APPLICATIONS OF BIVARIATE AND UNIVARIATE LOCAL LYAPUNOV EXPONENTS

YODIT SEIFU and N. REID *

North Carolina State University and University of Toronto

Abstract

Chaotic systems are characterized by sensitivity to initial conditions, and this property can be measured by global Lyapunov exponents, which are measures of the average divergence rate of initially close trajectories. Wolff (1992) introduced local Lyapunov exponents and used them to obtain two diagnostic plots for differentiating between stochastic and deterministic time series. We extend the definition of the local Lyapunov exponent and the diagnostic plots to accommodate time series that arise from bivariate maps and investigate the behavior of the local Lyapunov exponents and the corresponding diagnostic plots for some dynamical systems and stochastic time series. We consider the application of this diagnostic plots to some heart rate variability data.

KEY WORDS: Chaos, diagnostic plots, heart rate variability, local Lyapunov exponents.

AMS 1985 subject classification: 62M10, 62-09.

*This work was in part supported by the Natural Sciences and Engineering Council of Canada. We are very much obliged to R. Hughson and Y. Yamamoto for permission to use their data. We are also grateful to Colleen Cutler and John Kolassa for helpful discussions and to the reviewers for very useful comments on an earlier draft.

1 Introduction

Chaos is of particular interest to statisticians because a low-dimensional nonlinear deterministic system that is chaotic is associated with complex "random" behavior and forms of unpredictability. If apparent "random" observations arise as iterates of a low dimensional nonlinear system, then one could potentially try to determine or characterize the nonlinear system rather than assessing these observations using probabilistic models. It can thus become important to be able to discriminate between time series that arise as iterates of chaotic maps and those which are stochastic or arise from high-dimensional deterministic systems.

Chaos arises from the iteration of a nonlinear deterministic map or the evolution of nonlinear differential equations. In this paper we restrict our attention to chaotic time series that arise as iterates of discrete time maps. A necessary condition for a nonlinear deterministic system to be chaotic is the presence of sensitivity to initial conditions. Information about the system's sensitivity to initial conditions can be obtained from the Lyapunov exponents, which are a measure of the average rate of divergence for two initially close sequences of iterates: these sequences are also called trajectories. A dynamical system has as many Lyapunov exponents as the dimensions of the state space in which it is embedded, and the positivity of the largest one of these is an indication of the system's sensitivity to initial conditions. Whenever the equations governing the system are known these exponents can be computed: see Eckmann and Ruelle (1985) for more details. In the absence of knowledge of these equations, Nychka *et al.* (1992) implemented non-parametric regression techniques for estimating Lyapunov exponents.

Instead of looking at global Lyapunov exponents, for univariate maps, Wolff (1992) examined local tendencies of divergence. By looking at a number of trajectories that originate from a small part of the state space, called bins, he characterized the change in the trajectories' tendency to diverge from one part of the state space to another. As described by Wolff, this potentially provides both qualitative and quantitative information.

Let $x_1, x_2, \ldots$ be a time series which may have arisen as iterates of a univariate nonlinear map f and let

$$\lambda_{i,m}^x = \lambda^x(x_i, m, h) = \frac{1}{m}\frac{1}{n_i}\sum_{j \in S_i^x} \log\left|\frac{x_{i+m} - x_{j+m}}{x_i - x_j}\right|, \quad (i = 1, \ldots, n-m) \tag{1}$$

where $S_i^x = \{j : 0 < |x_i - x_j| \leq h\}$, n is the length of the time series, $n_i = \sharp(S_i^x)$, m is a positive integer representing lag, and $h > 0$. Wolff (1992) called $\lambda_{i,m}^x$ in equation (1) the local Lyapunov exponent (LLE) at x_i for lag m. In characterizing the properties of the LLE Wolff (1992) used a modified form of the $\lambda_{i,m}^x$ defined as in (1) with the term $|x_i - x_j|$ in the summand replaced by h. This modification makes it easer to determine the theoretical properties of the LLE. For $m = 1$, Wolff (1992) stated that the approximate expected difference between the modified $\lambda_{i,1}^x$ and the $\lambda_{i,1}^x$ in (1) is given by $\log 2 + 1/2 + O(h^2)$. This result is incorrect; the correct approximation is $1 + O(h^2 \log h)$ (Seifu, 1995).

Wolff (1992) suggested using LLEs in two types of diagnostic plots: a plot of the LLE against the state space (the LLE plot), and a scatter plot of pairs of averaged LLEs obtained from the first and second half of the time series (the LLE-LLE plot). Wolff (1992) examined the theoretical and empirical properties of LLE and LLE-LLE plots and demonstrated that to a certain extent these plots can be used to discriminate between stochastic time series and time series that arise as iterates of a univariate nonlinear map.

The LLE plot is constructed by plotting $\lambda_{i,m}^x$ against x_i. For small h and lag one, if the time series arises as a result of iterating a nonlinear deterministic map, we expect the LLE plots to show systematic change over the state space, because the logarithm of the absolute value of the derivative which the LLE estimates is changing with x_i. However, stochastic time series such as white noise will not exhibit such variation over the state space. Thus, if we indeed observe systematic change we have some evidence of the presence of nonlinear deterministic map which may be chaotic.

An LLE-LLE plot is obtained by first splitting up the time series into two parts and partitioning the state space into small intervals called bins. Then, using the first half of the series, the LLE at lag one for each data point is evaluated and the average of such exponents across data points belonging to a common bin is obtained. The same is done with the second half of the time series using the same choice of bins thus, obtaining two averaged LLE values per bin. Finally, a scatter plot of the pairs of averaged LLE values is constructed. The principle behind this graphical method is the same as described for the LLE plot above. If the system is dominated by a nonlinear dynamic system which may be chaotic, we expect the dynamic mechanism to prevail throughout the dynamical process and only to vary with respect to the change in the state space and not with respect to time. Hence, for a time series that arises as iterates of a nonlinear deterministic map, the LLEs arising from matching bins should have approximately the same value and we would expect the LLE-LLE plots to be approximately linear and to lie close to the line with zero intercept and slope one.

In this paper we study further the use of LLEs as diagnostic tools in the exploratory analysis of time series. We first extend the definition of LLE to accommodate time series that arise from bivariate maps, and investigate the behavior of the bivariate LLE-LLE plots for some nonlinear maps and stochastic time series. We then use these diagnostic techniques on some data on heart rate variability: the possible presence of nonlinear dynamics in heart rate variability is of current interest as it has implications for the understanding of feedback mechanisms in controlling heart rate. In Section 2 we present the definition of the bivariate LLE and study the behavior of the LLE and LLE-LLE plots for some stochastic time series and some time series that arise as iterates of a nonlinear deterministic map. In Section 3 we use LLE-LLE plots to study the heart rate variability data. Finally, in Section 4 we draw some conclusions.

2 Bivariate Local Lyapunov Exponents

2.1 Definition

Consider data of the form $\mathbf{x}_1 = (x_1, x_0)^T$, $\mathbf{x}_2 = (x_2, x_1)^T$,..., $\mathbf{x}_n = (x_n, x_{n-1})^T$ which may arise as $x_{t+1} = f(x_t, x_{t-1})$ where f is a nonlinear differentiable map which may be chaotic. Such maps could arise as an application of Takens's embedding theorem (Takens, 1981, Nychka *et al.*, 1992). Let

$$\widetilde{\lambda}^{\mathbf{x}}_{i,m} = \widetilde{\lambda}^{\mathbf{x}}(\mathbf{x}_i, m, h) = \frac{1}{m}\frac{1}{n_i}\sum_{j\in S_i^{\mathbf{x}}} \log\frac{||\mathbf{x}_{i+m} - \mathbf{x}_{j+m}||}{||\mathbf{x}_i - \mathbf{x}_j||}, \quad (i = 1, ..., n-m) \qquad (2)$$

where $||.||$ represents the Euclidean norm, $S_i^{\mathbf{x}} = \{j : ||\mathbf{x}_i - \mathbf{x}_j|| \leq h\}$, $n_i = \sharp(S_i^{\mathbf{x}}), m$ is a positive integer, and $h > 0$. We will call $\widetilde{\lambda}^{\mathbf{x}}_{i,m}$ in equation (2) the local Lyapunov exponent (LLE) at $\mathbf{x}_i$ for lag m. Thus, in computing $\widetilde{\lambda}^{\mathbf{x}}_{i,m}$, trajectories that are at most distance h away from $\mathbf{x}_i$ are chosen and the average of the logarithms of all the rates of divergence after m time steps apart is computed. This definition of the bivariate LLE is motivated by the algorithm of Wolf *et al.* (1985) for computing global Lyapunov exponents. Since our main interest is in analyzing the diagnostic property of the bivariate LLE-LLE plots, from here on we restrict our discussion to the bivariate LLE at lag one.

Let $\mathbf{x}_j \in S_i^{\mathbf{x}}$ and $\lambda_{ij}^{\mathbf{x}} = \log\left(\frac{||\mathbf{x}_{i+1} - \mathbf{x}_{j+1}||}{h}\right)$. Then, $\lambda_{ij}^{\mathbf{x}} \leq \log\left(\frac{||\mathbf{x}_{i+1} - \mathbf{x}_{j+1}||}{||\mathbf{x}_i - \mathbf{x}_j||}\right)$,

and an approximation for $\widetilde{\lambda}_{i,1}^{\mathbf{x}}$ is given by

$$\lambda_i^{\mathbf{x}} = \frac{1}{n_i} \sum_{j \in S_i^{\mathbf{x}}} \lambda_{ij}^{\mathbf{x}}. \tag{3}$$

For $m = 1$, the approximate expected difference between (2) and (3) is $0.5 + O(h^2 \log h)$: see the appendix for the derivation of this result. Simulation results for various stochastic time series and the Henon map show that, except for a vertical shift, the LLE plots obtained using expressions (2) and (3) have the same pattern. For $h = 0.1$, this vertical shift generally varies between 0.3 and 0.8. Since it is easier to statistically characterize the behavior of the LLE given by expression (3), henceforth all references to bivariate LLEs assume this form.

In this section we will first consider the behavior of the bivariate LLE for time series that arises from some familiar stochastic models and nonlinear deterministic maps. Then, we discuss how these properties lead to either the detection or undetection of linearity pattern in the corresponding bivariate LLE-LLE plots.

2.2 Behavior of bivariate local Lyapunov exponents

2.2.1 White noise

Let $x_1, x_2, ..., x_n$ be a random sample from a distribution with $EX^r = \mu_r'$ and $Var(X) = \sigma^2$. Furthermore, let $x^* = (x_1 - x_2)^2$. Then $\mu^* = EX^* = 2\sigma^2$ and $Var(X^*) = 2\mu_4' - 8\mu_3'\mu_1' + 6(\mu_2')^2 - 4\sigma^4$. Wolff (1992) shows that the quantity $\lambda_{ij}^x = \log|(x_{i+1} - x_{j+1})/h|$ needed in the definition of the univariate LLE is equal in distribution to $\frac{1}{2}\log x^* - \log h$. Hence, he obtains the approximation

$$E(\lambda_{ij}^{x}) + \log h \approx \frac{1}{2}\log(\mu^*) - \frac{Var(X^*)}{4(\mu^*)^2}. \tag{4}$$

In the bivariate setting, the distribution of $\lambda_{ij}^{\mathbf{x}}$ is not available in general models, as $\mathbf{x}_{i+1} - \mathbf{x}_{j+1}$ and $\mathbf{x}_i - \mathbf{x}_j$ are dependent. (For a discussion for the normal case see Seifu (1995).) However, by approximating $(x_i - x_j)^2$ by $h^2/3$ we can similarly obtain the approximate expected value of $\lambda_{ij}^{\mathbf{x}}$ as

$$E(\lambda_{ij}^{\mathbf{x}} \mid ||\mathbf{x}_i - \mathbf{x}_j|| < h) + \log h \approx \frac{1}{2}\log\left(h^2/3 + \mu^*\right) - \frac{Var(X^*)}{4(h^2/3 + \mu^*)^2}. \tag{5}$$

Table 1: *Approximation of the LLE for IID time series (h = 0.1). The simulation used 100 replications of time series of length 1000*

Distribution	Theoretical Univariate LLE	Theoretical Bivariate LLE	Simulated Univariate LLE	Simulated Bivariate LLE
$N(0,1)$	2.149	2.152	1.909	1.946
$U(-1,1)$	1.750	1.756	1.447	1.508
$Exp(1)$	1.399	1.404	1.631	1.698
$Gamma(5,1)$	2.804	2.804	2.659	2.675
$Beta(0.5,0.5)$	1.297	1.312	0.909	1.035
$Beta(1,0.5)$	1.046	1.070	0.654	0.824
$Beta(0.5,1)$	1.046	1.070	0.650	0.823

The approximate mean values presented in (5) and (4) suggest that for small h, the expected value of the univariate and bivariate LLEs are approximately equal. Table 1 presents the approximate expected values for the univariate and bivariate LLEs for h = 0.1. These approximate values are obtained via simulation and using expressions (5) and (4). We note that for both approximation methods the mean of the bivariate and the univariate LLEs are close in value. There is some discrepancy between the simulated and the approximated theoretical mean LLE values. For the univariate case this discrepancy was also observed by Wolff (1992). In the bivariate case there are added inaccuracies generated by the approximation $(x_i - x_j)^2 = h^2/3$ used to obtain (5).

In Fig. 1, a plot of $\mathbf{x}_i$ versus $\lambda^{\mathbf{x}}_{i,1}$ is presented for time series of length 2000 where the x_i's are a random sample from the standard normal distribution. As in the univariate case we will call this plot the LLE plot. As expected the LLE plot does not show systematic pattern and the LLE values are scattered around 2.

2.2.2 AR(k) model

Assume first that the time series $\{x_t\}$ comes from a first order autoregression: $x_{t+1} = \rho x_t + \varepsilon_{t+1}$, $(t = 1, 2, ..., n-1)$ where the ε_i's are IID with distribution $N(0, \sigma^2)$, $|\rho| < 1$ and x_0 is given. Let $|i - j| > 1$. Here and in subsequent models we assume that $\mathbf{x}_j = \mathbf{x}_i + \mathbf{h}$ where $\mathbf{h} = (h_2, h_1)^T$ and $||\mathbf{h}|| = h$. Then,

$$\begin{aligned} \lambda^{\mathbf{x}}_{ij} &= \frac{1}{2}\log\left\{\left(\frac{\rho h_2}{\sqrt{2}\sigma} + \frac{\varepsilon_{j+1} - \varepsilon_{i+1}}{\sqrt{2}\sigma}\right)^2 + \frac{h_2^2}{2\sigma^2}\right\} \\ &\quad + \frac{1}{2}\log\left(2\sigma^2\right) - \log h. \end{aligned} \tag{6}$$

Thus, under this assumption the distribution of $\lambda^{\mathbf{x}}_{ij}$ can be computed from noncentral chi-square distribution with noncentrality parameter $\{(\rho h_2)(\sqrt{2}\sigma)^{-1}\}^2$.

For a time series that comes from a stationary second order autoregression model: $x_{t+1} = b_1 x_t + b_2 x_{t-1} + \varepsilon_{t+1}$, $(t = 1, 2, ...n-1)$ where ε_i's are IID with distribution $N(0, \sigma^2)$. If

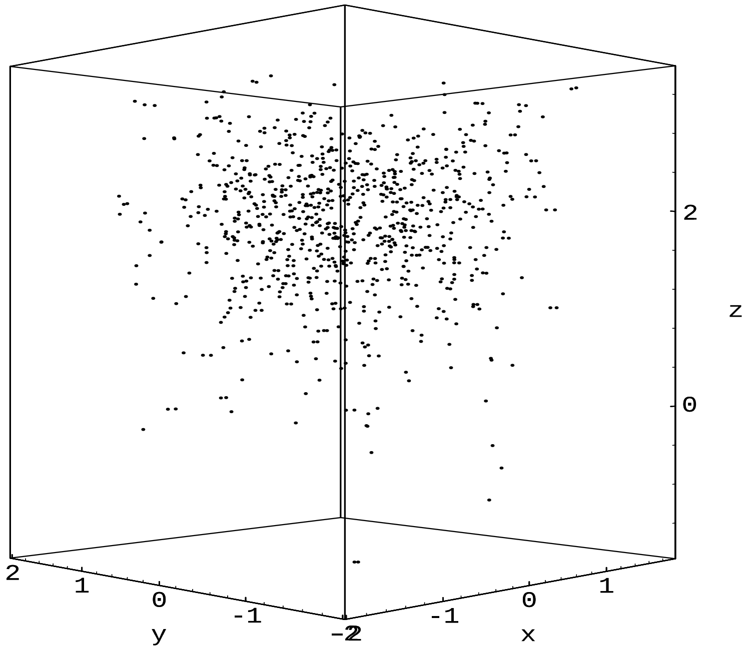

Figure 1: *LLE plot for white noise* ($N(0,1)$)*:* $(x, y, z) = (x_{i-1}, x_i, \lambda^{\mathbf{x}}_{i,1})$*, sample size=2000,* $h = 0.1$ *and 5% trimming is applied.*

$\mathbf{x}_1 = (x_1, x_0)$ is given, then

$$\lambda_{ij}^{\mathbf{x}} = \frac{1}{2}\log\left\{\left(\frac{b_1h_2 + b_2h_1}{\sqrt{2}\sigma} + \frac{\varepsilon_{j+1} - \varepsilon_{i+1}}{\sqrt{2}\sigma}\right)^2 + \frac{h_2^2}{2\sigma^2}\right\} + \frac{1}{2}\log\left(2\sigma^2\right) - \log h. \quad (7)$$

Hence, the $\lambda_{ij}^{\mathbf{x}}$ in (6) and (7) have the same approximate distributional properties where the noncentrality parameter for (7) is $\{(b_1h_2 + b_2h_1)(\sqrt{2}\sigma)^{-1}\}^2$.

From these approximate distributional properties we conclude that the expected value of the LLEs for time series that arises from first and second order autoregressive stationary models will not be state dependent.

For higher order autoregressive models, the dependence of x_t on more than two previous lag values makes it difficult to similarly obtain simplified expression for $\lambda_{ij}^{\mathbf{x}}$.

2.2.3 Nonlinear maps

Underlining our definition of the bivariate LLE is the existence of nonlinear bivariate maps which may be chaotic. In this subsection we examine the properties of the LLE for the Henon map which is a bivariate chaotic map.

The Henon map is given by $x_{t+1} = 1 - 1.4x_t^2 + 0.3x_{t-1}$ and,

$$\lambda_{ij}^{\mathbf{x}} = \frac{1}{2}\log\{(2.8x_ih_2 + 1.4h_2^2 - 0.3h_1)^2 + h_2^2\} - \log h. \quad (8)$$

In this case, the $\lambda_{ij}^{\mathbf{x}}$ and thus the LLE systematically varies with the state. In Fig. 2, the LLE plot for the Henon map is presented. This plot shows nonlinear structure which varies with the state.

The LLE and LLE-LLE plots are diagnostic plots that can be used to identify time series that arises as a result of iterating a low dimensional nonlinear deterministic map. Applications of this method implies that we do not know the dimension of the underlying system. Hence, we should also consider the $\lambda_{ij}^{\mathbf{x}}$ for nonlinear deterministic maps having dimension different from two.

The logistic map is a chaotic map and is given by $x_{t+1} = 4x_t(1 - x_t)$. Then

$$\lambda_{ij}^{\mathbf{x}} \quad = \quad \frac{1}{2}\log\{(8h_2 x_i)^2 + h_2^2\} - \log h. \tag{9}$$

Again note how $\lambda_{ij}^{\mathbf{x}}$ varies systematically with the state.

For higher dimensional maps the dependence of x_t on more than two previous lag values makes it difficult to obtain a simplified expression for $\lambda_{ij}^{\mathbf{x}}$. We run into such a difficulty when we consider the map, $x_{t+1} = x_t + 10.5\left[0.2x_{t-4}/\{1 + (x_{t-4})^{10}\} - 0.1x_t\right]$. This map is the discretized analog of the Mackey-Glass differential equation (McCaffrey *et al.*, 1992).

2.3 LLE-LLE plots

In Section 1, we described how LLE-LLE plots are constructed and used to discriminate between stochastic time series and time series that are iterates of a univariate nonlinear deterministic map. Similarly, bivariate LLEs can be used to obtain LLE-LLE plots. However, in the bivariate case the bins are two dimensional.

From the results of the previous subsection, for white noise and time series that arise from the AR(1) and AR(2) models, we do not expect the the averaged bivariate LLEs (bin average) obtained for the first and second half of the time series to systematically vary from bin to bin. Hence, we do not expect the bivariate LLE-LLE plots to show linearity behavior. In Fig 3 (a), (b) and (c), the LLE-LLE plots for the standard normal distribution, AR(1) model with parameter 0.7, and AR(2) model with parameters 1 and -0.5 are presented. For

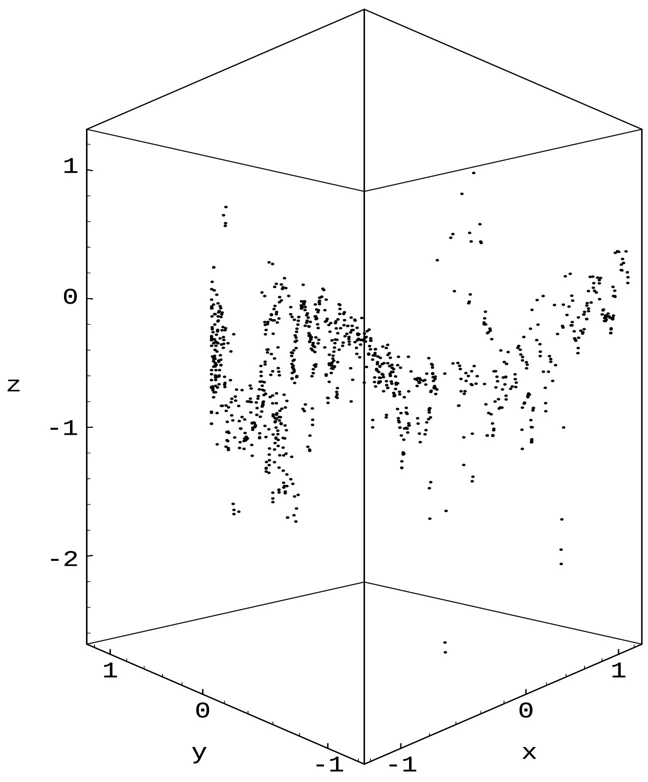

Figure 2: *LLE plot for the Henon map:* $(x, y, z) = (x_{i-1}, x_i, \lambda^{\mathbf{x}}_{i,1})$*, sample size=2000,* $h = 0.1$ *and 5% trimming is applied.*

each plot in Fig. 3 the sample size is 2000, 5% trimming was initially applied, $h = 0.1$, and the bin width and height over which we averaged the LLEs is taken to be the range divided by 50. As expected in all three cases no linearity pattern is detected.

In contrast, for the logistic and Henon map, we have shown that $\lambda_{ij}^{\mathbf{x}}$ varies systematically with the state. This implies that for time series generated from these models, for reasonable sample sizes, the averaged bivariate LLEs for the first and second half of the time series would vary systematically from bin to bin with the values obtained from matching bins being similar. Hence, we expect the bivariate LLE-LLE plots to show the linearity property. In Fig. 3 (f) and (d), the LLE-LLE plots for the Henon map and the logistic map are presented. As expected in both cases a strong linearity pattern is evident.

Furthermore, in Fig. 3, LLE-LLE plots for the discretized analog of the Mackey-Glass differential equation, the bilinear model $x_{t+1} = (-0.1 + 0.9\varepsilon_{t+1})x_t + \varepsilon_{t+1}$ (Tong, 1990) and the bilinear model $x_{t+1} = 0.3x_t - 0.4x_{t-1} + \varepsilon_{t+1} + 0.6x_t\varepsilon_t + 0.5x_{t-1}\varepsilon_t$ (Priestley, 1988) are presented. For the Mackey-Glass differential equation, because the nonlinearity of the time series is associated with lag five, instead of using the iterated time series $\{x_t\}$, the time series $\{v_t\}$ where $v_1 = x_1$, $v_2 = x_6$, $v_3 = x_{11}$,... was used. The LLE-LLE plot for the bilinear models show very little or no linearity pattern. In contrast, a strong linearity pattern is observed for the Mackey-Glass equation. This linearity pattern is not observed when looking at the LLE-LLE plot obtained from the time series $\{x_t\}$.

From these plots we note that the LLE-LLE plot has largely discriminated the time series that arose as iterates of a nonlinear deterministic map from the stochastic time series. We also note that even though the plots in Fig. 3 are only obtained for the deterministic case we still observe linearity for the case where the system is observed with stochastic perturbation, provided the standard deviation of the noise is small and satisfies distributional properties

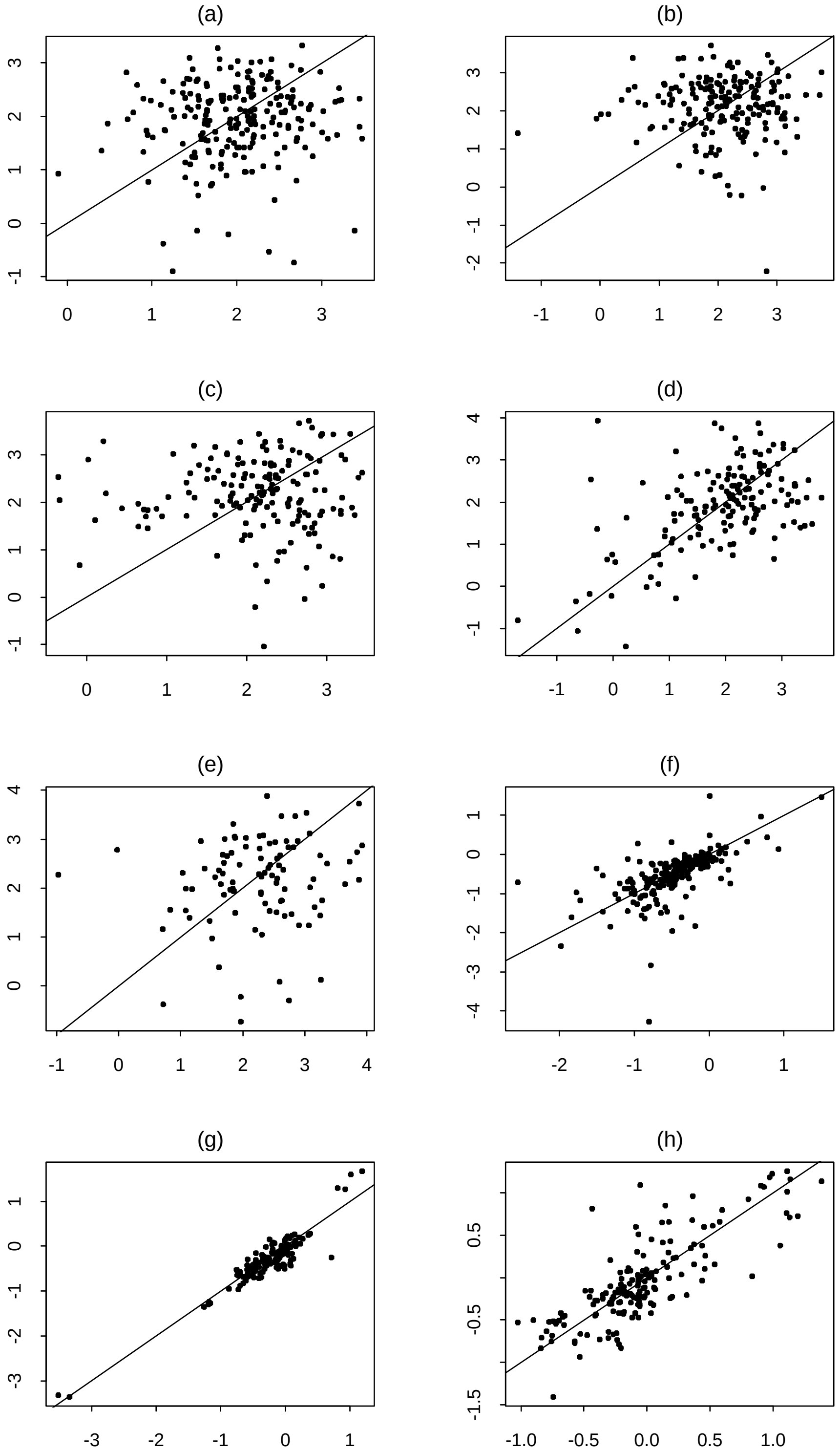

Figure 3: *Bivariate LLE-LLE plots for (a) white noise, (b) AR(1) model, parameter 0.7, (c) AR(2) model, parameters 1, -0.5, (d) bilinear model,* $x_{t+1} = (-0.1 + 0.9\varepsilon_{t+1})x_t + \varepsilon_{t+1}$*, (e) bilinear model,* $x_{t+1} = 0.3x_t - 0.4x_{t-1} + \varepsilon_{t+1} + 0.6x_t\varepsilon_t + 0.5x_{t-1}\varepsilon_t$*, (f) Henon map, (g) logistic(4) map, and (h) discretized analog of the Mackey-Glass delay differential equation. Sample size=2000,* $h = 0.1$ *and trimming of* 5%.

discussed in Chan and Tong (1994). For the Henon map observed with small levels of additive and systems noise, Seifu (1995) has shown that bivariate LLE-LLE plots can still detect linearity. Furthermore, this detection of linearity is improved by applying the MA(m) filter to the time series.

All the chaotic maps discussed in this section have been differentiable maps. However, we note that a chaotic map does not necessarily have to be differentiable. Consider the tent map which is given by

$$x_{t+1} = \begin{cases} 2x_t & 0 \leq x_t \leq 0.5 \\ 2(1 - x_t) & 0.5 \leq x_t \leq 1. \end{cases}$$

This map is not differentiable, but it is chaotic (Ruelle, 1987).

3 Analysis of Heart Rate Variability Data Using Local Lyapunov Exponents

3.1 Introduction

An electrocardiogram (ECG) is a recording of the electrical changes that occur in the heart during a cardiac cycle. Typically cardiologists identify five points on a cardiac cycle labeled P, Q, R, S and T. One method of studying beat-to-beat fluctuations of heart rate is to examine the variability of the R-to-R time intervals (RR intervals). In this section we use the univariate and bivariate LLE-LLE plots to investigate for possible presence of low dimensional nonlinear deterministic behavior in some heart rate variability (HRV) data which are in the form of RR intervals.

One of the most widely used method for discriminating between stochastic and chaotic data is the Grassberger-Procaccia (GP) algorithm which is used to estimate correlation dimension. See Cutler (1993) for a description and detailed discussion of the GP algorithm and the correlation dimension. One approach to detecting chaos in a given data set is to conclude if the GP algorithm converges to a finite value that there is some evidence that the dynamical system is chaotic (see Yamamoto *et al.* (1990) and references therein). In the study of HRV data some experimentalists have observed this convergence of the algorithm. This in turn has led them to suggest that there is some indication that human heart fluctuation has nonlinear low-dimensional deterministic dynamics (Yamamoto *et al.*, 1993). Noting that the GP algorithm gives finite value for some stochastic models and that there are many technical difficulties associated with the GP algorithm, Glass and Malt (1990) state that reliable conclusions about evidence of deterministic chaos cannot be drawn from such observations in physiological data. However, Yamamoto *et al.* (1993), by applying a number of other mathematical methods conclude that there is some evidence of deterministic and/or chaotic behavior, in some HRV data collected at Mt. Everest, above 6,000 meters.

The dynamics of HRV is very complex and interconnected with other physiological systems of the body. However, embedding theorems from dynamical systems (Takens, 1981, Nychka *et al.*, 1992) suggest that the investigation of univariate observations for the possible presence of underlying low-dimensional chaotic/nonlinear system is reasonable. In this section we investigate the HRV data sets collected in the Mt. Everest study (Yamamoto *et al.*, 1993) via LLE-LLE plots to see if these HRV data sets satisfy the model assumption $x_{t+1} = f(x_t, x_{t-1}, ..., x_{t-(d-1)})$ (possibly with some additive or system noise) where d is small and f is nonlinear. These data sets are part of a larger data set collected at Mt. Everest and were supplied to us by Professor Yamamoto. The data sets considered here comprise four pairs of readings taken from four individuals. On each subject readings were taken at both sea level and at an altitude of above 6000 m during sleep state. The sampling frequency was 100 Hz which implies that data was recorded in units of 10 ms. The HRV readings in each case were collected for about 30 minutes: the sample sizes of these data sets are between 800 and 3400 RR intervals. For each of the high altitude data sets Yamamoto *et al.* (1993) have concluded that there is evidence of deterministic dynamics. In contrast, no evidence of low dimensional deterministic dynamics was detected in the sea level data sets. We included these sea level data sets in our analysis for comparison.

In subsequent subsections data set *high one* through *high four* refer to data collected at high altitude (> 6000 m) on the four subjects, while data sets *sea one* through *sea four* refer to the corresponding sea level observations taken on these four individuals.

3.2 Preliminary analysis

For each data set, we initially looked at the time and phase plots. In Fig. 4 and 5, the time plots for these data sets are presented. The time plots for the sea level data sets show little evidence of structure such as cyclical dynamics, trend, etc. In contrast, the time plots for the high altitude data set show some evidence of cyclical dynamics. Furthermore, when looking at the phase plots of these data sets, we consistently found that the four high altitude data sets show more evidence of structure. In Fig. 6, based on 250 RR observations, the phase plots for these data sets are presented. Observe that for each subject the high altitude phase plot shows more structure.

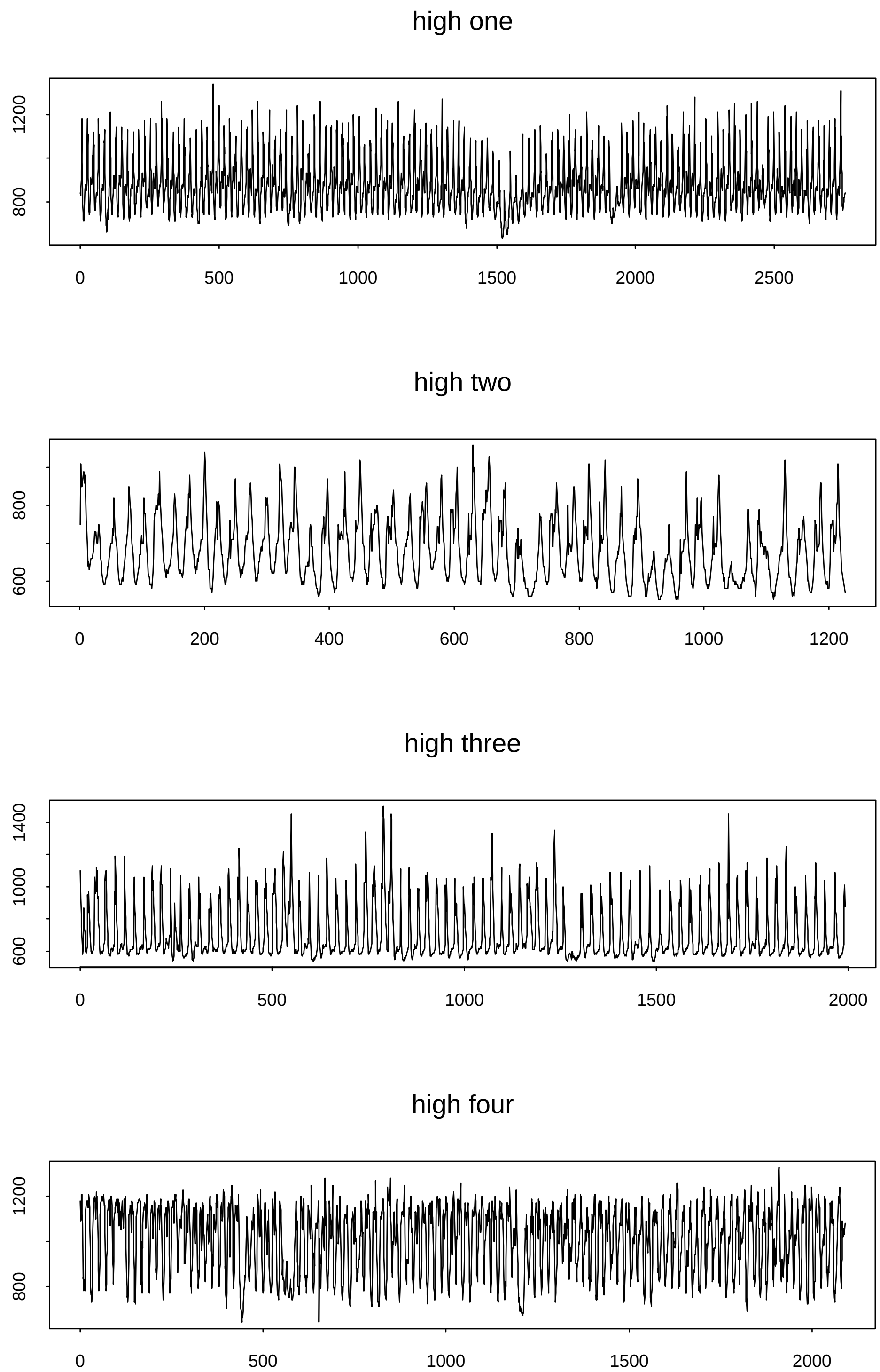

Figure 4: *Time plots for the data sets high one, high two, high three* and *high four.*

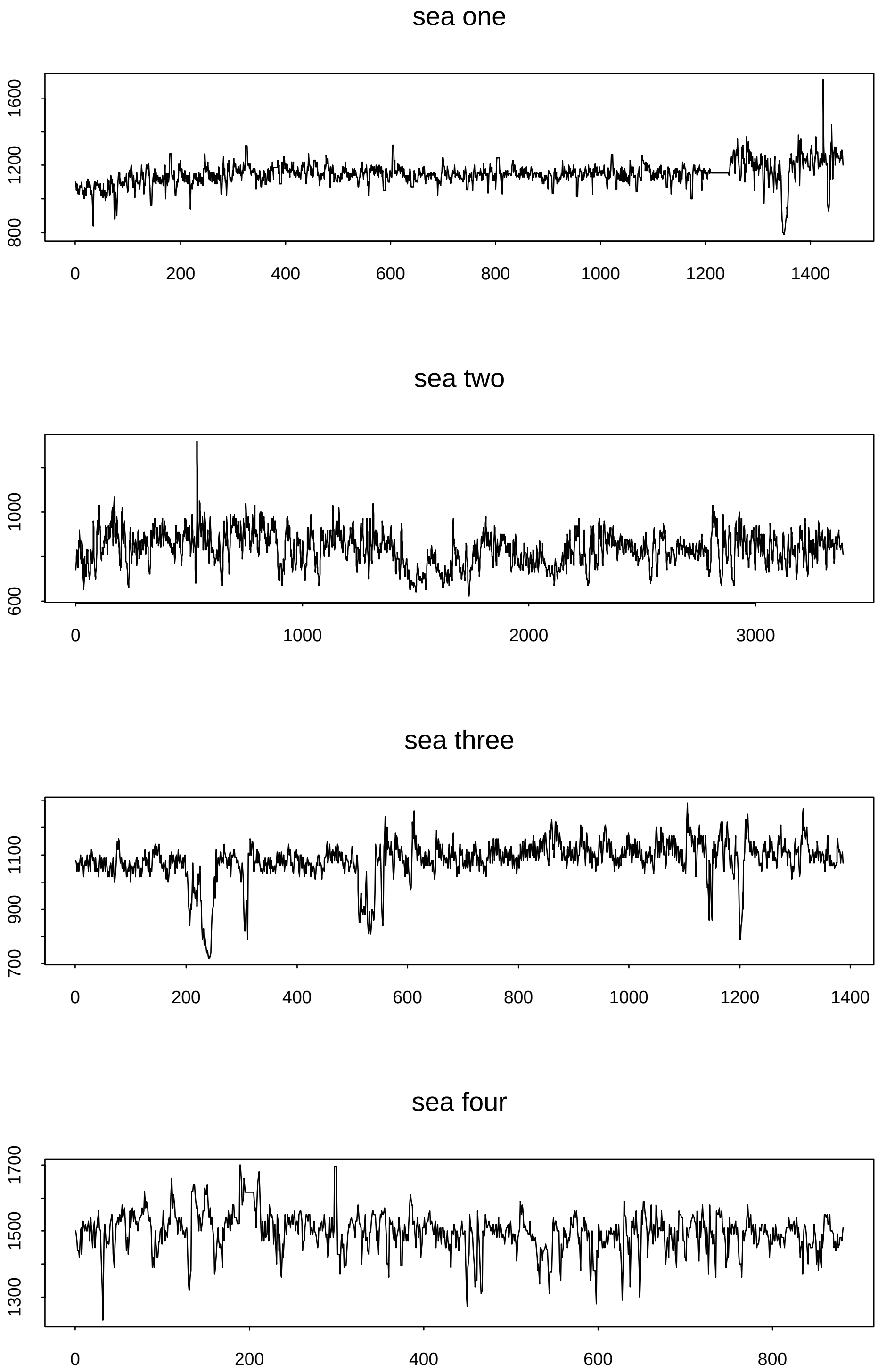

Figure 5: *Time plots for the data sets sea one, sea two, sea three* and *sea four.*

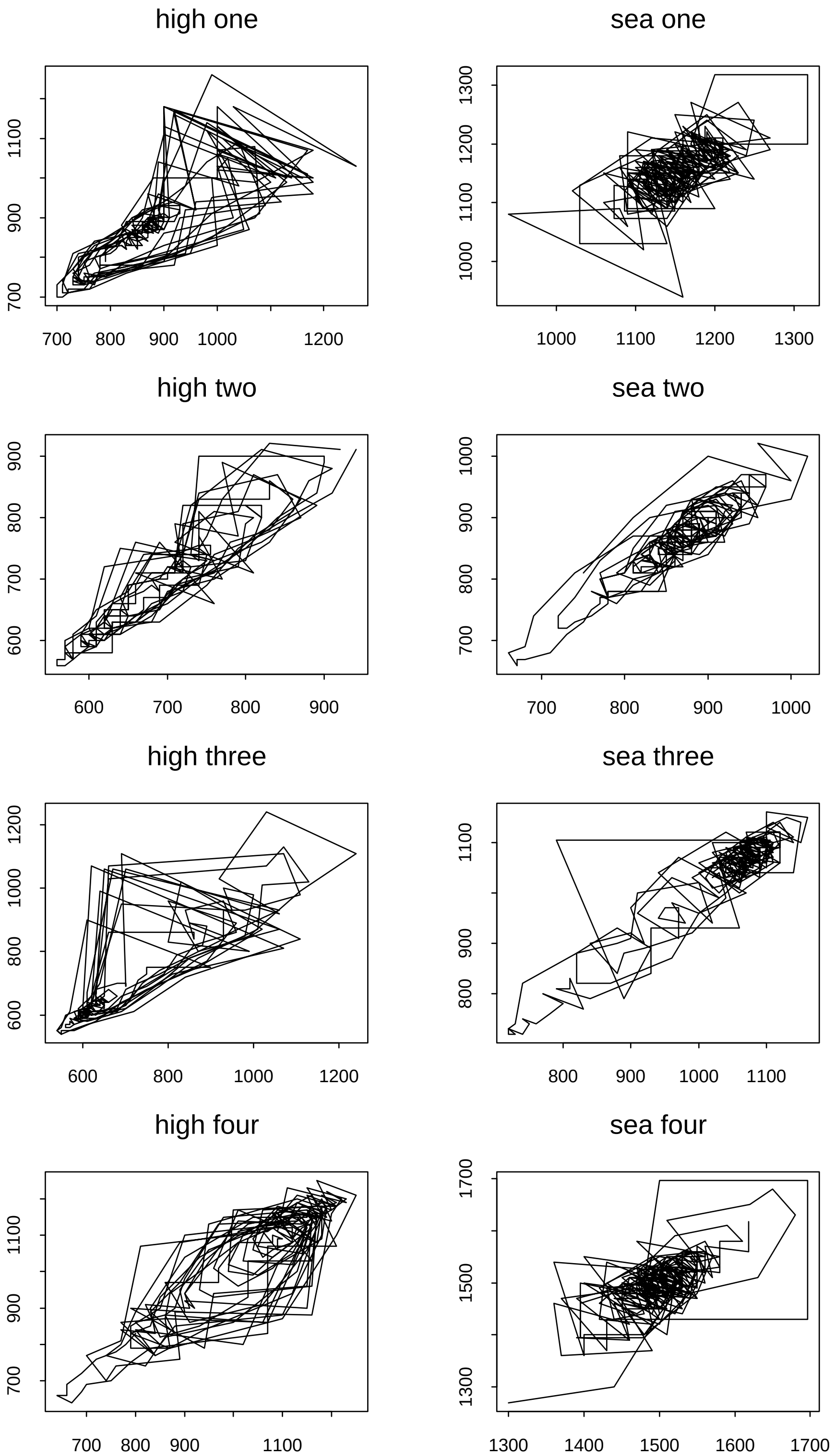

Figure 6: *Phase plots for subjects one, two, three and four in the Mt. Everest study (sea and high refer to data collected at sea level and high altitude).*

3.3 Analysis of HRV using LLE-LLE plots

Because the HRV readings are recorded only to 10 ms of accuracy, there are a large number of repeat observations in the data. For example, for the data set *high one*, the RR values range between 600 – 1300 hence, there are less than 80 distinct values. We “unrounded” the data by adding to each observation a randomly generated value from $unif(-5, 5)$ before constructing the LLE-LLE plots. The observations have also been rescaled by 1000.

In Fig. 7, the univariate plots obtained from the rescaled surrogate data for the subjects in the Mt. Everest study are presented. From the four LLE-LLE plots obtained from the data collected at sea level we observe very little or no linearity structure. In contrast, from the LLE-LLE plots obtained from the observation taken at high altitude, we observe that the LLE-LLE plot for subjects one and four appear to be linear while some linearity is detected for subject three. The corresponding correlation values computed for these plots are 0.871, 0.712 and 0.782 for subjects one, three and four, respectively. Fig. 8 shows the bivariate LLE-LLE plots obtained for the subjects in the Mt. Everest study. We again observe no or very little linearity from the LLE-LLE plots obtained for the sea level data sets. In contrast, among the LLE-LLE plots obtained at high altitude the LLE-LLE plots obtained for subjects one and three appear to be linear. The correlation values computed for these plots are 0.673 and 0.763, for subject one and three, respectively. Furthermore, note that in contrast to the univariate plot obtained for the data set *high four* the corresponding bivariate LLE-LLE plot shows very little linearity structure. This might be attributed to the increased sparsity of the data due to the increase in observational dimension. From the univariate and bivariate LLE-LLE plots obtained for the subjects in the Mt. Everest study, we conclude that there is some evidence of nonlinear deterministic behavior for the high altitude data set obtained from subjects one, three and four.

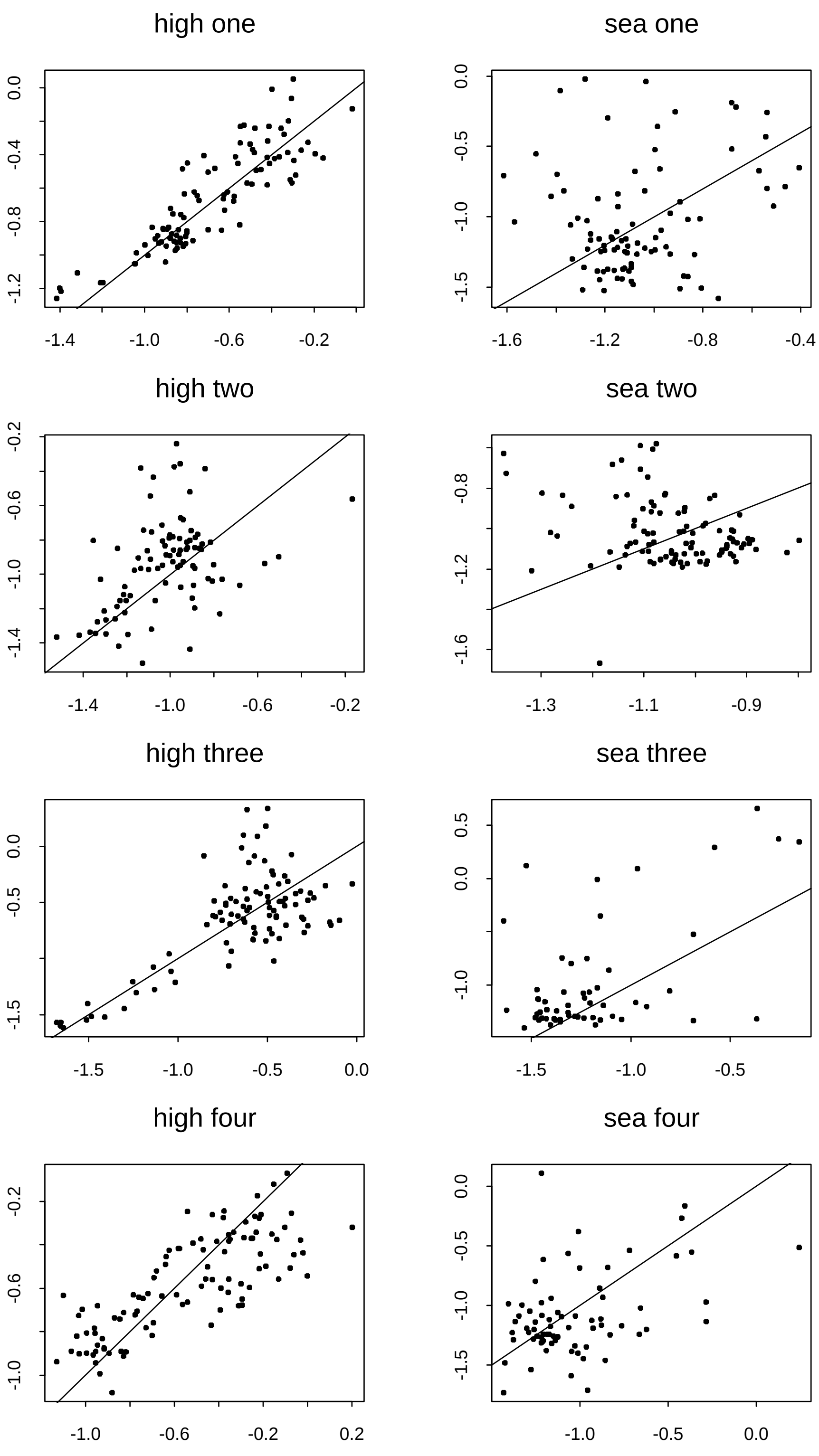

Figure 7: *Univariate LLE-LLE plots for subjects one, two, three and four in the Mt. Everest study (sea and high refer to data collected at sea level and high altitude).*

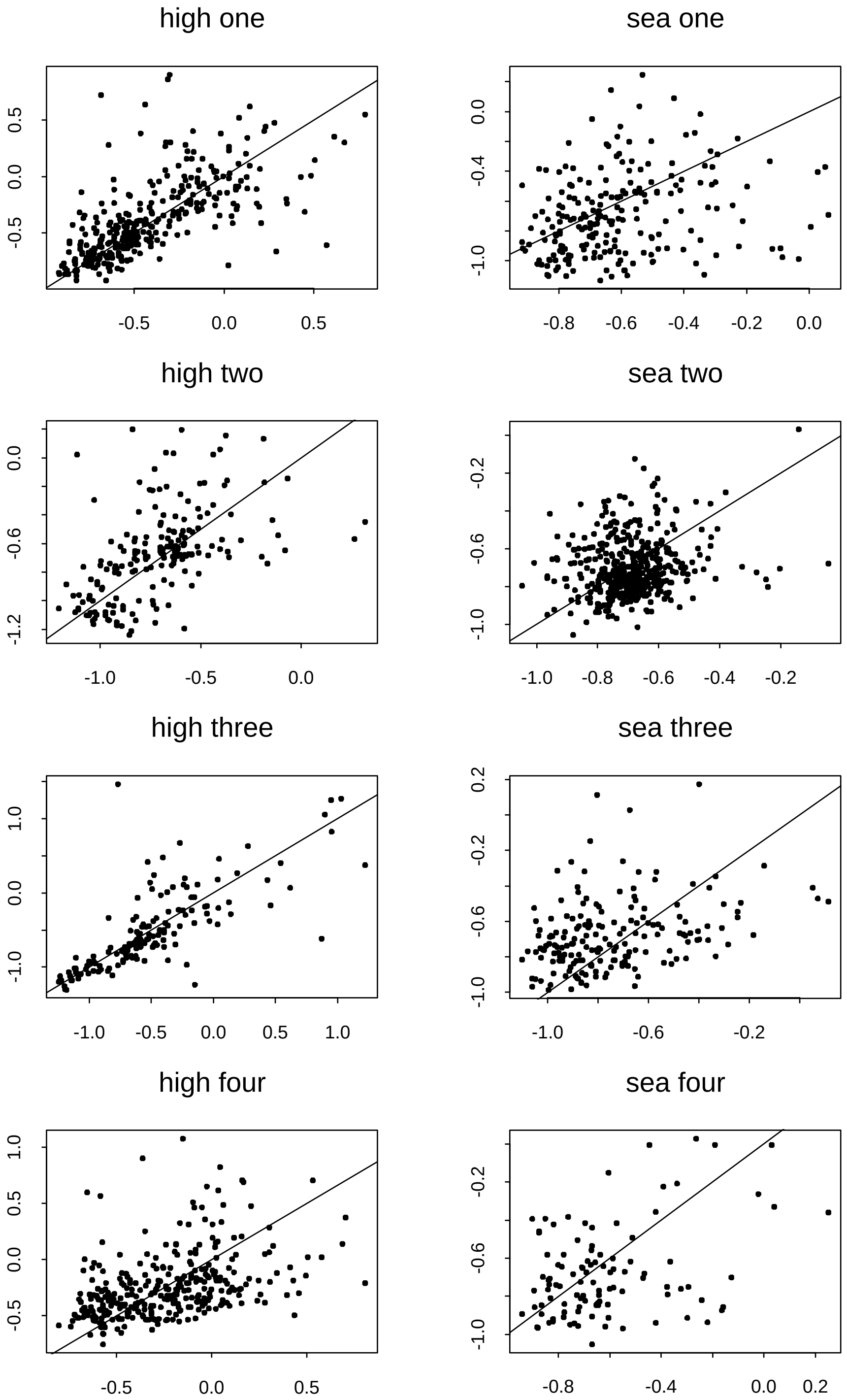

Figure 8: *Bivariate LLE-LLE plots for subjects one, two, three and four in the Mt. Everest study (sea and high refer to data collected at sea level and high altitude).*

It would be of interest to be able to construct inference on the slope estimates from the LLE-LLE plots. The analytical analysis, even asymptotic, is very difficult. Some simulation results are discussed in Seifu (1995), and work on this problem is ongoing. In addition, there is clearly considerable between-subject variability that had not yet been addressed.

3.4 Application of MA(m) filters on LLE-LLE plots

Seifu's (1995) investigation of the Henon, logistic and exponential map has shown that MA(m) filter can improve the LLE-LLE plots ability to detect nonlinearity, the optimal m being map dependent. Thus as a further analysis of the Mt. Everest data we filtered the scaled surrogate data and obtained LLE-LLE plots.

In Fig. 9 we see the filtered and non-filtered bivariate LLE-LLE plots obtained for the scaled surrogate data derived from the high altitude data sets. The filter applied in each case was the MA(3) filter. From these plots we note that for the filtered data there is stronger evidence of linearity structure in the LLE-LLE plots obtained for subjects one, two and three. In contrast, the filtered LLE-LLE plot obtained for subject four showed very little improvement over the LLE-LLE plot obtained for the non-filtered data.

For contrast we have also obtained the filtered and non-filtered bivariate LLE-LLE plots for the corresponding sea level data sets (Fig. 10). In each case, the filtered data showed very little evidence of linearity.

4 Conclusions

The univariate and bivariate LLE and LLE-LLE plots are useful diagnostic tools for bivariate and univariate nonlinear maps. However, their diagnostic nature is not limited only to such maps. Consider a difference equation $x_{t+1} = f(x_t, ..., x_{t-(d-1)})$ for some small d and

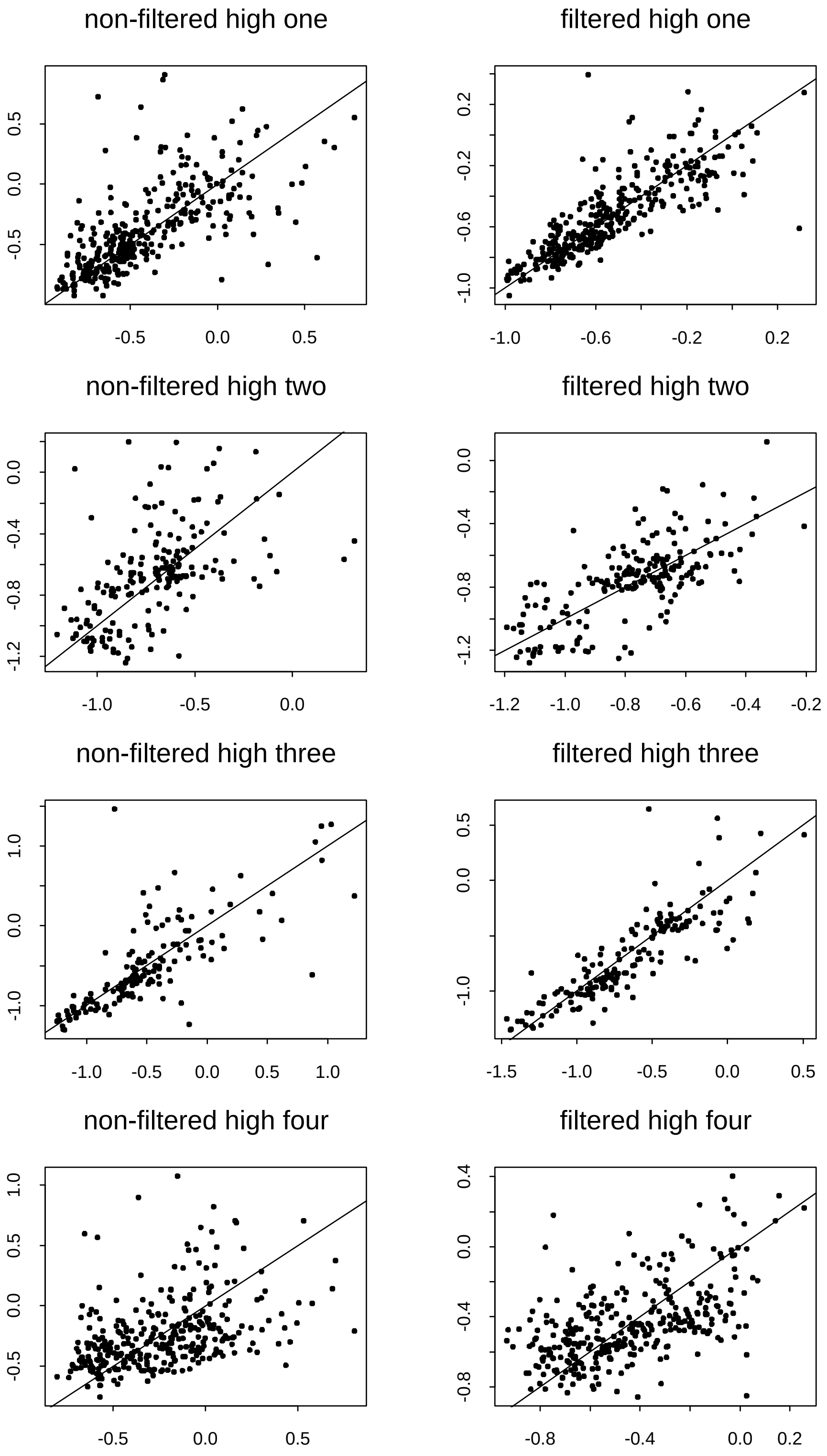

Figure 9: *Filtered and non-filtered bivariate LLE-LLE plots for high one, high two, high three and high four (the filter in each case is the MA(3) filter).*

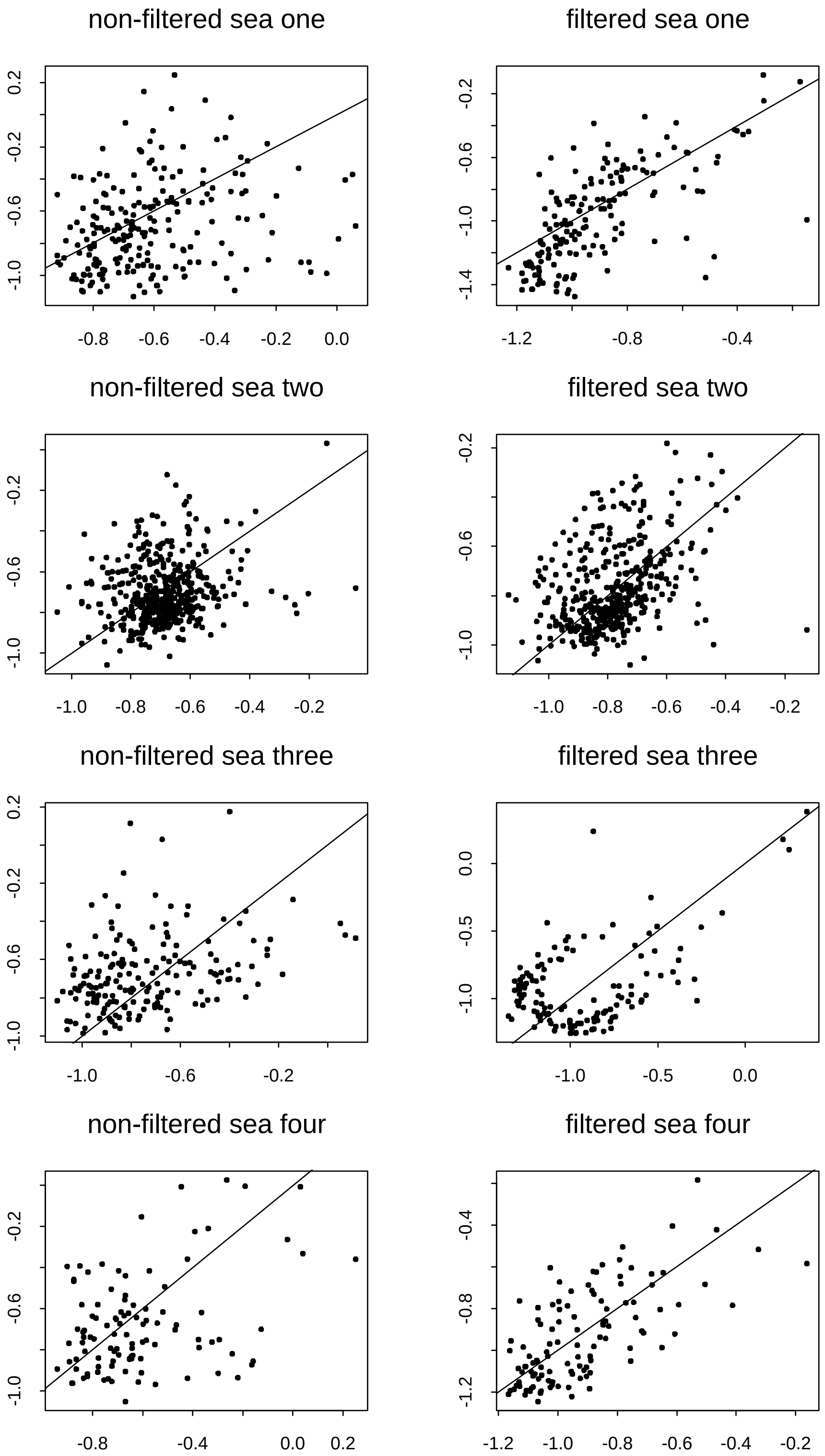

Figure 10: *Filtered and non-filtered bivariate LLE-LLE plots for sea one, sea two, sea three and sea four (the filter in each case is the MA(3) filter)*

let $\{x_t\}$ be the time series generated by iterating f. In the absence of any knowledge of d, we could consider univariate and/or bivariate plots at different lag values. For example, for the discretized analog of the Mackey-Glass differential equation of dimension five, the linearity structure is clearly seen when looking at the bivariate LLE-LLE plot constructed from the pairs (x_t, x_{t-5}). As well, for some d dimensional difference equations, the non-linearity contribution associated with x_{t-1} and x_{t-2} may be the strongest, in which case the univariate and/or bivariate LLE-LLE plots would be expected to show evidence of low dimensional nonlinear deterministic behavior.

In Section 2 and 3 we have visually searched for evidence of linearity in the LLE-LLE plots. Since the averaged LLE values are correlated and since we are only interested in detecting linearity relationship along the line with zero intercept and slope one, we have avoided fitting linear regression to the LLE-LLE plots.

In Section 3 we have jittered the data in order to eliminate the discreteness of the HRV data sets. To make sure that the results of the LLE-LLE plots are not dependent on a given jittering we have repeatedly obtained the surrogate data and obtained LLE-LLE plots for these data sets. The visually observed results were all consistent. For the univariate case, modification to the definition of the LLE in order to handle the discreteness of the HRV data sets was discussed by Seifu (1995). These univariate LLE-LLE plot results were very similar to Fig. 7.

Apart from MA(m) filters, in the literature, there are various other types of filters that are used for removing the noise and recovering the signal from a given time series. Further investigations in this area should include the effect of such filters on the LLE-LLE plots ability to detect nonlinearity.

A Appendix: Asymptotic Difference of Equation (1) and Equation (2)

Here we show that the approximate expected difference between $\widetilde{\lambda}^{\mathbf{x}}_{i,1}$ given in (2) and $\lambda^{\mathbf{x}}_i$ given in (3) is $0.5 + O(h^2 \log h)$. Let X and Y be two random variables with a joint probability distribution function $f_{XY}(x,y)$ with support in R^2. For our purpose we are interested in finding an expression for

$$E(\lambda^{\mathbf{x}}_{ij} - \widetilde{\lambda}^{\mathbf{x}}_{ij1} \mid ||\mathbf{x}_i - \mathbf{x}_j|| < h) + \log h = E(\log R | R < h) \quad \text{where} \quad R = \sqrt{X^2 + Y^2}.$$

Shifting to polar coordinates, the joint density of r and θ is $r f_{XY}(r\cos\theta, r\sin\theta)$, the marginal p.d.f. for r is

$$\begin{aligned} g_R(r) &= \int_0^{2\pi} r f_{XY}(r\cos\theta, r\sin\theta) d\theta \\ &= r\int_0^{2\pi} f_{XY}(r\cos\theta, r\sin\theta)\,|_{r=0}\, d\theta + r^2 \int_0^{2\pi} \frac{\partial f_{XY}(r\cos\theta, r\sin\theta)}{\partial r}\,|_{r=0}\, d\theta \\ &\quad + \frac{r^3}{2}\int_0^{2\pi} \frac{\partial^2 f_{XY}(r\cos\theta, r\sin\theta)}{\partial r^2}\,|_{r=0}\, d\theta + \frac{r^4}{6}\int_0^{2\pi} \frac{\partial^3 f_{XY}(r\cos\theta, r\sin\theta)}{\partial r^3}\,|_{r=0}\, d\theta \\ &\quad + O(r^5). \end{aligned}$$

Since over the interval $(0, 2\pi)$ the integrals of $\cos\theta$, $\sin\theta$, $\cos^3\theta$, $\sin^3\theta$, $\cos^2\theta\sin\theta$ and $\sin^2\theta\cos\theta$ are all zero, the coefficients of r^2 and r^4 in the expansion of $g_R(r)$ are zero. Furthermore, as the integral of $\cos\theta\sin\theta$ over this interval is zero the coefficient of r^3 is $\pi b/2$ where $b = \partial^2 f_{XY}(x,y)/\partial x^2\,|_{x=0,y=0} + \partial^2 f_{XY}(x,y)/\partial y^2\,|_{x=0,y=0}$. Hence,

$$g_R(r) = 2\pi r a + \frac{r^3}{2}\pi b + O(r^5)$$

where $a = f_{XY}(0,0)$. Now let $G_R(r) = P(R < h)$. Then

$$G_R(r) = G(0) + rg_R(0) + \frac{r^2}{2}g_R'(0) + \frac{r^3}{6}g_R''(0) + \frac{r^4}{24}g_R'''(0) + O(r^5) = r^2\pi a + \frac{r^4}{8}\pi b + O(r^5).$$

The conditional distribution for $(R \mid R < h)$ is given by

$$\begin{aligned}
g_h(r|r<h) &= \frac{g_R(r)}{G_R(h)} \\
&= \frac{2ra + br^3/2 + O(r^5)}{h^2a + h^4b/8 + O(h^5)} \\
&= \frac{2ra + br^3/2}{h^2a}(1 + O(h^2)) + \frac{O(r^5)}{h^2a}(1 + O(h^2)) \\
&= \frac{2r}{h^2} + O(h).
\end{aligned}$$

Hence,

$$\begin{aligned}
E(\log r \mid r < h) &= \int_0^h (\log r) g_h(r|r<h) dr \\
&= \frac{2}{h^2}\int_0^h r \log r \, dr + O(h)\int_0^h \log r dr \\
&= \log h - \frac{1}{2} + O(h^2 \log h).
\end{aligned}$$

Thus, the approximate expected difference between $\widetilde{\lambda}_{i,1}^{\mathbf{x}}$ given by (2) and $\lambda_i^{\mathbf{x}}$ given by (3) is $0.5 + O(h^2 \log h)$.

References

Chan, K.S. and Tong, H. (1994). A note on noisy chaos. *J. R. Statist. Soc. B,* 56, 301-311.

Cutler, C.D. (1993), Dimension estimation and models. In *Nonlinear Time Series and Chaos, Vol I* (ed. Tong, H.) World Scientific, Singapore 1-107.

Eckmann, J.-P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.,* 57, 617-656.

McCaffrey, D., Ellner, S., Gallant, A.R. and Nychka, D. (1992). Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *J. Amer. Statist. Assoc.,* 87, 682-695.

Nychka, D., Ellner S., McCaffrey, D. and Gallant, A.R. (1992). Finding chaos in noisy systems. *J. R. Statist. Soc. B,* 54, 399-426.

Priestley, M.B. (1988). *Non-linear Non-stationary Time Series Analysis.* Academic Press, San Diego.

Ruelle, D. (1987). *Chaotic Evolution and Strange Attractors.* Cambridge University Press, Cambridge.

Seifu, Y. (1995). *Bivariate Local Lyapunov Exponents: Results and Applications.* Unpublished Ph.D. Thesis, University of Toronto, Toronto.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980,* Lecture Notes in Mathematics, no. 890 (eds. Rand, D. and Young, L.- S.) Springer-Verlag, Berlin, 366-381.

Tong, H. (1990). *Nonlinear time series.* Oxford University Press, Oxford.

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.

Wolff, R.L. (1992). Local Lyapunov exponents: looking closely at chaos. *J. Roy. Statist. Soc. Ser. B,* 54, 353-371.

Yamamoto, Y., Hughson, R.L., Sutton J.R., Houston C.S., Cymerman, A., Fallen, E.L. and

Kamath, M.V. (1993). Operation Everest II: An indication of deterministic chaos in human heart rate variability at simulated extreme altitude. *Biol. Cybern.*, 69, 205-212.