CIMSA 2005 – IEEE International Conference on
Computational Intelligence for Measurement Systems and Applications
Giardini Naxos, Italy, 20-22 July 2005

Can We Estimate Atmospheric Predictability by Performance of Neural Network Forecasting? The Toy Case Studies of Unforced and Forced Lorenz Models

Antonello Pasini
CNR - Institute of Atmospheric Pollution
Via Salaria Km. 29.300, I-00016 Monterotondo Stazione (Rome), Italy
Phone: +39-06-9067-2274, Fax: +39-06-9067-2660, E-mail: pasini@iia.cnr.it

Vinicio Pelino
Met Service of the Italian Air Force, CNMCA
Aeroporto "De Bernardi", via di Pratica di Mare, I-00040 Pratica di Mare (Rome), Italy
Phone: +39-06-9129-3494, Fax: +39-06-9129-3254, E-mail: pelino@meteoam.it

***Abstract** – We present an analysis of the predictability for several regions on the attractor of the Lorenz-63 system, a simple non-linear model which mimics some features of the atmosphere, like its chaotic behaviour and the presence of preferred states or "regimes". In this framework, through a forecasting activity on the attractor, a multilayer perceptron shows its ability to recognise different values of predictability in various zones of the attractor, if compared with other estimations of local predictability, like the growth rates of the so called "bred vectors". Furthermore, following recent studies on the impact of weak imposed forcings on the Lorenz model, as a toy simulation of increased anthropogenic forcings on the climate system, we analyse the changes of predictability for a new scenario by neural network forecasting. Therefore, even if the present paper must be considered as a preliminary attempt at the use of neural networks for predictability assessments, this activity shows good results and opens perspectives of further improvements and applications.*

***Keywords** – neural networks, predictability, Lorenz model, bred vectors.*

I. INTRODUCTION

Since the very famous paper by Lorenz in 1963 [1], predictability has become a relevant topic in meteo-climatic studies. In particular, the discovery of the so called "deterministic chaos" led to change our perspective to medium- and long-range forecasting, both in the atmosphere and in the climate system. As a matter of fact, now a probabilistic approach is usually performed by ensemble integrations of deterministic models starting from perturbed initial conditions (see, for instance, [2] for an updated collection of papers on this topic). Inside this paradigm, however, two items of concern are the many feed-backs included in the models and the very high number of degrees of freedom of these dynamical systems that, for a standard general circulation model, usually reaches 10^7. This leads to run very complex and time-consuming models, so that their physical behaviour can be obscured by the mixing of feed-backs; furthermore, also the new approach in terms of ensemble integrations cannot be fully followed up, especially in climate simulations, when a very large amount of computer time is needed.

In this framework, a new interest is growing in simple dynamical systems which can mimic some features of the atmosphere and the climate system. Among them, the Lorenz-63 model [1] is the prototype of a large class of models of meteo-climatic relevance [3].

Some recent results supply us with the foundations for the research presented in this paper. In particular, in [4] the local predictability on the Lorenz attractor has been extensively studied by means of the so called bred-vector growth, that is to say through the rate of divergence of trajectories starting from very close points on the attractor itself. This research clearly shows the existence of regions of the attractor endowed with distinct local predictability values and also the possibility to predict regime changes.

Furthermore, the Lorenz model has been recently revisited in connection with studies about changes in atmospheric regimes induced by external forcings. In particular, Palmer and co-workers [5,6] showed that the observed climate change of the last decades "can be interpreted in terms of changes in the frequency of occurrence of natural atmospheric circulation regimes" [6]. In this framework, the Lorenz system represents a toy model in which we can recover a similar behaviour, if it is forced by a weak external forcing. From this viewpoint that forcing can be interpreted as the analogue of the increase of anthropogenic forcings in the real climate system. A further research showed that a weak forcing added to the original unforced Lorenz model can lead to a relevant increase in the frequency of occurrence of extremely persistent events [7]. At our knowledge, no study about predictability in the forced Lorenz model has been performed.

Here we explore the possibility of estimating local predictability on the attractors of both unforced and forced Lorenz models, through an evaluation of the performance of a neural network trained for forecasting trajectories some steps ahead on the attractors themselves. In doing so, we will compare our results with those of paper [4] for the original

Lorenz model and find some clear change of predictability in the case of a forced Lorenz system.

II. THE NEURAL TOOL

In the present paper we apply a neural network tool developed some years ago [8] and then extensively used in boundary layer investigations [9,10] and climate change studies [11,12].

The neural networks considered in this work are feed forward and endowed with a backpropagation training rule with both gradient descent and momentum terms; a usual quadratic cost function is chosen. A particular attention is paid to the form of transfer functions: we choose sigmoids whose arguments are normalised with respect to the number of connections converging to a single neuron of the hidden and output layer, respectively (this choice has shown its validity in order to avoid problems in cases of networks with many neurons [9]). An early stopping method is also used to furtherly prevent overfitting. Finally, our tool is endowed with particular schemes for the learning procedure in cases of "historical" data. For more details on this tool, see [8,9,11,12].

III. THE LORENZ MODEL

The Lorenz system is composed by the following equations:

$$\begin{cases} \dfrac{dx}{dt} = \sigma(y - x) \\ \dfrac{dy}{dt} = rx - y - xz \\ \dfrac{dz}{dt} = xy - bz \end{cases} , \qquad (1)$$

where our choice of the parameters ($\sigma = 10$, $b = 8/3$, $r = 28$, the same as in [1]) results in chaotic solutions. The integration of this system is performed *via* a 4[th]-order Runge-Kutta scheme with time step $\Delta t = 0.01$. Following [4], we first calculate the bred-vector growth $\delta\vec{x}$ (in 3 dimensions) over $n = 8$ time steps, starting from little perturbations $\delta\vec{x}_0$ around the points of Lorenz attractor previously found by integration of system (1). If we define the growth rate as

$$g = \frac{1}{n} \ln\left(\frac{|\delta\vec{x}|}{|\delta\vec{x}_0|} \right) , \qquad (2)$$

then we divide the bred-growth rates in four classes as in [4], even if with a slight different choice of thresholds (see Table 1 for our choice). A plot of the growth-rate classes on the Lorenz attractor gives clear evidence of the existence of regions of distinct predictability on the attractor itself (see Figure 1).

Thus, we go from blue points, where the perturbations are decaying, to red points, where a big spread between original and perturbed trajectories is evident.

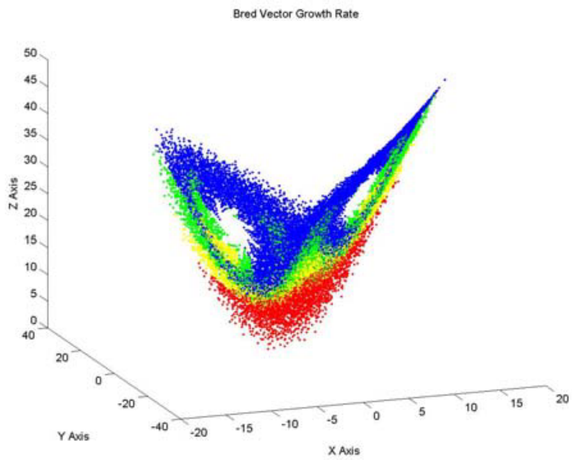

Fig. 1. Regions of different predictability on the Lorenz attractor.

IV. FORECASTING APPLICATION OF THE NEURAL MODEL IN THE LORENZ SYSTEM

The idea to forecast future states of the Lorenz system by neural networks is not new. Nevertheless, the known cases essentially deal with the prediction of the time series for a single variable of the Lorenz system (usually the x variable), with the attempt at reconstructing the complete dynamics under the conditions prescribed by the Takens theorem [13]: see, for instance, [14,15] and references therein. This permits to mimic the reconstruction of an unknown dynamics by observational data in a complex system. Here, instead, our aim is different. We consider the full dynamics of the Lorenz system (with 3 inputs, one for each variable, in this attempt) and try to estimate the predictability in several regions of its attractor by considering the neural forecasting performance. Furthermore, as shown elsewhere [16], the non-linear features of neural models allow us to explain the forecasting failures of the so called analogues method (first introduced by Lorenz [17]) in chaotic systems and to obtain better results.

Here we do not discuss how to choose the optimal topology for our networks. It is worthwhile to stress that we deal with 15 hidden neurons in a single hidden layer: this number is sufficient to obtain a good representation of the underlying function but not so big to determine some kind of overfitting.

Now, for each class of bred-growth rate, we train our neural model to make a single-step forecast from t_0 to $t_0 + n$ (here, of course, we choose $n = 8$, as in the bred-growth calculations). Attempts at performing 8 steps (to achieve the same forecast horizon) result in very poor performance, as already noted in other studies [10] and as known from theoretical considerations [18]. We divide our total set of Lorenz simulated data (20,000 input-target patterns) in a training set (80% of data) and a validation/test set (20%). As already noted, in this attempt we fix the topology of our networks by choosing 3 inputs (data co-ordinates in the Euclidean space), 15 neurons in a single hidden layer and 3 outputs (the 3-dimensional position after 8 time steps of integration). Furthermore, we choose to consider the 3D-Euclidean distance between output and target points as a measure of our forecast performance.

Table 1 shows that the neural model errors clearly depend on the bred-growth classes, giving us the best performance of neural forecast on the blue region and the worst one on the red points. As far as our performance on green and yellow classes is concerned, even if it is impossible to distinguish between them because each mean distance error falls inside the error bar associated to the other one, these values are however well separated from the performance values on the other two classes (blue and red). In this Table the mean distance error for each class is shown, together with error bars coming from ensemble runs of the model with different initial random weights, so that the network is able to widely explore the landscape of the cost function. These error bars represent $\pm 2\sigma$.

Table I. Performance of neural forecasts on the test set in terms of mean distance error.

Bred growth (class)	Bred-growth rate	Mean distance error in NN forecasts
Blue	$g < 0$	5.66 ± 0.15
Green	$0 \le g < 0.4$	6.58 ± 0.25
Yellow	$0.4 \le g < 0.64$	6.62 ± 0.24
Red	$g \ge 0.64$	8.36 ± 0.41

A closer look at our results, in terms of the distributions of distance errors for each class, allows us to appreciate that unimodal distributions (for blue and green classes) tend to split to quasi-bimodal distributions for yellow and red classes. In Figs. 2 and 3 we show these distributions for the blue and the yellow class, respectively. This fact implies a sensitivity of forecast performance to regions where a change of regime is possible and two close trajectories could evolve to opposite wings on the attractor. If one refers to Fig. 1, it is very clear, for instance, that the trajectories which end onto red points come from the middle of the attractor where they undergo a "bifurcation" towards the left or the right wing of the attractor itself.

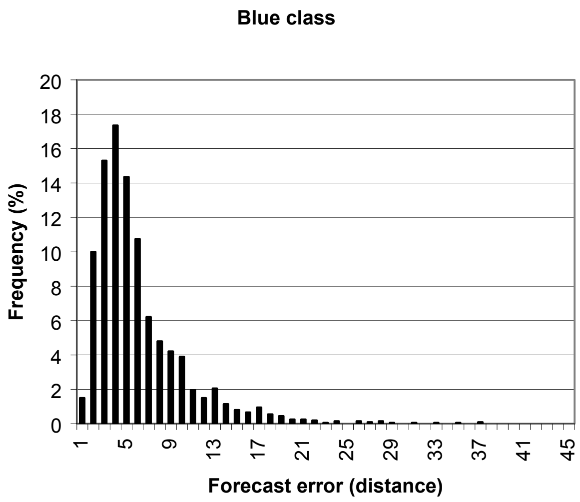

Fig. 2. Error distribution for the blue class.

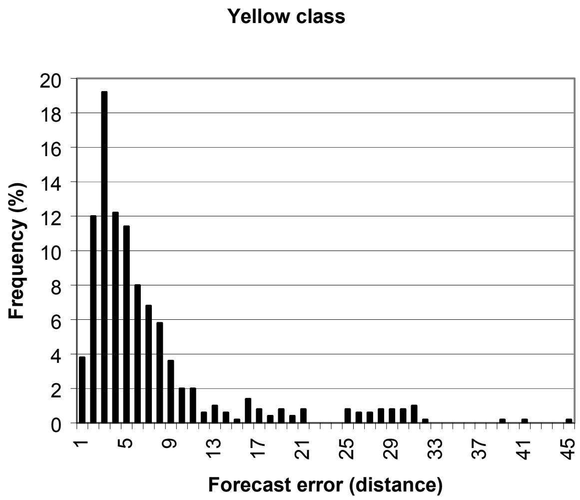

Fig. 3. Error distribution for the yellow class.

V. A PREDICTABILITY STUDY ON THE FORCED LORENZ SYSTEM *VIA* NEURAL NETWORKS

Once obtained a clear distinction among predictability classes on the classical Lorenz attractor by neural modelling, we move to consider a forced Lorenz model in order to see if any change in predictability can be seen in this case. We choose to add a weak constant forcing (= 2.5) to the right hand side of the second equation in system (1), as in [5,6] (case $f_0 = 2.5$ and $\theta = 90°$ in the formalism of Palmer and co-workers [5,6], that we are going to adopt here). This results in the following equations for the forced Lorenz system:

$$\begin{cases} \frac{dx}{dt} = \sigma(y-x) + f_0 \cos\theta \\ \frac{dy}{dt} = rx - y - xz + f_0 sin\theta \\ \frac{dz}{dt} = xy - bz \end{cases} , \quad (3)$$

The calculus of bred-vector growth allows us to depict a picture (here not shown) like Fig. 1, which results quite identical to that one, so showing that the attractor does not change substantially. As a matter of fact, a bred-growth analysis shows an increase (about 1%) in the total numbers of blue cases and a 2% decrease of the less numerous yellow and red ones, so suggesting a global increase of predictability on this new (very similar) attractor.

As a matter of fact, even the average forecast error of our neural model on all cases decreases slightly. Furthermore, we can appreciate a statistical quite significant forecast improvement specifically for the blue class (with a mean distance error = 5.41 ± 0.10). For the most critical red class a little increase in performance is found, even if it is not significant at the same degree of confidence. The forecast performance on the other two classes is very similar to the performance obtained for the analogous cases in the study of predictability on the unforced Lorenz model. Furthermore, some kind of bimodality in the distribution of forecast errors is present even now for yellow and red classes.

In any case, even if a general increase in neural performance can be observed in our runs, this is quite slight. In particular, this fact does not allow us to understand if this increased performance is due to a more frequent permanence of the system's state in regions of high predictability (blue points), or to a change in local predictability of single points in the Euclidean 3D-space, or, finally, to both of these phenomena.

At present, further work is in progress in order to analyse these results in more detail, with a particular attention to their relationships with the observed changes both in the frequency of occurrence of different regimes and in the persistence times of trajectories inside them.

VI. DISCUSSION

In this paper a clear relationship between local predictability in four regions of the attractors related to unforced and forced Lorenz systems and the forecasting performance of a neural network model has been established. Of course, this does not mean immediately that now we are able to answer "yes" to the question included in the title of this paper. First of all, because the Lorenz model is only a toy model which, anyway, mimics some features of the real atmosphere system, like its chaotic behaviour and the presence of preferred states, sometimes called "regimes". Secondly, because operationally we would like to obtain an estimate of predictability for future times, when observations are not still available, while in the present paper the recognition of distinct predictability regions are obtained by neural modelling just in comparison with the "observed" states/points of the Lorenz model.

Nevertheless, we think that the evidence of the ability of a neural network in recognising different values of predictability in several zones of the Lorenz attractor can be the cornerstone for the use of artificial intelligence methods in the approach to this problem. The feasibility of a fully operational estimation of predictability by neural networks will be investigated in further studies, even if we can anticipate that some hints about the right strategy to adopt for reaching this scope come just from an accurate analysis of our present results. For instance, with reference to Table 1, the increase in the amplitude of error bars in neural performance when going from low bred-growth rates to high ones leads to suppose that the spread of ensemble forecasts on different points can resemble this one and can become an index of the peculiar local predictability of that Lorenz state. Thus, an activity of ensemble neural forecasts can be envisaged for achieving a fully operational estimation of local predictability.

Anyway, this last activity is outside the scope of this paper. Here, on the contrary, it should be interesting to analyse in more detail our present situation, for instance to understand if we can link the forecasting performance of our neural model to dynamical properties of the attractors considered, in order to achieve a more strict parallelism between them. This is useful for leading to improvements in the paradigm presented here.

In particular, as stated in the previous Section, we are not able to solve some "puzzles" about our present situation. For instance, we are not able to discern if the slight increase in predictability (sometimes statistically not fully significant) shown by neural performance when including a weak forcing in the classical Lorenz model is due to an increased frequency of "blue states" or to local changes in predictability. An accurate analysis of these topics requires a deeper investigation of the dynamical properties of unforced and forced Lorenz attractors.

In the following Table II we present the results of calculations of some of the dynamical quantities relevant for better defining the properties of these attractors (refer to [19,20]). In particular, for the classical unforced Lorenz system and for three forced ones, we calculated the values of:

- the Positive Lyapunov Exponent (PLE);
- the fractal dimension of the attractor (DIM);
- the average value of bred-growth rate ($\overline{g}$).

Table II. Dynamical quantities for the classical unforced Lorenz system (f_0 = 0) and for several forced Lorenz systems with θ = 90°.

Quantity	$f_0 = 0$	$f_0 = 2.5$	$f_0 = 5.0$	$f_0 > 10.0$
PLE	0.9023	0.8877	0.7686	< 0
DIM	2.0618	2.0608	2.0532	0
$\overline{g}$	9.06×10^{-3}	8.90×10^{-3}	8.32×10^{-3}	< 0

As well known, when a Lyapunov exponent is positive the behaviour of the system is chaotic. On the other hand, the fractal dimension is a "measure" of some properties of the attractor (for instance, if the values of DIM are very similar, also the shape of the attractors are so). Finally, the average bred-growth rate is an index of the divergence of close trajectories, so that it is itself a "measure" of chaos.

As a first comment to this Table we can say that in a forced Lorenz model chaos disappears if $f_0 > 10$: in this case, after a more or less long transient period, we can recognise that the attractor becomes a fixed point. On the other hand, the forced Lorenz system studied in this paper has very similar properties, from a dynamical point of view, if compared with those of the unforced Lorenz system: in particular, the positive Lyapunov exponent and the average bred-growth rate are only slightly diminished (leading to a small increase in predictability) and the fractal dimension is quite the same of the classical Lorenz system.

Thus, the comparison of the cases $f_0 = 0$ and $f_0 = 2.5$ shows that the dynamical properties of their attractors are very similar and we are not able to solve our "puzzles". On the other hand, we are very distant from the point where chaos disappears. All this analysis leads us to consider, in future studies, a case which is more distinct from the unforced case. In doing so, we will be able to see higher differences in predictability, both from the dynamical and the neural performance points of view, just leading to test more strictly the parallelism between these two kinds of estimating predictability.

VII. CONCLUSIONS AND PERSPECTIVES

This study represents a first attempt at estimating predictability in some regions of the attractor of a dynamical system by neural networks. As already stated, our preliminary results show that neural modelling is able to distinguish regions of distinct predictability, on both the unforced and forced Lorenz attractors. Furthermore, some changes in neural performance, clearly related to changes in bred-vector growth and Lyapunov exponents, have been found in the case of a weak forcing added to an equation of the original Lorenz system. This simple toy case study can lead to envisage an increased predictability in future meteo-climatic scenarios that, as suggested by other studies [5-7], is probably due to a higher frequency of persistent situations, even if our work does not allow to exclude a change in local predictability at single points on the forced attractor.

Even if the trend in NN performance vs. bred-vector growth is very clear (refer, for instance, to Table 1), the NN performance itself is not very good and this fact does not enable us to achieve a deeper insight in the analysis of predictability results. This can be due both to the limited length of the patterns record and to the choice of a quite coarse time step of integration in the Runge-Kutta scheme, that leads to consider our forecasts as "long-range predictions". Therefore, further work is in progress both for performing NN runs with an increased number of input-target pairs for training and validation and for performing a finer numerical integration by decreasing the time-step value, so achieving just "medium-range predictions" by neural networks.

Moreover, referring to the problem under discussion in the previous Section, the little differences found in NN performance may be due to the too much weak forcing considered in the forced Lorenz system. Thus, adopting once more Palmer's formalism [5,6] and referring to the system (3), we think that it is worthwhile to increase the forcing to $f_0 = 5.0$ (and eventually to consider also $\theta \neq 90°$) in order to study a situation which is more clearly different from the case of unforced Lorenz system, even if still clearly chaotic.

Finally, further improvements can also be envisaged both by consideration of different input sets (for instance, truncated time series of delayed data) and by application of other NN architectures and training rules.

ACKNOWLEDGEMENTS

We would like to thank M. Peña and E. Kalnay for having gently allowed us to use their code for the calculation of bred vectors.

REFERENCES

[1] E.N. Lorenz, "Deterministic non-periodic flow", *J. Atmos. Sci.*, Vol. 20, pp. 130-41, 1963.
[2] ECMWF, *Proceedings of the ECMWF Seminar on Predictability*, ECMWF, Reading, UK, 2002.
[3] A. Pasini and V. Pelino, "A unified view of Kolmogorov and Lorenz systems", *Phys. Lett. A*, Vol. 275, pp. 435-46, 2000.
[4] E. Evans, N. Bhatti, J. Kinney, L. Pann, M. Peña, S.-C. Yang, E. Kalnay and J. Hansen, "RISE undergraduates find that regime changes in Lorenz's model are predictable", *Bull. Amer. Meteor. Soc.*, Vol. 85, pp. 520-4, 2004.
[5] T.N. Palmer, "A nonlinear dynamical perspective on climate prediction", *J. Clim.*, Vol. 12, pp. 575-91, 1999.
[6] S. Corti, F. Molteni and T.N. Palmer, "Signature of recent climate change in frequencies of natural atmospheric circulation regimes", *Nature*, Vol. 398, pp. 799-802, 1999.
[7] S. Khatiwala, B.E. Shaw and M.A. Cane, "Enhanced sensitivity of persistent events to weak forcing in dynamical and stochastic systems: Implications for climate change", *Geophys. Res. Lett.*, Vol. 28, pp. 2633-6, 2001.
[8] A. Pasini and S. Potestà, "Short-range visibility forecast by means of neural-network modelling: a case study", *Nuovo Cimento*, Vol. 24C, pp. 505-16, 1995.
[9] A. Pasini, V. Pelino and S. Potestà, "A neural network model for visibility nowcasting from surface observations: Results and sensitivity to physical input variables", *J. Geophys. Res.*, Vol. 106, No. D14, pp. 14951-59, 2001.
[10] A. Pasini and F. Ameli, "Radon short range forecasting through time series preprocessing and neural network modeling", *Geophys. Res. Lett.*, Vol. 30, No. 7, doi:10.1029/2002GL016726, 2003.
[11] A. Pasini, M Lorè and F. Ameli, "Influence of forcings and circulation patterns on mean temperatures at different scales: an analysis by neural network modeling", in *Proceedings of CIMSA 2004*, Boston, MD, IEEE, pp. 51-56, 2004.
[12] A. Pasini and F. Ameli, "Neural network modeling as a tool for climatic analyses of forcings/temperatures relationships at global and regional scales", in *Proceedings of the 85th Annual Meeting of the*

American Meteorological Society (Fourth Conference on Artificial Intelligence Applications to the Environmental Science), San Diego, CA, CD-ROM, 2005.
[13] F. Takens, "Detecting strange attractors in turbulence", in *Dynamical systems and turbulence* (D.A. Rand and L.-S. Young eds.), Lecture Notes in Mathematics, Vol. 898, Springer, pp. 366-81, 1981.
[14] K.A. de Oliveira, A. Vannucci and E.C. da Silva, "Using artificial neural networks to forecast chaotic time series", *Physica A*, Vol. 284, pp. 393-404, 2000.
[15] G. Boudjema and B. Cazelles, "Extraction of nonlinear dynamics from short and noisy time series", *Chaos, Solit. and Fract.*, Vol. 12, pp. 2051-69, 2001.
[16] A. Pasini, V. Pelino and S. Potestà, "On the cognitive behaviour of a multi-layer perceptron in forecasting meteorological visibility", in *Neural Nets - WIRN Vietri '97* (M. Marinaro and R. Tagliaferri eds.), Springer-Verlag, pp. 245-51, 1998.
[17] E.N. Lorenz, "Atmospheric predictability as revealed by naturally occurring analogues", *J. Atmos. Sci*., Vol. 26, pp. 636-46, 1969.
[18] A.F. Atiya, S.M. El-Shoura, S.I. Shaheen, and M.S. El-Sherif, "A comparison between neural-network forecasting techniques - Case study: river flow forecasting", *IEEE Trans. Neural Networks,* Vol. 10, pp. 402-9, 1999.
[19] A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano, "Determining Lyapunov exponents from a time series", *Physica D*, Vol. 16, 285-317, 1985.
[20] K. Briggs, "An improved method for estimating Lyapunov exponents of chaotic time series", *Phys. Lett. A*, Vol. 151, 27-32, 1990.