ANALYSIS OF NONLINEAR TIME SERIES (AND CHAOS)
BY BISPECTRAL METHODS

By

T. Subba Rao

IMA Preprint Series # 782
February 1991
ANALYSIS OF NONLINEAR TIME SERIES (AND CHAOS) BY BISPECTRAL METHODS

BY

T. Subba Rao
Department of Mathematics
University of Manchester Institute of Science and Technology
Manchester M60 1QD
ENGLAND

Abstract

Second order spectra have played a very significant role in the analysis of linear (Gaussian) time series. When the series is non-Gaussian (and nonlinear) it is important to study higher order spectra. Here we briefly consider the estimation of the bispectrum, and discuss the usefulness of this function in several situations, for example, when estimating the parameters of the signal in the presence of noise, for discriminating "deterministic" nonlinear models and chaotic models etc. We also consider bilinear models which were introduced recently in time series literature, and discuss their estimation and study the forecasts obtained in a specific example. Lastly we point out how this model can be generalised to deal with nonlinear long range dependence.

Key words: Bispectrum, chaos, bilinear models, nonlinear forecasts and nonlinear long range dependence.

This paper was presented at "Nonlinear Modeling and Forecasting" Conference organised by Santa Fe Institute, Sante Fe, New Mexico during September 17 - 21, 1990.
1. INTRODUCTION

One of the usual assumptions that is often made when analysing time series is that the time series is linear, and perhaps even Gaussian. If the process is Gaussian, then the second order covariances (and second order spectra) will contain all the useful information. If not one has to calculate the higher order spectrum to study the departures from linearity and Gaussianity. The simplest higher order spectrum is bispectrum. Bispectrum was originally proposed by Hasselman, Munk and Macdonald (1963) for investigating nonlinear interaction of ocean waves, and Godfrey (1965) used it for analysis of economic time series. Sato, Sasaki and Nakamura (1977) have analysed acoustic gear noise by bispectra. Lii and Rosenblatt (1982) have used it for deconvolution of seismic signals. In a series of papers, Lii, Rosenblatt and Van Atta (1976), Van Atta (1979), Helland, Lii and Rosenblatt (1979) have described how bispectrum could be used to study nonlinear spectral transfer of energy in turbulence.

Subba Rao and Gabr (1980) and Hinich (1982) have proposed tests for linearity and Gaussianity based on bispectrum, and possible applications seems to be unlimited. In this paper we briefly review a method of estimation of bispectrum and study possible applications to non-Gaussian signals, and especially we will study the phenomenon of "chaos." Our investigations in this new field are preliminary, and it seems, on the basis of the results we obtained, the estimated higher order spectra possibly could be used to distinguish between nonlinear deterministic stable systems and nonlinear deterministic chaotic systems. On the basis of these investigations it seems that for stable nonlinear deterministic systems the energy in the estimates of the spectrum and bispectrum seems to be in the low frequency range, where as for the "chaotic" systems the energy seems to be distributed over a wide band in the high frequency range and this may be consistent with "bifurcation" phenomenon of chaotic systems. In sections 2 and 3 we concentrate on frequency domain approach for analysing non-Gaussian time series and time domain approaches are considered in later sections. In section 4 we consider the properties of one nonlinear model called bilinear model and study their properties and aspects of forecasting. In the final section 5, we point out the possibility of extending the bilinear models to deal with long range dependence.

2. SECOND ORDER AND HIGHER ORDER SPECTRA.

Let \( x(t) \) be a discrete parameter, real valued, time series. Let the process \( \{x(t)\} \) satisfy the following conditions.

\[
(1) \quad E(x(t)) = \mu \text{ independent of } t \\
(11) \quad \text{Var}(x(t)) = \sigma^2 \text{ independent of } t \\
\text{Cov}(x(t), x(t+s)) = R(s) \text{ a function of } s \text{ only.}
\]

\[
(111) \quad \text{Cum}(x(t), x(t+s_1), x(t+s_2)) \\
= E(x(t) - \mu)(x(t+s_1) - \mu)(x(t+s_2) - \mu) \\
= C(s_1, s_2) \text{ is a function of } s_1 \text{ and } s_2 \text{ only}
\]
The time series \( \{x(t)\} \) which satisfies the above three conditions is said to be third order stationary.

We note that \( R(s) \) and \( C(s_1, s_2) \) satisfy the following symmetry conditions.

\[
R(s) = R(-s), \quad C(s_1, s_2) = C(s_2, s_1) = C(-s_1, -s_2 - s_1) = C(s_1 - s_2, -s_2)
\]

The second order spectrum \( f(\omega) \) and the third order spectrum \( f(\omega_1, \omega_2) \) are defined as follows:

\[
f(\omega) = \frac{1}{2\pi} \sum_{\omega} R(s)e^{i\omega}, \quad -\pi \leq \omega \leq \pi
\]

\[
f(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{\omega_1} \sum_{\omega_2} C(s_1, s_2)e^{i\omega_1 - i\omega_2}
\]  

(2.1)

In view of the above symmetries, we have

\[
f(\omega) = f(-\omega) \quad \text{and} \quad f(\omega_1, \omega_2) = f(\omega_2, \omega_1) = f(-\omega_1, -\omega_2) = f(-\omega_1, -\omega_2, \omega_2) = f(-\omega_1, -\omega_2)
\]

(2.2)

and because of (2.2), the bispectrum \( f(\omega_1, \omega_2) \) is completely specified on any one of the twelve sectors, including the boundaries (see Subba Rao and Gabr, 1984, Fig 1.3, page 13) Bispectrum is usually complex valued. It is easy to evaluate the theoretical forms of the spectrum and the bispectrum if we know the model the series \( x(t) \) satisfies. At least this is true in the case of linear models. For example, if \( x(t) \) satisfies the finite parameter autoregressive moving average model of order \( (p,q) \) of the form

\[
x(t) + a_1 x(t-1) + \ldots + a_p x(t-p) = e(t) + b_1 e(t-1) + \ldots + b_q e(t-q)
\]

where \( \{e(t)\} \) is a sequence of independent, identically distributed random variables with \( E(e(t)) = 0, \) \( \text{var}(e(t)) = \sigma^2, \) \( \mu_3 = E(e^3(t)) \), then we know

\[
f(\omega) = \frac{\sigma^2}{2\pi} \cdot |\Gamma(\omega)|^2,
\]

(2.3)

\[
f(\omega_1, \omega_2) = \frac{\mu_3}{(2\pi)^2} \Gamma(e^{-1}) \Gamma(e^{-2}) \Gamma(e^{-1} + e^{-2})
\]
where

\[ \Gamma(e^{-i\omega}) = \frac{(1 + b_1 e^{-i\omega} + \ldots + b_q e^{-iq\omega})}{1 + a_1 e^{-i\omega} + \ldots + a_p e^{-ip\omega}} \]

In general, if \( x(t) \) satisfies the linear representation

\[ x(t) = \sum_{u=0}^{\infty} g_u e(t-u), \]

then \( f(\omega) \) and \( f(\omega_1, \omega_2) \) are given as above with the transfer function

\[ \Gamma(e^{-i\omega}) = \sum_{u=0}^{\infty} g_u e^{-i\omega u}. \]

We observe that if the time series \( \{x(t)\} \) is Gaussian, then \( f(\omega_1, \omega_2) = 0 \) for all \( \omega_1 \) and \( \omega_2 \). If \( x(t) \) is linear, but not Gaussian, then

\[ \frac{|f(\omega_1, \omega_2)|^2}{f(\omega_1) f(\omega_2) f(\omega_1 + \omega_2)} = \text{constant for all } \omega_1 \text{ and } \omega_2. \quad (2.4) \]

In other words the bispectrum can be used for testing departure from Gaussianity and linearity, and this is the basis of the Subba Rao and Gabr's test (1980) and Hinich's test (Hinich, 1982).

We now consider, briefly, the estimation of \( f(\omega) \) and \( f(\omega_1, \omega_2) \). For details of estimation see Priestley (1981) and Subba Rao and Gabr (1984).

**Estimation of \( f(\omega) \).**

Let \( x(1), x(2), \ldots, x(n) \) be a sample from \( \{x(t)\} \). Then the natural estimates of \( \mu, R(s) \) and \( C(s_1, s_2) \) are given by

\[ \hat{x} = \frac{1}{N} \sum x(t), \quad R(s) = \frac{1}{N} \sum_{t=1}^{N+s} (x(t) - \hat{x})(x(t+s) - \hat{x}), \quad s \geq 0 \]

and\[ \hat{C}(s_1, s_2) = \frac{1}{N} \sum_{t=1}^{N-\gamma} (x(t) - \hat{x})(x(t+s_1) - \hat{x})(x(t+s_2) - \hat{x}), \]

where \( \gamma = \max (s_1, s_2) \) \( s \geq 0, s \geq 0 \). A form of the spectral estimate of \( f(\omega) \) is

\[ \hat{f}(\omega) = \frac{1}{2\pi} \sum_{\tau=-N+1}^{N-1} \hat{\lambda}(\tau) R(\tau) \cos \omega \tau, \]

4
where $M = M(N)$ and $\lambda(.)$ is a lag window generator. If $\lambda(s) = 0$, for $|s| \geq 1$, $M$ corresponds to the truncation point. We assume that the function $\lambda(s)$ is a bounded, even and square integrable such that $\lambda(0) = 1$. The integer $M$ is chosen such that $M \to \infty$, $N \to \infty$, $M \to N$. It is well known that

$$\text{Var} \left( \hat{f}(\omega) \right) = \frac{M}{N} f^2(\omega) \int \lambda^2(s) ds, \quad \omega \neq 0, \pi$$

In table 1, we give some standard lag window generators.

### Table 1. Lag Window Generators

<table>
<thead>
<tr>
<th>Window</th>
<th>Function $\lambda(s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deniell Window</td>
<td>$\lambda_D(s) = \frac{\sin \pi s}{\pi^2}$</td>
</tr>
<tr>
<td>Tukey - Hanning Window</td>
<td>$\lambda_T(s) = \begin{cases} 0.54 + 0.46 \cos \pi s &amp;</td>
</tr>
<tr>
<td>Parzen Window</td>
<td>$\lambda_P(s) = \begin{cases} 1 - 6s^2 + 6</td>
</tr>
<tr>
<td>Bartlett - Priestley Window</td>
<td>$\lambda_{BP}(s) = \frac{3}{(\pi s)^2} \left{ \frac{\sin \pi s}{\pi s} - \cos \pi s \right}$</td>
</tr>
</tbody>
</table>

### Estimation of $f(\omega_1, \omega_2)$

Let $K_O(\theta_1, \theta_2)$ be a bounded and a non-negative function satisfying

1. $$\int \int K_O(\theta_1, \theta_2) d\theta_1 d\theta_2 = 1$$
2. $$\int \int K_O^2(\theta_1, \theta_2) d\theta_1 d\theta_2 < \infty, \quad \int \int K_O(\theta_1, \theta_2) d\theta_1 d\theta_2 < \infty$$
   \hspace{1cm} (i=1, 2)
3. $$K_O(\theta_1, \theta_2) = K_O(\theta_2, \theta_1) = K_O(\theta_1, -\theta_2) = K_O(-\theta_1, \theta_2) = K_O(-\theta_1, -\theta_2).$$
Then the bispectral estimate $\hat{f}(\omega_1, \omega_2)$ is given by

$$
\hat{f}(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum \frac{\lambda(\tau_1/M, \tau_2/M) \hat{C}(\tau_1, \tau_2)}{N} e^{-i\tau_1 \omega_1} e^{-i\tau_2 \omega_2}
$$

where $M$, which is a function of $N$, is a window parameter chosen such that $M^2/N \to 0$ as $M \to \infty$, $N \to \infty$. Then (see Brillinger and Rosenblatt, 1967 and Rosenblatt and Van Ness, 1965).

$$
\text{Var}(\hat{f}(\omega_1, \omega_2)) = \frac{N^2}{N^2} \int \frac{V_2}{2\pi} f(\omega_1) f(\omega_2) f(\omega_1 + \omega_2), (\omega_1 < \omega_2 < \omega_1)
$$

where $V_2 = \int \int \lambda^2(u_1, u_2) du_1 du_2 = (2\pi)^2 \int \int K^2(\theta_1, \theta_2) d\theta_1 d\theta_2$

One choice for the two dimensional window $\lambda(u_1, u_2)$ is to consider product of three one dimensional windows of the form

$$
\lambda(u_1, u_2) = \lambda(u_1) \lambda(u_2) \lambda(u_1 - u_2)
$$

An optimum window has been obtained by Subba Rao and Gabr (1984), which produces the smallest mean square error.

One can also estimate the bispectrum using Fast Fourier transforms (see Huber et al. 1971).

3. APPLICATIONS OF BISPECTRA

As pointed out earlier, second order spectral analysis is not always sufficient to draw all the information contained in the series, if the series happens to be non-Gaussian. In fact, if one stops at second order analysis, it is possible to draw wrong conclusions as the following example shows.

Consider the nonlinear process $\{x(t)\}$ satisfying the model

$$
x(t) = e(t) + \beta e(t-1) e(t-2)
$$

where $\{e(t)\}$ are i.i.d random variables. We can easily show that Cov $(x(t), x(t+s) = 0, s \neq 0$ and this implies that the time series $\{x(t)\}$ is white noise when in fact $\{x(t)\}$ is highly correlated. If we compute the third order cumulants, Cum $(x(t), x(t+s_1), x(t+s_2))$, we find these cumulants are not zero for some values of $s_1$ and $s_2$, confirming that the underlying structure is nonlinear. Also, it may be important to mention here that some series which may look like nonstationary, may in fact be nonlinear.
We now consider a simple nonlinear model recently introduced called Bilinear model which has been studied extensively (see Granger and Andersen, 1978 Subba Rao 1981, Subba Rao and Gabr, 1984). Let $x(t)$ satisfy the equation

$$x(t) = a x(t-1) + b x(t-1) e(t-1) + e(t)$$

It is well known that the second order properties of this model are similar to the ARMA (1, 1) model or in other words the second order covariances cannot distinguish between a linear model and a bilinear model.

**Test for Linearity and Gaussianity**

If $x(t)$ satisfies the linear relation, then we know that $f(\omega_1, \omega_2)$ is given by (2.3), and this function is zero if the series is Gaussian. If it is linear and non-Gaussian, then the ratio given by (2.4), is constant for all $\omega_1$ and $\omega_2$. Using these properties, Subba Rao and Gabr (1980) and later Hinich (1982), Brockett et al. (1988) have constructed statistical tests for testing linearity and Gaussianity.

To illustrate these, we consider two well known data sets.

**Canadian Lynx data.** Consider the annual number of Canadian Lynx trapped in the Mackenzie river district of North West Canada for the years 1821 - 1934;

The bispectral density function of the logarithmically transformed data is computed using the optimum lag window (with $M = 16$) and the modulus is plotted in Fig 1 for the frequencies $\omega_1, \omega_2 = 0.0(0.01\pi) 0.26\pi$. There is a clear peak at $\omega_1 = \omega_2 = 0.20\pi$ which corresponds to the periodicity of 10 years. Since the bispectral values are non zero, it is clear that the process is non-Gaussian. It may be interesting to observe that the graph is smooth with a single dominant peak at the low frequency, and this is in contrast to what we will observe for chaotic models. It is reasonable to conclude that Canadian Lynx data is nonlinear, but not chaotic.

**Sunspot numbers.** For our second illustration, we consider the annual Wolfer sunspot numbers for the years 1700-1955 giving 256 observations. The modulus of the bispectral density function is plotted for frequencies $\omega_1, \omega_2 = 0.0 (0.01\pi) 0.27\pi$ in Fig 2. There is a clear peak at the frequency at $\omega_1 = \omega_2 = 0.18\pi$ corresponding to approximately 11 years. Once again we can confirm that the data is nonlinear, but not chaotic.

**Higher Order Cumulants and Noise.** In many practical situations, the signal $x(t)$, which we will assume to be non-Gaussian and perhaps nonlinear, is corrupted by noise $N(t)$. Actually, for each $t$, we observe $Z(t)$,

$$Z(t) = x(t) + N(t),$$  \hspace{1cm} (3.1)
where the noise $N(t)$ is assumed to be independent of $x(t)$ and is stationary up to the $r$th order.

In view of the relation (3.1), we have the $r$th order cumulant of $Z(t)$,

$$\text{Cum} \left( Z(t), Z(t+s_1), \ldots, Z(t+s_{r-1}) \right) = \text{Cum} \left( x(t), x(t+s_1) \ldots x(t+s_{r-1}) \right)$$

$$+ \text{Cum} \left( N(t), N(t+s_1), \ldots, N(t+s_{r-1}) \right).$$

If we take Fourier transforms both sides we obtain the relation between higher order cumulant spectra

$$f_z(\omega_1, \omega_2, \ldots, \omega_{r-1}) = f_x(\omega_1, \omega_2, \ldots, \omega_{r-1}) + f_N(\omega_1, \omega_2, \ldots, \omega_{r-1}) \quad (3.2)$$

Our main interest is when $r = 2$ and $3$. When $r = 2$, we have the relation between the second order spectra

$$f_z(\omega) = f_x(\omega) + f_N(\omega)$$

and when $r = 3$, we have the relation between bispectra

$$f_z(\omega_1, \omega_2) = f_x(\omega_1, \omega_2) + f_N(\omega_1, \omega_2). \quad (3.3)$$

In general if $N(t)$ is Gaussian (and $x(t)$ is non-Gaussian) we have the important relationship between the high order spectra

$$f_z(\omega_1, \omega_2, \ldots, \omega_{r-1}) = f_x(\omega_1, \omega_2, \ldots, \omega_{r-1}), \quad r > 2.$$  

This extremely useful relation suggests that if one wishes to study the properties of the signal (which is non-Gaussian and may satisfy a nonlinear relation of the form $x_{t+1} = f(x_t)$ as is common in chaotic models) one could obtain all the information about the signal by looking at the higher order cumulants of the observations ($Z(t)$). For example, if one wishes to obtain the dimensions of the state $x_{t+1} = f(x_t)$ in the presence of noise, canonical correlations of the higher powers of $Z_t$, say $Z_t^2$, may be useful. It is similar to the approach followed by Broomhead and King (1986) where they performed covariance analysis for extracting dimensions of the state. Recently Subba Rao and da Silva (1990) have proposed higher order canonical correlations approach for determining the order of bilinear systems.

We now consider two specific examples.

Let $x(t)$ satisfy an autoregressive model of order $p$ of the form

$$x(t) + a_1 x(t-1) + \ldots + a_p x(t-p) = e(t),$$

where $\{e(t)\}$ is a sequence of independent, identically distributed random variables. The object is to estimate $(a_1, a_2, \ldots, a_p)$ when a sample of
\{Z(t)\} is available. It is possible to estimate these parameters of the signal using third order moments of \{Z(t)\} (for details see, Parzen, 1966).

The other example we consider is the estimation of the frequencies of the series \{x(t)\}, when

\[ x(t) = \sum_{i=1}^{K} A_i \cos(\omega_i t + \phi_i), \]

where the independent random phases \{\phi_i\} are distributed uniformly over the interval \((-\pi, \pi)\) and the coloured noise \(N(t)\) is Gaussian, and stationary.

For a given sample \((Z(1), Z(2), \ldots, Z(m))\), the phases can be considered to be deterministic, and it has been shown by Subba Rao and Gabr (1988), that the modulus of bisppectrum of \(Z(t)\) will show clear peaks corresponding to the frequencies of the signal.

**Bispectrum and Chaos.** Trajectories generated by some nonlinear deterministic difference equations (or differential equations) look random (stochastic), and this phenomenon is referred to as "chaos." It is interesting to note that that only nonlinear equations (not all nonlinear equations) produce this phenomenon. So it is interesting to discriminate between deterministic chaos and random process. It is well known that second order covariances (or second order spectra) do not provide adequate information of these nonlinear models (and this feature is similar to bilinear models) and as such they cannot be used for discrimination purposes (see Eubank and Farmer, 1989, p 77 and Brock, 1986). Since second order spectra are not adequate, it is necessary to calculate the higher order spectra, and the simplest higher order spectrum is bispectrum. In this paper, our object is to report the preliminary empirical results we obtained when we calculated bispectra from the samples generated from three well known chaotic models. These results have yet to be confirmed theoretically.

**Tent Map.** Let the time series \(\{x(t)\}\) satisfy the model

\[ x(t+1) = \begin{cases} 
\frac{x(t)}{a} & 0 \leq x(t) \leq a \\
\frac{(1-x(t))}{1-a} & a \leq x(t) \leq 1 
\end{cases} \]

where \(0 \leq a \leq 1\). If \(x(0)\) is uniformly distributed over the interval, \((0,1)\), then \(x(t)\) for all \(t\), is uniformly distributed over the interval \((0,1)\). Sakai and Tokomaru (1980) have shown that the second order covariances of the tent model is similar to the linear model

\[ x(t+1) = (2a-1) x(t) + e(t+1), \]

where \(\{e(t)\}\) are independent, identically distributed. We have generated 250 observations with \(a = 0.25\) and \(x(0) = 0.6\), the plot is given in Figure 3 and the spectrum and bispectrum is calculated from the sample. The spectrum looks like the spectrum of a linear model confirming Sakai and Tokomaru (1980), and the normalised bispectral density ratio (2.4), is plotted in Figure 4. It is clear that the process cannot be linear, and also we see a ridge along the line \(\omega_1 + \omega_2 = 1.2\pi\) which seems to be
common in chaotic models.

**Logistic Map.** Let the series be generated from the model \( x(t+1) = ax(t)(1-x(t)) \), \( t = 1, 2, \ldots, 500 \). It is well known that we observe interesting trajectories when \( 3 < a \leq 4 \). More unstable trajectories (or chaos) seems to occur for \( a \approx 3.5 \). In order to see this we have calculated the bispectrum for various values of \( a \) using the product windows with the truncation point \( M = 32 \) as described earlier. The modulus of the bispectrum \( |f(\omega, \omega)| \) for several values of \( a = 3(0.1)3.9 \) are plotted in Figure 5. For values of \( a \), \( 3 \leq a \leq 3.4 \), the values of the modulus of the bispectrum are very small (the values are smaller than \( 10^{-7} \) (see Figures 7 and 9). Though we find a ridge along the line \( \omega_1 + \omega_2 = \pi \) in the bispectra, they are not significant because their values are very small. When \( a = 3.5 \), we find three dominant peaks, one at the frequency corresponding to the period 4 units (this is when period doubling starts, Stewart 1989, p 159). The bispectral values obtained when \( a > 3.5 \) are several times larger than the corresponding values when \( a < 3.5 \). Also interesting is to compare these values with the values we obtained for \( a \leq 3.4 \). These are several times bigger. As the value of \( a \) increases, we find several ridges in the modulus of the bispectrum, a phenomenon which we have observed consistently in chaotic models.

**Henon Map.** Now consider the map

\[
x(t) = 0.3 \, x(t-2) - 1.4 \, x^2(t-1) + 1. \quad (t = 1, 2, \ldots, 512).
\]

The plot of the data and the modulus of the bispectrum are given in Figures 20 and 21 respectively. The shape of the bispectrum is similar to the shapes we observed for the logistic map for values of \( a = 3.5 \) to 3.9.

From these calculations, we feel that the higher order spectra can distinguish between nonlinear deterministic stable models and nonlinear deterministic models which produce "chaos." On the basis of this empirical evidence, we can conclude that Canadian Lynx data and Sunspot numbers are not chaotic, (ie) they are nonlinear and nondeterministic. Another important feature is that for the chaotic models, the "energy" in the spectrum and the bispectrum is distributed in the higher frequency range, and for the stable nonlinear models the energy is in the low frequency range (for example Sunspot data and Canadian Lynx data).

In the following we briefly consider one nonlinear model introduced recently in time series, and study its usefulness for forecasting purposes.

**4. BILINEAR TIME SERIES MODEL AND NONLINEAR PREDICTION**

Once we decide that the series is nonlinear, it is important to see whether we can find a finite parameter nonlinear model to describe the series. One such model (Granger and Andersen, 1978, Subba Rao 1981, Subba Rao and Gabr, 1984) is bilinear model, whose analytic properties have been extensively studied. Let \( x(t) \) satisfy the difference equation
\[ x(t) + \sum_{j=1}^{p} a_j x(t-j) = e(t) + \sum_{j=1}^{q} b_j e(t-j) + \sum_{i=1}^{P} \sum_{j=1}^{Q} a_{ij} x(t-1)e(t-j) \] (4.1)

The above model is called Bilinear model and is denoted by BL(p,q,P,Q). It is linear in \( x(t), e(t) \) but not jointly. For \( p = P, q = 0, Q = 1 \), Subba Rao (1981) has shown that one can write an equivalent state space form and then it is easy to evaluate the moments of the process \( x(t) \). The solution of the equation can be written in the form of Volterra series. In Figures 22 and 23, we show plots of the observations generated from the bilinear models. It may be observed that when the coefficients of the nonlinear part of the model tend towards the nonstationary region, the trajectories generated by these bilinear models may produce behaviour similar to "chaotic" models. The theoretical properties of the bilinear models have been extensively investigated (see Brockwell and Liu, 1988, Subba Rao, 1981, Bhaskara Rao et al. (1983), Tuan and Tran (1981), Terdik and Subba Rao (1989)). The parameters of the model are estimated by minimising the least squares criterion \( \sum e^2(t) \), and this minimisation is done using iterative techniques (see Subba Rao and Gabr, 1984). We now consider a real example.

**Sunspot Numbers.** We consider the first 221 observations (from the year 1700 to 1920) for fitting the bilinear model, and then calculate one step ahead predictions for the subsequent 35 years. The following model is fitted to the data (see Subba Rao and Gabr 1984, p 198 for details)

\[
x(t) = 1.5x(t-1) + 0.7 x(t-2) - 0.12 x(t-9)
= 6.8 - 0.146 x(t-2)e(t-1) + 0.01 x(t-8)e(t-1)
- 0.01 x(t-1)e(t-3) - 0.01 x(t-4)e(t-3)
+ 0.004 x(t-2)e(t-1) + 0.002 x(t-3)e(t-2) + e(t). \quad (4.2)
\]

The adequacy of the fit is checked by testing for the independence (and normality too) of the residuals. The \( m \)th step ahead forecast is calculated from the conditional expectation

\[
\hat{X}(t_0, m) = E(x(t_0 + m)/x(s), s \leq t_0).
\]

These forecasts together with the true values are given (for \( m = 1 \)) in Table 2. This model seems to have the smallest mean sum of squares, and also error variances calculated up to 5 steps compared to other nonlinear models (see Table 6.1 of Subba Rao and Gabr (1984), p 201).

5. **LONG RANGE DEPENDENT BILINEAR MODEL**

A time series \( \{x(t)\} \) is said to be short range memory model if its autocovariances \( R(s) \) satisfy the condition \( \sum |R(s)| < \infty \). All the standard
stationary linear models satisfy this condition. A process \( \{x(t)\} \) for which \( \sum |R(s)| = \infty \), is said to be long range dependent. This definition of the long range dependence in terms of the second order Covariances is because of the fact that the process so far considered in the literature is assumed to be Gaussian. If the process is not Gaussian this condition has to be replaced by conditions involving higher order cumulants which takes care of nonlinear dependence. Recently there is an enormous interest in this type of model because of its application in hydrology, meteorology, economics and many related fields. Also these models are closely related to Fractals. One way of generating the long range dependent Gaussian time series is from the model (see Hosking (1981) and Granger and Joyeux, (1980)),

\[
x(t) = (1-B)^{-\delta} e(t), \quad \delta \in (0,1/2)
\]

and \( \{e(t)\} \) is a Gaussian white noise. The spectral density function of this process is of the form

\[
f(\lambda) = \frac{\sigma^2}{2 \pi} \left( 4 \sin^2 \frac{\lambda}{2} \right)^{-\delta} \text{and as } \lambda \to 0, f(\lambda) \to c \lambda^{-2\delta}.
\]

The other types of linear long range memory models are generated from the models of the form

\[
(1 + a_1 B + \ldots + a_p B^p) (1-B)^{-\delta} x(t) = (1 + b_1 B + \ldots + b_q B^q) e(t),
\]

where \( \delta \in (-1/2,1/2) \) and the usual assumptions on the polynomials \( Z^p + a_1 Z^{p-1} + \ldots + a_p \) and \( Z^q + b_1 Z^{q-1} + \ldots + b_q \) are imposed. The condition (that \( \delta \in (-1/2,1/2) \)) ensures both stationarity and invertibility of the model. Recently attempts have been made to extend the definition of long range dependence to non-Gaussian situations but mainly restricted to linear models with non-Gaussian variables \( \{e(t)\} \). Here our object is to extend this definition to nonlinear models. Since we know that bilinear models have covariances similar to the linear models, and they decay to zero exponentially, it is natural to generate long range nonlinear models through these bilinear models; and the covariances and the third order moments of these processes will decay at hyperbolic rate in contrast to bilinear models. We can define two types of nonlinear models (for further details see Subba Rao, Terdik and Bhaskara Rao, (1991). For example, they can be generated as follows:

\[
x(t) = a x(t-1) + b x(t-1) w(t-1) + w(t)
\]

Let \( w(t) = (1-B)^{-\delta} e(t) \), where \( e(t) \) is a Gaussian White noise. We can write

\[
\Gamma(j + \delta)
\]

for \( \delta \in (-1/2,1/2) \), \( w(t) = \sum_{j=0}^{\infty} \psi_j e(t-j) \), where \( \psi_j = \frac{\Gamma(j + \delta)}{\Gamma(\delta) \Gamma(j + 1)} \), as \( j \to \infty \),

\[
\psi_j \to \frac{1}{\Gamma(\delta)}.
\]

Let \( x(t) \) satisfy

\[
x(t) = a x(t-1) + b x(t-1) w(t-1) + w(t)
\]

12
We can define the above as long range bilinear model LR BL \(1,0,1,1\). The solution of the above bilinear model can be written in the Volterra series form (see Subba Rao and Gabr, 1984 p. 154). The first term of which will be linear in \(w(t)\), second will be quadratic in \(w(t)\) and so on. Equivalently we can write a Wiener–Ito representation. For example, when we take first two terms of Volterra expansion only we have

\[
x(t) = x_1(t) + x_2(t),
\]

where \(x_1(t)\) and \(x_2(t)\) satisfy the difference equations

\[
x_1(t) = a x_1(t-1) + w(t), \quad |a| < 1,
\]

\[
= \sum_j a^j w(t-j),
\]

\[
x_2(t) = a x_2(t-1) + b x_1(t-1) w(t-1)
\]

\[
= \sum_j a^j b x_1(t-j-1) w(t-j-1)
\]

We can show that the covariances and higher order moments decay to zero at a much slower rate (hyperbolic rate) than the usual bilinear model. The other nonlinear model is of the form

\[
x(t) = ax(t-1) + bx(t-1) e(t-1) + e(t),
\]

and

\[
y(t) = (1-B)^{\delta} x(t).
\]

For this model, the covariances decay at the same rate as the usual bilinear model and the spectrum and the bispectrum will have singularities in the neighbourhood of the origin.

The properties and the usefulness of the above models will be discussed in a later paper.

**ACKNOWLEDGEMENTS**

I wish to thank Professors M Casdagli, S Eubank and D Farmer for inviting me to present the paper at the conference which helped me to learn a lot on "chaos." The last section of this paper was as a result of the inspiration provided by the first two weeks of the meeting on Program on Time Series organised by the Institute of Mathematics and Application, Minneapolis, Minnesota, July 2 – 27, 1990, and the author is very grateful to the organisers for the invitation and to NSF for financial assistance to attend this meeting and the excellent facilities offered. Lastly I must thank Dr Jingsong Yuan for the computations reported in this paper.

13
REFERENCES


<table>
<thead>
<tr>
<th>I</th>
<th>X(I)</th>
<th>SRL</th>
<th>FULL AR</th>
<th>SUBSET AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>222</td>
<td>26.1</td>
<td>24.55</td>
<td>24.60</td>
<td>19.07</td>
</tr>
<tr>
<td>225</td>
<td>14.2</td>
<td>11.43</td>
<td>13.13</td>
<td>18.69</td>
</tr>
<tr>
<td>224</td>
<td>5.8</td>
<td>10.08</td>
<td>13.74</td>
<td>11.39</td>
</tr>
<tr>
<td>226</td>
<td>44.3</td>
<td>37.52</td>
<td>35.25</td>
<td>32.82</td>
</tr>
<tr>
<td>227</td>
<td>63.9</td>
<td>75.67</td>
<td>67.83</td>
<td>68.32</td>
</tr>
<tr>
<td>228</td>
<td>89.0</td>
<td>77.86</td>
<td>74.57</td>
<td>74.11</td>
</tr>
<tr>
<td>229</td>
<td>77.8</td>
<td>73.93</td>
<td>67.11</td>
<td>67.14</td>
</tr>
<tr>
<td>230</td>
<td>64.9</td>
<td>75.04</td>
<td>72.20</td>
<td>71.42</td>
</tr>
<tr>
<td>231</td>
<td>35.7</td>
<td>57.03</td>
<td>49.87</td>
<td>48.73</td>
</tr>
<tr>
<td>232</td>
<td>21.2</td>
<td>29.34</td>
<td>16.12</td>
<td>17.56</td>
</tr>
<tr>
<td>233</td>
<td>11.1</td>
<td>9.37</td>
<td>14.05</td>
<td>14.27</td>
</tr>
<tr>
<td>234</td>
<td>5.7</td>
<td>6.68</td>
<td>11.49</td>
<td>11.28</td>
</tr>
<tr>
<td>235</td>
<td>8.7</td>
<td>10.94</td>
<td>13.28</td>
<td>14.23</td>
</tr>
<tr>
<td>236</td>
<td>36.1</td>
<td>21.84</td>
<td>24.12</td>
<td>23.90</td>
</tr>
<tr>
<td>237</td>
<td>79.7</td>
<td>63.83</td>
<td>57.50</td>
<td>57.25</td>
</tr>
<tr>
<td>238</td>
<td>114.4</td>
<td>105.82</td>
<td>98.80</td>
<td>97.95</td>
</tr>
<tr>
<td>239</td>
<td>109.6</td>
<td>100.99</td>
<td>114.89</td>
<td>115.36</td>
</tr>
<tr>
<td>240</td>
<td>88.8</td>
<td>85.10</td>
<td>86.85</td>
<td>85.86</td>
</tr>
<tr>
<td>241</td>
<td>67.8</td>
<td>58.95</td>
<td>60.38</td>
<td>60.34</td>
</tr>
<tr>
<td>242</td>
<td>47.5</td>
<td>38.76</td>
<td>44.48</td>
<td>44.04</td>
</tr>
<tr>
<td>243</td>
<td>30.6</td>
<td>24.79</td>
<td>27.11</td>
<td>29.44</td>
</tr>
<tr>
<td>244</td>
<td>16.3</td>
<td>20.28</td>
<td>17.92</td>
<td>19.95</td>
</tr>
<tr>
<td>245</td>
<td>9.6</td>
<td>16.20</td>
<td>14.69</td>
<td>15.31</td>
</tr>
<tr>
<td>246</td>
<td>33.2</td>
<td>16.58</td>
<td>21.63</td>
<td>21.55</td>
</tr>
<tr>
<td>247</td>
<td>92.6</td>
<td>73.04</td>
<td>59.19</td>
<td>59.94</td>
</tr>
<tr>
<td>248</td>
<td>151.6</td>
<td>140.92</td>
<td>119.53</td>
<td>120.44</td>
</tr>
<tr>
<td>249</td>
<td>136.3</td>
<td>148.47</td>
<td>157.61</td>
<td>158.52</td>
</tr>
<tr>
<td>250</td>
<td>134.7</td>
<td>121.87</td>
<td>104.42</td>
<td>103.54</td>
</tr>
<tr>
<td>251</td>
<td>83.9</td>
<td>87.77</td>
<td>104.82</td>
<td>106.93</td>
</tr>
<tr>
<td>252</td>
<td>69.4</td>
<td>50.79</td>
<td>45.94</td>
<td>41.80</td>
</tr>
<tr>
<td>253</td>
<td>31.5</td>
<td>30.73</td>
<td>40.88</td>
<td>49.52</td>
</tr>
<tr>
<td>254</td>
<td>13.9</td>
<td>7.70</td>
<td>11.57</td>
<td>9.15</td>
</tr>
<tr>
<td>255</td>
<td>4.4</td>
<td>14.83</td>
<td>4.20</td>
<td>11.58</td>
</tr>
<tr>
<td>256</td>
<td>38.0</td>
<td>7.65</td>
<td>22.63</td>
<td>18.31</td>
</tr>
</tbody>
</table>

$\hat{\sigma}^2(1)$ 123.77 190.89 214.07
Figure 7. Bisection of sample (logistic) with a = 3.0

Figure 6. Sample from logistic map with a = 3.0

Figure 5. Plot of f(x, y) against a of the Logistic Map
Figure 20.
Sample from Henon Map

Figure 21.
Bispectrum of Sample from Henon Map
**Figure 22.**

Bilinear Series from \( x(t) = 0.4 \ x(t-1) \ e(t-1) \)

+ \( e(t) \)

**Figure 23.**

Bilinear Series from \( x(t) = 0.8 \ x(t-1) - 0.4 \ x(t-2) \)

+ \( 0.6 \ x(t-1) \ e(t-1) + 0.7 \ x(t-2) \ e(t-1) \)

+ \( e(t) \)