Nonlinear prediction as a way of distinguishing chaos from random fractal sequences

A. A. Tsonis* & J. B. Elsner†

* Department of Geosciences, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA
† Department of Meteorology, Florida State University, Tallahassee, Florida 32306, USA

NONLINEAR forecasting has recently been shown to distinguish between deterministic chaos and uncorrelated (white) noise added to periodic signals[1], and can be used to estimate the degree of chaos in the underlying dynamical system[2]. Distinguishing the more general class of coloured (autocorrelated) noise has proven more difficult because, unlike additive noise, the correlation between predicted and actual values measured may decrease with time—a property synonymous with chaos. Here, we show that by determining the scaling properties of the prediction error as a function of time, we can use nonlinear prediction to distinguish between chaos and random fractal sequences. Random fractal sequences are a particular class of coloured noise which represent stochastic (infinite-dimensional) systems with power-law spectra. Such sequences have been known to fool other procedures for identifying chaotic behaviour in natural time series[9], particularly when the data sets are small. The recognition of this type of noise is of practical importance, as measurements from a variety of dynamical systems (such as three-dimensional turbulence, two-dimensional and geostrophic turbulence, internal ocean waves, sandpile models, drifter trajectories in large-scale flows, the motion of a classical electron in a crystal and other low-dimensional systems) may over some range of frequencies exhibit power-law spectra.

In the past, the identification of chaotic dynamics from time series relied heavily on the estimation of the dimension of the underlying attractor[3-6], but this approach has certain drawbacks. First, the algorithms involved[7,8] require a large number of data points, which are often not available. Second, even when enough points are available, a finite dimension may not be indicative of deterministic chaos. For small data sets, random fractal sequences (fractional brownian motions) may show an anomalous scaling which can be interpreted as a finite dimension[9,10]. Fractional brownian motions (FBM) are random processes and thus are dictated by an infinite number of degrees of freedom. They have power spectra of the form $P(f)=Cf^{-a}$, where $a=2H+1$ with $0<H<1$. The exponent H is related to the fractal dimension, D_f, of the trail of the fractional brownian motion through $D_f=1/H$. Irrespective of the relevance of such sequences to the natural world, these drawbacks mean that the methods have often required subjective judgement about whether an attractor of a given dimension exists, and other approaches have been sought.

One of these approaches is nonlinear prediction[11-13] which is becoming indispensable in the study of chaotic dynamical systems. The idea behind using nonlinear prediction as a signature of chaos is simple. Chaotic systems obey certain rules. The limited predictive power of chaotic dynamical systems is because they are sensitive to initial conditions and because we cannot have perfect measurements (which require an infinite amount of information). One would therefore expect chaos to be characterized by a decrease of the correlation between predicted and actual values as prediction time increases. This property can be used to differentiate between chaos and additive uncorrelated noise. Additive noise produces a fixed amount of error regardless of the prediction time, as has been demonstrated[1] for deterministic maps and data from the natural world (the same conclusions were reached by Wolpert and Miall[14] using the same data and a neural network as the prediction method). If a system is chaotic, then the decrease of predictive power with prediction time is equivalent to the presence of a positive Lyapunov exponent. The decrease with time of the correlation between predicted and actual values can indeed be used to infer the largest positive Lyapunov exponent[2]. Those correlation functions do not require extremely large samples and have thus become a popular alternative to dimension estimates.

But how does nonlinear prediction perform when it comes to random fractal sequences? If random fractal sequences 'fool' the algorithms that characterize the structure of chaotic attractors, they may also fool prediction approaches whose signatures are based on the existence of those structures. In fact, as we will show, the correlation between predicted and actual values of fractional brownian motions also decreases with prediction time. In this case nonlinear prediction may have the same drawback as dimension estimates, and therefore may not be a definitive way of identifying chaos. Sugihara and May[1] speculated that the relationship between correlation and prediction time might scale differently but did not offer a solution to this central problem. Here we present theoretical arguments together with computer simulations that settle this issue.

If the predicted value at some time t is assumed to be a random variable x_t and the actual value a random variable y_t, then the Pearson's correlation coefficient between these two distributions, $r(t)$, ranges in magnitude between zero and one for uncorrelated and identical distributions, respectively. The correlation coefficient is defined as

$$r(t) = (\langle x_t y_t \rangle - \langle x_t \rangle \langle y_t \rangle)/\sigma(x_t)\sigma(y_t) \tag{1}$$

where the triangular brackets denote the average over a series of predictions and σ denotes the standard deviation. For stationary chaotic signals the correlation can take the form[2]

$$r(t) = 1 - (s^2(0)\, e^{2Kt})/2\sigma^2(y_t) \tag{2}$$

where $s(0)$, $\sigma(y_t)$ and K are positive constants. For stationary processes $\sigma(y_t)$ is considered independent of the prediction time, t. Equation (2) dictates that the correlation should decrease exponentially with prediction time. Chaotic systems generally have positive Lyapunov exponents and exponential divergence of nearby trajectories.

If x is a FBM then $x_t = N(0, \propto t^{2H})$—that is, it is normally distributed with zero mean and a variance proportional to t^{2H}—and $\mathrm{cov}\,(x_t, y_t) = \langle x_t y_t \rangle$. In nonlinear prediction it is common to consider that

$$y_t = f\left(\sum_i w_i x_{t-i}\right) = f(z_t) \tag{3}$$

where f is some nonlinear function ranging between zero and one, and w_i are coefficients. If we expand f into a Taylor series we get

$$\begin{aligned}\langle x_t y_t \rangle &= \mathrm{cov}\,(x_t, y_t)\\ &= \mathrm{cov}\,(x_t, f(0) + z_t f'(0) + z_t^2 f''(0)/2 + \ldots)\\ &= \mathrm{cov}\,(x_t, f(0)) + f'(0)\,\mathrm{cov}\,(x_t, z_t)\\ &\quad + f''(0)\,\mathrm{cov}\,(x_t, z_t^2)/2 + \ldots\end{aligned} \tag{4}$$

where $f(0)$, $f'(0)$ and $f''(0)$ represent some constants. If we consider only terms of second order or lower, then according to the problem in hand we have $\mathrm{cov}\,(x_t, f(0)) = 0$ and $\mathrm{cov}\,(x_t, z_t^2) = \mathrm{E}(x_t z_t z_t) = 0$ (ref. 15). This allows us to write equation (4) as

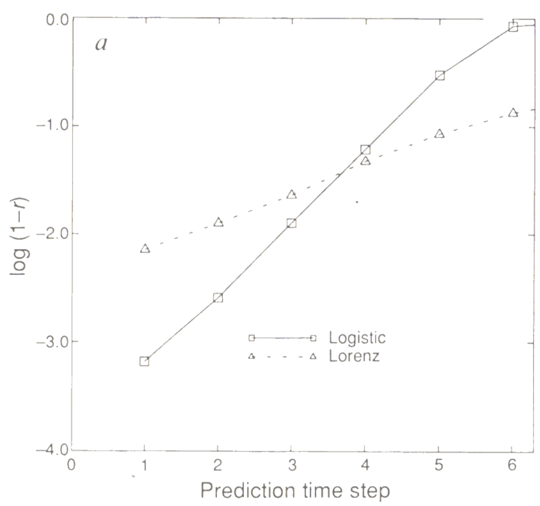
$$\langle x_t y_t \rangle = A\,\mathrm{cov}\,(x_t, z_t) \tag{5}$$

where $A = f'(0)$. Recalling that if $X = N(0, \min(t_1, t_2))$ then $\mathrm{cov}\,(X_{t1}, X_{t2}) = \min\,(t_1, t_2)$ (ref. 15), and considering equation (3) we have

$$\begin{aligned}\langle x_t y_t \rangle &= A\,\mathrm{cov}\left(x_t, \sum_i w_i x_{t-i}\right)\\ &= A \sum_i w_i\,\mathrm{cov}\,(x_t, x_{t-i})\\ &= A \sum_i w_i \min\,[t^{2H}, (t-i)^{2H}]\\ &= A \sum_i w_i (t-i)^{2H}\end{aligned} \tag{6}$$

Consequently, the correlation coefficient takes the form

$$r(t) = B \sum_i w_i (1 - i/t)^{2H} \tag{7}$$



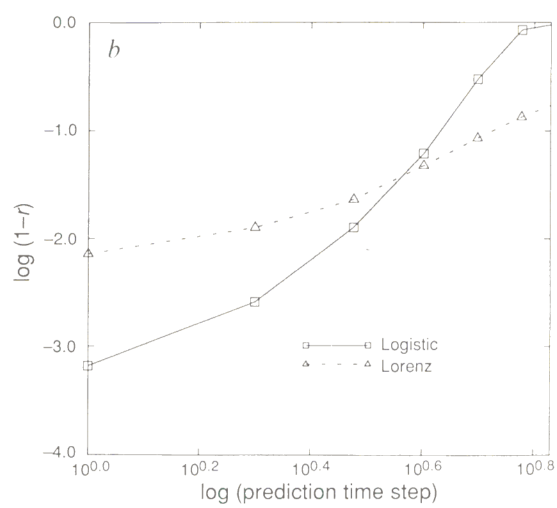


FIG. 1 Nonlinear prediction can be used to distinguish chaotic signals from random fractal sequences. *a*, Logarithm of $1 - r(t)$ against prediction time step, t; *b*, logarithm of $1 - r(t)$ against log t for single realizations of the x coordinate of the Lorenz system, and the logistic map. These simulations show the expected scaling from the theoretical arguments developed in the text. For short prediction times, scaling is observed in the semi-log plot, indicating that the curve of $1 - r(t)$ against t curve for the chaotic signals is exponential.

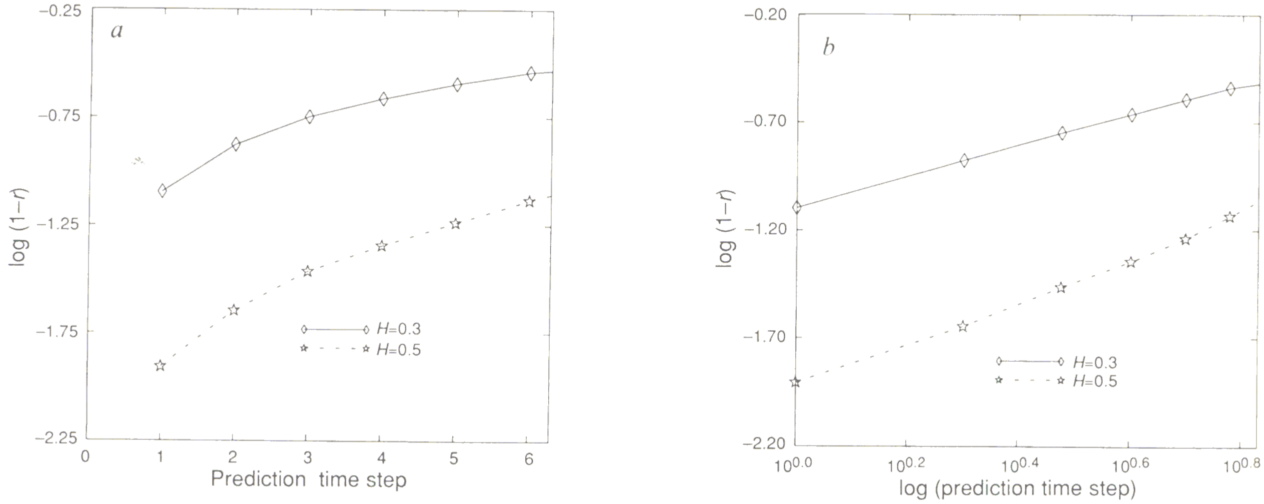


FIG. 2 As Fig. 1, but showing single realization results for two FBMs with $H=0.3$ and 0.5. Here scaling is observed in the log–log plot indicating that for the FBMs the curve of $1-r(t)$ against t is a power-law curve, as expected.

where B is some constant. Considering that in our case $\operatorname{cov}(x_t, z_t^3) = E(x_t z_t z_t z_t) = 3\operatorname{cov}(x_t, z_t)\operatorname{Var}(z_t)$ and that $\operatorname{cov}(x_t, z_t^4) = E(x_t z_t z_t z_t z_t) = 0$, addition of the next two higher-order terms in the expansion modifies equation (7) to

$$r(t) = \left[B + C\left(\sum_i w_i^2 (t-i)^{2H} + \sum_{i<j}\sum w_i w_j (t-j)^{2H}\right)\right] \times \sum_i w_i (1-i/t)^{2H} \qquad (8)$$

where C is some other constant. Equations (7) and (8) do not include any exponential terms. They represent curves determined by a synthesis of power-law terms. This is a direct consequence of the power-law characterization of FBM. Including additional terms in the expansion of f does not change this property. The above arguments, and especially equations (2) and (7) or (8), show that we should expect differences in the scaling of the correlation coefficient with prediction time between chaotic signals and FBMs. For chaotic systems the logarithm of $1-r(t)$ should be a linear function of prediction time step t, and for FBMs it should be a linear function of log t.

These scaling differences are emphasized by computer simulations. Figure 1a is a plot of $\log(1-r(t))$ against t, and Fig. 1b is a plot of $\log(1-r(t))$ against $\log t$ showing single realization results for the x coordinate of the Lorenz system, and for the logistic map $x_{n+1} = 4x_n(1-x_n)$. The integration step for the Lorenz system was 0.03. The prediction method was similar to that used in ref. 1, which is a simple variant of the Farmer and Sidorowich interpolative approach[11]. For each example we take 1,000 values from the corresponding time series, use the first 500 as a database, and make predictions on the last 500 values. Correlation coefficients are thus based on sample sizes of ~500. Note that for the Lorenz time series, the prediction time step is the integration step. Figure 2 is similar to Fig. 1 but shows single realization results for two FBMs with $H=0.3$ and 0.5. In our simulations the embedding dimension is always $2D_f+1$. The time delay used to define the coordinates of the embedding space was $\tau=1$. The results show a decrease of the correlation coefficient with prediction time in all cases. In accordance with our theoretical arguments, however, Figs 1 and 2 indicate that for short prediction times, $\log(1-r(t))$ is a linear function of t for the chaotic signals and a linear function of $\log t$ for the FBMs. Although the role of τ in nonlinear prediction needs to be better understood, preliminary experiments with different values of τ indicated that the expected scaling is preserved; other prediction techniques, such as neural networks[13], give

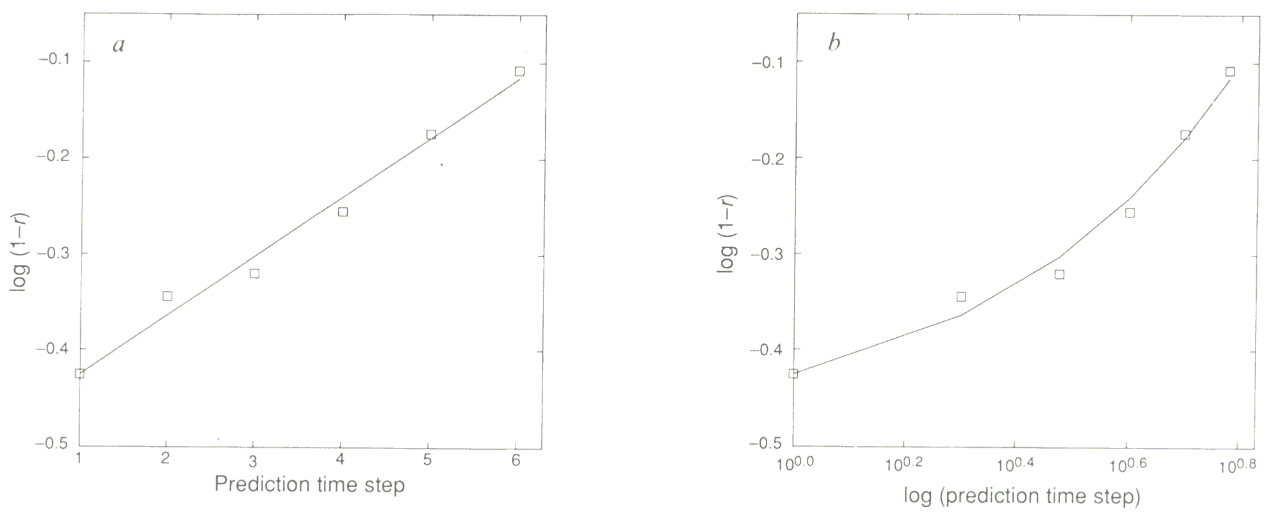


FIG. 3 Nonlinear prediction on a time series of the Southern Oscillation Index. a, $\log(1-r(t))$ against t; b, $\log(1-r(t))$ against $\log t$. For short prediction time steps, scaling is observed in the semi-log plot indicating that the curve of $1-r(t)$ against t is exponential, and thus the time series is a chaotic signal.

similar results but seemed to yield less accurate predictions than the interpolative approach.

For periodic signals with additive uncorrelated noise, the correlation coefficient is independent of the prediction time. A fall in the correlation with prediction time does not indicate a periodic signal with additive noise, but it may indicate a chaotic signal with or without noise (in the chaotic case, additive noise produces a constant offset in the correlation which would not obliterate the fall-off). Our work provides the means to go one step further and decide whether such a fall-off is indicative of a random fractal sequence. A direct approach will be to check the scaling behaviour of $\log(1-r(t))$ against t and of $\log(1-r(t))$ against $\log t$. For example, Fig. 3 shows results from a time series of the Southern Oscillation Index, which is derived from the mean sea-level pressure difference between the Tahiti and Darwin stations. The record is 1,248 values long (monthly values from January 1883 to December 1986) and is related to the El Niño which is hypothesized to be chaotic[16]. A dimension of around five has been suggested for this type of data[17]. Figure 3a is a semi-log plot and Fig. 3b a log-log plot of $1-r(t)$ against t. The solid lines indicate best fits. For the semi-log plot the best fit is linear, whereas for the log-log plot it appears to be non-linear. According to our theory this indicates that the time series is indeed chaotic. In some cases, noise, data imperfections and data length may make the identification of the actual scaling difficult. In such cases nonlinear prediction could be used statistically. For example, if a time series is suspected to be a FBM, then prediction results could be compared with the average prediction properties of the family of FBMs in question. □

Received 12 November 1991; accepted 5 May 1992.

1. Sugihara, G. & May, R. M. *Nature* **344,** 734–741 (1990).
2. Wales, D. J. *Nature* **350,** 485–488 (1991).
3. Tsonis, A. A. & Elsner, J. B. *Nature* **333,** 545–547 (1988).
4. Essex, C., Lookman, T. & Nerenberg, M. A. H. *Nature* **326,** 64–66 (1987).
5. Frank, G. W., Lookman, T., Essex, C. & Nerenberg, M. A. H. *Physica* D**46,** 427–438 (1990).
6. Schaffer, W. M., Olsen, L. F., Truty, G. L., Fulmer, S. L. & Graser, D. J. in *From Chemical to Biological Organization* (eds Marcus, M., Muller, S. C. & Nicolis, G.) (Springer, New York, 1988).
7. Grassberger, P. & Procaccia, I. *Physica* D**9**, 186–208 (1983).
8. Grassberger, P. & Procaccia, I. *Phys. Rev. Lett.* **50,** 346–349 (1983).
9. Osborne, A. R. & Provenzale, A. *Physica* D**35,** 357–381 (1989).
10. Theiler, J. *Phys. Lett.* A**155,** 480–492 (1991).
11. Farmer, J. & Sidorowich, J. J. *Phys. Rev. Lett.* **62**, 845–848 (1987).
12. Casdagli, M. *Physica* D**35,** 335–356 (1989).
13. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. *Nature* **323,** 533–536 (1986).
14. Wolpert, D. M. & Miall, R. C. *Proc. R. Soc.* B**242,** 82–86 (1990).
15. Anderson, T. W. *An Introduction to Multivariate Statistics* (Wiley, New York, 1958).
16. Valis, G. K. *Science* **232,** 243–245 (1986).
17. Hense, A. *Beitr. Phys. Atm.* **60,** 34–47 (1987).

ACKNOWLEDGEMENTS. We thank J. D. Farmer, K. P. Georgakakos and J. H. Beder for discussion and comments, S. Nicholson for providing the Southern Oscillation Index data, and J. J. Sidorowich for his review and advice. Parts of this work were supported by the NOAA TOGA programme on prediction.