European Symposium on Computer Aided Process Engineering – 15
L. Puigjaner and A. Espuña (Editors)

State space analysis, a tool for progress

Javier Villa Briongos[a], Jesús Guardiola Soler[b]

[a]Investigador contratado, PDI
Universidad Complutense de Madrid, Facultad de Ciencias Químicas, Department of Chemical Engineering, Avda. Complutense s/n, 28040 (Madrid) España

[b†]Profesor Titular de Universidad
Universidad de Alcalá, Facultad de Químicas, Department of Chemical Engineering, Crtra. Madrid-Barcelona, km 33,6 (Campus), 28871-Alcalá de Henares (Madrid), España

Abstract
This contribution is aimed to give a fast view of State Space Analysis by presenting some codes written as m-files that can serve to understand the know-how methodology behind this advanced analytical tool, and yet little known. Mutual Information Function, reconstruction of the attractor system and correlation dimension allow to identify the time series under study as periodic, random or chaotic. Moreover, they can be used to characterize the underlying dynamics when the time series is collected from a physical system. Some version of the m-files can be free downloaded from:
http://www2.uah.es/dep_qaiq/caos/texto_en.htm

Keywords: State space analysis, signal processing, non-linear analysis

1. Introduction

During the last decade deterministic chaos analysis has become a tool of common use in Chemical Engineering for studying complex systems like fluidized beds or bubble columns, so it is not surprisingly to find works where time, frequency domain and state space analysis they are complementary used to characterize the underlying dynamics of those systems. Thus, statistical tools derived from the deterministic chaos theory, they have been successfully applied for monitoring and control (Van Ommen et al., 2000; Villa et al., 2003), modeling (Van Wachem et al., 1999) and scale up process (Schouten et al., 1999). Moreover, its use can be extended to any system that shows a non-linear behavior.
However, despite there are available software suitable for state space analysis, the know-how methodology behind the algorithms involved in deterministic chaos analysis is far from being known, i.e. the source codes are not accessible by the user, that fact shrinks the freedom of the researcher to model new sort of experiments and applications for those techniques. Furthermore a deeper knowledge of those techniques becomes necessary for a reliable development of strategies and control systems.

[†] Author/s to whom correspondence should be addressed: jesus.guardiola@uah.es

This paper is therefore addressed to give the reader a key to the know-how methodology for state space analysis by providing a set of programs suitable for computing the Mutual Information Function, *I*, the reconstruction of the attractor system and the study of the correlation dimension, D_2. Other properties like the Kolmogorov entropy has been left out of the paper due to length limitations. The fact the code has been implemented as m-file, helps to the understanding of the deterministic chaos theory. Moreover, since those codes have been written with teaching purposes can therefore, be optimized.
In order to present the general applicability of those tools they have been applied over several gas-solid fluidized bed time series obtained from different sources like pressure, and out-bed acoustics signals.
The algorithms have been successfully used in fluidization engineering over bed surface fluctuations time series by Villa Briongos and Guardiola (2003, 2004) and some versions of them can be free downloaded from:
http://www2.uah.es/dep_qaiq/caos/texto_en.htm

2. Mutual Information Function

Mutual information function, *I*, is based on the uncertainty concept developed by Shanon and Weaver (1949). According to that, the uncertainty associated with any measure depends on probability from all possible outcomes (eq. 2).
$I(\tau)$ gives the average of bits that can be predicted correctly for a sample τ in the near future, with the knowledge about a "*X*" measure. On equation 2, the entropy *H* is representing the probability for all possible outcomes.
So when dealing with *deterministic signals* the connection between successive measurements is repeated over fundamental frequency intervals, the future behavior is therefore completely predictable (fig. 1a). For *random processes* there is not a deterministic connection among the successive measures and the knowledge over past history, it does not have consequences on the future dynamic behavior of the signal. In contrast, *deterministic chaotic processes* are an intermediate case, having characteristics that are between both deterministic and random behaviors.
As depicts Figure 1a, the Mutual Information Function provides information about the evolution time of the process, as the same as the autocorrelation function does. However, the fact the mutual information function does not assumes any functional relationship between the data points, makes it more appropriated for studying non-linear dynamics (Daw and Hallow, 1993; Karamavruς and Clark, 1997). Figure 1 is a good example of that, it shows how the autocorrelation coefficient, *r* (fig.1b), does not accounts for the complexity showed by the power spectral density function (fig.1d), the harmonics present within the signal is hiding non-linearities leading to large values of autocorrelation coefficient. In contrast the mutual information function (fig.1c) detects the indeterministic behavior existing within the signal showing low values of *I* and pointing out the random component.
To compute the Mutual Information Function, the entropy concept is the central point of the uncertainty principle, and is computed from:

$$H(X) = -\sum_{i=1}^{N} P(x_i)\log_2 P(x_i) \qquad (1)$$

where X is the original time series and $P(x_i)$ is the probability obtained from the time series. For computing $P(x_i)$, the data (x_i) are introduced into equals bins of the same size. According to that, for any time series (signal) the probability of any value $x(t_i)$ falls into a specific bin, x_i, is $P(x_i)$.

$H(X)$ is the average entropy or information in bits that contain the time series X. Finally the mutual information function between a time series, X, and a delayed version, X + τ, is defined like:

$$I(X, X+\tau) = H(X) + H(X+\tau) - H(X, X+\tau) \qquad (2)$$

H(X+ τ) and H(X, X+ τ) are given by:

$$H(X+\tau) = -\sum_{j=1}^{M} P(x_j)\log_2 P(x_j) \qquad (3)$$

$$H(X, X+\tau) = -\sum_{i}^{N}\sum_{j}^{M} P(x_i, x_j)\log_2 P(x_i, x_j) \qquad (4)$$

$P(x_i, x_j)$ is the join probability, representing the probability that any measure of X fall into bin x_i for the first and any measure of X + τ fall into bin x_j for the second.

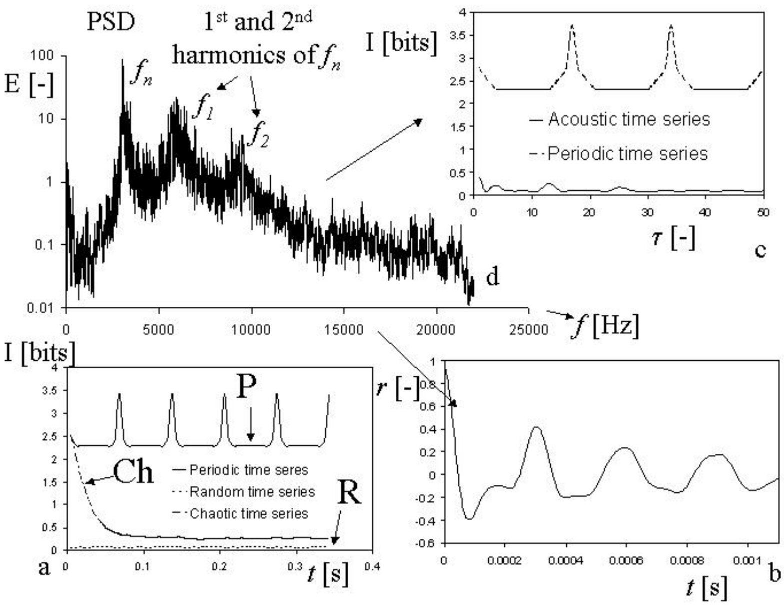

Figure 1. a) Mut. Inf. Func. of periodic (P), random (R) and chaotic time series (Ch); b,c,d) Mut. Inf. Func., autocorrelation coefficient and PSD of fluidized bed acoustic signals.

The uncertainty of a join event $H(X, X{+}\tau)$ is less or equal than the sum of the individual uncertainties. However, occasionally $H(X, X{+}\tau)$ might be a little greater than the sum, due to round off errors yielding a mutual information function, I, slightly negative.
The join probability is computed as:

$$P\left(x_i, x_j\right) = \frac{\text{number of data from } X \text{ and } X+\tau \text{ into bin } x_{i,j}}{\text{number of total data in both time series}} \tag{5}$$

So when the events are independent $P(x_i, x_j) = P(x_i)P(x_j)$, getting the mutual information function therefore, a zero value which means maximum independence between the compared signals.

3. Reconstruction of the attractor

An attractor characterizes a dynamical system. For experimental time series (i.e. pressure fluctuations signals), its embedding into a phase space of embedding dimension m (Takens, 1981) is a useful way to reconstruct the attractor.
The embedding of time series by means of the delay method (Takens, 1981), uses the original time series, x(t), and provides a new coordinate system by using delay versions of x(t). According to Takens (1981), each point X(i) within the phase space of embedding dimension m will be given as:

$$X(i) = \{x(t_i), x(t_i + \tau), \ldots, x(t_i + (m-1)\tau)\} \tag{6}$$

where τ is an integer number of temporal distance into the time series (time step), $x(t_i)$ is the measure at time t_i and $x(t_i + \tau)$ is the delay measure at time $t_i + \tau$. Since the embedding dimension is m, each point X(i) is defined by m coordinates.

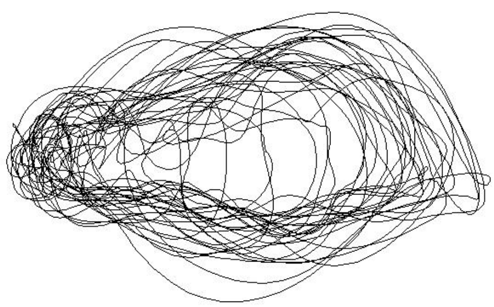

Figure 2. Reconstructed attractor from a fluidized bed pressure time series where the bed was performing a slugging regime, m =19

However, due to the practical implementation of the method of delays proposed by Takens has some problems (Sauer et al., 1991), the statistical approach to that method, suggested by Broomhead and King (1986) for reconstructing the attractor, it is preferred. The method has been broadly applied for reconstructing experimental attractors in fluidization engineering (Schouten and Van den Bleek, 1992; Daw and Hallow, 1993; Villa Briongos and Guardiola, 2003).

4. Correlation dimension

Since its description by Grassberger and Procaccia (1983), the method of estimating the correlation dimension, D_2, as a measure of the local structure of a strange attractor has been extensively used. This parameter is obtained from the spatial correlation between random points on the reconstructed attractor. Thus, due to the exponential divergence of trajectories, most pairs of points (X_i, X_j) of the time series in which i ≠ j will be dynamically uncorrelated. Nevertheless, the fact the points lie on the attractor means they will be spatially correlated, and the correlation integral, $C\ (l)$, measures this spatial correlation, expressing the probability of finding pairs of points (X_i, X_j) where i ≠ j on the attractor within the specific distance l. Thus, for small values of l, $C\ (l)$ scales as $C(l) \propto l^{D_2}$, where D_2 is the correlation dimension and is related to the number of degrees of freedom of the system, measuring the homogeneity of the attractor.

The literature describes several methods of computing the correlation integral (Ellner, 1988; Schouten et al., 1994). Nevertheless, these methods are all based on the well known algorithm of Grassberger and Procaccia (1983).

Here the use of that algorithm is followed by application of an embedding dimension chosen by a previous false neighbor analysis (Abarbanel, 1996). Additionally, the interpoint distance is computed using the maximum norm (Eckman and Ruelle, 1985).

Table 1 shows the comparison between the correlation dimension estimated with the proposed tool, and the results found in literature for a set of well-known theoretical time series. The little difference (less than 5%) existing between the correlation dimension values confirms the reliability of the algorithm used.

Table 1. Correlation Dimension for several theoretical time series.

	N, points	τ	m	D_2 (exp)	D_2 (ref)	Author
Lorenz model (X-coordinate)	10 500	1	10	1.97	2.05	Grassberger and Procaccia (1983)
Henon Model (X-coordinate)	10 500	1	3	1.25	1.25	
Logistic Equation (μ = 4)	10 500	1	4	0.96	1	
Logistic Equation (μ = 3.569956)	10 500	1	2	0.44	0.50	

5. Conclusions

The high velocity performance of the algorithms makes possible their application i.e. for dynamic monitoring and control purposes, and the use of CAD utilities for state space analysis.

Mutual information function reveals the deterministic component exiting in a signal, making possible to characterize the type of signal in terms of its predictability. So the signal could be stochastic, chaotic or periodic.

The attractor characterizes the dynamics, providing a picture about their complexity visualizing its trajectories. Moreover, correlation dimension quantify the strangeness of the attractor in terms of the homogeneity of the reconstructed attractor.

Literature

Abarbanel, H.D.Y., 1996, Analysis of Observed Chaotic Data, Springer – Verlag New York, Inc.

Broomhead, D.S., G.P. King, 1986, Qualitative Dynamics from Experimental Data, Physica. 20D, 217.

Daw, C.S. and J.S. Hallow, 1993, Evaluation and Control of Fluidization Quality Through Chaotic Time Series Analysis of Pressure Drop Measurements, AIChE Sym. Ser. 89, No. 296, 103.

Eckman, J.P. and D. Ruelle, 1985, Ergodic Theory of Chaos and Strange Attractors, Rev. Mod. Phys. 57, 617.

Ellner, S., 1988, Estimating Attractor Dimensions From Limited Data: A New Method, With Error Estimates, Phys. Rev. Let. A 133, 128.

Grassberger, P. and I. Procaccia, 1983, Measuring the Strangeness of Strange Attractors, Physica 9D, 189.

Karamavruς, A.I. and N.N. Clark, 1997, Local Differential Pressure Analysis in a Slugging Bed Using Deterministic Chaos Theory, Chem. Eng. Sci.. 52, 357.

Sauer, T., J.A. Yorke and M. Casdagli, 1991, Embedology, J. Stat. Phys. 65, 579.

Schouten, J.C. and C.M. van den Bleek, 1992, Chaotic Hydrodynamics of Fluidization: Consequences For Scaling and Modeling of Fluid Bed Reactors, AIChE Sym. Ser. 88, No. 289, 70.

Schouten, J.C., F. Takens and C.M. van den Bleek, 1994, Estimation of The Dimension of A Noisy Attractor, Phys. Rev. E. 50, 1851.

Schouten, J.C., R.C. Zijerveld and C.M. Van den Bleek, 1999, Scale-up of Bottom Bed Dynamics and Axial Solids-Distribution in Circulating Fluidized Beds of Geldart-B Particles, Chem. Eng. Sci., 54, 2103.

Shanon, C.E. and W. Weaver, 1949, The Mathematical Theory of Communication, The University of Illinois press: Urbana.

Takens, F., 1981, Detecting Strange Attractors in Turbulence, Lecture notes in Mathematics, D.A. Rand and L.S. Young Eds. (Springer, Berlin), 366.

Van Ommen, J.R., M.-O. Coppens, C.M. Van den Bleek and J.C. Schouten, 2000, Early Warning of Agglomeration in Fluidized Beds by Attractor Comparison, AIChE J. 46, 2183.

Van Wachem, B.G.M., J.C. Schouten, R. Krishna and C.M. Van den Bleek, 1999, Validation of the Eulerian Simulated Dynamic Behaviour of Gas-Solid Fluidised Beds, Chem. Eng. Sci. 54, 2141.

Villa, J., J.R. Van Ommen and C.M. Van den Bleek, 2003, Early detection of foam formation in bubble columns by attractor comparison, AIChE J. 49, 2442.

Villa Briongos, J. and J. Guardiola, 2003, Free Top Fluidized Bed Surface Fluctuations as a Source of Hydrodynamic Data, Powder Technol. 134, 133.

Villa Briongos, J. and J. Guardiola, 2004, Using Free Bed Surface Fluctuations in 3D Fluidized Bed Dynamic Characterization, AIChE J. 50, 3060.

Acknowledgements

The authors would like to thank Prof. J.M. Aragon and M.C. Palancar of UCM for their assistance with the acoustic time series processing. Moreover, the financial support from project PPQ2003-08742 is kindly acknowledged.