



Deterministic prediction of surface wind speed variations

G. V. Drisya¹, D. C. Kiplangat^{1,2}, K. Asokan³, and K. Satheesh Kumar¹

¹Department of Futures Studies, University of Kerala, Thiruvananthapuram, Kerala, India

²Department of Statistics & Actuarial Science, Dedan Kimathi University of Technology – Box 657-10100, Nyeri, Kenya

³Department of Mathematics, College of Engineering, Thiruvananthapuram, Kerala, India

Correspondence to: K. Satheesh Kumar (kskumar@keralauniversity.ac.in)

Received: 11 June 2014 – Revised: 4 September 2014 – Accepted: 14 October 2014 – Published: 19 November 2014

Abstract. Accurate prediction of wind speed is an important aspect of various tasks related to wind energy management such as wind turbine predictive control and wind power scheduling. The most typical characteristic of wind speed data is its persistent temporal variations. Most of the techniques reported in the literature for prediction of wind speed and power are based on statistical methods or probabilistic distribution of wind speed data. In this paper we demonstrate that deterministic forecasting methods can make accurate short-term predictions of wind speed using past data, at locations where the wind dynamics exhibit chaotic behaviour. The predictions are remarkably accurate up to 1 h with a normalised RMSE (root mean square error) of less than 0.02 and reasonably accurate up to 3 h with an error of less than 0.06. Repeated application of these methods at 234 different geographical locations for predicting wind speeds at 30-day intervals for 3 years reveals that the accuracy of prediction is more or less the same across all locations and time periods. Comparison of the results with f-ARIMA model predictions shows that the deterministic models with suitable parameters are capable of returning improved prediction accuracy and capturing the dynamical variations of the actual time series more faithfully. These methods are simple and computationally efficient and require only records of past data for making short-term wind speed forecasts within practically tolerable margin of errors.

Keywords. Meteorology and atmospheric dynamics (general or miscellaneous)

1 Introduction

Wind is widely recognised as a clean, economically viable and eco-friendly source of electric power. Unlike power produced from coal or nuclear energy, wind power production is safe for the environment since it does not produce any greenhouse gases or harmful by-products. Wind is produced by the uneven heating of Earth's surface by the Sun and is therefore an inextinguishable source of energy. The generation of wind power has increased steadily over the last few years all over the world, and as of 2011 the worldwide installed capacity of wind power stands at 237 GW. It is estimated that, by 2020, more than 12 % of the total demand for electricity could be met from wind energy resources (GWEC, 2012).

Nowadays in many countries, wind energy is being connected to existing electric power grids along with traditional sources. Wind-powered electricity must be used as soon as it enters the grid, and to determine the additional amount of power to order from traditional sources to meet the demand at the grid, it is important to be able to predict wind power to the order of several minutes to a few hours in advance (Hering and Genton, 2010). Predictions of wind power, of the order of a couple of hours to a day ahead, are also crucial in liberalised electricity markets where expected power production and market prices are used in devising best bidding strategy with minimum possible risk (Gomes and Castro, 2012). Short-term predictions, ranging from seconds to a few minutes, are useful in operation control of wind turbines and improvement of the power quality of wind farms (Wang and Hui, 2012).

It is clear from the above discussion that predicting wind power, in the range of at least a few hours ahead, is important both for optimising the performance of wind turbines and for maintaining a cost-effective power distribution system. Wind

power is a function of wind speed, so an accurate prediction of wind speed leads to improved predictions of wind power in a given wind farm. For a range of wind speeds, the amount of wind energy produced from a wind turbine is proportional to the cube of wind speed, so any small improvement in short-term predictions of wind speed can significantly improve predictions of wind power (Hering and Genton, 2010). However, with its dependence on topography, climate, seasonal changes, temperature, pressure and a host of other factors and its highly variable and random nature, wind speed is one of the most difficult meteorological parameters to predict. The literature on the various methods of predicting wind speed has grown extensively in recent years, especially in the wake of large-scale deployment of wind farms across the globe.

The simplest among the prediction schemes for wind speed is the method of persistence which is based on the assumption that over very small time intervals wind speed does not change appreciably. Its usage is very limited, but is still used in the industry for making very short-term predictions (Soman et al., 2010). The physical models, which utilise data of various atmospheric parameters to build up complex mathematical models, furnish another classical way to forecast wind speed. They are useful in identifying recurring patterns and making long-term predictions when weather conditions are stable, but the prohibitive computational volume involved in solving such models renders them unreliable for short-term predictions (Potter and Negnevitsky, 2006; Candy et al., 2009). Models which use statistical methods for wind speed predictions are also popular in the literature. They include moving average models such as ARMA, ARIMA and its variants fitted to the time series of wind speed (Kamal and Jafri, 1997; Cadenas and Rivera, 2007; Kavasseri and Seetharaman, 2009) and models based on probability distribution of wind speed (Hennessey Jr., 1977; Celik, 2004; Mathew et al., 2011; Jiang et al., 2013). These models are fairly good in very short-term predictions, but do not improve significantly on prediction error compared to the elementary method of persistence. Models based on artificial neural networks, which emulate the parallel distributed processing of human nervous system to adapt by learning from past data, have also been developed by many researchers for making short-term predictions of wind speed and power (Mohandes et al., 1998; Cadenas and Rivera, 2007; Bilgili et al., 2007; Monfared et al., 2009). In general, these models outperform the time series models in short-term predictions, but their performance edge is not maintained across all locations universally (Soman et al., 2010). Recently researchers have also begun to use hybrid models, which combine different approaches for better forecasting results, such as mixing physical and statistical models or short-term and medium-term models (Soman et al., 2010; Liu et al., 2014; Haque et al., 2013). The central idea of physical approach is to incorporate the physical considerations of local topography into the numerical weather prediction scheme by modelling the local wind profile possibly considering the atmospheric stability.

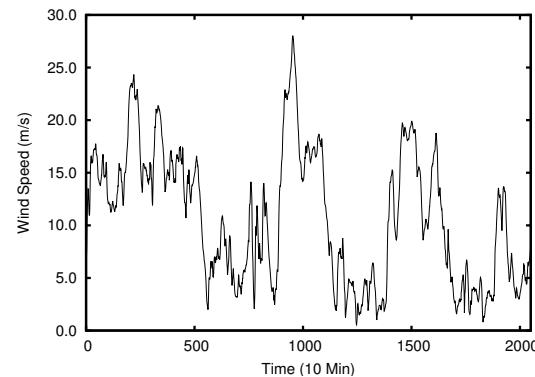


Figure 1. Time series of wind speed at location given by latitude: 34.98420° N, longitude: 104.03971° W at 80 m.

For example, Cassola and Burlando (2012) use Kalman filtering technique applied to the output of a numerical weather prediction model to improve the accuracy of wind speed forecasts and wind energy output predictions significantly.

As a matter of fact, since the occurrence of wind is highly uncertain in time and space, no single technique can be used universally across all locations and timescales for predicting wind speed, and there is always scope for new methods. In a previous work (Sreelekshmi et al., 2012) we had carried out a detailed analysis of the time series of daily mean wind speed at Thiruvananthapuram, India, which revealed strong evidence for the existence of an underlying system which is deterministic, low-dimensional and chaotic. This means that the apparent random fluctuations found in wind speed data could originate from the chaotic dynamics of the underlying system, and not necessarily due to the system being stochastic as assumed in most of the aforementioned prediction theories, and this could also explain why wind speed predictions become erroneous beyond a certain time limit. However, provided wind speed dynamics is chaotic in a given location, we can use existing non-linear prediction schemes developed for chaotic time series to make more accurate short-term predictions about wind speed. In this work we apply the methods of non-linear dynamics for forecasting wind speeds at various locations to get fairly accurate predictions up to 3 h. For the analysis we have used the wind speed data of 10 min resolution of the period from January 2004 to January 2007 for 234 locations available from National Renewable Energy Laboratory (<http://www.nrel.gov>), USA.

2 Analysis of the data

Figure 1 shows a plot of part of the wind speed data at location given by latitude of 34.98420° N and longitude of 104.03971° W at 80 m. We start with a detailed analysis of the underlying dynamics of wind speed data from this particular site and then move on to other locations. The irregular, random character of the data evident in the figure is typical

of systems which are stochastic, but – as we have shown previously (Sreelekshmi et al., 2012) – these fluctuations could also arise from an underlying system which is deterministic, low-dimensional and chaotic. Since the stochastic versus deterministic character of wind speed has to be ascertained on a per location basis due to the highly uncertain nature of wind from location to location, we will now carry out a brief analysis of the data, referring the reader to Sreelekshmi et al. (2012) for the finer details of the features of chaotic systems and of the methods applied. The wind speed prediction schemes to be described later in coming sections are based on the results of this analysis.

The first step in the analysis of time series by methods of dynamical system theory is reconstructing the state space dynamics of the original system using the given time series data (Packard et al., 1980). This is done by constructing from the time series $x(t)$ a new vector time series $\mathbf{x}(t)$ given by

$$\mathbf{x}(t) = (x(t), x(t - \tau), \dots, x(t - (m - 1)\tau)), \quad (1)$$

where τ is a suitable multiple of the sampling time, called delay. Takens' embedding theorem and its extensions (Takens, 1981; Sauer et al., 1991; Sauer and James, 1993) assert that the dynamics of $\mathbf{y}(t)$ in the reconstructed phase space will be topologically identical to the dynamics of the original system. In general, if the cloud of points generated by $\mathbf{y}(t)$, called the attractor, fills out the m -dimensional phase space for all values of m , the time series may be considered to be generated by a stationary stochastic process. On the other hand, if the attractor occupies a region of small dimension, for all sufficiently large values of m , it may be an indication that the original system is deterministic and chaotic. On a chaotic attractor, nearby trajectories diverge with time exponentially fast, the rates of which are quantified by the Lyapunov exponents in the principal directions. A positive Lyapunov exponent is considered a signature of chaos. Due to the exponential divergence of trajectories, a chaotic attractor usually has a complex structure with a non-integral dimension. A chaotic time series exhibit broadband spectrum (cf. Fig. 7) and other characteristics of random time series when analysed using linear stochastic tools, so a very detailed analysis is often necessary to distinguish a chaotic time series from a stochastic data. The space–time separation plot given in Fig. 2 of the time series helps us identify the temporal correlations within the time series and is useful in estimating a reasonable delay. The diurnal variations are evident in this figure. In order to reduce its modulation effects we removed the average diurnal variation by carrying out an epoch analysis as discussed by Kumar et al. (2004), and further analysis was carried out on this detrended time series.

In practical applications of embedding theorem, an optimal choice of the delay τ and the embedding dimension m are important. We have used the method of autocorrelation (Kantz and Schreiber, 2003) as well as the method of mutual information (Fraser and Swinney, 1986) to arrive at a proper choice of τ . In Fig. 3a and b, which respectively plot

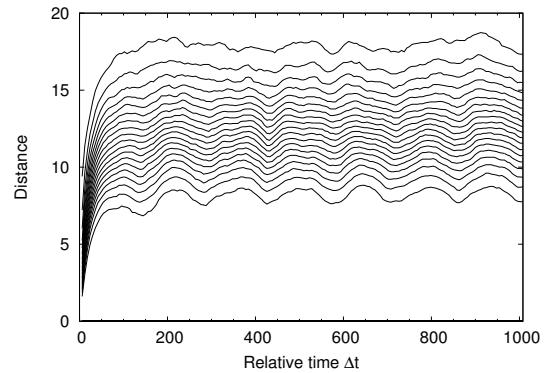


Figure 2. The space–time separation plot for the time series for $m = 14$ and $\tau = 85$. Each point in the plot represents a pair of points on the trajectory with their relative separation in time along the horizontal axis and separation in space along the vertical axis. The diurnal variations are evident in this figure.

the autocorrelation and mutual information of the wind speed data as a function of τ , the first minimum of the autocorrelation curve is observed at around $\tau = 85$ suggesting an optimal value $\tau = 85$ and the mutual information almost level off by $\tau = 75$, suggesting the value of τ can optimally be taken around this value. The value $\tau = 85$ is taken for further analysis. To determine the embedding dimension m , we used the method of false neighbours (Kennel et al., 1992; Kantz and Schreiber, 2003) which is based on the idea that, in a large enough embedding dimension, the fraction of false neighbours, which arise due to crossing of trajectories in a lower than true dimension, would be negligibly small. The fraction of false neighbours for the wind speed data is shown in Fig. 4, which suggests that $m = 14$ would be an optimal choice for the embedding dimension as the fraction attains a minimum around $m = 14$. However precise knowledge of m is desirable only to exploit the determinism in the dynamics with minimal computational effort, and a large value of m will add redundancy and thus degrade the performance of many algorithms such as those for predictions (Kantz and Schreiber, 2003). It may also be noted that the parameter values obtained here are consistent with those in our previous analysis of wind speed data from a different geographical location (Sreelekshmi et al., 2012).

The dimension of the attractor gives a quantitative measure of the self-similarity of the attractor and also gives an idea of how large or small a region is occupied by the attractor within the embedding space. A standard dimension estimate for time series data is the correlation dimension, introduced by Grassberger and Procaccia (1983), which proceeds by first computing the correlation sum defined by (Hegger et al., 1999)

$$C(\epsilon, m) = \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N \Theta(\epsilon - \|y_i - y_j\|), \quad (2)$$

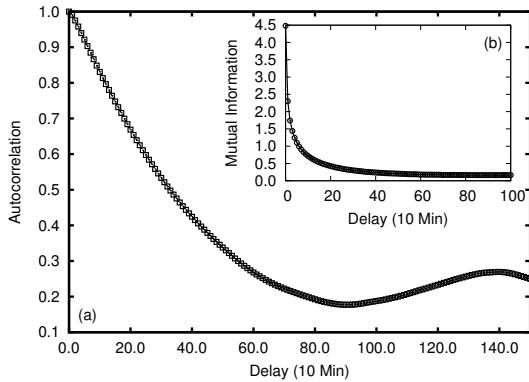


Figure 3. (a) The autocorrelation function of wind speed data. (b) Mutual information of the wind speed data as a function of delay (τ).

where $\Theta(a) = 1$ if $a > 0$, $\Theta(a) = 0$ if $a \leq 0$, and then the local slopes

$$D_2(\epsilon, m) = \frac{d \ln C(\epsilon, m)}{d \ln \epsilon}, \quad (3)$$

which are estimates of the correlation dimension. Figure 5 plots $D_2(\epsilon, m)$ versus ϵ for the wind speed data with the previous choice of delay and for embedding dimensions ranging from 14 to 16. The curves exhibit convergence onto a plateau for a range of values of ϵ , and the value corresponding to the plateau, $D_2 = 1.656 \pm 0.008$, is an estimate of the correlation dimension for the given data. The low dimensionality of the attractor is an indication of the deterministic character of the underlying dynamics exhibiting chaotic behaviour.

A chaotic system must have at least one positive Lyapunov exponent; to check for this, one usually computes the largest Lyapunov exponent in the system which, if found positive, is considered strong evidence for chaos. We have used the Kantz algorithm (Kantz, 1994) to estimate the largest Lyapunov exponent, which proceeds by computing the stretching factor $S(\Delta n)$, involving a reference point y_{n_0} and its neighbours y_n in the embedding space over a neighbourhood $U(y_{n_0})$ of y_{n_0} , defined by

$$S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{\|U(y_{n_0})\|} \times \sum_{y_n \in U(y_{n_0})} \|y_{n_0+\Delta n} - y_{n+\Delta n}\| \right). \quad (4)$$

For the wind speed data, Fig. 6 plots variations of $S(\Delta n)$ with Δn for $m = 14, 15$ and shows a linear growth in the range of $0 < \Delta n < 20$, the slope of which gives an estimate of the maximum Lyapunov exponent, in this case 0.15 ± 0.006 . The positive value of the largest Lyapunov exponent is another piece of evidence for the chaotic dynamics of the underlying system.

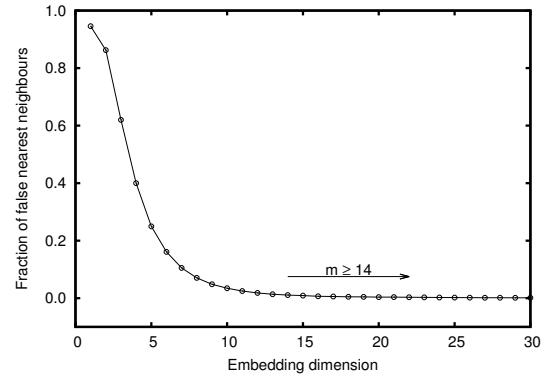


Figure 4. The fraction of false nearest neighbours as a function of the embedding dimension m for the time series of wind speed with $\tau = 85$, showing that any $m \geq 14$ can be considered optimal.

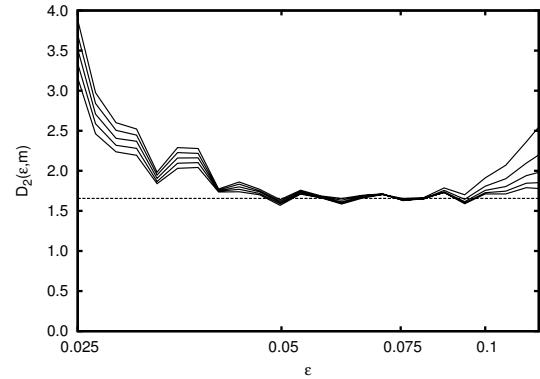


Figure 5. The local slopes $D_2(\epsilon, m)$ for the wind speed time series with $\tau = 85$ and m ranging from 14 to 16 showing a plateau for small values of ϵ and giving an estimate of $D_2 = 1.656 \pm 0.008$.

The frequency decomposition of the time series, the power spectrum, is broadband and exhibits exponential decay as shown in Fig. 7, and this is an indication of the chaotic behaviour of the time series. The first part of the spectrum decays abruptly with an estimated value of -120.29 for the exponent while the second and third parts exhibit slow decay with values -15.80 and -4.07 for the exponent. This could possibly be caused by qualitatively different mechanisms at work in the dynamics of the underlying system.

3 Predictions

Due to the random nature of wind speed data, most of the wind speed predictions assume that the data are a realisation of a stochastic process. However, as we have shown in the previous section, the cause of randomness in the wind speed data can also be low-dimensional chaos, in which case there is a fundamental limit on long-term predictions. However, accurate short-term predictions can still be made, taking advantage of simple deterministic relationships which may

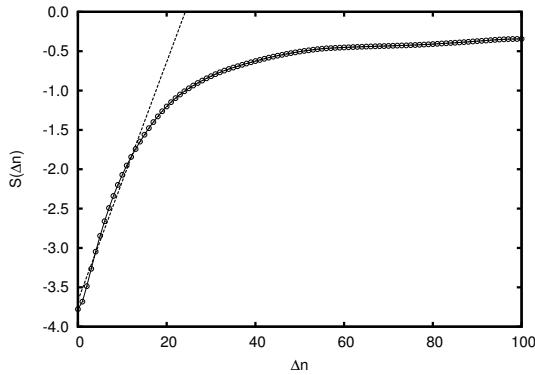


Figure 6. The curve of $S(\Delta n)$ for embedding dimensions $m = 14, 15$. The maximum Lyapunov exponent of the time series is the slope of the dashed line 0.15 ± 0.006 .

exist within the data, involving a few degrees of freedom. The extent of predictability depends on the local factors affecting wind speed, and in the data we considered accurate predictions up to 3 h could be made using non-linear prediction tools.

Given a time series x_1, x_2, \dots, x_n , the forecasting methods try to predict values a few time steps ahead, namely x_{n+k} for $k = 1, 2, \dots$. Non-linear forecasting methods are based on construction of delay vectors \mathbf{x}_n from the time series using Eq. (1);

$$\mathbf{x}_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}), \quad (5)$$

for a suitable integral delay τ and embedding dimension m . If deterministic rules govern the system, we expect a functional relation between \mathbf{x}_{n+1} and \mathbf{x}_n :

$$\mathbf{x}_{n+1} = \mathbf{F}(\mathbf{x}_n) \quad (6)$$

which in delay embedding reduces to

$$x_{n+1} = f(x_n). \quad (7)$$

If the dynamics is chaotic, \mathbf{F} will be non-linear and we try to approximate \mathbf{F} or f in various ways to predict an estimate \hat{x}_{n+k} for the actual value x_{n+k} . In chaotic systems these approximations are facilitated by an important property, that a sufficiently long time series produces a sequence of vectors that are dense on the attractor, so that new vectors will be arbitrarily close to some of those already observed. The approximation schemes used for predictions of chaotic time series can be broadly classified into local methods and global methods.

The local approximation schemes try to approximate f locally, probably by a different function in each time step, by looking for vectors in the past which are close to \mathbf{x}_n in the embedding space and using their future for prediction. The simplest of these is the zeroth-order approximation, which uses the average of the futures of the neighbours of \mathbf{x}_n in

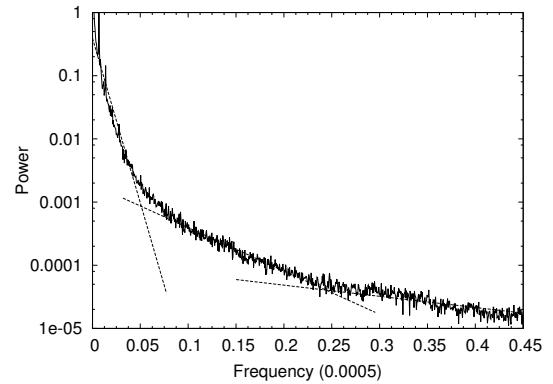


Figure 7. The power spectrum of wind speed time series as a function of frequency at location latitude: 34.98420° N longitude: 104.03971° W. Broadband and exponential decay of the power with frequency are typical characteristics of chaotic signals.

an ϵ neighbourhood $U_\epsilon(\mathbf{x}_n)$. If there are N neighbours in $U_\epsilon(\mathbf{x}_n)$, the prediction is simply (Kantz and Schreiber, 2003)

$$\hat{x}_{n+1} = \frac{1}{N} \sum_{\mathbf{x}_j \in U_\epsilon(\mathbf{x}_n)} x_{j+1}. \quad (8)$$

A better method is the local first-order (LFO) approximation, where – instead of taking the average of the neighbours in $U_\epsilon(\mathbf{x}_n)$ – a linear model is fitted to these neighbours, so that the prediction takes the form (Kantz and Schreiber, 2003)

$$\hat{x}_{n+1} = \mathbf{A}_n \mathbf{x}_n + \mathbf{b}_n. \quad (9)$$

These local linear models, one for each time step, together generate a non-linear model globally.

The global models of prediction try to approximate \mathbf{F} by a single function on the whole attractor. One of the popular global models used in predictions of chaotic time series is the radial basis function (RBF) model introduced by Lowe and Broomhead (1988). In this, the approximating function \mathbf{F} is taken as a linear superposition of a set of radial basis functions $\Phi_i(r)$, with $r > 0$, which are typically bell-shaped with maximum at $r = 0$ and rapidly decaying towards zero with increasing r . For a set of suitably chosen points \mathbf{y}_i , called the centres, which are reasonably well distributed on the attractor, the model assumes the form (Kantz and Schreiber, 2003)

$$\mathbf{F}(\mathbf{x}) = \alpha_0 + \sum_{i=1}^p \alpha_i \Phi(\|\mathbf{x} - \mathbf{y}_i\|). \quad (10)$$

The basis functions Φ are modelled using Gaussians with their number and width kept fixed throughout the model. This makes the estimation of the constants α_i a linear problem which can be solved using least-squares method.

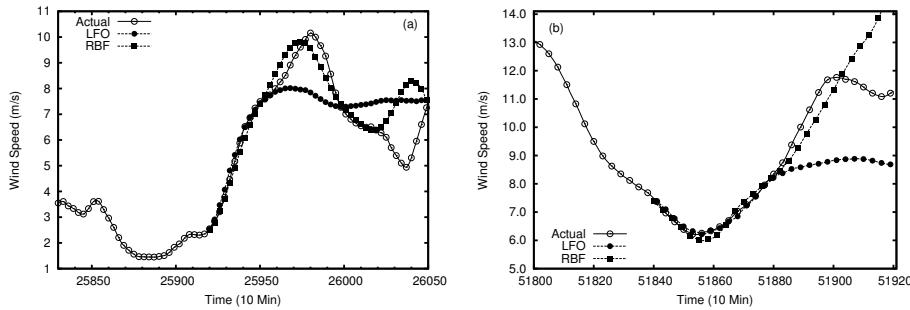


Figure 8. Comparison of predicted values with the actual values for LFO and RBF. The symbols are plotted only at every 30 min for legibility. (a) Latitude: 42.31925° N, longitude: 98.60197° W $m = 5, \tau = 8$; (b) latitude: 43.51076° N, longitude: 99.47652° W, $m = 12, \tau = 2$.

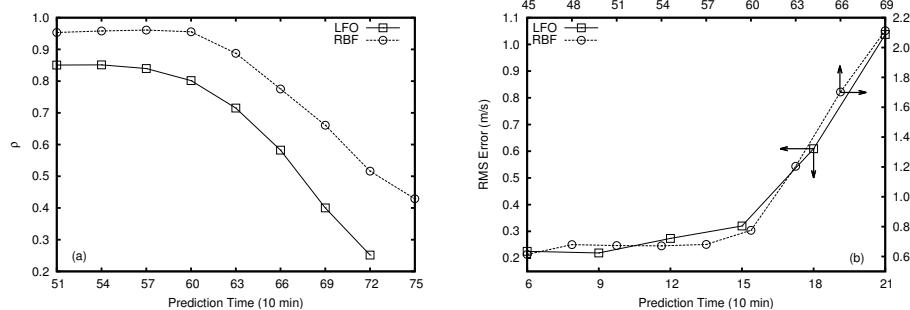


Figure 9. (a) The exponential decay of the correlation coefficient between the predicted and actual for every three prediction time steps. (b) The exponential growth RMSE prediction for every three prediction time steps. The arrows show axes for the respective symbols of data points.

The advantage of LFO method is its flexibility, but it may not yield desirable performance on parts of the phase space where the points do not span the available space dimensions. On the other hand, global models have the advantage of providing the structure and properties of the underlying system as it can yield closed expressions for the full dynamics. These models can effectively describe the observed process in regions of the space which have been visited by the data, but outside this area the shape of the model depends heavily on the chosen function (Hegger et al., 1999).

Figure 8a and b show typical results of wind speed prediction using the above methods, made at a couple of locations for suitable choices of the embedding parameters, and their comparison with actual data. Each prediction uses the available wind speed data for the location up to a specific point of time for modelling, and then employs the model for predicting future values. It is seen from these figures that, provided we use appropriate embedding parameters, the deterministic methods can predict wind speed with remarkable accuracy up to 3 h, with the RBF method giving fairly accurate predictions for another 15 h in both the cases. A method for determining the optimal embedding parameters for prediction is discussed in the next section.

Figure 9a quantifies the similarity of predicted values with the measured data for a typical case, by plotting the statistical

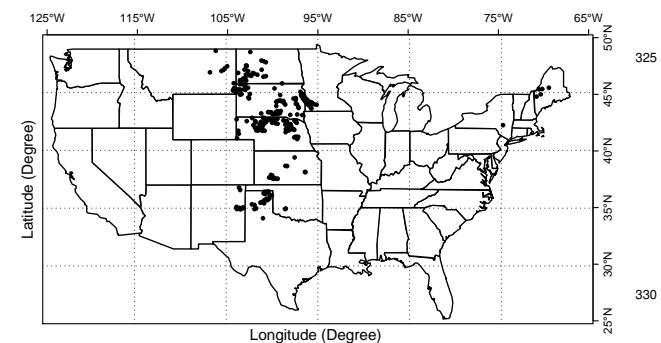


Figure 10. The geographic locations, denoted by filled circles, where wind speed data were analysed for performance of deterministic model prediction.

coefficient of correlation between the predicted and actual values as a function of the number of time steps into the future. The correlation coefficients were calculated cumulatively, at the end of every three prediction time steps, using all the predicted values available up to that time and the corresponding measured values. The exponential deterioration of the correlation with increasing prediction time is a characteristic feature of deterministic chaos (Sugihara and May, 1990) and provides further evidence of the fact that the erratic

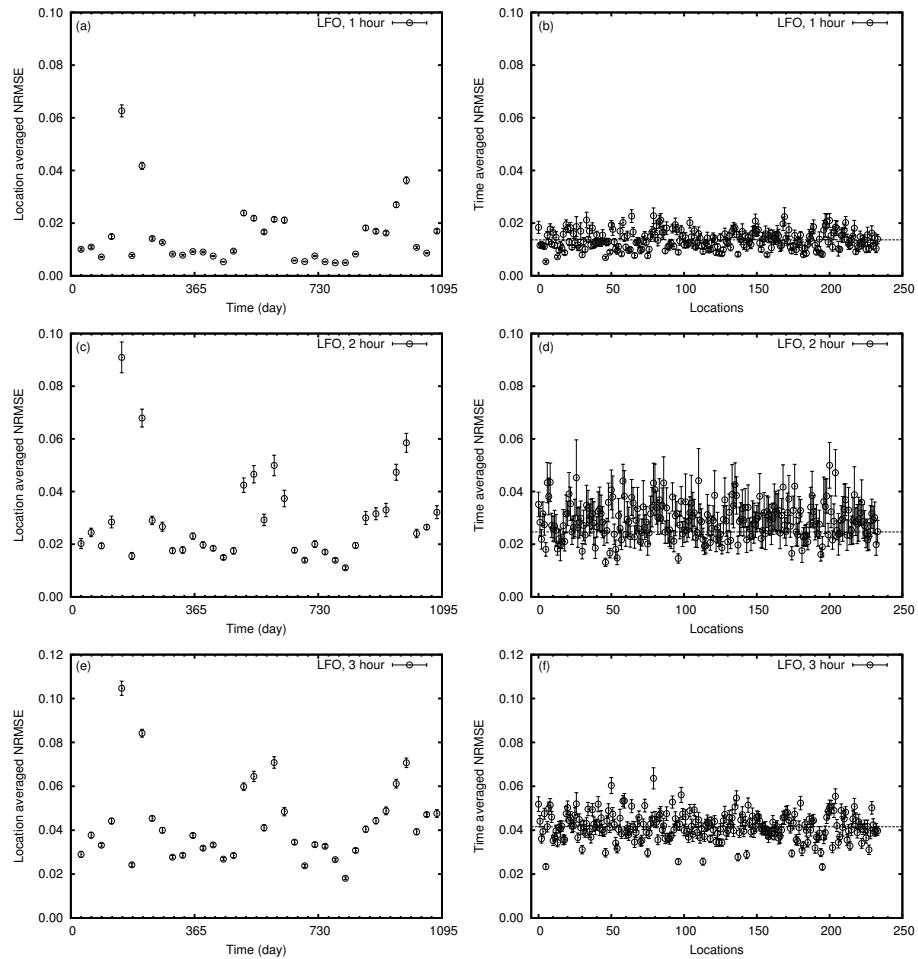


Figure 11. (a) NRMSE of 1 h predictions using LFO, for the period from 2004 to 2006 at an interval of 30 days, averaged over 234 locations. (b) NRMSE of 1 h predictions using LFO for the 234 locations, calculated at an interval 30 days and averaged over the period from 2004 to 2006. (c) and (e) are location-averaged NRMSE for 2 and 3 h predictions respectively. (d) and (f) are time-averaged NRMSE for 2 and 3 h predictions respectively. The error bars in the figures are with respect to the standard error of the mean. The horizontal dotted lines in (b), (d) and (f) represent the mean of the respective time-averaged NRMSE values over the entire set of locations.

fluctuations in wind speed data are caused by the chaotic dynamics of the underlying system and are not an artefact of uncorrelated additive noise. Figure 9b shows how the root mean square error (RMSE) between the predicted and measured values, again calculated cumulatively every three time steps, propagates as we predict further into the future. The exponential growth of the prediction error further substantiates the chaotic nature of the data.

4 Statistical analysis of prediction errors

To demonstrate the wider applicability of the deterministic methods for making short-term wind speed forecasts, we now carry out a statistical analysis of the prediction errors for forecasts made at a total of 234 geographical locations. For the analysis we have considered 10 min interval wind speed data for 3 years from 2004, available from National

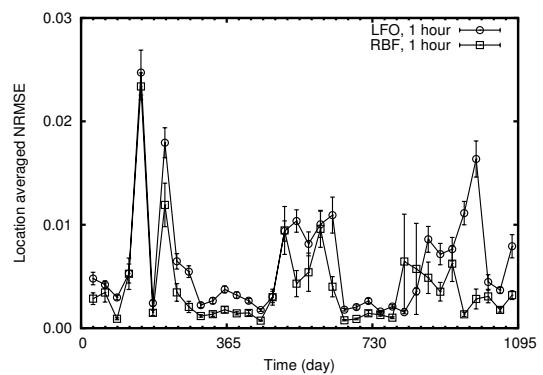


Figure 12. NRMSE, with standard error, of 1 h prediction for the period of 3 years from 2004 over 234 locations using LFO and RBF. The lines connecting the symbols are to guide the eye.

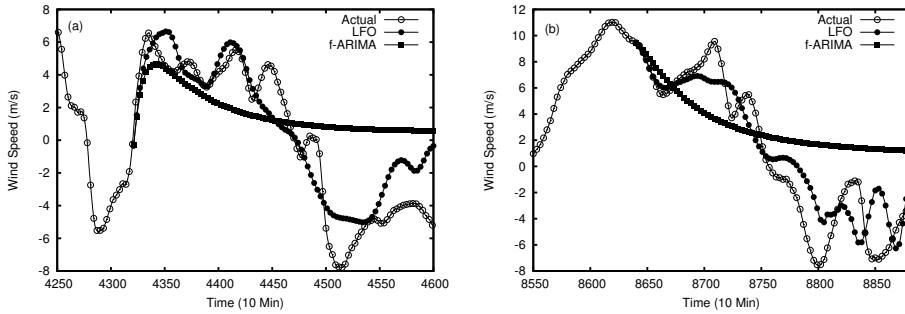


Figure 13. Comparison of predicted values with the actual values for LFO and f-ARIMA. The symbols are plotted only at every 30 min for legibility. Latitude: 34.98420° N, longitude: 104.03971° W, (a) $m = 13$, $\tau = 3$ and (b) $m = 8$, $\tau = 7$.

Renewable Energy Laboratory, USA, for the 234 locations depicted in Fig. 10.

The optimal choice of the embedding parameters m and d is a major factor affecting the accuracy of prediction. Since the dynamics of wind speed varies over locations, these parameters have to be determined for each location separately. However, the embedding parameters suggested by the auto-correlation function or the fraction of false neighbours need not always give the most accurate predictions (Domenico et al., 2013), and a systematic procedure for determining the most suitable parameters for predictions using model data is still elusive. For the present analysis, to fix the optimal parameters at each location, we have used a test run procedure as described below. Given a time series of n points x_1, \dots, x_n , we want to predict the next k values x_{n+1}, \dots, x_{n+k} . From the given data set, we take the first $n - 3$ data points to form a model for making a test prediction of the next three data points for various values of m and d using one of the deterministic algorithms described earlier. These predicted values $x_{n-2}^p, x_{n-1}^p, x_n^p$ are then compared with the actual data points x_{n-2}, x_{n-1}, x_n to find the RMSE. The values of m and d which yield the minimum RMSE for these three predicted values are selected as the parameter values for the given data set and used for the prediction of values of x_{n+1}, \dots, x_{n+k} .

We use spatial averages of prediction errors over various locations as well as time averages at each location to assess the performance of these methods. Since the range of values of wind speed varies over locations, we have chosen as a measure of the prediction error the root mean squared error normalised over the range of the observed data (NRMSE) given by

$$\text{NRMSE} = \sqrt{\frac{\sum_{j=1}^{n+k} (x_i - x_j^p)^2}{k}} / (x_{\max} - x_{\min}), \quad (11)$$

where x_i^p are the predicted values. For each location we have calculated NRMSE for 1, 2 and 3 h predictions at intervals of 30 days for a 3-year period from 2004 to 2006. Figure 11a depicts NRMSE with an error bar for 1 h predictions averaged over the locations (location-averaged NRMSE), computed at

30-day intervals and plotted for a 3-year period. Figure 11b shows NRMSE for 1 h prediction for each location averaged over a 3-year time period (time-averaged NRMSE) where the horizontal dotted line shows the mean 0.0136 of these values over the locations. Similar estimates of location and time-averaged errors for 2 h predictions are given in Fig. 11c and d and for 3 h predictions in panels e and f. The mean of the time-averaged NRMSE over the locations, indicated in each figure by a dotted line, is 0.0299 for 2 h predictions and 0.0415 for 3 h predictions. This shows that the prediction accuracy observed in the typical forecasts shown in Fig. 8 is more or less maintained across all locations and various time periods consistently.

Between the deterministic methods, the RBF consistently maintains a slight performance edge over LFO for short-term predictions up to 1 h, as is clear from Fig. 12, which compares the location-averaged prediction errors for 1 h predictions by the two methods.

5 Comparison with f-ARIMA

Among the various statistical methods used in wind speed prediction, ARIMA is a popular model which gives reasonably accurate predictions of wind speed at many locations (Kamal and Jafri, 1997; Cadenas and Rivera, 2007; Kavasseri and Seetharaman, 2009). An ARIMA (p, d, q) model combines an autoregressive(AR) process of order p , a moving average (MA) process of order q and a differencing operator of order d into a single model. It has the general form (Box et al., 2013)

$$\Phi(B)\Delta^d x_t = c + \Theta(B)\epsilon_t, \quad (12)$$

where ϵ_t is a white noise process and $\Phi(B)$ and $\Theta(B)$ are respectively the autoregressive and moving average operators defined by

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ \Theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q, \end{aligned} \quad (13)$$

for suitably chosen constants θ_i and ϕ_i and non-negative integers p and q . B is the backward-shift operator so that

$Bx_t = x_{t-1}$ and $\Delta = 1 - B$ is the differencing operator and in the general ARIMA model d is an integer.

f-ARIMA is a generalisation of ARIMA where the parameter d is allowed to have a fractional value with the operator $(1 - B)^d$ interpreted to have the binomial expansion (Hosking, 1981)

$$(1 - B)^d = 1 - dB + \frac{d(d-1)}{2}B^2 + \dots \quad (14)$$

The possibility of wide range of choices for the parameters p , q , and d and the constants ϕ_i and θ_i gives the model great flexibility and wider applicability.

One of the features that distinguishes a f-ARIMA process from an ARIMA process is that the former is characterised by a slow decay in its autocorrelation function compared to the latter. This feature makes f-ARIMA model an attractive choice for data sets that exhibit long-range correlations such as the wind speed data (Kavasseri and Seetharaman, 2009).

General characteristics of the predictions by f-ARIMA and how they compare with the predictions by LFO can be seen from Fig. 13a and b. The performance of f-ARIMA is comparable to LFO initially, but its predictions deviate from actual values and level off to a steady value after a brief period. In contrast, while the accuracy of prediction of LFO also falls off gradually after 3 h, it nevertheless captures the essential dynamics of the original time series even further.

For comparing the performance of LFO versus f-ARIMA, we have elected to generate the best possible predictions by both the methods, by experimenting with various values of the parameters which determine the accuracy of prediction. Thus, for a model data set x_1, \dots, x_n , we would generate several trial predictions for the next k data points using various parameter values, compare each of them with the actual observed data x_{n+1}, \dots, x_{n+k} , and choose the one that gives the least prediction error. For the LFO method this might yield better predictions than would be obtained with the embedding parameters selected by the procedure described in the last section. While the latter procedure would be useful in real-world applications where there are no future data to compare the predictions with, it need not always give the optimum parameter values giving the most accurate predictions. In fact, we have observed that the LFO predictions obtained here (Fig. 14a, b) are marginally better than those in Fig. 11a and b, with the location-averaged NRMSEs being smaller by 0.8 % and 1.8 % on the average for 1 and 3 h predictions respectively.

Figure 14 shows the results of a statistical analysis of the performance of LFO and f-ARIMA, with optimum parameter values, over all the 234 locations described earlier. The panels a, b and c compare NRMSE for 1, 3 and 6 h predictions averaged over all locations computed in intervals of 30 days for a 3-year period. The prediction accuracy of LFO is noticeably better than that of f-ARIMA across all locations and all time periods. For low-resolution wind speed data of the kind considered in this work, the accuracy and the longevity

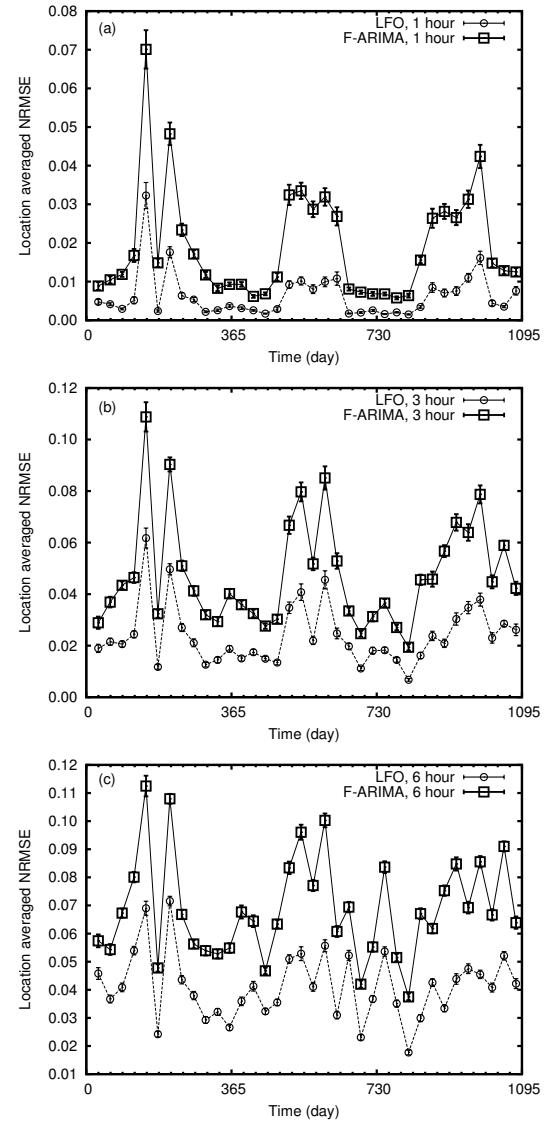


Figure 14. Location-averaged NRMSE, with standard error, of (a) 1 h, (b) 3 h and (c) 6 h predictions for the period of 3 years from 2004 over 234 locations using LFO and f-ARIMA.

of the predictions obtained by the deterministic methods are therefore a significant improvement over existing methods.

6 Conclusions

In this work we demonstrate the suitability of deterministic methods in making short-term forecasts of wind speed based on past data. These methods are applicable in situations where the underlying dynamics of wind speed is chaotic leading to random-like fluctuations in the time series of wind speed. We have applied a couple of chaotic time series prediction tools (one local method and one global method) on records of wind speed data of 10 min resolution from a total

of 234 different geographical locations, at each location making 1, 2 and 3 h predictions at intervals of 30 days for a period of 3 years. The predictions are very accurate for up to 1 h and fairly accurate for up to 3 h. A statistical analysis of the prediction errors from these locations reveals that the average prediction error is 1.36 % of the range of wind speed for 1 h predictions, 2.99 % for 2 h predictions and 4.15 % for 3 h predictions.

We have also compared the efficiency of the deterministic methods with predictions by f-ARIMA at each of the above 234 locations on the basis of 6 h predictions at intervals of 30 days for a period of 3 years. It is seen that, compared to f-ARIMA, the deterministic methods give better prediction accuracy for longer periods of time and capture the dynamics of the fluctuations in the original data more faithfully. These prediction methods are simple and computationally efficient alternatives for short-term wind speed forecasts.

Acknowledgements. The authors are grateful to University Grants Commission (UGC), New Delhi, for their financial assistance (no. F. 42-30/2013(SR)). They wish to thank National Renewable Energy Laboratory (<http://www.nrel.gov>), USA, for making their data available, Manoj Changat, Department of Futures Studies, Univ. of Kerala for his help and support and campus computing facility of University of Kerala under DST-PURSE programme for providing computational facilities. Authors are indebted to the referees for their constructive comments and useful suggestions which helped them to improve the original manuscript substantially.

Topical Editor V. Kotroni thanks three anonymous referees for their help in evaluating this paper.

References

- Bilgili, M., Sahin, B., and Yasar, A.: Application of artificial neural networks for the wind speed prediction of target station using reference stations data, *Renew. Energ.*, 32, 2350–2360, 2007.
- Box, G. E., Jenkins, G. M., and Reinsel, G. C.: *Time series analysis: forecasting and control*, John Wiley & Sons, 2013.
- Cadenas, E. and Rivera, W.: Wind speed forecasting in the South Coast of Oaxaca, Mexico, *Renew. Energ.*, 32, 2116–2128, 2007.
- Candy, B., English, S. J., and Keogh, S. J.: A comparison of the impact of quikscat and windsat wind vector products on met office analyses and forecasts, *IEEE T. Geosci. Remote*, 47, 1632–1640, 2009.
- Cassola, F. and Burlando, M.: Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output, *Appl. Energ.*, 99, 154–166, 2012.
- Celik, A. N.: A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey, *Renew. Energ.*, 29, 593–604, 2004.
- Domenico, M. D., Ghorbani, M. A., Makarynskyy, O., Makarynska, D., and Asadi, H.: Chaos and reproduction in sea level, *Appl. Math. Model.*, 37, 3687–3697, 2013.
- Fraser, A. and Swinney, H.: Using mutual information to find independent coordinates for strange attractors, *Phys. Rev. A*, 33, 1134–1140, 1986.
- Gomes, P. and Castro, R.: Wind Speed and Wind Power Forecasting using Statistical Models: AutoRegressive Moving Average (ARMA) and Artificial Neural Networks (ANN), *International Journal of Sustainable Energy Development*, 1, 1/2, 41–50, 2012.
- Grassberger, P. and Procaccia, I.: Measuring the strangeness of strange attractors, *Physica D*, 9, 189–208, 1983.
- GWEC: Global Wind Energy Outlook, available at: http://www.gwec.net/wp-content/uploads/2012/11/GWEO_2012_lowRes.pdf (last access: 11 June 2014), 2012.
- Haque, A. U., Mandal, P., Meng, J., and Negnevitsky, M.: Wind speed forecast model for wind farm based on a hybrid machine learning algorithm, *International Journal of Sustainable Energy*, 1–14, 2013.
- Hegger, R., Kantz, H., and Schreiber, T.: Practical implementation of nonlinear time series methods: The TISEAN package, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 9, 413–435, 1999.
- Hennessey Jr., J. P.: Some aspects of wind power statistics, *J. Appl. Meteorol.*, 16, 119–128, 1977.
- Hering, A. S. and Genton, M. G.: Powering up with space-time wind forecasting, *J. Am. Stat. Assoc.*, 105, 92–104, 2010.
- Hosking, J. R.: Fractional differencing, *Biometrika*, 68, 165–176, 1981.
- Jiang, Y., Song, Z., and Kusiak, A.: Very short-term wind speed forecasting with Bayesian structural break model, *Renew. Energ.*, 50, 637–647, 2013.
- Kamal, L. and Jafri, Y. Z.: Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan, *Sol. Energy*, 61, 23–32, 1997.
- Kantz, H.: A robust method to estimate the maximal Lyapunov exponent of a time series, *Phys. Lett. A*, 185, 77–87, 1994.
- Kantz, H. and Schreiber, T.: *Nonlinear time series analysis*, Vol. 2000, Cambridge University press Cambridge, 2003.
- Kavasseri, R. G. and Seetharaman, K.: Day-ahead wind speed forecasting using f-ARIMA models, *Renew. Energ.*, 34, 1388–1393, 2009.
- Kennel, M. B., Brown, R., and Abarbanel, H. D.: Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Phys. Rev. A*, 45, 3403–3411, 1992.
- Kumar, K. S., Kumar, C., George, B., Renuka, G., and Venugopal, C.: Analysis of the fluctuations of the total electron content (TEC) measured at Goose Bay using tools of nonlinear methods, *J. Geophys. Res.-Space*, 109, A02308, doi:10.1029/2002JA009768, 2004.
- Liu, D., Niu, D., Wang, H., and Fan, L.: Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm, *Renew. Energ.*, 62, 592–597, 2014.
- Lowe, D. and Broomhead, D.: Multivariable functional interpolation and adaptive networks, *Complex Systems*, 2, 321–355, 1988.
- Mathew, T., George, S., Sarkar, A., and Basu, S.: Prediction of mean winds at an Indian coastal station using a data-driven technique applied on Weibull Parameters, *The International Journal of Ocean and Climate Systems*, 2, 45–54, 2011.
- Mohandes, M. A., Rehman, S., and Halawani, T. O.: A neural networks approach for wind speed prediction, *Renew. Energ.*, 13, 345–354, 1998.
- Monfared, M., Rastegar, H., and Kojabadi, H. M.: A new strategy for wind speed forecasting using artificial intelligent methods, *Renew. Energ.*, 34, 845–848, 2009.

- Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S.: Geometry from a time series, *Phys. Rev. Lett.*, 45, 712–716, 1980.
- Potter, C. W. and Negnevitsky, M.: Very short-term wind forecasting for Tasmanian power generation, *IEEE T. Power Syst.*, 21, 965–972, 2006.
- Sauer, T. and James, A. Y.: How many delay coordinates do you need?, *Int. J. Bifurcat. Chaos*, 3, 737–744, 1993.
- Sauer, T., Yorke, J. A., and Casdagli, M.: Embedology, *J. Stat. Phys.*, 65, 579–616, 1991.
- Soman, S. S., Zareipour, H., Malik, O., and Mandal, P.: A review of wind power and wind speed forecasting methods with different time horizons, in: North American Power Symposium (NAPS), 2010, 1–8, IEEE, 2010.
- Sreelekshmi, R. C., Asokan, K., and Satheesh Kumar, K.: Deterministic nature of the underlying dynamics of surface wind fluctuations, *Ann. Geophys.*, 30, 1503–1514, doi:10.5194/angeo-30-1503-2012, 2012.
- Sugihara, G. and May, R. M.: Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, *Nature*, 344, 734–741, 1990.
- Takens, F.: Detecting strange attractors in turbulence, in: *Dynamical systems and turbulence*, Warwick 1980, 366–381, Springer, 1981.
- Wang, X. and Hui, L.: Multiscale prediction of wind speed and output power for the wind farm, *Journal of Control Theory and Applications*, 10, 251–258, 2012.