ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/234877884

Extreme events in surface wind: Predicting turbulent gusts

Article · December 2004

DOI: 10.1063/1.1846492

CITATIONS
7

READS
542

4 authors, including:

Detlef Holstein
temporarily no affiliation
9 PUBLICATIONS **229** CITATIONS
SEE PROFILE

Mario Ragwitz
Fraunhofer Institute for Energy Infrastructures and Geothermal Systems
162 PUBLICATIONS **4,471** CITATIONS
SEE PROFILE

Nikolay K. Vitanov
Bulgarian Academy of Sciences
228 PUBLICATIONS **4,611** CITATIONS
SEE PROFILE

All content following this page was uploaded by Nikolay K. Vitanov on 21 May 2014.

The user has requested enhancement of the downloaded file.

EXPERIMENTAL CHAOS

8th Experimental Chaos Conference

Florence, Italy 14 – 17 June 2004

EDITORS

Stefano Boccaletti
Istituto Nazionale di Ottica Applicata, Florence, Italy

Bruce J. Gluckman
George Mason University, Fairfax, VA

Jürgen Kurths
Universität Potsdam, Potsdam, Germany

Louis M. Pecora
Naval Research Laboratory, Washington, DC

Riccardo Meucci
Istituto Nazionale di Ottica Applicata, Florence, Italy

Oleg Yordanov
American University in Bulgaria, Blagoevgrad, Bulgaria

SPONSORING ORGANIZATION
Office of Naval Research

AMERICAN INSTITUTE OF PHYSICS

Melville, New York, 2004
AIP CONFERENCE PROCEEDINGS ■ VOLUME 742

Extreme events in surface wind: Predicting turbulent gusts

Holger Kantz,[1*] Detlef Holstein,[1] Mario Ragwitz,[2] Nikolay K. Vitanov[1,3]

[1]Max Planck Institute for the Physics of the Complex Systems,
Noethnitzer Str. 38, 01187 Dresden, Germany
[2]Fraunhofer Institute for Systems and Innovation Research,
Breslauer Str. 48, 76139, Karlsruhe, Germany
[3]Institute of Mechanics, Bulgarian Academy of Sciences
Akad. G. Bonchev Str., Block 4, 1113, Sofia, Bulgaria

* E-mail: kantz@mpipks-dresden.mpg.de.

Abstract

The potential to create extreme events is an inherent property of complex systems. Since our highly structured society is particularly sensitive to extreme events such as larger power failures in electric networks, stock market crashes, epedemics caused by new types of viruses, flash floods by summer storms, their potential predictability is of highest relevance. In this contribution we assume a physical point of view and concentrate on a specific phenomenon, namely on turbulent wind gusts. We show how a rather general model, namely a continuous state Markov chain, can be employed for data driven predictions of strong wind gusts. A Markov chain can represent arbitrary finite memory processes within the range of deterministic chaotic systems on the one extreme to uncorrelated white noise on the other, but its particular strenght lies in between: Nonlinear stochastic processes. Clearly, the modelling of the turbulent flow at a single site by a Markov chain is an approximation, whose accuracy will be discussed in the talk. From a statistical point of view, the focus on the prediction of extreme events implies the usage of unconventional cost functions, such that our predictor does not neccessarily perform well on "normal" bulk events, but has a surprisingly good performance on extreme events.

Flashfloods, earthquakes, stock market crashes and alike are extreme events in complex systems. The complexity of the underlying dynamics on the one hand makes it hard to predict such events, since the dynamics is usually non-preriodic with many aspects of randomness. On the

CP742, *Experimental Chaos: 8th Experimental Chaos Conference*, edited by S. Boccaletti et al.
© 2004 American Institute of Physics 0-7354-0226-4/04/$22.00

other hand, complexity enables the system to generate extreme events, which are large deviations from the average behaviour of the system. Without dwelling on this aspect too much, one can state that the presence of many more degrees of freedom than conservation laws, that nonlinear feedback loops, sensitive dependence on perturbations of a system's state or parameters, long transient times and other properties of complex systems allow the state vector to deviate grossly from its mean, hence, typical observables might assume exceptionally large values from time to time, which are extreme events. In view of the potential damage of extremes in our environmental systems, and moreover considering the fact that many evolving systems are much more shaped by extremes than by the majority of bulk events, the prediction of individual extreme events is of utmost relevance.

In this paper we will restrict ourselfs to the phenomenon of strong turbulent bursts in the local velocity of the surface wind field. We will motivate a prediction scheme for such data, demonstrate its performance, and discuss the relation to extreme events.

Many time series data sets contain a very strong aperiodic, seemingly random component. Nonetheless, one might suspect nontrivial temporal correlations to be hidden inside the data, which might be used, e.g., for the prediction of future observations. With reference to the paradigm of deterministic chaos, in recent years numerous data sets have been tested for the presence of deterministic structures. The corresponding set of methods, sometimes called nonlinear time series analysis[1], have proven their power in many physical laboratory experiments such as mechanical devices, electric resonance circuits, well controlled hydrodynamical convection experiments, simple chemical oscillations and many more. In many cases, deterministic model equations were contructed from observed data, which, in principle, could be used for short range predictions. However, even more data sets remain where such an approach fails.

Nonstationarity is one reason for the failure of such deterministic prediction schemes on data from field measurements. The other reason lies in the fact that a restricted description by a few marcoscopic observables in most systems involves stochastic forces, which represent the ignored (and typically experimentally unobservable) microscopic degrees of freedom.

A typical example of this scenario is hydrodynamic turbulence. Turbulence is one of the most complicated dynamical phenomena we are aware of, involving a huge range of time and length scales. From the modelling point of view, turbulence is found as solutions of the well known deterministic Navier Stokes partial differential equations. Using the current values of the velocity field in a whole volume around a certain point as initial conditions, numerical solutions [2] of these equations would supply accurate short term predictions of the velocity of a turbulent fluid at this point. However, experimental measurements of the intial conditions everywhere in the bulk are typically unavailable. This lack of information can then only be compensated by a suitable ensemble average over guessed initial conditions, which would lead to a probabilistic description.

In the following, we will not assume that we know any model equation of the system underlying a given experimental data set. Instead, we derive a data driven stochastic model for recurrent time series data, which will be applied to measurements of the wind speed at a given point. We will be able to make predictions of extreme events to follow in a probabilistic way.

The predictive power will be verified and characterised by statistical methods. As a particularly astonishing result, we will be able to predict the sudden increase of the wind speed by more than 3 m/s inside the next 2 s time interval correctly in 70% of such events, at the cost of only 10% false alarms. This might be of benefit for the save operation of wind turbines, whose rotor blades suffer from strong turbulent gusts. Predicted gusts could be made innocent by a small change of the pitch angle of these blades, since through aerodynamical effects, a rather small change of this angle can lead to a strong reduction of the mechanical load on the blade.

The concept of embedding, first proven by Takens[3] forms the basis for the analysis of time series data with a dominantly low-dimensional deterministic component [3, 4]. In this framework, a sequence of scalar time series measurements equidistant in time, $\{v_n\}$, $n = 1, \ldots, N$, is converted into a sequence of m-dimensional vectors, which are composed of successive time series elements, $\mathbf{v} = (v_n, v_{n-1}, v_{n-2}, \ldots, v_{n-m+1})$. For ideally deterministic low dimensional data, two such vectors at successive times are related to each other by a unique deterministic map, if $m > 2D_f$, where D_f is the dimension of the attractor on which the dynamics lives. This map can be extracted from observed data through the time evolution of the neighbouring points in the reconstructed phase space[5]. Time delay embedding builds the basis of nonlinear time series analysis and has been applied successfully for modelling and prediction of future measurements in many physical laboratory experiments[1].

Evidently, many time series recordings from natural phenomena do not represent a low-dimensional deterministic system. Then, a much more natural hypothesis is that deterministic nonlinear feedback loops together with stochastic inputs act on the system's variables and generate an aperiodic time evolution. Mathematically, this is described by a vector valued stochastic differential equation, called Langevin equation in the physical literature[6, 7], $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\xi(t)$ with white noise $\xi(t)$. The corresponding stochastic process is a Markov process, i.e., the probabilities to observe future states of the system depend exclusively on the current state, without influence from the past.

Starting from a given (experimental) time series we assume that the measurements v_n are obtained by a time-discrete sampling of some scalar measurement function applied to the state vectors $\mathbf{x}(t_n)$ of some underlying vector valued Markov process. If we had full knowledge of this process and of the state vectors, we could compute the probability to find a value v' in the next measurement as $p(v_{n+1} = v'|\mathbf{x}(t_n))$. Without requiring any knowledge of this Markov process, we want to construct a stochastic process in v whose properties are compatible with the observed data $\{v_n\}$. Unfortunately, we cannot expect that the recorded observable represents a one variable continuous state Markov chain of finite order [7]. In general, we have to expect the presence of infinite memory. However, in many applications correlations decay fast with the time lag. For these cases we will approximate the stochastic dynamics of v by a continuous state Markov chain, whose order m_0 is a model parameter we have to choose. The present state of the underlying Markov process $\mathbf{x}(t_n)$ is hence replaced by the last m measurements of the observable v, so that the probability to observe v' in the next measurement is approximated by $p(v_{n+1} = v' \mid v_n, v_{n-1}, \ldots, v_{n-m_0+1})$. The vector $(v_n, v_{n-1}, \ldots, v_{n-m_0+1}) = \mathbf{v}_n$ is formally identical to a delay vector used in the time-delay embedding approach outlined above (compare

also [8]). Nonstationarity can be taken into account, if we assume that a slow time dependence of the conditional probability $p(v_{n+1}|\mathbf{v}_n)$ can be traced back to the modulation of some parameters $\mathbf{a}(t)$. Then, $\mathbf{a}$ has to be included as additional condition, i.e., we should employ $p(v'|\mathbf{x}(t_n), t) = p(v'|\mathbf{x}(t_n), \mathbf{a}(t)) \approx p(v_{n+1}|\mathbf{v}_n, \mathbf{a}(t))$. Since $\mathbf{a}(t)$ is typically unknown we replace it by an appropriate increase of the dimension $m > m_0$ of the conditioning vector, which is called overembedding, $p(v_{n+1}|\mathbf{v}_n, \mathbf{a}(t)) \approx p(v_{n+1} \mid v_n, v_{n-1}, \ldots, v_{n-m+1})$. Overembedding is well justified in the framework of Takens's theory for deterministic systems [9], whereas it can hold for stochastic Markov chain modelling only in an approximate sense, namely, that formally infinitely many additional time lags have to be included, where we assume that a finite number yields a good approximation [10]. Indeed, quantities which can be thought of as being responsible for non-stationarity in the wind data treated below, such as surface and air temperatures or pressures typically fluctuate at low amplitudes on long time scales.

In order to implement the method we have to choose a value m of the number of time steps representing the memory. Moreover, we have to assume that the conditional probabilities are smooth in their arguments $\mathbf{v}_n$, i.e., that similar conditions give rise to a similar probability distribution. Geometrically, the current state vector $\mathbf{v}_n$ is a point in an m dimensional space. Neighbouring points in this space represent similar state vectors. Due to the assumed continuity, the observed "futures" v_{k+1} of close neighbours $\mathbf{v}_k$ form a random sample of the unknown distribution $p(v_{n+1} \mid \mathbf{v}_n)$. Thus we can estimate different moments of this empirically obtained sample which can be employed for different types of predictions.

We denote by $\Phi_\epsilon(\mathbf{v}_n)$ a neighbourhood of small diameter ϵ around the vector $\mathbf{v}_n$. The number of vectors $\mathbf{v}_k$ in this neighbourhood taken from the past of the time series, $k < n$, is denoted as $|\Phi_\epsilon(\mathbf{v}_n)|$. For these vectors we inspect the future values v_{k+1} and denote the number of these values which are in the interval $[v', v' + \Delta v']$ as $N(v', \Delta v')$. Then an estimate of the conditional probability at the first future step is

$$p(v_{n+1} = v' \mid \mathbf{v}_n)\Delta v' \approx \frac{N(v', \Delta v')}{|\Phi_\epsilon(\mathbf{v}_n)|} \tag{1}$$

The optimal prediction $\hat{v}$ of v_{n+1}, in the maximum likelihood sense, i.e., which in an ensemble average minimizes the root mean squared prediction error, is given by the first moment of this estimated conditional probability,

$$\hat{v}_{n+1} = \int dv'\, v' p(v' \mid \mathbf{v_n}) \approx \frac{1}{|\Phi_\epsilon(\mathbf{v}_n)|} \sum_{\mathbf{v}_k \in \Phi_\epsilon(\mathbf{v}_n)} v_{k+1}\,. \tag{2}$$

This prediction scheme is identical to what has been proposed before for deterministic data [5] (called Lorenz' method of analogues), but has now gained a novel justification for non-deterministic and non-stationary time series. Our probabilistic approach supplies additional information, if we pose probabilistic questions, whose answers involve more features of the estimated conditional probabilities than only its mean. One example is the prediction of the probability of an increment $\Delta v_{n+1} = v_{n+1} - v_n$ larger than g in the next time step. It is

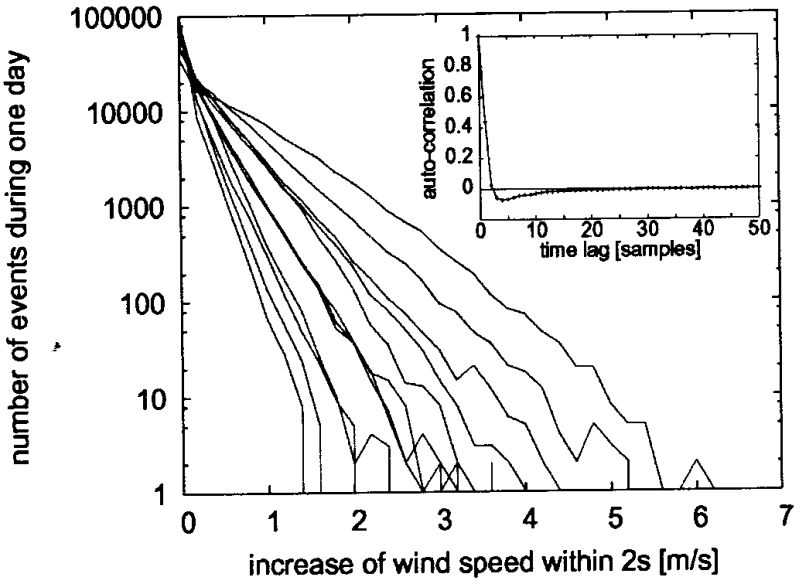

Figure 1: Histogram of the relative frequency of wind gusts within 2s (h=16) for 12 different days. The results below are obtained from the most turbulent day no.191 (topmost curve). Inset: the auto-correlation function of increments $\Delta v_{n+1} = v_{n+1} - v_n$ on day 191. On larger lags it is statistically consistent with zero.

straightforwardly given by the following integral of the probability distribution,

$$p(\Delta v_{n+1} > g) = \int_{v_n+g}^{\infty} dv'\, p(v' \mid \mathbf{v}_n) \tag{3}$$

which can be directly estimated from the data exploiting Eq.(1).

We illustrate the method by its application for predictions of the probability of a turbulent wind gust to arrive at a measurement device. Strong gusts are rare events, as it can be seen in Fig.1. Whereas the wind speed is strongly correlated in time, the increments Δv_n are almost δ-correlated (the positive correlation at time lag 1 is a signature of the inertia of the cup anemometer in use), with a weak anti-correlation over less than one second. Conventional prediction schemes based on second order statistics (e.g., autoregressive models) would not possess significant predictive power.

Since wind speeds are correlated, predictions refering to time intervals covering h time steps, $h > 1$, require a suitable generalisation of Eq.(3). The relevant probability is straightforwardly given by the relative number of ϵ-neighbours $\mathbf{v}_k$ whose future fulfils our criterion to be a gust,[1] as illustrated by Fig. 2. As a result, for every time step n we create a prediction $\hat{p}_n^{(\mathrm{gust})}$ of

[1] We defi ne a gust of strength g to be a situation where the maximum wind speed in the following time window of

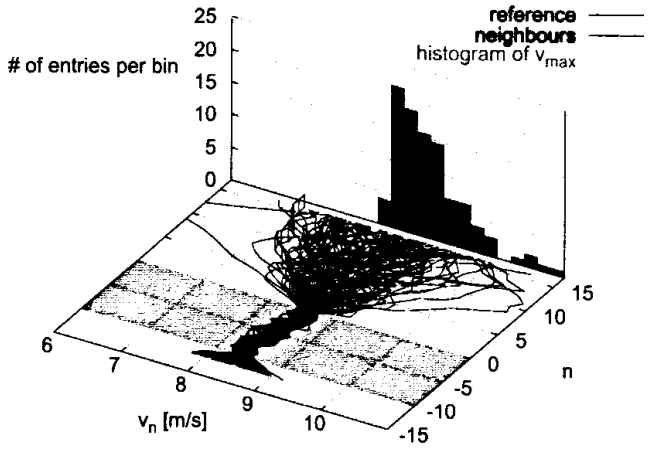

Figure 2: Sketch of the probabilistic prediction scheme. The red time series segments of the wind velocity are the elements of the neighbourhood of the reference trajectory (printed in blue). The neighbourhood is defined through the time indices ranging from -10 to 0 (green region), where the differences between the values of the reference trajectory and each neighbour are required to be smaller than some ϵ. Outside the interval [-10,0] the data are not constrained, and consequently the neighbouring trajectories spread much for positive n. This illustrates the lack of determinism. However, the distribution of future events such as the maximun value of each neighbour depends strongly on the values of the reference trajectory and hence on the sample of the neighbours thus selected. The vertical part of the figure shows the histogram of the maximum wind velocity between the timesteps $n = 1$ and $n = 16$. From this histogram one can easily obtain the probability of a gust for arbitrary gust intensities g.

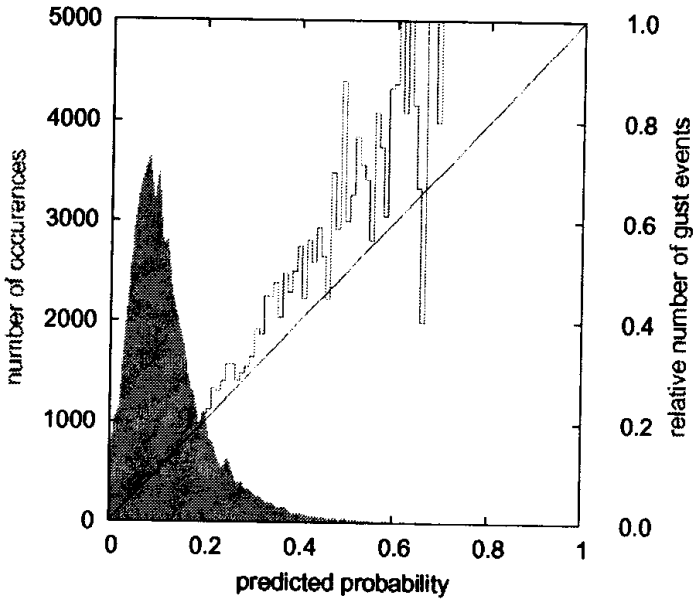

Figure 3: Checking the correctness of predictions by a reliability plot for 50000 predictions during a 24 h record of the wind velocity in Lammefjord. We select sub-samples of events, where the predicted probability for a gust was close to x. If the prediction is, in the probabilistic sence, perfect, then a fraction of x of the events inside this sample should be gust events. In fact, the relative number of gust events found empirically (red curve) is in reasonable approximation identical to the value x, represented by the diagonal. The green histogram shows the number of events, where the value x was predicted. Most predictions yield low probabilities. Only very few high probability events are in the data set. This agrees with the fact that gusts are rare events.

the probability of a turbulent gust to occur within the following time window of h time steps.[2]

A verification of the predicted probabilities is contained in what sometimes is called the reliability plot. We create sub-samples according to the predicted probabilities and verify that the relative number of gust events inside this sample agrees with the predicted probability, i.e., we construct the sample $\Sigma_x = \{n : \hat{p}_n^{(\mathrm{gust})} \in [x, x + \Delta x]\}$ and compute

$$r(x) = \frac{\text{Number of gust events in } \Sigma_x}{|\Sigma_x|} \,. \tag{4}$$

h measurements exceeds the last observation by more then g m/s. Our prediction scheme has similar performance when using other gust defi nitions.

[2]For the Ljammefjord 8 Hz data [11] we chose $m = 10$ and variable ϵ requiring at least 40 neighbours in $\Phi_\epsilon(\mathbf{v}_n)$, when $h = 16$ (= 2s). Preliminary studies suggest to increase the embedding window m when increasing the prediction horizon h.

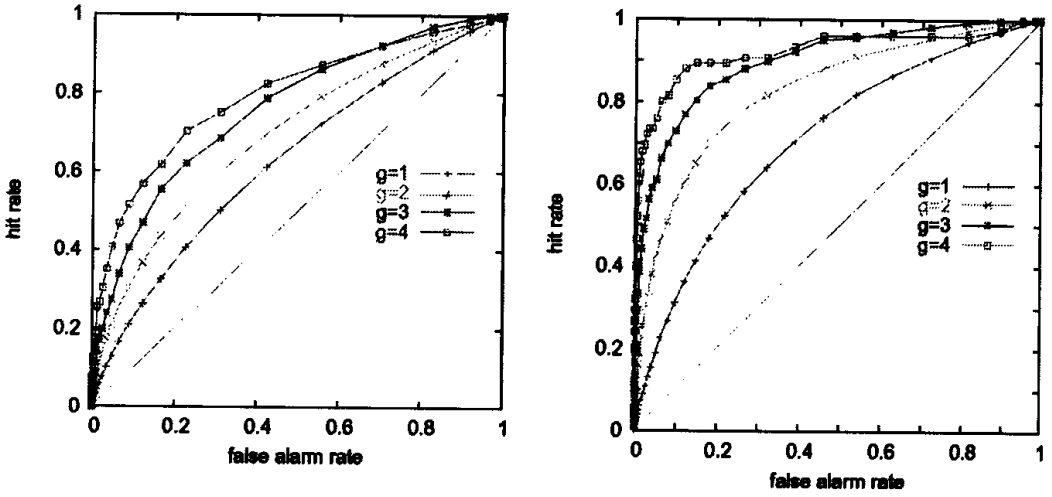

Figure 4: Relative number of correctly predicted gusts (hit rate) versus the relative frequency of false alarms for gusts exceeding 1, 2, 3, and 4 m/s in 2s. left panel: univariate data taken at a height of 20m. right panel: bivariate data taken at heights of 20 and 30m, predicting the wind speed at 20m.

In fact, our predictions pass this test, as the histogram close to the diagonal in Fig. 3 demonstrates.

In order to quantify the predictive power of our scheme in a different, maybe more practical way, we generate a "warning" at every time when the predicted probability exceeds a critical value, $\hat{p}_n^{(\mathrm{gust})} > p_c$. In Fig.4 we report the relative number of correctly predicted gusts of different strength g as a function of false alarms (which are tuned by varying p_c from 0 to 1) (ROC statistics [12]). If warnings were given randomly without correlation to the data, the three curves would coincide with the diagonal. Instead, we see that a significant percentage of the strong gusts (increase of 3 m/s and more) is correctly predicted at the cost of about 10% false alarms. The rate of correct predictions can be enhanced considerably, if we extend the scheme to bivariate data, using as inputs the wind speed at two different heights measured on the same mast. The probable reason for the improvement lies in the fact that wind speeds at different heights are correlated, and that the wind field in higher layers is usually ahead of the wind field in lower layers. Consequently, when predicting the speed at 30m above ground using the same bivariate inputs does only lead to a negligible improvement with respect to the univariate scheme.

Our continuous state m-th order Markov chain approach for the stochastic modelling of wind speed data is very different from previous stochastic approaches based on first-order Markov chains, where a discrete set of wind states is considered and global transition matrices [13] are extracted from the data. Our states are vectors formed by real numbers, and the transition probabilities are extracted from the data on-line for every given actual state separately,

not fitting some prescribed distribution[14]. For the prediction of extreme events, probabilistic forecasting is more significant than deterministic predictions such as Eq.(2), which is consistent with the fact that these wind speed data cannot be assumed to have a deterministic time evolution. The method exploits nonlinear higher order temporal correlations in wind speed data without computing them explicitly. Introducing a threshold on the predicted probabilities, extreme events can be correctly forecasted at a reasonable rate. Clearly, this scheme can be employed for other data with local Markov property as well.

Coming back to the issue of extreme events, we want to point out that our prediction scheme performs the better the more extreme the events to be predicted are. The larger the increase g of the gusts to be predicted, the higher is the rate of "hits". It has to be clarified by future work whether this means that extreme events are better predictable than bulk events due to the existence of clear precursors, or whether the enhanced predictability is a statistical effect. In any case, the findings reported here are encouraging in the sense that in seemingly random data there might be sufficient structure for a better than average predictability.

References

[1] H. Kantz, T. Schreiber. *Nonlinear Time Series Analysis.* (Cambridge University Press, Cambridge, 1997)

[2] P. Moin, K. Manesh. *Annu. Rev. Fluid Mech.* **30**, 539 (1998).

[3] F. Takens. *Lecture Notes in Mathematics* **898**, 366 (1981).

[4] T. Sauer, J. Yorke, M. Casdagli. *J. Stat. Phys.* **65**, 579 (1991).

[5] J. D. Farmer, J. J. Sidorovich. *Phys. Rev. Lett.* **59**, 845 (1987).

[6] H. Risken. *The Fokker- Planck Equation* (Springer, Berlin, 1989).

[7] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[8] F. Paparella, A. Provenzale, L. A. Smith, C. Tarrica, R. Vio. *Phys. Lett. A* **235**, 233 (1997).

[9] R. Hegger, H. Kantz, L. Matassini, T. Schreiber. *Phys. Rev. Lett.* **84**, 4092 (2000).

[10] H. Kantz, M. Ragwitz. *Int. J. Bif. Chaos* (to appear June 2004).

[11] Lammefjord data obtained from the Risø National Laboratory in Denmark, `http://www.risoe.dk/vea`, see also `http://www.winddata.com`.

[12] J. A. Hanley and B. J. McNeil. *Radiology* **143**, 29 (1982).

[13] A. D. Sahin, Z. Sen. *J. Wind Eng. Ind. Aerodynamics*. **89**, 263 (2001).

[14] F. Y. Ettoumi, H. Sauvageot, A. -E. -H. Adane. *Renewable Energy* **28**, 1787 (2003).

[15] We thank to the Alexander von Humboldt Foundation and to NCSR of Ministry of Education and Science of Bulgaria, contract # MM 1201/02 (N.K.V.) and to the German Ministry of Economy (M. R.) for the support of our research.

View publication stats