Detection of nonlinear dynamics in short, noisy time series

Mauricio Barahona* & Chi-Sang Poon†‡

* Physics Department and † Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

THE accurate identification of deterministic dynamics in an experimentally obtained time series[1–5] can lead to new insights regarding underlying physical processes, or enable prediction, at least on short timescales. But deterministic chaos arising from a nonlinear dynamical system can easily be mistaken for random noise[6–8]. Available methods to distinguish deterministic chaos from noise can be quite effective, but their performance depends on the availability of long data sets, and is severely degraded by measurement noise. Moreover, such methods are often incapable of detecting chaos in the presence of strong periodicity, which tends to hide underlying fractal structures[9]. Here we present a computational procedure, based on a comparison of the prediction power of linear and nonlinear models of the Volterra–Wiener form[10], which is capable of robust and highly sensitive statistical detection of deterministic dynamics, including chaotic dynamics, in experimental time series. This method is superior to other techniques[1–6,11,12] when applied to short time series, either continuous or discrete, even when heavily contaminated with noise, or in the presence of strong periodicity.

Consider the usual description of a dynamical system as a 'black box' with input x_n and output y_n, at time $n = 1, \ldots, N$ in multiples of the sampling time τ. The discrete Volterra series is then a Taylor-like polynomial expansion of y_n in terms of $x_n, x_{n-1}, \ldots, x_{n-\kappa+1}$, where κ is the memory of the system. To overcome the computational intractability of the Volterra series, Wiener introduced an orthogonal formulation[13] assuming that x_n is a gaussian white noise sequence over an infinite time period. Korenberg's recasting of the expansion[14], on which our approach is based, relaxes these restrictions thereby allowing its application to finite and arbitrary signals.

For a dynamical system, either strictly autonomous or reformulated as such, we propose a closed-loop version of the Volterra series in which the output y_n feeds back as delayed input (that is, $x_n \equiv y_{n-1}$). Within this framework, we analyse univariate time series by using a discrete Volterra–Wiener–Korenberg series of degree d and memory κ as a model to calculate the predicted time series y_n^{calc}:

$$y_n^{\text{calc}} = a_0 + a_1 y_{n-1} + a_2 y_{n-2} + \cdots + a_\kappa y_{n-\kappa} + a_{\kappa+1} y_{n-1}^2 + a_{\kappa+2} y_{n-1} y_{n-2} + \cdots + a_{M-1} y_{n-\kappa}^d = \sum_{m=0}^{M-1} a_m z_m(n) \quad (1)$$

where the functional basis $\{z_m(n)\}$ is composed of all the distinct combinations of the embedding space coordinates[15] $(y_{n-1}, y_{n-2}, \ldots, y_{n-\kappa})$ up to degree d, with a total dimension $M = (\kappa + d)!/(d!\kappa!)$. Thus, each model is parametrized by κ and d, which correspond to the embedding dimension and the degree of nonlinearity of the model, respectively. The coefficients a_m are recursively estimated through a Gram–Schmidt procedure[14] from linear and nonlinear autocorrelations of the data series itself. The calculations can be readily performed on a workstation.

The short-term prediction power of a model is then measured by the standard one-step-ahead prediction error

$$\varepsilon(\kappa, d)^2 \equiv \frac{\sum_{n=1}^{N} (y_n^{\text{calc}}(\kappa, d) - y_n)^2}{\sum_{n=1}^{N} (y_n - \bar{y})^2} \quad (2)$$

where $\bar{y} = 1/N \sum_{n=1}^{N} y_n$ and $\varepsilon(\kappa, d)^2$ is in effect a normalized variance of the error residuals. We now search for the best

‡ To whom correspondence should be addressed.