2004 International Symposium on Nonlinear
Theory and its Applications (NOLTA2004)
Fukuoka, Japan, Nov. 29 - Dec. 3, 2004

Global reconstructions of equations of motion from data series, and validation techniques

G. Gouesbet† and S. Meunier-Guttin-Cluzel†

†Laboratoire d'Electromagnétisme et Systèmes Particulaires,
CORIA, UMR CNRS 6614, Université de Rouen et INSA de Rouen,
BP12, Avenue de l'Université, Technopôle du Madrillet,
76801 Saint-Etienne-du-Rouvray
Email: gouesbet@coria.fr, meunier@coria.fr

Abstract– This paper provides a short review on the topic of global reconstructions of equations of motion from data series, and validation techniques. Any technique used to characterize an attractor may also be used to validate a model and, therefore, we also provide a review on characterization methods. This paper may be viewed as an introduction to a recently published book on the topic of "chaos and its reconstruction" [1].

1. Introduction

Reconstructing equations of motion of time-continuous dynamical systems from time series defines the topic of global vector field reconstruction or flow phenomenological modeling. Reconstructing equations of motion of discrete-time dynamical systems defines the topic of global map reconstruction or map phenomenological modeling. Time-delay systems, not discussed in this paper, have also been considered [2]. Once a phenomenological model has been obtained, it has to be validated. Any technique used to characterize an attractor provides an opportunity for a validation technique. Accordingly, we briefly review the topic of global reconstruction of equations of motion from data series, and validation techniques.

2. Basic background and overview of problems

We consider non-linear dynamical systems. One topic of interest is, given a time series, to forecast the future, this may be done by local or global modeling. Global models may be viewed as local models in which all the data pertain to a single neighborhood which is the entire available phase space. In other words, a local model provides an atlas, i.e. a collection of local charts while a global model uses only a single chart. Global models, when successful, compress the data more efficiently than local models, are computationally faster, and provide a better knowledge of the vector field (or map) generating the data.

In an extreme case, we possess a single scalar time series. Relying on a redundancy principle, according to which the evolution of any variable is influenced by the evolution of the variables which interact with it, we may provide an embedding of the time series in a phase space which will support the model. Many kinds of embeddings may be used such as time-delay embeddings, embeddings with principal components, differential embeddings. Our favorite embedding is the standard embedding, i.e. a particular differential embedding in which all non-linearities are reported in only one component of the vector field.

Another issue is noise and data preparation because, most usually, a time series is not ready for a global reconstruction. Noise reduction procedures may rely on filtering. But this may be a dangerous procedure because the filter possess its own dynamics. In another approach, we may work directly in a phase space with a trajectory adjustment technique. Data preparation may also include the removal of drifts which are often unavoidable in experiments. A particularly bad situation may occur in observational astronomy when the SNR is poor, and data cannot be regularly sampled at equal space time intervals. Then, beside filtering, interpolations may have to participate to data preparation.

Once we have an embedding with possibly prepared data, reconstructions of equations of motion eventually require us to approximate one or several unknown functions by modeling functions, in the framework of the theory of nonlinear approximation of functions. Models will contain parameters. An important issue is then to avoid over-parameterization, relying on a parsimony principle telling us that the best model is the smallest model (Occam's razor principle). This requires to quantify the size of the model. According to our bibliography, the most popular way to choosing the model size is by using a criterion due to Rissanen. The model is then generated under minimum description length constraints.

Following a terminology used by L.A. Aguirre, a white-box modeling is a modeling from first principles. Black-box modeling is modeling from time series, without any prior knowledge of the system. This is the case for global models discussed above. In gray-box modeling, some amount of prior white information is poured into the black. Expectedly, this should result in a better global model, including more parsimony. The a priori information may for instance be used to delete terms which, in the model, are not compatible with the information.

Therefore, for gray-box modeling, any extra information is welcome. The worst black-box case is when we only possess a single time series. The landscape is whiter when all variables are measured or in the hybrid case when several variables, but not all of them, are measured. In these last cases, we apparently retrieved more information than required if we invoke the redundancy principle. But it is a fact that all variables are not exactly equivalent both in practice (some variables are of easier use for modeling) and possibly in principle in the case of an equivariant system, i.e. a system presenting some kind of symmetry. Indeed, an equivariant system usually displays two kinds of variables , i.e. equivariant variables (containing information on symmetries such as variables x and y of the Lorenz system) and invariant variables (variables which do not contain any information on symmetries).

3. Global vector field reconstructions with standard embeddings

We provide an example of a global vector field reconstruction with a standard embedding, and a polynomial model. Fig 1 provides a time series of the electrode current generated by a copper electrodissolution experiment, (data from J. Hudson and Z. Fei, University of Virginia, USA). Fig 2 represents a projection of the attractor (in a 3-D phase space) generated by the experimental data. Fig 3 represents similarly a projection of the attractor (in a 3-D phase space) generated by a phenomenological model built on the experimental data. The model is "visually" validated.

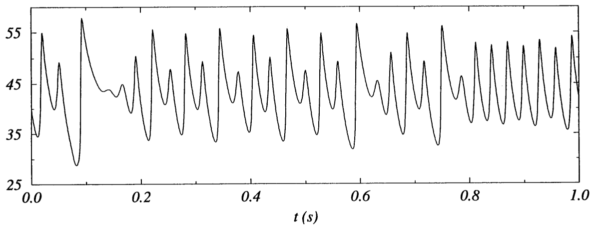

Fig 1

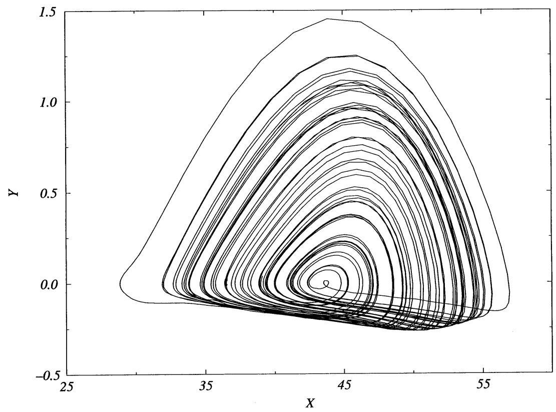

Fig 2

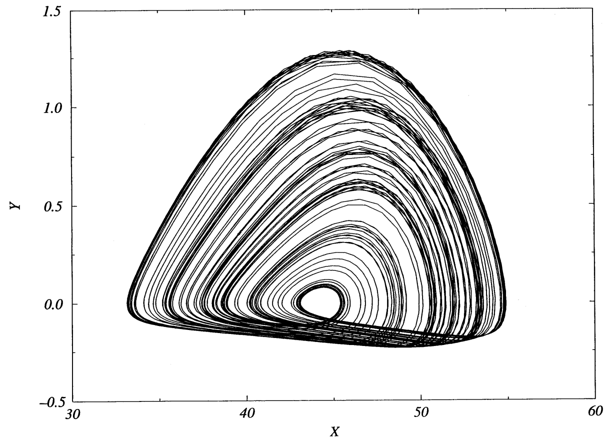

Fig 3

4. Global map reconstructions

We similarly present the results of a global map reconstruction for a thermoionic diode with forcing (data from T. Mausback et al, Kiel University). Fig 4 displays a first return map based on experimental data. Fig 5 displays the associated first return map based on a global model. The model is "visually" satisfactory.

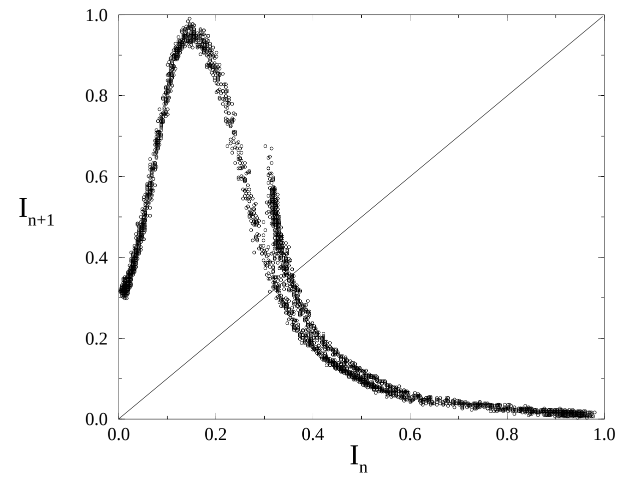

Fig 4

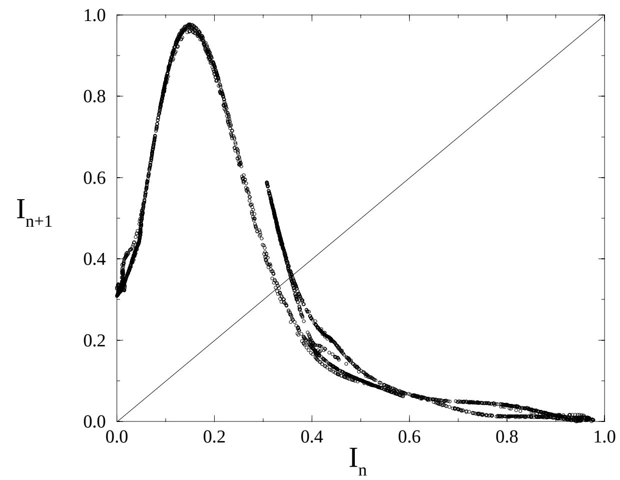

Fig 5

5. Validation techniques

When the original system is known, we may formally compare expansion coefficients for the original vector field and for the reconstructed vector field. Otherwise, we may also use visual validations (as above), records of probabilities of visit, generalized dimensions and entropies, checks for diffeomorphisms, Lyapunov exponents, synchronization, topological approaches (templates and populations of unstable periodic orbits). As an illustration, we choose to discuss generalized dimensions to have the opportunity to point out a somewhat controversial issue.

Fig 6 displays a comparison between D_q-spectra for an original Rössler attractor and for an associated model (named ISRS). People accustomed to dimension calculations would claim that results for negative values of q are rubbish because they would require an unaffordable amount of data. Yet, similar data are displayed in fig 7 after improvements of the reconstruction algorithm [3]. The influence of the improvements is well exemplified. This points out to the fact that requirements for the dimension computations are not the same whether we deal with characterization of attractors or with validation of models.

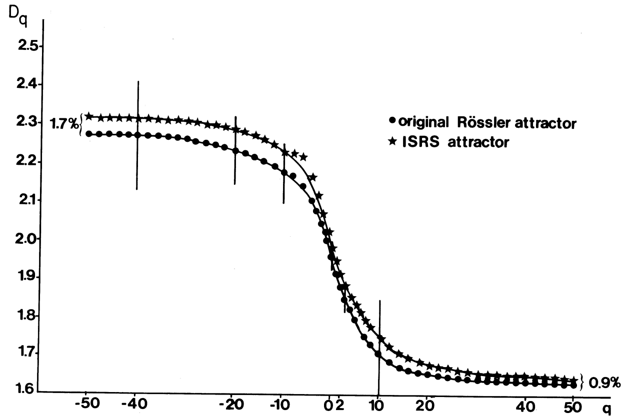

Fig 6

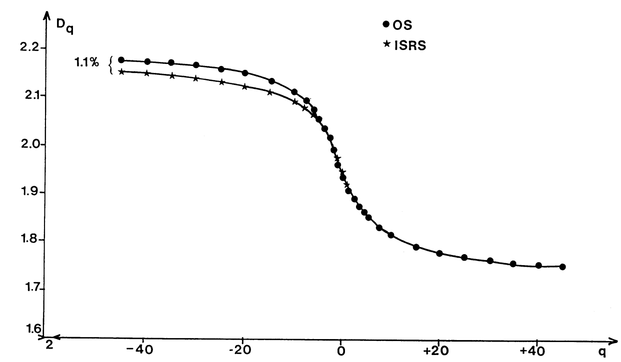

Fig 7

6. Applications of phenomenological modeling.

It is possible to forecast without reconstructing equations. For instance, it may be sufficient to exploit a library of past patterns in a time series. But forecasting is one of the immediate applications of local or global models.

Another important application, particularly of global models, is to provide surrogate data. For instance, a time series may be too unclean, too short, or exhibit significant drifts preventing a secure evaluation of some quantities such as fractal dimensions. If a global model can be obtained from these data, we may generate long time series of clean data without any drift. These data may be used as surrogate data for a better characterization of the attractor under study.

Phenomenological modeling also produces an efficient compression of information. We may replace a large set of data by the formal representation of a simple mathematical function which may flow more efficiently through a communication channel. With respect to this issue, global models perform dramatically better than local models. Starting from a huge amount of data, we may end with a formal representation of a vector field or of a map.

Global models may also be used to control chaotic processes such as in Ref [4] where chaos is suppressed and the system is driven to a periodic orbit. They may also be used to secure the transmission of information along communication channels.

Furthermore, a successful global reconstruction of equations of motion provides many deep insights on the underlying dynamics. When it relies on an embedding technique, it defines the minimal embedding dimension, that is to say the minimal number of coordinates required to describe the dynamics. If we thereafter decide to build a physical model from first principles, then we know from the beginning that our physical model should not use more variables than the phenomenological model.

Also, a successful global reconstruction provides us with a strong test to decide whether or not the system under study is deterministic. Such a test is more convincing than the evaluation of the fractal dimensions which has been used in earlier stages, and led to controversial results.

Global models may also be used for signal detection, i.e. identifying the presence of a deterministic contribution in a signal, and classification, i.e. assigning a detected signal to a particular class.

7. Conclusion.

We provided a brief review of the contents of the book of ref [1]. The topic of global reconstruction is now a mature field, with many applications. Much work remains to be done and we therefore hope that this paper could attract newcomers.

References

[1] G. Gouesbet, S. Meunier-Guttin-Cluzel, O. Menard, editors, “Chaos and its reconstruction”, Nova Science Publishers, Inc. New York, 2003 (317 pages, about 800 references).

[2] B. P. Bezruchko, A.S. Karavaev, V.I. Ponomarenko, M.D. Prokhorov, “Reconstruction of time-delay systems from chaotic time series”, *Physical Review E*, 64, 056216-1, 056216-6, 2001.

[3] G. Gouesbet, J. Maquet, “Construction of phenomenological models from numerical scalar time series”, *Physica D*, 59, 202-215, 1992.

[4] L.A. Aguirre, S.A. Billings, “Close-loop suppression of chaos in nonlinear driven oscillators”, *Journal of Nonlinear Science*, 5, 189-206, 1995.