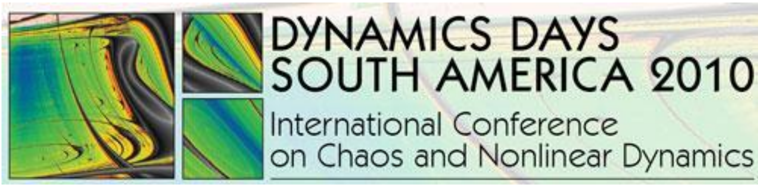

INPE – National Institute for Space Research
São José dos Campos – SP – Brazil – July 26-30, 2010

IMPROVING FORECASTING CAPABILITIES FOR TIME SERIES ANALYSIS

H. Carli[1], *L.G.S.Duarte*[2], *L.A.C.P.da Mota*[3]

[1]UERJ,Rio de Janeiro,Brazil,hcarli@gmail.com
[2]UERJ,Rio de Janeiro,Brazil,lgsduarte@gmail.com
[3]UERJ,Rio de Janeiro,Brazil,lacpdamota@gmail.com

keywords: Applications of Nonlinear Sciences; Chaotic Dynamics; Time series analysis.

1. TIME SERIES ANALYSIS AND PREVIOUS RESULTS

For any observed system, physical or otherwise, one generally wishes to make predictions on its future evolution. Sometimes, very little is known about the system. Suppose that in particular the dynamics behind the phenomenon being studied is unknown, and one is given just a time series[1] of one (or a few) of its parameters.

If the time series is the only source of information on the system, prediction of the future values of the series requires a modelling of the system's (perhaps nonlinear) dynamical law through a set of differential equations or through discrete maps. However, it is even possible that we do not know whether the measured quantity is the only relevant degree of freedom (usually it is not) of the dynamical problem, nor how many of them there are.

For time series originated from low dimensionality chaotic systems, we have the non-linear analysis apparatus at our disposal and we will not be concerned with stochastic processes[3].

Methods for dealing with such a problem fall mainly into two categories: local or global methods. We will focus on global methods. We are also going to Suppose that the system can be modelled by a set of differential equations of low dimensionality. What we would like to obtain is some kind of global map that, given any point of the state space, could calculate a subsequent point of the trajectory. If we have known the set of differential equations (SDE) that models the system, we could find a solution (starting from an initial condition) by making a numerical integration through some map obtained from the SED (probably a Runge-Kutta map, a Taylor series one or an expansion in some function basis). For practical purposes (computers can not work with the infinity) a truncation must occur at some order of the series expansion. However, if the truncation order is low, we can run away from the real solution in a few time steps (even if each time step is very small). For chaotic systems it is not used (in general) a Runge-Kutta expansion of degree less than four. This implies that the map generated present polynomials of high degree.

That means that the mathematical problem can be very complex. For example, for the Lorenz system, a fourth-order Taylor expansion would be equivalent to a problem of determining 168 coefficients a very elaborated task. Therefore, despite the fact that the global approach has many attractive features, such as the fact that, once it is determined it is applicable to the whole series[4], one sees that the effective use of it can be difficult to achieve in practice. So, there is a clear demand for procedures that can, without increasing the degree of the global mapping, enhance the accuracy of such mappings.

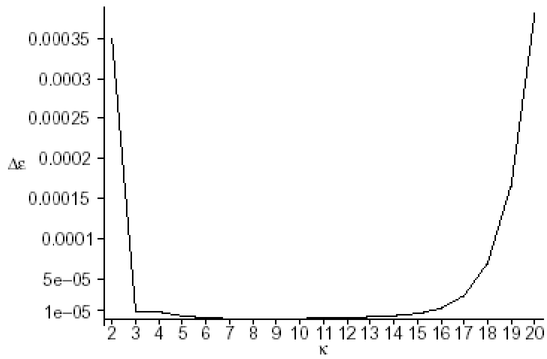

Figure 1 – Example of the "Plateau" formed for the Lorenz System Time Series

In a few words, our method [1] is based on the realization that the expansion with which we will be working is a "real" one, i.e., it is finite. So the known (accepted) fact that the next-order term will be negligible in relation to the previous one faults at some stage. Our method detects where this assumption starts to fail and uses this information to improve the forecasting. Basically, we analyze the so-called plateau that is formed because of the balance between this tendency of the expansion to have ever smaller terms and the "reality" of our expansion, i.e., the truncation of the series causes this tendency to be, eventually, overcome and the relative size of the next term to the previous one, with increasing number of terms, starts to rise again. A picture of such a behavior

[1] A time series is a set of numbers that are the possible outcome of measurements of a given quantity, taken at regular intervals.The relevance of performing Time Series[2] analysis can be equated to the fact that these Series mentioned above can come from a great diversity of branches of knowledge. There are extensive studies in the area of Physics, Economics, Meteorology, Oceanography, Stock Exchange, Medicine, etc.

[3] Indeed, the first thing one has to ask when working with a Time Series is whether the series represents a causal process or it is stochastic. In the Time Series analysis frame we also have tools to deal with that fundamental question.

[4] In the case of Local mappings, we have to determine a mapping for each entry of the series.

is displayed on figure 1, where this difference $\Delta\epsilon$ is plotted against the number of terms (k) in the expansion.

2. IMPROVING THE SITUATION

Here, we present an algorithm which is concerned with improving this forecast capabilities of our above mentioned improved method [1] even further. The algorithm uses more elaborated steps to extract the better way of deciding at which stage to stop the approximating procedure to have a better forecast capability.

In a nutshell, we are looking into creating an algorithm that, based on a previous analysis of the particular Time Series, actually on the study of the particular section of the Time Series under study, could determine how far one can go "inside" the plateau to obtain the best result possible where forecasting is concerned. So far, while working in the scope of [1], we stop our expansion in the very beginning of the Plateau, without further examining if we could extract more information from the terms already "inside" the plateau, etc.

We have obtained preliminary results that indicate that one can achieve a significantly better result for some cases. For instance, in the case of "well behaved" time series, such as the one coming out of a Lorenz system, one can find a factor of five (or above) in the improvement for the forecasting capabilities of this new improved algorithm.

References

[1] Barbosa, L. M. C. R.; Duarte, L. G. S.; Linhares, C. A.; da Mota, L. A. C. P. Physical Review E, vol. 74, Issue 2, id. 026702

[2] Abarbanel H., Brown R., Sidorowich J., Tsimring L. (1993) The analysis of observed chaotic data in physical systems *Review of moderns Physics* **65(4)**, *1331.*;Casdagli M. Nonlinear Prediction of Chaotic Time Series *Physica D* **35 (3)**, *335.*;Farmer and Sidorowich Predicting chaotic time series *Physical Review Letters* **59**, *845.*

[3] Lorenz E. (1963) Deterministic Nonperiodic Flow *Journal of the Atmospheric Sciences* **20** - *130.*

[4] Takens F. (1981) Detecting strange attractors in turbulence *Springer Lectures Notes in Math.*,**898**, *366.*