E C O N O M I C S   B U L L E T I N

Searching for chaos on low frequency

Nicolas Wesner
MODEM Université Paris X Nanterre

Abstract

A new method for detecting low dimensional chaos in small sample sets is presented. The method is applied to financial data on low frequency (annual and monthly) for which few observations are available.

Citation: Wesner, Nicolas, (2004) "Searching for chaos on low frequency." *Economics Bulletin*, Vol. 3, No. 1 pp. 1−8
Submitted: October 27, 2003. **Accepted:** January 27, 2004.
URL: http://www.economicsbulletin.com/2004/volume3/EB−03C10007A.pdf

1 Introduction

Since the works of Bachelier (1900), the orthodox academic point of view in finance has been to consider fluctuations in stock prices as random and unpredictable. In the same way, the theory of real business cycles (Kydland and Prescott 1982) states that fluctuations in real aggregates are due to random productivity shocks. Thus, the cornerstone of modern finance: the informational efficiency hypothesis (Fama 1965) as well as the neoclassical approach in macroeconomics rely on this mathematical representation. The concept of deterministic chaos, which implies the generic possibility that an apparently random phenomenon is actually generated by a deterministic process, has renewed the debate on randomness of prices dynamics and the theory of endogenous business cycles (Grandmont 1985).

It has motivated numerous authors for detecting chaos in economic and financial time series. The first tools used such as the Lyapunov exponent (Wolf et al. 1984), which measures sensitive dependence on initial conditions, or the correlation integral (Grassberger et Procaccia 1983), which measures spatial correlations in the phase space, were designed for very large data sets and perform poorly with small samples. Since, researchers have come up with numerous methods for observing or measuring nonlinear determinism in relatively shorter time series (see Aleksic 1990, Kennel et al. 1992, Wayland et al. 1993, Cao 1997 and Zbilut et al. 1998). However, here again, those methods perform poorly with very small sample sets (under 500 observations). By the fact, almost only long period of time and high frequency data (daily or weekly) have been analyzed.

The aim of this paper is to present a very simple method for detecting low dimensional deterministic structure in small sample sets and to apply it to low frequency (monthly and yearly) financial time series for which few observations are available.

2 Methodology

2.1 *Reconstruction of dynamics by the method of time delays*

Let $s(t)$ $(t = 1, ..., N)$ denotes an observable process generated by an unknown or unobservable system. Following Brock (1986), $s(t)$ is said to have a deterministic explanation if:

$$F : \mathrm{M} \longrightarrow \mathrm{M} \tag{1}$$
$$x(t+1) = F\left(x(t)\right),$$

where F is an unknown smoth map, and M is a d-dimensional manifold, and:

$$h : \mathrm{M} \longrightarrow \mathrm{IR} \tag{2}$$
$$s(t) = h\left(x(t)\right),$$

where h is an unknown smooth map.

Then, according to Takens theorem (1981), for an adequate choice of parameters m and τ, there exists a function G such as:

$$G : \mathrm{IR}^{m^2} \longrightarrow \mathrm{IR}^{m^2} \tag{3}$$
$$y(t+\tau) = G\left(y(t)\right),$$

with $y(t) = (s(t), s(t-\tau), ..., s(t-(m-1)\tau))$, *m being the embedding dimension and* τ *the time delay usually fixed to one.*

G is a diffeomorfism to F and is called topologically conjugate to F. That is G conserves the same properties as F and in particular the continuity properties of the trajectories in the phase space. Thus, if $s\,(t)$ is a deterministic time series, then, for any pair of points $(y(i), y(j))$, for an adequate choice of m, there exist arbitrary small $\alpha, \delta > 0$ so that:

$$if \parallel y\,(i) - y\,(j) \parallel < \alpha, \; then, \parallel G(y\,(i)) - G(y\,(j)) \parallel < \delta, \tag{4}$$

where $\|a - b\|$ *being the distance (according to a given norm) between the vectors a and b.*

This property means that images of close points are close in the phase space. This "phase space continuity" is characteristic of deterministic processes, it can be used to distinguish random numbers from chaotic dynamics and to make short term prediction of chaotic time series (Farmer and Sidorowich 1987).

2.2 A measure of determinism

The measure of determinism presented here relies on the reconstruction scheme described above. The basic idea is to measure the extent to which the time series considered verify the property of "phase space continuity", by observing the dynamics of nearest neighbors (see Fernandez Rodriguez et al. 2003 for a recent application of a method based on nearest neighbors to the prediction of exchange rates). For instance, it is assumed that nearest neighbors whose images are nearest neighbors satisfy the continuity property (4). That is, $y(i)$ and $y(j)$ satisfy the continuity property (4) if: they are nearest neighbors:

$$y(j) = \arg \min_{s \neq i, = m, \ldots, N} \{\|y(i) - y(s)\|\}, \tag{5}$$

and if their images are nearest neighbors:

$$y(j+1) = \arg \min_{s \neq i+1, = m, \ldots, N} \{\|y(i+1) - y(s)\|\}. \tag{6}$$

Indeed for those points:

$$\| y(i) - y(j) \| = r(i) \ \textit{and} \ \| y(i+1) - y(j+1) \| = r(i+1), \tag{7}$$

where $r(s)$ is the minimum distance between $y(s)$ and another vector in the phase space.

Thus, if $y(i)$ and $y(j)$ satisfy (5) and (6) it is possible to choose arbitrary small α and δ for which $y(i)$ and $y(j)$ verify (4).

Numerical experiments show that, for short time series, the proportion of points satisfying those properties grows with the embedding dimension. So the measure of determinism D proposed here is defined as follows :

$$D = \frac{\textit{number of pairs of points } y(i) \text{ and } y(j) \text{ satisfying (5) and (6)}}{(N-m+1)m} \tag{8}$$

where N is the number of observations and m is the embedding dimension.

The quantity D is calculated for different values of m. For each vector $y(i)$ $(i = m, ..., N)$, only the first neighbor is considered, that is the vector $y(j)$ $(j \neq i = m, ...N)$ which minimizes the Eulidean distance like in equation (5). In a second step, the pairs of nearest neighbors whose images are nearest neighbors are counted. For deterministic time series, this measure of determinism is expected to be significantly different from zero for a value of m sufficiently large.

Except for special cases like the tent map, doubling map or logistic map where $D = 1/(2m)$ for large N (all of those applications have symmetric invariant density on the interval (0,1) and two equally probable pre-images for each state), it is difficult to derive theoretical values of D. Nevertheless, for independent and stationary data, the probability that a pair of points verify accidentally equations (5) and (6) decreases when N grows, so D is expected to be close to 0 for N sufficiently large.

2.3 Comparative analysis

To illustrate its efficacy for small data sets, the measure of determinism D was calculated for time series of 150 observations. Data under investigation are: white noise, colored noise (with a correlation coefficient of 0.95) and chaotic time series generated by the Logistic map, Hénon map, and Lorenz and Mackey Glass systems. Results were compared to those obtained by the method of false nearest neighbors (Kennel et al. 1992).

The false nearest neighbors approach is widely used and is known to be not strongly dependent of the number of observations. False neighbors are defined as points apparently lying close together due to projection that are separated in higher embedding dimensions.

Nearest neighbors $y(i)$ and $y(j)$ are declared false if:

$$\frac{|s(i+1) - s(j+1)|}{\|y(i) - y(j)\|} > R_{tol} \tag{9}$$

or if:

$$\frac{\|y(i) - y(j)\|^{2} - |s(i+1) - s(j+1)|^{2}}{R_A^2} > {A^2}_{tol}, \tag{10}$$

where:

$$R_A^2 = \frac{1}{N}\sum_{k=1}^{N}\left[s(k) - \langle s\rangle\right]^{2}, \;\; \langle s\rangle \;\; \textit{is the mean of } s(t). \tag{11}$$

For a deterministic process, the percentage of false nearest neighbors should drop to zero or some acceptable small numbers by increasing the embedding

Table 1: Percentage of false nearest neighbors.

	1	2	3	4	5	6
white noise	51.2%	21%	11.2%	13.7%	30.1%	22.1%
colored noise	11%	0%	0%	0%	0%	0%
Hénon	10.1%	2.9%	0%	0%	0%	0%
Logistic	0%	0%	0%	0%	0%	0%
Lorenz	6.9%	0%	0%	0%	0%	0%
Mackey Glass	60%	64%	56%	45%	30%	39%

Table 2: The values of the measure of determinism D.

	1	2	3	4	5	6
white noise	0	0.04	0.051	0.046	0.056	0.05
colored noise	0.047	0.094	0.097	0.073	0.067	0.054
Logistic map	0.5	0.248	0.186	0.136	0.11	0.085
Hénon map	0.11	0.265	0.213	0.157	0.137	0.108
Lorenz	0.127	0.289	0.23	0.187	0.156	0.124
Mackey Glass	0	0.06	0.086	0.102	0.142	0.13

dimension. For the following application as in most studies, R_{tol} is set to 10 and A_{tol} to 2.

Results displayed in Table 1 show that, the method of false neighbors is unable to discriminate between the Mackey Glass process and white noise. Moreover, it fails to distinguish colored noise from deterministic process. The measure of determinism presented here is more effective. Indeed, for chaotic processes, the quantity D becomes superior to 0.1 for a sufficient high value of m, while for random numbers, it never exceeds 0.1 (see table 2). In addition, values obtained for the logistic map are very close to theoretical values $D = 1/(2m)$. Finally, it should be noticed that the low values of D obtained for the Mackey Glass system suggest that the method would be unable to discriminate chaotic processes with higher dimension from random numbers and particularly from colored noise.

The poor results obtained by the false nearest neighbors approach are not surprising. Indeed it is well known that temporal correlations in time series cause difficulties to traditional methods (Theiler 1986). This problem can be treated, notably in using the method of surrogate data (Theiler et al. 1992). Nevertheless, according to numerous authors (see Osborne 1989), the lack of a sufficient number of observation has the same effect as the presence of temporal correlation: it produces a spurious dimension estimate of the dimension and prevents one from distinguishing correctly random numbers from chaotic time series. The measure of determinism presented here overcomes this shortcoming and is specially designed to detect low dimensional chaos in small sample sets. The specific task for which this method is designed can find many applications in economics.

3 Application to financial time series

The presence of nonlinear determinism in asset prices dynamics is of great importance in finance since it indicates a certain degree of predictability, and thus is susceptible to invalidate the informational efficiency hypothesis. Indeed, numerous authors have studied the possibility of exploiting nonlinear dependencies for the forecast of stock prices fluctuations (Wesner 2001, Fernandez-Rodriguez et al. 1999, LeBaron 1992). Here again, all those works have only studied relatively long time series, that is weekly or daily observations.

The measure of determinism presented in the previous section was calculated for time series of six major stock prices indices. Data series under investigation are : Dow Jones annual returns from 1896 to 2002 (104 observations), and monthly returns of the Nasdaq, the S&P500, the Nikkei and the FT-SE100 indexes over the period [1985:1-2003:5] (220 observations).

Table 3: The values of the measure of determinism D for stock prices indices data.

	1	2	3	4	5	6
Dow Jones yearly	0	0.048	0.078	0.79	0.06	0.055
Nasdaq monthly	0	0.039	0.068	0.039	0.058	0.047
Nikei monthly	0	0.027	0.051	0.054	0.05	0.048
FT-SE 100 monthly	0.005	0.034	0.058	0.053	0.044	0.054
S&P500 monthly	0	0.025	0.052	0.043	0.048	0.056

Results (Table 3) show that in all cases, as for random numbers, the quantity D does not exceed 0.1. In summary, this application does not provide evidence of the presence of low dimensional chaos in the dynamics of those stock price indices. The results are in adequation with the informational efficiency hypothesis.

4 Conclusion

A quantitative measure of determinism in a time series was proposed. The main advantage of the method is that it works well for very short time series, is very simple, requires few computer resources and does not contain subjective parameters. The method was applied to stock prices indices data on low frequency.

Although this application has not produced positive results on the presence of low dimensional chaos in stock prices dynamics, it shows the possibilities given by the method. Indeed, it could be applied to economic or financial time series on low frequency or over short periods of time, for which relatively few observations are available.

References

[1] Aleksic Z. (1991) "Estimating the embedding dimension", *Physica D* **52,** 362-368.

[2] Bachelier L. (1900) "Théorie de la speculation" *Annales Scientifiques de l'Ecole Normale Superieure*, III-17, 21–86.

[3] Brock W.A. (1986) "Distinguishing Random and Deterministic Systems: Abriged Version", *Journal of Economic Theory* **40**, 168-195.

[4] Cao L. (1997) "Practical method for determining the minimum embedding dimension of a scalar time series", *Physica D* **110,** 43-50.

[5] Fama E. (1965) "Random walks in stock market prices", *Financial Analysts Journal* **21**(5), 34-109.

[6] Farmer J.D. and Sidorowich J.J. (1987), "Predicting chaotic time series", *Physical Review Letters* **59**, n°8, 845-848.

[7] Fernandez-Rodriguez F., Sosvilla-Rivero S., Garça-Artiles M.D. (1999) "Dancing with bulls and bears : Nearest-neighbour forecast for the Nikkei index", *Japan and the World Economy* **11**, 395-413.

[8] Fernandez-Rodriguez F., Sosvilla-Rivero S., Andrada-Felix J. (2003) "Nearest-neighbour Predictions in Foreign Exchange Markets", in Shu-Heng Chen and Paul Wang (eds), *Computational Intelligence in Economics and Finance* (Berlin: Physica Verlag), 279-375.

[9] Grassberger P. and Procaccia I. (1983), "Chararcterization of strange attractors", *Physical Review Letters* **50**, 189-208.

[10] Grandmont J.M. (1985) "On Endogenous Competitive Business Cycles", *Econometrica* **53**, 995-1045.

[11] Kennel M. , Brown R. and Abarbanel H. (1992) "Determining embedding dimension for phase-space reconstruction using a geometrical construction", *Physical Review A* **45**, 3403.

[12] LeBaron B. (1992) "Nonlinear Forecast for the S&P Stock Index", in *Nonlinear Modelling and Forecasting, SFI Studies in the Sciences of Complexity*, Eds. Casdagli Eubank, Vol.XII., 381-393.

[13] Lucas R. (1978) "Asset Prices in an Exchange Economy", *Econometrica* **46**, 1429-1455.

[14] Osborne A..R. and Provenzale A.(1989) "Finite correlation dimension for stochastic systems with power-law spectra", *Physica D* **35,** 357-381.

[15] Takens F. (1981), "Detecting strange attractors in fluid turbulence", In Rand D.A., Young L.S. (Eds.), *Dynamical System and Turbulence, Lecture notes in mathematics*, Springer Verlag, Berlin, 366-381.

[16] Theiler J. (1986) "Spurious dimension from correlation algorithm applied to time series data", *Physical Review A* **34**, 2427-2432.

[17] Theiler J. Galdrikian B., Longtin A., Eubank S., Farmer J.D. (1992) "Using Surrogate Data to Detect Nonlinearity in Time Series", *Nonlinear Modelling and Forecasting, SFI Studies in the Sciences of Complexity*, Eds. Casdagli Eubank, Vol.XII., 163-188.

[18] Wayland R., Bromley D., Pickett D. and Passamante A. (1993) "Recognizing determinism in a time series", *Physical Review Letters* **70**, n°5, 580-587.

[19] Wesner N. (2001) "Prévision non linéaire et efficience informationnelle : une application à la dynamique du CAC40", *Journal de la Société Française de Statistique*, tome 142, n°3, 89-109.

[20] Wolf A., Swift J.B., Swinney H.L. and Vastano J.A. (1985) "Determining Lyapunov exponents from a time series", *Physica D* **16,** 285-317.