV, Wunsch DC. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. IEEE Trans Neural Networks 1998;9(6):1456–70.

[23] Sitte R, Sitte J. Analysis of the predictive ability of time delay neural networks applied to the S&P 500 time series. IEEE Trans Systems Man Cyberet Part C 2000;30(4):568–72.

[24] Akpinar EK, Akpinar S. An analysis of the wind energy potential of Elazig Turkey. Int J Green Energy 2004;1(2):193–207.
[25] Tregidgo D. MEng Thesis, Imperial College London, London, 2004.
[26] Gautama T, Mandic DP, Hulle MMV. 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.
[27] Gautama T, Mandic DP, Hulle MMV. A differential entropy based method for determining the optimal embedding parameters of a signal. Proc. Int Conf Acoustics Speech Signal Process (ICASSP) 2003;6:29–32.
[28] Gautama T, Mandic DP, Hulle MMV. The delay vector variance method for detecting determinism and nonlinearity in time series. Physica D 2004;190(3–4):167–76.
[29] Takens F. Detecting strange attractors in turbulence. Dynamical Systems Turbulence 1980;898:336–81.
[30] Sauer T, Yorke JA, Casdagli M. Embedology. J Stat Phys 1991;65(3/4):579–616.
[31] Aihara K. Chaos engineering and its application to parallel distributed processing with chaotic neural networks. Proc. IEEE 2002;90(5):919–30.
[32] Gautama T, Mandic DP, Hulle MMV. A novel method for determining the nature of time series. IEEE Trans Biomed Eng 2004;51(5):728–36.
[33] Wan E. Time series prediction using a neural network with embedded tapped delay-lines. In: Weigend A, Gershenfeld N, editors. Predicting the future and understanding the past. SFI studies in the science of complexity, Proc. vol. XVII, Addison-Wesley; 1993. p. 195–217.
[34] Mandic DP, Chambers JA. Toward an optimal PRNN-based nonlinear predictor. IEEE Trans Neural Networks 1999;10(6):1435–42.