Symbolic Dynamics Applied to Velocity Time-series in Wind Farms

K. Karamanos[1], I.S. Mistakidis[2], S.I. Mistakidis[3] and I.E. Sarris[4,]

[1] *Complex Systems Group, Institute of Nuclear and Particle Physics, NCSR Demokritos, GR-15310, Aghia Paraskevi, Attiki, Greece.*
[2] *Department of Physical Sciences and Applications, Hellenic Army Academy, Vari, Attiki, Greece.*
[3] *Zentrum für Optische Quantentechnologien, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany.*
[4] *Department of Energy Technology, Technological & Educational Institute of Athens, Ag. Spyridona 17, 12210 Athens, Greece.*

Abstract
The development of standards for wind farms, presupposes the correct description of wind potential and this can be done with the field measurements of wind flow by cup anemometers. The utilization of new concepts, coming from the world of Cybernetics of Nonlinear Science and Complex Systems could open the road to uncover information hidden in both the mean polar velocity and the mean angle time-series. In particular, with the use of block entropies, it is shown that we can achieve a better and deeper understanding of the phenomenon of filtered turbulence, produced by time-series of the average wind velocity logged every ten minutes. The present analysis allows in principle a characterization of the experimental time-series in terms of the complexity for selected stationary windows of the signal, as well as the underlying mechanisms of the filtered turbulence.
Keywords wind farm, anemometer, wind velocity measurements, filtered turbulence, symbolic dynamics, block-entropies.

1 Introduction.

In the context of energetic problems of modern societies, a proposed solution has been the use of alternative energy sources. One particular realization of this idea is for instance the installation and usage of wind farms for electricity. A cluster of wind turbines in the same site used to produce energy is called wind farm. Investment in a wind farm consults temporal measurements of the wind potential on the prospective site by using suitably located towers. The wind velocity, the pressure and temperature, are collected in the measurement tower by cup anemometers and meteorological instruments [1]. The height of the tower is up to the hub height of the planned wind turbines, and the data are recorded frequently, i.e. contains the averaged quantities every ten minutes, and for at several months. These data, in principle, may allow the developer to decide if the investment in a wind farm is economically feasible for the selected site.

Understanding and quantify aspects of their behaviour and nature is one of the most important and challenging problem of modern engineering [2]. The long recordings of the averaged wind velocity time-series from the anemometer, are a kind of filtered turbulent temporal data set with intermediate properties that may be analysed. The most important problem in this framework is the prediction of the production of electric energy in wind farms [2]. The problem of prediction is exactly, what connects these studies with cybernetics and nonlinear science.

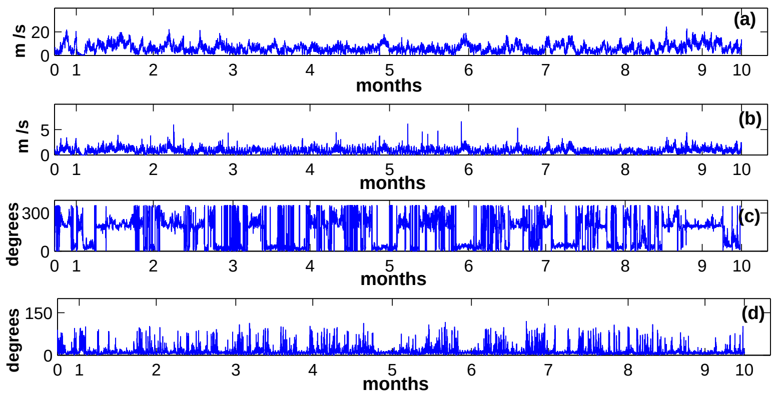

Fig.1 Time-series of the velocity field between 23 March 2005 and 16 December 2005. The collected data have been obtained from the mountains of Peloponnesus, Greece. Shown are the distributions over time (obtained from the anemometer) of (a) the mean polar velocity in m/s, (b) the standard deviation of the mean velocity, the distribution of (c) the mean angles (degrees) and (d) the standard deviation of the angles.

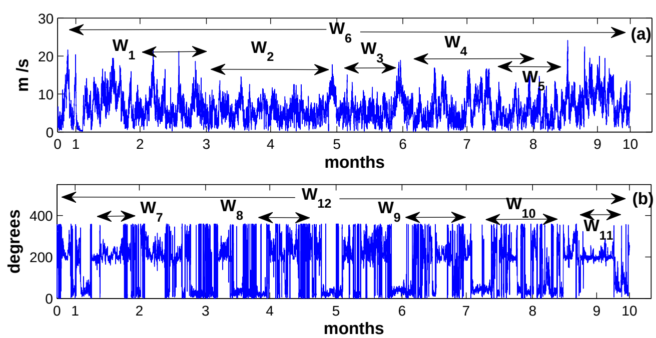

Fig.2 Segments of experimental time-series (a) for the mean value of the velocity and (b) the mean angle of the anemometer depicted in Fig.1. The vertical lines indicate the months during the measurements.

The development of standards for wind farms presupposes the correct description of wind potential and this can be done only with the field measurements of wind force by anemometers. Herein, we utilize new concepts coming from the world of Nonlinear Physics, to quantify and understand the relative complexity of wind signals of filtered or mild turbulence. Although field measurements of filtered turbulence by mechanical purpose anemometers presuppose enormous and violent averagings on the molecular nature of turbulence, also averagings in time (months, epochs), not to mention the inertial phenomena in the measuring apparatus. We intend to show that a coherent and self-consistent mathematical description of turbulence within the arsenal of Nonlinear Physics is not only possible, but also beneficial for physicists and engineers.

In this paper, one set of wind speed data from one measurement tower situated in the mountains of the region Achaia, Peloponnesus, Greece, in the form of polar velocity and angle was analyzed. The data set consists of some thousands of wind speed values, recorded over every ten minutes by a cup anemometer, covers the period March 2005 to December 2005. More specifically, the wind polar speed and angle was measured using cup anemometers and the 38500 recordings logged into a digital anemograph logging equipment system.

The complexity and variability of the data is analyzed here by the use of nonlinear techniques, and the information hidden in both the mean polar velocity and the mean polar angle time series are discussed.

More precisely, the temporal evolution of nonlinear characteristics is studied by applying a recently proposed technique [4-5]. The original continuous time, of the mean polar velocity/angle data are projected to symbolic sequence and a block entropy analysis by the novelty of lumping follows [4-5].

The paper is articulated as follows : In Sec 2, we recall basic facts about symbolic sequences, and the block entropy analysis by lumping. Sec 3, will be devoted to the application of the entropy analysis by lumping to anemometer recordings. Finally, in the last Sec 4, we draw the main conclusions and discuss future plans.

2 Symbolic dynamics

A way to examine transient phenomena is to analyze the original time series (anemometer recordings) into a sequence of distinct time windows (epochs). The basic aim is to discover a clear difference of dynamical characteristics as time evolves by employing techniques from the toolbox of symbolic dynamics. In particular here we employ the notion of block entropy analysis by lumping [4-10]. Towards this direction, within a stationary time window, the block entropy serves as a measure of "complexity" of the signal. The lower the value of entropy, the more "ordered" it is. In the following, in order to proceed with the analysis of the experimental data we will briefly review the concepts of symbolic dynamics and

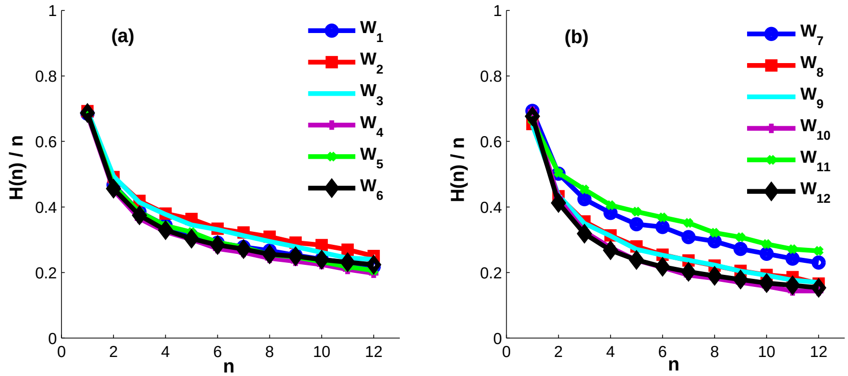

Fig.3 Block-entropy per letter as a function of the word length for the various stationary windows shown in Figs.2 (a) and (b). We observe a reduction of the block-entropy per letter which can be interpreted as a sign of complexity reduction of the respective time-window of the signal.

the notion of block entropies [3-18].

We restrict ourselves to the simplest possible coarse graining of the recording. This is given by choosing a threshold C and assigning the symbols "1" and "0" to the signal, depending on whether it is above or below the threshold (binary partition or bipartition). In this way, each stationary time window of the original aiolic time-series for a given threshold is transformed into symbolic sequences, which contains "linguistic" or "symbolic dynamics" characteristics.

More specifically, the block entropies, depending on the word-frequency distribution, are of special interest, extending Shannon's classical definition of the entropy of a single state to the entropy of a succession of states. Thus, each entropy takes a large (small) value if there are many (few) kinds of patterns, i.e it decreases while the organization of patterns is increasing. In this manner we can argue that the block entropy constitutes a measure the complexity of a stationary signal.

In particular, we estimate the block entropy by lumping [4-5]. Lumping is the reading of the symbolic sequence by "taking portions", as opposed to gliding [10-17] where one has essentially a "moving frame". In general, the basic novelty of the analysis by lumping is that, unlike the Fourier transform or the conventional entropy by gliding, it gives results that can be related to algorithmic aspects of the sequences. It is useful to transform the initial raw data of the anemometer recording into symbolic sequences taking values in the alphabet $\{0, 1\}$, according to the rules $A_i = 1$ if $A(t_i) > E[A(t_i)]$ and $A_i = 0$, if $A(t_i) < E[A(t_i)]$. Here the quantities $A(t_i)$ denote the values of the measured mean velocity/polar angle at

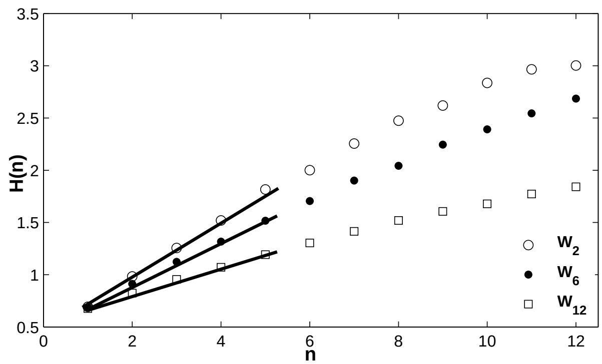

Fig.4 The observed scaling of the Block Entropy $H(n)$ (Eq. 6) as a function of the word length n. The experimental values show a linear best fit for small word lengths with a very good precision. The slope of the line gives the Kolmogorov-Sinai entropy which in 1D coincides with the Lyapunov exponent. This scaling is consistent with the corresponding theoretical predictions, see Nicolis and Gaspard [13].

time t_i and $E[A(t_i)] =< A(t_i) >$ is the mean value in the particular time windows.

So, let us consider a subsequence of length N, selected out of a very long (theoretically infinite) symbolic sequence. We stipulate that this subsequence is to be read in terms of distinct "blocks" of length n, i.e :

$$\cdots \underbrace{A_1 ... A_n}_{B_1} \underbrace{A_{n+1} ... A_{2n}}_{B_2} ... \underbrace{A_{jn+1} ... A_{(j+1)n}}_{B_{j+1}} \cdots . \tag{1}$$

We call this reading procedure "lumping" and we shall implement it in the following to the experimental time-series of the filtered turbulence. It is also useful to mention the following quantities which give the information content of the sequence :

– The dynamical (Shannon-like) block-entropy for blocks of length n has the form :

$$H(n) = \sum_{(A_1,...,A_n)} p^{(n)}(A_1, ..., A_n) \ln p^{(n)}(A_1, ..., A_n), \tag{2}$$

where the probability of occurrence of a block $A_1, ..., A_n$, denoted $p^{(n)}(A_1, ..., A_n)$, is defined by the fraction (when it exists) in the statistical limit as follows :

$$\frac{No\ of\ blocks A_1, ..., A_n encountered\, when\, lumping}{total\ No\ of\, blocks\, when\, lumping}, \tag{3}$$

Table 1 The Kolmogorov-Sinai (KS) entropy h, the estimated error δh and the corresponding percentage h/ln(2) (see Fig.5) in respect to the maximum value of the KS entropy for the different windows (see Figs. 2(a),(b))

Window No.	h	δh	$h/\ln 2(\%)$
W1	0.2192	0.0350	31.6
W2	0.2784	0.0163	40.2
W3	0.2588	0.0407	37.3
W4	0.2040	0.0128	29.4
W5	0.2301	0.0152	33.2
W6	0.2065	0.0194	29.8
W7	0.2611	0.0576	37.7
W8	0.1879	0.0414	27.1
W9	0.1781	0.0611	25.7
W10	0.1245	0.0398	18.0
W11	0.3140	0.0535	45.3
W12	0.1277	0.0201	18.4

starting from the beginning of the sequence. However, the associate entropy per letter reads :

$$h^{(n)} = \frac{H(n)}{n}. \tag{4}$$

– On the other hand, the entropy of the source (a topological invariant), defined in the limit (if it exists) reads :

$$h = \lim_{n\to\infty} h^{(n)}, \tag{5}$$

which is the discrete analog of metric or the Kolmogorov-Sinai entropy. Therefore, in order to determine the abundance of long blocks one is led to examine the scaling properties of $H(n)$ as a function of n (see also Fig.4).

3 Results in terms of symbolic dynamics.

To begin with, in Fig. 3 we depict the block entropy by lumping per letter as a function of the word length for the selected time windows that we present in Fig. 2. We note that a complete absence of structure in the signal, would lead to an horizontal line in the block entropy diagram. As one can observe this is not the present case. However, one important conjecture, due essentially to Ebeling and Nicolis [13-14] states that the most general (asymptotic) scaling of the block entropies takes the form

$$H(n) = y + nh + gn^{\mu_0}(\ln n)^{\mu_1}, \tag{6}$$

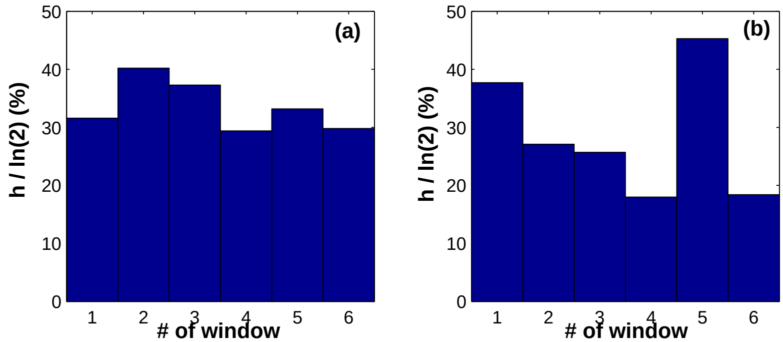

Fig.5 The normalized Kolmogorov-Sinai entropy (taken as the normalized slope of the linear part of the Block entropy H(n)), for the respective time-windows of (a) the mean velocities (W_1-W_6), and (b) the polar angles (W_7-W_{12}).

where y, h and g are constants and μ_0 and μ_1 are the corresponding non-scaling constant exponents.

In the following, we attempt to examine the behavior of eq.(6) for each of the twelve stationary windows under study, which are depicted in Fig. 2. Hence, in Fig. 4 we present the typical variation of the block entropy by lumping $H(n)$ as a function of the word length n for three representative windows. This study reveals that if we restrict ourselves to the first five values of $H(n)$, a linear scaling is observed with a great precision. In this manner, we next perform a least square method for this region and we estimate the slope h. Note, that the associated correlation coefficients r, are close to 1 with a precision better than 10^{-4}. Working similarly for the rest of the 9 time windows, we conclude that for $n < 6$ the same behavior is observed, i.e. the equation for the scaling of the block entropy by lumping, is transformed to the remarkably simple linear relation

$$H(n) = y + nh. \tag{7}$$

This means that $g = 0$, for $n < 6$. In Fig 2(a),2(b) we isolate 6 time windows for the mean velocity and 6 time windows for the polar angle respectively, which present a good overall stationary behavior according to our tests. Their corresponding KS entropy is given in Table 1. The KS entropy (complexity) of the whole window W_6 (mean velocity) is of the order of 30%. This shows that the underlying mechanism of filtered turbulence has an underlying organized molecular basis, as it does not correspond to a completely random process (Bernoulli shift). The maximum value of KS entropy for the mean velocity is about 41% for the window (W_2), while the minimum value is about 30% (see also Fig.5). A close inspection of the W_2 window, shows that it has a structure more reminiscent of noise and this is confirmed by its high value of KS entropy. The KS entropy of the whole window for the mean angle (W_{12}) is about 19% and it is quite low. The highest value is achieved in W_{11} which is of the order of 46%.

We note that when $g = 0$ and $h > 0$, long words are penalized exponentially. We focus on the quantity h, namely the Kolmogorov-Sinai entropy defined as the slope of eq. (7). We notice that for a one-dimensional process the Kolmogorov-Sinai entropy coincides with its Lyapunov exponent. The Lyapunov exponent under these conditions gives a measure of the chaoticity (or dynamical randomness) of the signal. For a two-letter alphabet, the Kolmogorov-Sinai entropy h takes values from zero to $\ln 2$ (see the discussion in Sec 2), so that one can normalize dividing by $\ln 2$ and obtain the respective percentage. Hence, it is important to note that the linear part of the scaling helps for a classification and categorization of the recordings. The question which arises naturally, is whether this is an independent algorithmic law of nature. This seems to be an open problem for the moment. However, our results strongly support this hypothesis.

Remark : We restrict ourselves to the region $n < 6$, because the maximum statistical accuracy for the block entropies by lumping is of the order of lnL, where L is the total number of points (the size of the window). In our case L, is of the order of 2800, so that $n < 8$ and due to the underestimation of the higher entropies, we have enough statistical precision for $n < 6$.

4 Conclusions.

In this paper, we made an attempt to apply techniques and methods from Cybernetics, Nonlinear Physics and General systems to the temporal unfolding of the phenomenon of filtered turbulence in wind farms as measured by cup anemometers. The tremendous molecular averaging implied by this simplistic procedure, has been equilibrated by the high specialization and the novelty of the techniques used in General systems.

Block entropy analysis by lumping, as introduced by Karamanos et al. [4-5], is for the first time used for the understanding and categorization of time series of filtered turbulence in anemometer recording. A first linear region has been revealed, as it has been already happened for EM preseismic precursors [6], cardiac signals of coronary patients [9], and DNA strands in oligonucleotide basis ACGT [8]. Hence, it is important to note that the linear part of the scaling helps us for a complete classification and categorization of the recordings. As it has been already pointed out, the question which arises naturally, is whether this is an independent algorithmic law of nature. This seems to be an open problem for the moment. However, our results strongly support this hypothesis.

Future projects include the enhancement of data and recordings, the cross-checking of many complexity measures, the month-to-month monitoring of the dynamics and the application of prediction techniques, with special use to wind parks.

References

[1] Probst, O. and Cárdenas, D. (2010), "State of the Art and Trends in Wind Resource Assessment", Energies, 3, 1087-1141.

[2] Huang, Z. and Chalabi, Z.S. (1995), "Use of time-series analysis to model and forecast wind speed", J*ournal of Wind Engineering and Industrial Aerodynamics*, Vol.56, No.2-3, pp.311-322.

[3] Nicolis, G., Rao, G., Rao, J. and Nicolis, C. (1988), Generation of spatially asymmetric, information-rich structures in far from equilibrium systems, in: Christiansen.
P. L., and Parmentier, R. D. (Eds.). (1989), "Structure, Coherence, and Chaos in Dynamical Systems", R. D., Manchester University Press.

[4] Karamanos, K., and Nicolis, G. (1999), "Symbolic dynamics and entropy analysis of Feigenbaum limit sets", *Chaos, Solitons and Fractals*, Vol.10, No.7, pp.1135-1150.

[5] Karamanos, K. (2001), "Entropy analysis of substitutive sequences revisited", *Journal of Physics A: Mathematical and General*, Vol.34, No.43, pp.9231.

[6] Karamanos, K., Peratzakis A., Kapiris P., Nikolopoulos S., Kopanas J. and Eftaxias K. (2005), "Extracting preseismic electromagnetic signatures in terms of symbolic dynamics", *Nonlinear Processes in Geophysics (NPG)*, Vol.12, pp.835-848.

[7] Karamanos, K., Dakopoulos, D., Aloupis., K., Peratzakis, A., Athanasopoulou, L., Nikolopoulos, S., Kapiris, P. and Eftaxias, K. (2006), "Preseismic electromagnetic signals in terms of complexity", *Physical Review E*, Vol.74, No.1, 016104, pp.21.

[8] Karamanos, K., Kotsireas, I.S., Peratzakis, A., and Eftaxias, K. (2006), "Statistical compressibility analysis of DNA sequences by generalized entropy-like quantities: towards algorithmic laws for biology?", *WSEAS Transactions on Systems*, Vol.5, No.11, pp.2503-2509.

[9] Karamanos, K., Nikolopoulos, S., Hizanidis, K., Manis, G., Alexandridi, A. and Nikolakeas, S. (2006), "Block entropy analysis of heart rate variability signals", *Int. J. Bif. Chaos*, Vol.16, No.7, pp.2093-2101.

[10] Khinchin, A.I. (1949), Mathematical foundations of statistical mechanics, Dover Publications.

[11] Mistakidis, I., Karamanos, K., and Mistakidis, S. (2013), "Statistical versus optimal partitioning for block entropies", *Kybernetes*, Vol.42, No.1, pp.35-54.

[12] Nicolis, G. (1995), Introduction to nonlinear science, Cambridge University Press.

[13] Nicolis, G., and Gaspard, P. (1994), "Toward a probabilistic approach to complex systems", *Chaos, Solitons and Fractals*, Vol.4, No.1, pp.41-57.

[14] Ebeling, W., and Nicolis, G. (1992), "Word frequency and entropy of symbolic sequences: A dynamical perspective", *Chaos Solitons and Fractals*, Vol.2, pp.635-650.

[15] Steuer, R., Molgedey, L., Ebeling, W. and Jimenez-Montano, M.A. (2001), "Entropy and optimal partition for data analysis", *Eur. Phys. J. B*, Vol.19, pp.265-269.

[16] Nicolis, G., Nicolis, C., and Nicolis, J. S. (1989), "Chaotic dynamics, Markov partitions, and Zipf's law", *Journal of Statistical Physics*, Vol.54, No.3, pp.915-924.

[17] Nicolis, J. (1991), "Chaos and information processing: A heuristic outline", *World Scientific*.

[18] Nikolopoulos, S., Kapiris, P., Karamanos, K., and Eftaxias, K. (2004), "A unified approach of catastrophic events", *Natural Hazards and Earth System Science*, Vol.4, No.5/6, pp.615-631.

[19] Sammis, C. G., and Sornette, D. (2002), "Positive feedback, memory, and the predictability of earthquakes", *Proceedings of the National Academy of Sciences of the United States of America*, Vol.99(Suppl 1), pp.2501-2508.

Corresponding Author
Author can be contacted at: sarris@teiath.gr.