Nonstationarity Detection

Bunty Byarugaba Kiremire

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.

Johannesburg, April 2007

Declaration

I declare that this dissertation is my own, unaided work, except where otherwise acknowledged. It is being submitted for the degree of Master of Science in Engineering in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Signed this ____ day of ______________ 20___

____

Bunty Byarugaba Kiremire

Abstract

This dissertation is an investigation into nonstationarity detection. Current methods are looked into and a novel method is proposed. An index was developed during the course of the study called the Stationarity Index. The method of nonstationarity detection proposed is based on this index and it gives a measure that varies directly proportionally to changes in the dynamic behavior of the signal. This allows not only for the detection of nonstationarity but also serves as a means of determining relatively how much it is changing. The Stationarity Index maps changes in the dynamics to a range of $0 - 100$ in the similar manner to the mapping of measures to percentages. This allows for the comparison of the variation of dynamics between separate signals. This ability sets the test apart from current tests. Most current tests simply determine if nonstationarity is present or not, and do not allow for the kind of measure that allows relative comparisons between signals. The index is then used to successfully distinguish between stationary and nonstationary chaotic signals. It is then used in the analysis of electrocardiogram (ECG) and electroencephalogram (EEG) profiles. It is able to clearly pick out the changes in the dynamics that are the result of the onset of partial seizures. It shows potential for use in the identification of certain dynamic traits in ECG and EEG profiles. Its ability to discriminate between stationarity and nonstationarity also shows potential for use in the segmentation of nonstationary biomedical signals into partially (quasi) stationary segments.

Acknowledgements

I would like to acknowledge Professor Tshilidzi Marwala and thank him for being a fantastic role model and source of inspiration for myself and countless others. I would like to thank Vincent Nelwamondo and Anthony Gigudu for their invaluable help with the ideas of the research and their wisdom. I would also like to thank the members of the research group, Gregory Hulley and Bodie Crossingham, with whom I could bounce off ideas as well as enjoy pleasant conversation.

Funding and support from the following organisation is gratefully acknowledged:

- National Research Foundation

Special thanks go to my family for their support and encouragement, as well as to my friends who have made the hard work and effort a more pleasant experience.

Contents

Declaration **i**

Abstract **ii**

Acknowledgements **iii**

Contents **iv**

List of Figures **vii**

1 Introduction **2**

2 Background **5**

2.1 Problem Statement 5

2.2 Statistical and Frequency Based Methods 6

2.2.1 Statistical Measures 6

2.2.2 The Short Time Fourier Transform and the Power Spectral Density 6

2.2.3 Wavelet Transforms 8

2.2.4 The Discrete Wavelet Transform 9

2.2.5 The Discrete Gabor Transform 10

2.3 Phase Space Based Methods 11

2.3.1 Recurrence quantification analysis 14

2.3.2 Recurrence Time Analysis 14

2.3.3 Nonstationarity Detection by the Loss of Recurrence 16

2.3.4 Space Time Separation Plots 16

2.3.5 Nonlinear Cross Prediction Analysis 17

2.3.6 Additional Methods of Analyzing Non-Stationarity 18

2.4 Conclusion 18

3 The Cross Correlation Integral 19

3.1 The Cross Correlation Integral 19

3.2 Current Nonstationarity Detection Methods based on the Cross Correlation Integral 20

3.2.1 Cross Correlation Sum Analysis 20

3.2.2 Fel'dshtein's Application of the Cross Correlation Integral . . 21

3.3 The Development of a Nonstationarity Detection Technique based on the Cross Correlation Integral 21

3.3.1 The Theiler Window 22

3.3.2 The Stationarity Index 26

3.3.3 Methodology 28

3.4 Conclusion 29

4 Results I 31

4.1 Results of Nonstationarity Tests on Chaotic Signals 31

4.2 Conclusion 37

5 Results II **38**

5.1 Results of Nonstationarity Test on ECG and EEG Signals 38

5.2 Conclusion 46

6 Conclusion **47**

A Papers **49**

References **50**

List of Figures

2.1 Henon signal coordinates for 1^{st} phase vector 12

2.2 Henon signal coordinates for 2^{nd} phase vector 12

2.3 Henon signal coordinates for 3^{rd} phase vector 13

2.4 Phase plot of Henon map 13

2.5 Full phase plot of the Henon signal 14

2.6 Nearest Neighbors plot 15

4.1 Henon signal 32

4.2 Stationarity Index plot of Henon signal: wrt. the 1st ref window . . 33

4.3 Stationarity Index plot of Henon signal 3D plot 33

4.4 Lorenz signal Stationary Conditions 34

4.5 Stationarity Index plot of a Lorenz signal Stationary Conditions: wrt. the 1st ref window 34

4.6 Stationarity Index plot of a Lorenz signal Nonstationary Conditions: wrt. the 1st ref window 35

4.7 Nonstationary Lorenz Section 1 35

4.8 Nonstationary Lorenz Section 2 36

4.9 Nonstationary Lorenz Section 3 36

4.10 Stationarity Index plot of a Lorenz signal Nonstationary Conditions 3D plot 37

5.1 Stationarity Index plot of an ECG signal 39

5.2 Stationarity Index plot of an ECG signal: wrt. the 1st ref window . 40

5.3 Stationarity Index plot of an ECG signal: wrt. the 28th ref window 41

5.4 Stationarity Index plot of an EEG signal 42

5.5 Stationarity Index plot of the above EEG signal wrt. the 1st ref window 42

5.6 Stationarity Index plot of ECG signal 1: wrt. the 1st ref window . . 43

5.7 Stationarity Index plot of ECG signal 2: wrt. the 1st ref window . . 43

5.8 Stationarity Index plot of ECG signal 3: wrt. the 1st ref window . . 44

5.9 Stationarity Index plot of ECG signal 4: wrt. the 1st ref window . . 44

5.10 Stationarity Index plot of ECG signal 5: wrt. the 1st ref window . . 45

5.11 Stationarity Index plot of ECG signal 6: wrt. the 1st ref window . . 45

Chapter 1

Introduction

The detection of nonstationarity has been noted as one of the most vexing problems in time series analysis. This is because though there is a clear definition of stationarity and nonstationarity, there is no clear method classifying a time signal as stationary [1; 2]. Most methods of time series analysis assume stationarity [3] because these techniques, e.g. the Fourier transform, require that the signals analysed be stationary in order to yield a meaningful analysis. Nonstationarity implies that the dynamics of the signal are changing with time. Thus an analysis of the signal at a given time may not give useful information on the future behavior of the signal. Most signal analysis techniques assume stationarity because this assumes that signal characteristics are unaffected by time and thus an analysis of these characteristics can be taken to give useful information that is time invariant. However stationarity is not an inherent property of signals and it is because of this that the detection of stationarity in time series data is very useful in statistical analysis. Furthermore, the detection of stationarity is very useful in the construction of statistical models such as the stock market, as stationarity implies that the models do not change with time, and therefore are reliable. Tests for the detection of stationarity are also applicable in fault detection.

Most biomedical signals are nonstationary by nature [4]. Though they are nonstationary most processing techniques, e.g. the Short Time Fourier transform (STFT), require that these signals be segmented into quasi or partially stationary segments. It would therefore be important to determine the extent to which a signal is stationary in order to classify a signal as quasi stationary or not. Doing so would aid the process by which medical signals are segmented into quasi-stationary sections that can then be analysed by traditional methods. Furthermore different psychotropic drugs appear to have their own 'electroencephalogram (EEG) portraits' [5]. Therefore changes in EEG and electrocardiogram (ECG) profiles can be linked to the presence of various conditions. The analysis and comparison of the dynamics of the profiles may aid in the identification of the respective conditions. It would be of use, therefore, to be able to compare similarities in the dynamics of any two signals.

This dissertation presents research aimed at developing a test for nonstationarity that not only detects nonstationarity but also allows the level of nonstationarity to be quantified. Several current methods of nonstationarity detection are looked into and a new method is proposed. The Stationarity Index (SI) is developed. This is an index that allows the unique use of the auto and cross correlation integrals in the development of a test for nonstationarity. It is unique in that it gives a means of quantifying the extent to which a signal is nonstationary. The index reflects the similarities in the dynamics of a signal as they change with time. This also allows for a comparison of how consistent the dynamics are and can be used to give an idea of the relative stationarity of signals, a concept that is novel and has not been seen in the literature reviewed. The index is successfully used in the analysis of electrocardiogram (ECG) and electroencephalogram (EEG) profiles as a means of examining the extent to which the dynamics of any two sections of a signal are similar. It can also serve as a means of determining the occurrence of an event.

The structure of the thesis is as follows: In Chapter 2 a brief outline of previous research is provided, several current methods of nonstationarity detection are discussed along with their assumptions and limitations. Fundamental principles are introduced, these are principals on which the developed tests are based.

In Chapter 3 the cross correlation integral is presented. Tests based on this integral are described. The Stationarity Index as well as its use in testing for nonstationarity are introduced. The differences between this test and current tests are discussed. Assumptions upon which the test is based, possible weaknesses and possible advantages are also discussed.

In Chapter 4 the results of Stationarity Index tests on chaotic signals are presented. In these cases the stationarity of the signals is already known. This is because their behavior is determined by mathematical equations and though their behavior may seem random and highly sensitive to initial conditions there are fixed rules that govern them [6]. The Stationarity Index tests on these signals serve as a bench mark to give some indication of the tests ability to discriminate against nonstationarity.

In Chapter 5 the results of Stationarity Index tests on electrocardiogram (ECG) and electroencephalogram (EEG) signals are presented. These signals have disturbances in them caused by the onset of partial epileptic seizures. Biomedical signals are generally nonstationary in nature [4], therefore the Stationarity Index is tested on them to determine whether it is useful in their analysis.

In Chapter 6 the findings of the dissertation are summarised and areas for further research are identified.

Chapter 2

Background

2.1 Problem Statement

For any given time duration a process is considered stationary if the probabilistic structure of the process does not vary for that period of time [7]. A random process $x(t)$ is declared strictly stationary if the joint probability distribution $P\Big(x(t_i), x(t_i+\delta_1), x(t_i+\delta_2), ..., x(t_i+\delta_n)\Big)$ of any n number of samples (where δ is the distance in time between successive points) is independent of t_i, the time at which the sample occurs [8]. Given a certain process $x(t)$ there is a probability distribution associated with it. For a stationary process this probability distribution does not change with time. As long as the process $x(t)$ is in operation it will follow the probability distribution that is associated with it. Therefore in a simplistic example, $x(t) = 3$ may have a probability of 0.5 and $x(t) = 5$ may have a probability of 0.3. The probabilities associated with these values will always be consistent, no matter when the samples are taken. If the process is nonstationary the probability distribution that is associated with the process will vary with time. Therefore at time t_1 the process may have a value of $x(t_1) = 3$ and this value may have a probability of 0.5. At time t_2 the process may have a value of $x(t_2) = 3$ but the probability of this is now 0.1 for example. This is an example of nonstationary behavior. By definition a signal is stationary when the joint probability distribution of its samples is time invariant.

As previously stated in Chapter 1 the detection of nonstationarity has been noted as one of the most vexing problems in time series analysis. To ideally determine the stationarity of a time series one would have to determine the joint probability distribution of a process $x(t)$ at any given time t_i even though there is only a single value of the time series $x(t_i)$ available at this time.

There are traits that time series exhibit that serve as good indicators of whether they are stationary or not. These traits form the basis of current tests for the detection of nonstationarity. They fall into two general categories, statistical and frequency based methods; and phase space methods. They shall be expounded on in the following paragraphs. Statistical and frequency based methods are the more traditionally used methods of nonstationarity detection. Phase space based methods are more recently developed and show greater potential for use with regard to the detection of nonstationarity.

2.2 Statistical and Frequency Based Methods

2.2.1 Statistical Measures

Statistical measures, such as the mean or standard deviation of a time series, have been used as a means of classifying signals as stationary or not [3]. Usually a time series is windowed or divided into sections. Then the measure (mean or standard deviation etc...) of the time series in a given window is calculated. This measure is compared to the measure calculated in other windows of the same time series. If the difference or the average of the difference crosses some predetermined threshold value, the time series is deemed nonstationary. Such methods usually require a priori information on the time series in order to determine what the threshold should be. They are not effective as the statistic is usually arbitrarily related to the natural geometric properties of the attractor thus the results when tested on various forms of stationarity vary greatly [9]. The arbitrary relationship between the measures and the geometric properties of the attractor means that the measures are not directly related to the dynamics of the signal and are therefore not an accurate measure of the extent to which the dynamics remain consistent.

2.2.2 The Short Time Fourier Transform and the Power Spectral Density

The Short Time Fourier Transform (STFT) can also be used as a means of determining the extent to which a signal is stationary [8]. Again the signal is segmented into windows. Each window is broken down to determine its frequency components and these frequency components are compared. A time series is deemed nonstationary if predetermined frequency thresholds are crossed. Again a priori information on the signal is required to determine where the thresholds should be.

The STFT is a Fourier transform that is done on different windowed portions of a signal [10; 11]. The windows are of a fixed width. For a time signal $f(t)$ the STFT is given as [12],

$$STFT(t,\omega) = \int f(\tau)g^*(\tau - t))e^{-j\pi\omega}d\tau \tag{2.1}$$

where the function g corresponds to a compactly supported window (Compactly supported in that the function g exists within a finite time period); $e^{-j\pi\omega} = \cos\pi\omega - j\sin\pi\omega$, with $j = \sqrt{-1}$. The $STFT$ is a variable of both time and frequency. Essentially the windowing function g sections the signal into different time sections and the Fourier transform of each section is determined. Thus the STFT is a $3D$ plot with frequency on one axis and time on the other, and the magnitude of the frequency components on the third axis showing how the different frequency components change over time. The signal $f(t)$ can be reconstructed via the inversion formula as shown in *Equation (2.2)* [12]. This implies that given the $STFT(t,\omega)$ for all t and ω, $f(t)$ can be completely recovered. By observing how the Fourier transforms vary from window to window, it is possible to determine if a signal is stationary or not. Vast redistributions of the frequency spectrum indicate non-stationarity.

$$f(t) = \frac{1}{2\pi g(0)}\int STFT(t,\omega)exp(jt\omega)d\omega \tag{2.2}$$

The Power Spectral Density (PSD) [4] is used in a very similar manner to determine the nonstationarity of a signal. It essentially examines the power of a signal as a function of its frequency components. If there are significant shifts in the distribution of the power across the different frequency components then a signal is deemed nonstationary.

These methods are however biased against the discovery of dynamic structure at certain scales of time and space [13]. This is due to the fact that they split the signal into sections and find the Fourier transform of the different sections. The window size is fixed in order to do this and this limits the extent to which the dynamic frequency characteristics of the signal can be captured.

2.2.3 Wavelet Transforms

Wavelet transforms are very similar to Fourier transforms except the signal is not decomposed into scaled sines and cosines, it is decomposed into the scaled sum of wavelets [11; 13; 14]. Consider a given time-signal $E(t)$ given by *Equation (2.3)*, where ψ(t) is the equation for the wavelet shape as shown in *Equation (2.4)* [11; 13; 14]. The wavelet shape is translated and scaled differently for each value of a, and b, where a represents the scale of each wavelet and b represents the value by which each wavelet is translated. A typical example of a wavelet is the Mexican Hat function with the equation as shown in *Equation (2.5)*. *Equation (2.4)* is the generalized equation for the wavelet shape where as *Equation (2.5)* is the equation specific to the Mexican Hat wavelet. The wavelet coefficients $c_{a,b}$ can be found using the wavelet transform which performs a correlation of the specific wavelet with the signal to determine its coefficient as shown in *Equation (2.6)*, where * denotes the complex conjugation.

$$E(t) = k \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \frac{c_{a,b}}{a^2} \psi_{a,b}(t) da db \tag{2.3}$$

$$\psi(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \tag{2.4}$$

$$\psi(t) = (1 - t^2) e^{-(1/2)t^2} \tag{2.5}$$

$$c_{a,b} = \int_{-\infty}^{+\infty} E(t) \psi^*_{a,b}(t) dt \tag{2.6}$$

Time-domain wavelets are simple, oscillating, amplitude functions of time [13]. So are the sine and cosine waves of Fourier analysis. However, unlike sine and cosine waves, which are precisely localized in frequency but extend infinitely in time (sines and cosines have definite single frequencies, e.g., 50Hz, constant for all time), wavelets are relatively localized in both time and frequency (compactly supported). A wavelet should have [13]:

(a) a zero mean amplitude,

(b) finite energy over its time course, and

(c) relatively little lower frequency energy compared with its higher frequency energy.

The wavelet shape can be chosen to fit the characteristics of the signal being analyzed and there are various types of wavelets such as the Haar wavelet, the Daubechies D4 wavelet and the Coifman 30 Wavelet [13; 14]. Wavelets can be used to analyse signal stationarity as they can be used to determine how the different frequency components of a signal dynamically behave. As such they have been used to analyze heart rate variability [15]. Additionally wavelets have been used for noise filtering, preprocessing neuroelectric data for input to neural networks, neuroelectric waveform compression, spike and transient detection, component and event detection, and time-scale/space-scale analysis of neuroelectric EEG waveforms. The Time-scale and space-scale analysis make them well suited for stationarity analysis as they enable the precise study of the small-scale and large-scale dynamical behavior [13].

2.2.4 The Discrete Wavelet Transform

The discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translations. They obey some defined rules [13; 14; 16]. Several different wavelets are created with the same shape but that are of different scales and translations. This is in order to create a mutually orthogonal set of wavelets. They are mutually orthogonal in that, though they are the same shape, the result of the different scales and translations is that none of the wavelets in the orthogonal set correlate with each other at all [13]. The correlation of each of these wavelets with the time signal gives an indication of how the different frequency components of the signal vary with time. Variations in the different frequency components are an indication of non-stationarity. Each scale, a, corresponds to a wavelet with a specific time window and a specific amplitude, and b is the value by which each wavelet is translated. Any time-signal $E(t)$ can be represented as a summation of scaled and translated wavelets as shown in *Equation (2.7)* where ψ(t) is the equation for the wavelet shape with $\psi(t) = 2^{a/2}\psi\big(2^a t - b\big)$ [13; 14; 16]. Once more the wavelet transform can be used to determine the wavelet coefficients $c_{a,b}$ of each of the wavelets in the orthogonal set applied *Equation (2.8)* [13; 14; 16].

$$E(t) = \sum_b \sum_a c_{a,b}\psi_{a,b}(t) \tag{2.7}$$

$$c_{a,b} = \int E(t)\psi_{a,b}(t)dt \tag{2.8}$$

The benefit of the discrete wavelet transform is that the orthogonal set of wavelets chosen is only a small subset of the wavelets that would have been used in the continuous wavelet transform. Because of this, much redundant signal information is done away with, greatly speeding up computation and reducing its complexity [13]. It has also been shown that, though the set of wavelets being used in the decomposition and analysis of the signal is greatly reduced, none of the signal characteristics are lost [13]. However, the ability to do scale and time independent pattern recognition, by simply looking for invariant patterns in the distribution of DWT coefficients, is lost. This makes the discrete wavelet transform less suited to Image Analysis than the continuous wavelet transform [13].

2.2.5 The Discrete Gabor Transform

The Discrete Gabor Transform is similar to the Continuous Wavelet Transform in that it can be used for time-frequency domain analysis. It does this by decomposing a time signal into the summation of scaled Gabor Wavelets [17; 18; 19; 20]. For a discrete-time signal φ(n) [17; 18; 19; 20]:

$$\varphi(n) = \sum_{m} \sum_{k=<K>} a_{m,k} g\big(n - mN\big) e^{j2\pi kn/K} \tag{2.9}$$

Thus the time signal $\varphi(n)$ can be decomposed into the scaled sum of windowed sinusoids. The $g(n)$ sequence is known as the synthesis window or the elementary signal, m represents the amount by which each window is translated, and $a_{m,k}$ is the coefficient for each window. In the above equation coefficients $a_{m,k}$ and $e^{j2\pi kn/K}$ are periodic with a period of K. In addition $k =< K >$ corresponds to a finite interval of K successive integers [19]. The Gabor transform is used to find the Gabor coefficients $a_{m,k}$ as shown in *Equation (2.10).* The $w(n)$ sequence is known as the analysis window and is a summation over all n integers [19].

$$a_{m,k} = \sum_{n} \varphi(n) w^*(n - mN) e^{-j2\pi kn/K} \tag{2.10}$$

Gabor transforms can effectively be used in the analysis of stationarity as they can be used for the characterization and estimation of non-stationary signals. They can also be used on non-stationary signals for time varying filtering [21] and they have been used in image processing [22] and signal processing [21]. As with STFTs and PSDs continuous wavelet transforms, discrete wavelet transforms and Gabor transforms require a priori information about the signal to determine what thresholds must be crossed for a signal to be classified as nonstationary. There is also no clear way of quantifying the level of nonstationarity present. With any two given signals there is no mapping that allows the general comparison of one signal to another to determine how much more nonstationary one is to another.

2.3 Phase Space Based Methods

Most effective methods of nonstationarity detection involve the quantification of nearest neighbours in phase space (state space, Poincare recurrence points) [9; 23; 24]. A time series is mapped into phase space by time delay embedding. Here phase space is an ideal space in which the coordinate dimensions represent the variables that are required to describe a system. The variables are the time delayed values of the time series. Given a scalar time series $\{x(i), i = 1, 2, \ldots\ N_x)$ where N_x is the total number of samples, vectors are constructed of the form $\vec{x_i} = [x(i), x(i+L), \ldots, x\big(i+(m-1)L\big)]$, with m being the embedding dimension and L being the time delay. Thus $\{\vec{x_i}, i = 1, 2, \ldots, N_{\vec{x}}\}$ represents a trajectory in m-dimensional space [23], where $N_{\vec{x}}$ is the total number of vectors. For example in three dimensional phase space with a time delay of $L = 1$ the time series $x(i)$ gets mapped from $\{x(i), i = 1, 2, \ldots\}$ to $\vec{x_i} = [x(i), x(i + 1), x(i + 2)]$ a series of vectors known as the trajectory - in three dimensional phase space - of the time series $x(i)$. *Figure 2.1* gives an example of the Henon Map time series reconstructed into a three dimensional phase space trajectory ($m = 3$) with a time delay of $L = 1$.

A close view of the Henon signal is shown in *Figure 2.1.* The first three points are taken as the x, y and z coordinates of the phase space vector respectively. These points can be seen as the three highlighted dots in *Figure 2.1.* Embedding in $m = 3$ dimensions means that points are grouped together in groups of three. A time delay of $L = 1$ means that the points will follow successively in the time series [9; 23; 24]. The next phase space vector is now taken by starting with the $2nd$ point of the time series. The $2nd$ point is now the x coordinate and the following becomes the y coordinate and the one that follows that becomes the z coordinate as shown in *Figure 2.2.* In this way the phase space vectors are formed. With the $m \ll N_x$ where m is the embedding dimension and N_x is the number of points in the time series there will be almost as many vectors in phase space as the data points of the

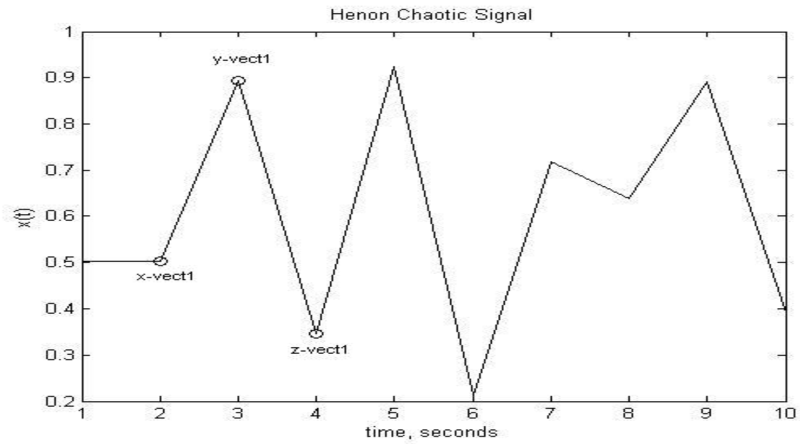


Figure 2.1: Henon signal coordinates for 1^{st} phase vector

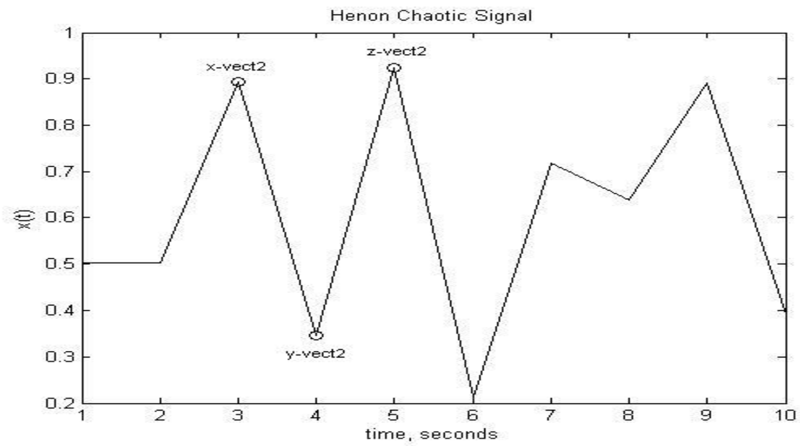


Figure 2.2: Henon signal coordinates for 2^{nd} phase vector

time series $N_x \approx N_{\vec{x}}$. In the example with $m = 3$ and $L = 1$ there will be two less vectors in phase space than the data points of the time series.

Figure 2.4 shows the phase space vectors as they are formed. *Figure 2.4(a)* shows the phase space representation with the first 10 vectors formed. *Figure 2.4(b)* shows the first 100 vectors formed, and finally *Figure 2.4(c)* shows the first 1000 vectors formed. The trajectory has began to fill out and a consistent recurrent pattern can be seen. For stationary time series these trajectories tend to be consistent, and the vectors remain in the same general region of phase space.

For any given point in phase space (where $\vec{X}$ is the vector from the origin to that point) the nearest neighbours are all the points that fall such that $\|\vec{X} - \vec{x(i)}\| < \varepsilon$, where ε is some minimal tolerance distance as shown in *Figure 2.6*. In the figure a

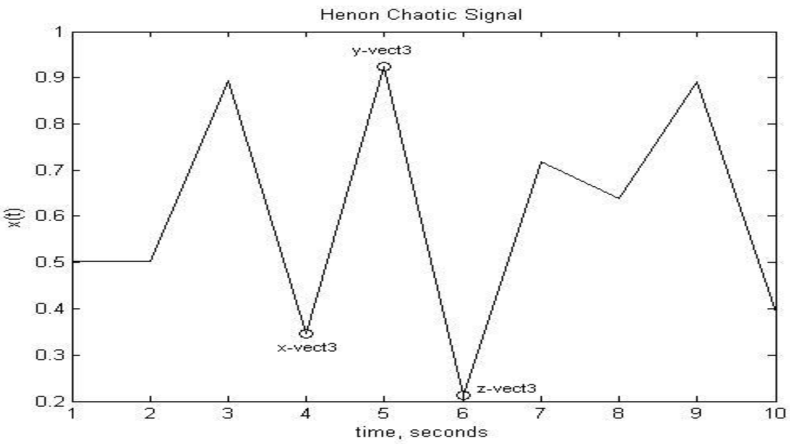
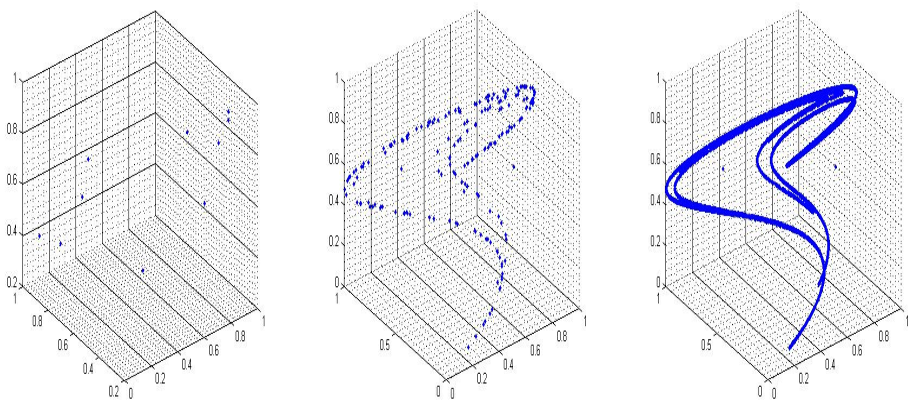


Figure 2.3: Henon signal coordinates for 3^{rd} phase vector



(a) First 10 vectors in phase space

(b) First 100 vectors in phase space

(c) First 1000 vectors in phase space

Figure 2.4: Phase plot of Henon map

phase space plot of the Henon map is shown for the first few vectors. The circle in the figure represents the sphere defined by the error tolerance. The vector pointed to in the diagram is the reference vector. Any vector that falls within a sphere that has the reference point as its center and a radius of ε is considered a near neighbour. Several methods make use of the above as a basis for nonstationarity detection and they shall be briefly discussed in the following paragraphs.

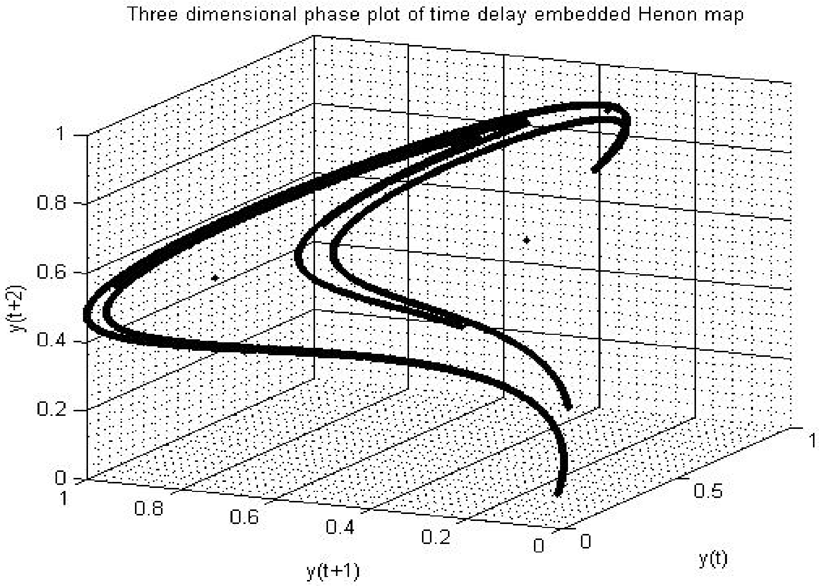


Figure 2.5: Full phase plot of the Henon signal

2.3.1 Recurrence quantification analysis

Recurrence quantification analysis has been used [25] to employ recurrence plots for the detection of nonstationarity. A recurrence plot is based on the phase space reconstructions of time series. If a time series reconstructed into phase space is represented by $N_{\vec{x}}$ vectors the recurrence plot is an $N_{\vec{x}}$ by $N_{\vec{x}}$ plot. A point is placed on the plot at position (a, b) if the vectors a and b of the phase space reconstruction fall within an error tolerance ε of each other. Recurrence quantification analysis combines recurrence plots and principal component analysis to make probabilistic evaluations. Metadynamical recurrence plots have been used for the detection of nonstationarity [1] and involve the analysis of the recurrence of metastates. Stationarity is seen by consistent patterns in the plots. This consistency shows that the plots are displaying consistent behavior.

2.3.2 Recurrence Time Analysis

Recurrence time analysis techniques [23; 24] are an extension of recurrence quantification analysis. They deal with the recurrence times of vectors in phase space. This method for the detection of nonstationarity was developed by Gao [23]. It is based on the recurrence times of recurrence points of the second type (true recurrence points). Nearest neighbours or recurrence points fall into two categories: sojourn

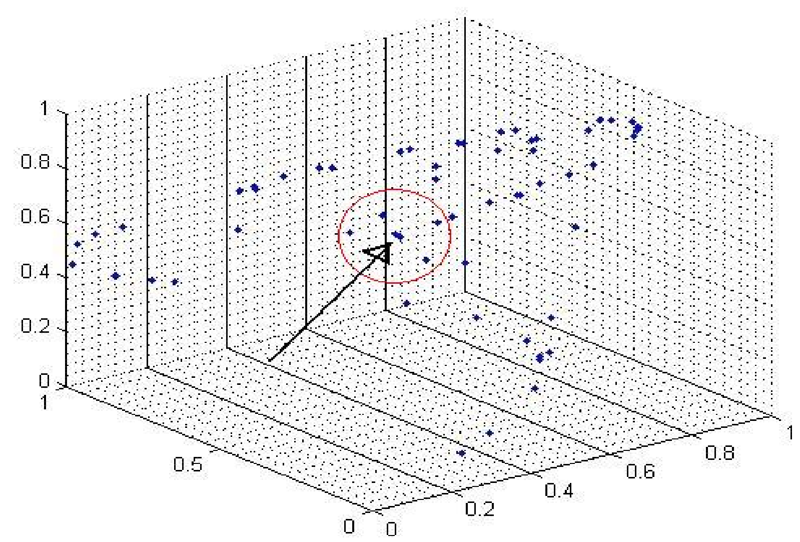


Figure 2.6: Nearest Neighbors plot

points (false recurrence points or recurrence points of the first type - these are the result of high autocorrelation in the signal) and true recurrence points (recurrence points of the second type) [23; 24; 26]. The analysis of true recurrence points has been found to yield more robust representations of the dynamics of the signal being considered and recurrence plots have been developed that only consider true recurrence points [26]. The method is founded on the observation that nonstationarity causes the average recurrence times of the second type to change with time.

The test is executed as follows: Given the phase space reconstruction of a time series and a fixed error tolerance, all the true recurrence points for each point in phase space are determined. Then the recurrence times for each of these phase space points are calculated such that, $T_2(j) = t_{j+1} - t_{j,j} = 1, 2, \ldots N_x$, where N_x is the total number of points in the data set. These recurrence times are essentially the times between successive recurrences to a point in phase space where j denotes the jth return to the reference point. Since the trajectory in phase space may visit some regions more than others the dependency of visiting frequency on phase space location is removed by normalizing $T_2(j)$ by its mean for every reference point. The normalized $T_2(j)$ are then grouped together according to j and the mean $\overline{T_2}(j)$ is computed for each group. Now j represents the number of recurrences to any point in phase space and $\overline{T_2}(j)$ represents the consistency of recurrence times to that point. $\overline{T_2}(j)$ does not vary with j for stationary time series, remaining almost constant at a value of one. In other words if a time series is stationary then the mean time of recurrence to any given point remains the same. $\overline{T_2}(j)$ varies with j for nonstationary time series.

Each point that is considered in phase space is considered to be first in the time series. The phase space plot is looped in such a way that it begins with the point in consideration and ends with it. Thus the reference point $x(ref)$ becomes the point at which time $t = 0$, the following points $x(ref + i), i = 1, 2, 3 \ldots N_x$ become the points at which time $t(ref + i) = t(i), i = 1, 2, 3 \ldots$. In this way the whole data set is considered when searching for the recurrence points to any point in phase space allowing for a better calculation of the mean recurrence time to that point. It is interesting to note that the above method does not require the time series to be broken up into segments.

2.3.3 Nonstationarity Detection by the Loss of Recurrence

This method was proposed by Rieke et al [27]. The variance of mean recurrence times from frequency distributions yielded by stationary conditions gives an indication of the overall stationarity of the signal. The previous method uses a normalization technique that causes the recurrence times to fall at a value of 1 if they are consistent. This method develops frequency distributions that represent how often each point in phase space recurs. Frequency distributions obtained from different sections of the signal are compared and variations in this distribution indicate nonstationarity. A priori knowledge of the frequency distribution expected under stationary conditions is necessary.

2.3.4 Space Time Separation Plots

Further techniques of nonstationarity detection include space time separation plots [28]. These are plots of the separation in time versus the separation in phase space of the phase space vectors of a time series. Analysis of this relationship gives a reflection of the consistency of the dynamics of the time signal. Consistency in the relationship between the separation in phase space and the separation in time of the vectors is an indication of stationarity. Yu et al. similarly makes use of the probability distributions of space time separation plots [7] to analyse stationarity. A very similar method is proposed by Kennel [9]. It is a statistical test using the distribution of points in reconstructed phase space for the analysis of stationarity. The main difference between this method and the one proposed by [7] is the definition of the time index used. The time index defined in [7] results in the density distributions as a function of time index being horizontal for stationary series. The time index defined in [9] results in the density distributions sloping triangularly downwards for stationary time series. Both tests simply give an indication of whether nonstationary behavior is present or not, but a clear way of quantifying the exact extent to which it is present is not possible.

2.3.5 Nonlinear Cross Prediction Analysis

The following is a test for nonstationarity developed by Schreiber [3], and is based on the similarities between different parts of the time series. It checks the extent that one segment of a time series can be predicted using another segment as a database. The cross prediction error is defined as $\gamma_{x,y}$ and is used as a statistic to compare segments. Take a segment of a time series x_i with a phase space reconstruction $\{\vec{x_i}, i = 1, 2, \ldots, N_{\vec{x}}\}$ where $N_{\vec{x}}$ is the total number of vectors in phase space. In addition take another segment of the same time series y_i with a phase space reconstruction $\{\vec{y_i}, i = 1, 2, \ldots, N_{\vec{y}}\}$. For each $\vec{y_i}$ a prediction is made one step into the future. Therefore given the phase space vector $\vec{y_i} = [y_{i-m+1}, \ldots, y_n]$ where m is the embedding dimension, an estimation for the actual time series value y_{i+1} is made. The dynamics relating $\vec{x_i}$ to x_i are used to make the approximation *Equation (2.11)*.

$$\vec{y}^{x}_{i+1} = \frac{1}{|\bigcup^{x}_{\varepsilon}(\vec{y_i})|} \sum_{\vec{x}_i \in \bigcup^{x}_{\varepsilon}(\vec{y_i})} x_{i+1} \tag{2.11}$$

Where $\bigcup^{x}_{\varepsilon}(\vec{y_i}) = \vec{x_i} : \|\vec{x_i} - \vec{y_i}\| < \varepsilon$ and ε is the error tolerance in phase space, a neighborhood region of $\vec{y_i}$. The algorithm basically takes both segments x and y of the time series and transforms them into phase space. It then takes a vector $\vec{y_i}$ in phase space and finds all the vectors of $\vec{x_i}$ that are within a neighborhood ε of $\vec{y_i}$. It then takes the values x_{i+1} in time of x_i that immediately follow the phase space vectors found near $\vec{y_i}$. These time series values are then averaged to find the approximation of y_{i+1}. The cross prediction error is then given by *Equation (2.12)* [3] where $\vec{y}^{x}_{i+1}$ is the value that is predicted to follow $\vec{y_i}$ and y_{i+1} is the actual value. Inconsistency in this error is an indication of nonstationary behavior.

$$\gamma(x, y) = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left(\vec{y}^{x}_{i+1} - y_{i+1}\right)^2} \tag{2.12}$$

2.3.6 Additional Methods of Analyzing Non-Stationarity

Conforto and D'Alessio [29] developed a parametric method for the spectral estimation of non-stationary signals, also comparing their method to previously existing methods including the Short-Time Fourier Transform. They found their method to show improved frequency and amplitude estimation. Mukhopadhyay and Sircar [30] proposed a parameter estimation method for non-stationary signals using an ARMA model for modeling non-stationary signals.

In other work, Sircar and Syali [31] used the weighted sums of complex amplitude modulated signals to model non-stationary signals. Tambouratzis and Antonopoulos-Domis applied artificial intelligence neural networks to estimate on-line transit time. They observed the transit time for both constant (stationary) and Oscillating (non-stationary) flow regimes and proposed a non-stationary correlator for on-line transit time estimation [32].

An interpolation technique, that allows very accurate measurements of the instantaneous frequencies and amplitudes of non-stationary signals that have stationary components was proposed by Andria et al [33]. Nandi and Kundu [34] used a generalized multiple fundamental frequency model for the analysis of non-stationary signals, and Burri et al. [15] proposed a methodology to use wavelet transforms for the analysis of heart rate variability (stationarity analysis.)

With all of the stated methods there is no mapping that allows the direct comparison of the levels of nonstationarity of any two given signals. The tests, in general, are solely for the detection of nonstationarity, in order to determine whether it is present or not.

2.4 Conclusion

This chapter gives a background of all the methods of nonstationarity detection reviewed in the literature except those based on the correlation integral. Statistical and frequency based methods are presented as well as phase space methods. **Chapter 3:** presents the correlation integral (autocorrelation integral) and the cross correlation integral. Methods of nonstationarity detection, based on the correlation integral, reviewed in literature are then presented. The application of the cross correlation integral in the development of the Stationarity Index test is then discussed in detail. This is followed by the methodology applied.

Chapter 3

The Cross Correlation Integral

3.1 The Cross Correlation Integral

The correlation integral is a tool widely used in the analysis of time series. It is a phase space based analysis technique and it is widely used in experimental and numerical investigation due to its ease of computation [35]. It has, in the last decade or so, been used for a wide variety of applications such as the analysis of data from disciplines such as medicine, to astrophysics and finance [36; 37; 38]. It has been used for the detection of nonstationarity in time series [1; 39]. Time lags are used to reconstruct the time series in phase space. The coordinates of each vector in phase space are the time lags of the actual time series i.e. $\vec{x_i} = [x(i), x(i+L), \dots, x\big(i+(m-1)L\big)]$. The auto correlation integral is defined as follows [1; 39]:

$$\begin{aligned} C_m = P\Big(\|\vec{x}_i^m - \vec{x}_j^m\| < \varepsilon\Big) = \\ \frac{1}{N^2}\sum_{i,j=1}^{N} \theta\Big(\varepsilon - \|\vec{x}_i^m - \vec{x}_j^m\|\Big) \end{aligned} \tag{3.1}$$

where $P\Big(\|\vec{x}_i^m - \vec{x}_j^m\| < \varepsilon\Big)$ is the probability that $\|\vec{x}_i^m - \vec{x}_j^m\| < \varepsilon$ and θ is the Heaviside step function. The Heaviside step function, θ, equates $\Big(\varepsilon - \|\vec{x}^m(i) - \vec{x}^m(j)\|\Big)$ to one if it is positive and to zero if it is negative. The auto correlation integral represents the probability of finding any two points in the phase space reconstruction that are near neighbours or within a radius ε of each other. The function basically counts up by a single digit every time two points fall within a distance ε of each other and divides by the total number of possible combinations N^2. For two separate time series $x(i)$ and $y(i)$ the cross correlation integral is defined as follows [1; 39]:

$$C_m(x,y) = P\Big(\|\vec{x}_i^m - \vec{y}_j^m\| < \varepsilon\Big) = \frac{1}{N^2}\sum_{i,j=1}^{N}\theta\Big(\varepsilon - \|\vec{x}_i^m - \vec{y}_j^m\Big) \tag{3.2}$$

It represents the probability of finding points in the phase space reconstruction of x that are closer to the points in the reconstruction of y than a distance of ε. It follows that if any two signals have the same underlying dynamics then the values of $C_m(x,x)$, $C_m(x,y)$, and $C_m(y,y)$ should approximately be the same for any dimension m *Equation (3.3) Equation (3.4)* [1]. To follow are current methods of nonstationarity detection based on the correlation integral (auto-correlation integral) and the cross correlation integral.

$$C_m(x,x) - C_m(x,y) \cong 0 \tag{3.3}$$

$$C_m(y,y) - C_m(x,y) \cong 0 \tag{3.4}$$

3.2 Current Nonstationarity Detection Methods based on the Cross Correlation Integral

3.2.1 Cross Correlation Sum Analysis

Kantz uses cross correlation sum analysis to evaluate fractal measures. The cross correlation integral is used to determine the correlation dimension which is a measure of the fractal dimension [40; 41; 42]. The method allows for the comparison of the overlap in phase space trajectories and presents a criterion to determine the precision that two sets are the realization of the same measure. If two sections of the same time series can be considered as two sets that are realizations of a different measure the time series can be considered nonstationary.

3.2.2 Fel'dshtein's Application of the Cross Correlation Integral

Fel'dshtein [39] applies the cross correlation integral for the detection of nonstationarity. A quantity is defined as shown in *Equation (3.5)* [39]:

$$\Delta_{xy} = |lnC_{xy}(\varepsilon) - lnC_{yy}(\varepsilon)| \tag{3.5}$$

where $C_{xy}(\varepsilon)$ is the cross correlation integral of the section x of the time series with section y for a given error tolerance ε. $C_{yy}(\varepsilon)$ is the autocorrelation of section y of the time series for a given error tolerance ε. Another quantity is defined *Equation (3.6)* [39]. A time series is deemed stationary if δ_{xy} is nonzero [39].

$$\delta_{xy} = \int_{\varepsilon_{min}}^{\varepsilon_{max}} \Delta_{xy}(\varepsilon) d\varepsilon \tag{3.6}$$

3.3 The Development of a Nonstationarity Detection Technique based on the Cross Correlation Integral

A test is developed for the detection of nonstationarity (based on *Equation (3.3)* and *Equation (3.4)*) for any two segments of the time series that are long enough to contain the general dynamics of the signal. The equations are used to derive the stationarity index and this index is used to analyse the stationarity of a given time series. The test is based on the notion that a stationary time series should have the same underlying dynamics and thus cross correlations of any two sections of the series should be approximately the same. There is great power in a test such as this for the quantification of stationarity. Firstly time delay embedding gives a representation of the joint probability distribution of successive points in the time series, and secondly the cross correlation integral gives a measure of the probability that points in the phase space trajectory of a time series fall near each other. This gives a measure of the similarities between the two trajectories. Therefore combined, the two give a representation of similarities of the joint probability distributions of the time series from which the phase space trajectories were constructed. With time delay embedding with a dimension of $m = 3$ and a time delay of $L = 1$, the phase space plot is a representation of the joint probability distribution of $x(i), x(i+1)$ and $x(i+2)$. Therefore the cross correlation integral of different sections of a time series, by virtue of the fact that it is a representation of the probability that any points in the trajectories are nearest neighbours, is also a representation of the extent to which the joint probability distribution of $x(i), x(i+1)$ and $x(i+2)$ in the time series remains consistent. This allows for the development of a test for nonstationarity based on the fundamental definition of stationarity.

3.3.1 The Theiler Window

For a finite sequence of time series data where $N_x < 2\tau^{m/2}$, N_x being the number of points in the sequence, τ being the autocorrelation time (defined further in this passage) and m being the embedding dimension, it is found that anomalous structure is exhibited in its correlation integral [43]. For both stochastic and dynamical data with limited data sets with high autocorrelations, the correlation integral exhibits anomalies that inhibit it from being calculated accurately [43]. If N_x is large enough or the autocorrelation α is very small then the effect of the autocorrelation is negligible, however if N_x is not sufficiently large or α is very close to 1 the effect of the autocorrelation cannot be ignored and the correlation integral is no longer an accurate representation of the dynamics of the signal. With a high autocorrelation there is a strong interdependency between the recorded samples. An example would be sampling changes in temperature at a rate of 1000 times per second. Because the temperature would change slowly compared to the sampling rate, there would be many repeated readings, effectively like taking the same reading and repeating it several times. No new information would be added. For such a time series it would appear that there are many vectors in phase space that are near each other but they are effectively the same vector repeated multiple times. The solution would therefore be to only consider points or vectors in phase space that are far enough from each other that they are effectively independent vectors e.g. sampling the temperature at a rate of once every minute. The vectors considered, should however, not be so far from each other that information on the dynamics is lost. Ideally, for the accurate calculation of the correlation integral, the time series should have a zero autocorrelation. The negative effects of high autocorrelation in the samples, on the correlation integral of finite data windows, are dealt with by using a Theiler window. In other words, the Theiler window is a certain number of vectors that should be skipped in the calculation of the cross correlation integral. This is to insure that these vectors are not nearest neighbours in phase space simply because they are formed by data points in the time series that are close in time. This minimises the effect that high autocorrelation in the data has on the calculation of the cross correlation integral. To illustrate this, the correlation integral is rewritten as shown in *Equation (3.7)* [43]. It is generalized a step further to give *Equation (3.8)* [43]:

$$C(\varepsilon, N_{\vec{x}}) = \frac{2}{N_{\vec{x}}^2} \sum_{n=1}^{N_{\vec{x}}} \sum_{i=1}^{N_{\vec{x}}-n} \theta\Big(\varepsilon - \|\vec{x}_{i+n} - \vec{x}_i\Big) \tag{3.7}$$

$$C(\varepsilon, N_{\vec{x}}, W) = \frac{2}{N_{\vec{x}}^2} \sum_{n=W}^{N_{\vec{x}}} \sum_{i=1}^{N_{\vec{x}}-n} \theta\Big(\varepsilon - \|\vec{x}_{i+n} - \vec{x}_i\Big) \tag{3.8}$$

W is the Theiler window or the number of vectors (from the reference vector) to be skipped before vectors are considered in the calculation of the cross correlation integral. In the generalized correlation integral *Equation (3.8)* the ith and $ith + n$ vectors of $\vec{x}$ are considered, yet the first pair of vectors to be considered are W vectors from each other. In the original correlation integral *Equation (3.7)* the vectors $\vec{x}_i, \vec{x}_{i+1}, \vec{x}_{i+2} \ldots \vec{x}_{i+(N_{\vec{x}}-1)}$ are all considered. In *Equation (3.8)* the first $W - 1$ vectors are skipped. This window of vectors that are not considered is known as the Theiler window [43]. Theiler carried out investigations into the correlation integral and the following is a summary of the proof demonstrating the anomalous behavior. Though it is carried out on stochastic data it was found to be a general feature of autocorrelated data and was seen in dynamical data as well [43].

Given a finite data set N_x of autocorrelated Gaussian noise with mean $\mu = 0$, variance σ^2 and autocorrelation α, then $\langle x_i \rangle = \mu = 0$ and $\frac{\langle x_{i+n} . x_i \rangle}{\sigma^2} = \alpha^2$. For $\alpha < 1$ the autocorrelation time τ, with the sample time $\Delta t = 1$, is given by $\tau \equiv \frac{1}{ln(1/\alpha)}$. $\alpha \approx 1$ yields $\tau \approx \frac{1}{1-\alpha}$. The autocorrelation time τ is the characteristic decay time of the autocorrelation of a point in the time series to the points that follow. It is safe to assume that any two samples that have a time duration of $\geq \tau$ between them are are independent of each other.

Statistical methods are used on the stochastic data to obtain an expression for the correlation integral $C(\varepsilon, N_{\vec{x}}, W)$ which is now a function of the error tolerance ε, the length of the segment in phase space $N_{\vec{x}}$ and the Theiler window W. The L_∞ metric is used for convenience but numerical experiments with other metrics such as Euclidean (L_2) and taxicab (L_1) yielded similar results [43]. The L_∞ metric specifies that the distance between two vectors is the maximum of the differences between the components of those two vectors. *Equation (3.8)* can be rewritten as follows, replacing the Heaviside step function with its expected value [43]:

$$C(\varepsilon, N_{\vec{x}}, W) = \frac{2}{N_{\vec{x}}^2} \sum_{n=W}^{N_{\vec{x}}} (N_{\vec{x}} - n) P(|\vec{x}_{i+n} - \vec{x}_i| \leq \varepsilon) \tag{3.9}$$

Where $P(|\vec{x}_{i+n} - \vec{x}_i| \leq \varepsilon)$ is the probability that $|\vec{x}_{i+n} - \vec{x}_i| \leq \varepsilon$. Therefore given the L_∞ metric [43]:

$$P(|\vec{x}_{i+n} - \vec{x}_i| \leq \varepsilon) = P(|x_{i+n+k} - x_{i+k}| \leq \varepsilon \quad \text{for all} \quad 0 \leq k < m) \tag{3.10}$$

$$= \prod_{0 \leq k < m} P(|x_{i+n+k} - x_{i+k}| \leq \varepsilon) \tag{3.11}$$

$$= [P(|x_{i+n} - x_i| \leq \varepsilon)]^m \tag{3.12}$$

Where m is the embedding dimension.

$$P(x_{i+n} = a, x_i = b) = Ae^{-(a^2 - 2\alpha^n ab + b^2)/B} \tag{3.13}$$

Equation (3.13) is the joint probability distribution for correlated Gaussian variables. The constants $B = 2\sigma^2(1 - \alpha^{2n})$ and $A = \frac{1}{2\pi\sigma^2(1-\alpha^{2n})^{1/2}}$. It follows that [43]:

$$P(|x_{i+n} - x_i| \leq \varepsilon) = A \int_{-\infty}^{\infty} \int_{a-\varepsilon}^{a+\varepsilon} e^{-(a^2 - 2\alpha^n ab + b^2)/B} dbda \tag{3.14}$$

$$= erf\Big(\frac{\varepsilon}{2\sigma(1 - \alpha^n)^{1/2}}\Big) \tag{3.15}$$

where erf is the error function. The above result is substituted into *Equation (3.9)* yielding the following which is the basis of the analysis:

$$C(\varepsilon, N_{\vec{x}}, W) = \frac{2}{N_{\vec{x}}^2} \sum_{n=W}^{N_{\vec{x}}} (N_{\vec{x}} - n)\Bigg(erf\Big(\frac{\varepsilon}{2\sigma(1 - \alpha^n)^{1/2}}\Big)\Bigg)^m \tag{3.16}$$

As $\alpha \to 0$ *Equation (3.16)* reduces to:

$$C(\varepsilon, N_{\vec{x}}, W) = \frac{2}{N_{\vec{x}}^2} \sum_{n=W}^{N_{\vec{x}}} (N_{\vec{x}} - n)\Big(erf(\frac{\varepsilon}{2\sigma})\Big)^m \tag{3.17}$$

For $x \ll 1$ $erf(x) \propto x$ and saturates at unity for $x \gg 1$. Therefore for $\varepsilon \ll 2\sigma$:

$$C(\varepsilon, N_{\vec{x}}, W) = \frac{2}{N_{\vec{x}}^2} \sum_{n=W}^{N_{\vec{x}}} (N_{\vec{x}} - n)\Big(\frac{\varepsilon}{2\sigma}\Big)^m \tag{3.18}$$

The sum will consist of error functions with arguments very near $\varepsilon/2\sigma$. The above is the case for which $\alpha \to 0$ in which the correlation integral exhibits no anomalies.

Equation (3.16) is not so readily simplified with $\alpha \to 1$ or with $N_{\vec{x}}$ not sufficiently large. For a small delay n between considered samples the argument of the error function is noticeably larger than $\varepsilon/2\sigma$ and $1-\alpha^n$ is noticeably less than unity. This effect is magnified with greater embedding dimension. $\varepsilon/2\sigma$ is a good approximation to the argument of the error function for most terms (those with large delay n from the reference vector), as in the case in which the autocorrelation of the time series $\alpha \to 0$. However there are a few terms with a small delay n from the reference vector, for which this is not the case and these few can actually dominate the sum. The effect is that the values of the correlation integral will be distinctly different from the values obtained in ideal cases with zero autocorrelation in the time series. For $n = W$ the first term of the right hand side of *Equation (3.16)* is [43]:

$$\frac{2}{N_{\vec{x}}^2}(N_{\vec{x}} - W)\left(erf(\frac{\varepsilon}{2\sigma(1-\alpha^W)^{1/2}})\right) \tag{3.19}$$

For large autocorrelation α this first term can be large and dominate the sum. It is found [43] that for an error tolerance of $\varepsilon > \sigma\sqrt{\pi}(2/N_x)^{1/m}$ the first term losses its significance and the correlation integral begins to look like that corresponding to the $alpha \to 0$ or $N_x \to \infty$ limits. Theiler [43] also proposes an algorithm to 'toss out those over contributing early terms'. He achieves this by having a window of $W > \tau.ln(m/2)$ where τ is the autocorrelation time and m is the embedding dimension. Therefore there is a delay window and any vectors in this window are not considered in the correlation integral. He points out that as long as $W > \tau.ln(m/2)$, where τ is the autocorrelation time of the time series, the exact choice of W is not important. These results are used to develop algorithms that chose the error tolerance ε and the Theiler window W based on the autocorrelation of the time series itself. This is all to enable the correlation function to count only the ‘accidentally’ close pairs of vectors and not those close to each other because of a high autocorrelation, i.e. not vectors close in phase space simply because they are close in time.

3.3.2 The Stationarity Index

The stationarity index (SI) is proposed here for the development of a nonstationarity detection technique. It is the basis of the test and is defined as follows:

$$SI = \frac{C_m(x,x) - C_m(x,y)}{C_m(x,x) + C_m(x,y)} 100. \tag{3.20}$$

This index is an indication of the nonstationarity of a signal and serves as a means of quantifying the extent to which a signal is nonstationary. This is due to the fact that a stationary signal should have the same underlying dynamics. Therefore the values of $C_m(x,x)$ and $C_m(x,y)$ for any two sections, x and y, of this signal should be approximately equal. Differences would indicate nonstationarity and thus the defined index would increase the greater the differences in the underlying dynamics of sections x and y. The above defined index therefore ranges from values of 0 to 100, 100 being the case where the $C_m(x,y) = 0$, such as in cases where the two signals are so different that no points of the two signals, in phase space, fall within a distance ε of each other. Extremely stationary signals would give a stationarity index of close to zero where $C_m(x,x) - C_m(x,y) \cong 0$, whereas extremely nonstationary signals would give stationarity indices of close to 100. It is quite unique in that it maps the similarities in the dynamics of a signal onto a scale of 1 *to* 100. Therefore given the first window of a signal as a reference it compares the similarities of the dynamics of following windows of equal size and maps this behavior onto a scale of 0, for approximately identical behavior, to 100 for non-identical behavior. Thus, different signals of the same size and with the same window length can be directly compared to observe how the dynamics of each vary as compared to the others. Even with signals of different sizes the index should give some idea of how the changes in their dynamics compare to each other.

Now *Equation (3.20)* is basically a simple normalization of *Equation (3.3)*, but the normalization gives insight into the behavior of the signals. For example, given that $C_m(x,x) = 0.5$ (meaning that the probability that any two vectors of signal ’x’ are near neighbours is 0.5), $C_m(x,y)$ could be anything from $\leq C_m(x,x)$ to $\geq C_m(x,x)$.

This SI normalization scales the relative behavior of $C_m(x,x)$ and $C_m(x,y)$ to a range of $0-100$. Similarities are mapped to low indices and differences are mapped to high indices. Say for example that $C_m(x,x) = 0.9$ and $C_m(x,y) = 0.8$. In another example we could have $C_m(x,x) = 0.1$ and $C_m(x,y) = 0$. In both cases, $C_m(x,x) - C_m(x,y) = 0.1$ but this value doesn't give much information on the variance of the dynamics of the signal in either case. The SI in the first example however is $\frac{0.1}{1.7} \times 100 = 5.88\%$, and in the second example the SI is $\frac{0.1}{0.1} \times 100 = 100\%$. The fact that $C_m(x,y)$ is low doesn't tell us much about the variation in the dynamics of the signal until we see what $C_m(x,x)$ is. The SI gives an idea of how $C_m(x,y)$ is varying relative to $C_m(x,x)$, and thus gives insight into the consistency of the signals dynamics.

From a theoretical perspective, the advantage of this method over the previous methods based on the cross correlation integral is that this method provides a mapping of the stationarity of each signal. The ratio *Equation (3.20)* of the difference between $C_m(x,x)$ and $C_m(x,y)$ divided by their sum maps cases in which $C_m(x,x)$ and $C_m(x,y)$ are identical to zero and cases in which $C_m(x,y) = 0$ to one. Therefore for any two time series windows the index scales $C_m(x,y)$ in such a way that it falls between a range of $0 \rightarrow 1$. It is therefore a measure that simply maps similarities in time series windows: 'Identical dynamics' are mapped to zero and 'completely different dynamics' are mapped to 1. The multiplication by a factor of 100 allows the index to be viewed as a percentage. With the index it is therefore possible, not simply to detect nonstationarity, but to also have some measure of the extent to which a signal is nonstationary. It also allows for comparisons between signals to be made in a similar manner that percentages can be used to compare between measures. An example is two students both getting 80% in two different tests. The first student may have gotten $\frac{40}{50}$ and the second student may have gotten $\frac{72}{90}$. The measure simply gives an idea of the ratio between what they could answer correctly and what they were given to answer. The knowledge of such a ratio gives insight into patterns in measures that are useful. In a similar manner the SI index presents the potential of giving insight into patterns in the variances of dynamics that is useful.

The are several assumptions made with this method. The first is that for two time series of the same lengths, windowed in the same manner, the index is an accurate reflection of the how the dynamics of the signals are changing and when compared directly a higher index in one indicates greater changes in the dynamics. The second assumption made is that variations in the Theiler window will not affect the index to the extent that the comparison of the SI for two different signals would be completely inaccurate. The third assumption is that slight variations in the length of each signal will not affect the index to the extent that the comparison of the SI for two different signals would be completely inaccurate. For the purposes of this investigation the lengths of each of the signals compared are kept the same. Further testing is required to test these assumptions, however this testing is not covered in the scope of this dissertation and shall be carried out in future work.

3.3.3 Methodology

The following procedure was implemented on all signals tested. This was done by designing and applying several functions all coded in the matlab programming language. Each signal was normalized to fall between a range of $0 \rightarrow 1$ and then divided into 50 nonoverlapping segments (windows) equal in size. The number of windows is chosen to allow a suitable number of comparisons between windows as well as to allow each window to contain enough of the signal to adequately capture the dynamics that the signal displays in that region. Further testing would be necessary do determine what window sizes are optimal. The first segment was used to calculate the Theiler window and the error tolerance ε that would be used for each cross correlation integral. The stationarity index between the first segment and every other segment, including itself, was calculated. The result is a plot of the stationarity index with respect to the first nonoverlapping segment e.g. Figure 5.2. The second segment was taken as a reference and the stationarity index between this segment and every other segment, including itself, was calculated. The result is now a plot of the stationarity index with respect to the second segment. This process was repeated for every segment of the time series. The overall resultant plot is a function of two variables, i.e., the stationarity index as a function of the reference segment and the compared segments.

The stationarity index method was tested firstly on chaotic signals to test its ability to detect nonstationary behavior. The first two chaotic signals are known stationary signals, the Henon and the Lorenz chaotic signals. The third chaotic signal was a modified version of the Lorenz, modified in such a way that its behavior was nonstationary. The signals all consisted of $60,000$ data points and each segment was $1,200$ data points in length. The Lorenz signal was considerably more autocorrelated than the Henon. This means that there is more interdependency between the samples of the series and thus the Theiler window required is larger. Therefore less points are considered in the correlation integral and there would therefore be more variations in this figure.

The stationarity index was then tested on both electrocardiogram (ECG) and electroencephalogram (EEG) signals. ECGs are signals generated by the heart beat whereas EEGs are electrical signals generated by the brain. Both EEGs and ECG are affected by the onsets of seizures. EEGs more sensitively reflect changes that are the result of the onset of a seizure than ECGs do. ECG dynamics may change with the onset of a partial seizure but they may not immediately return to normal behavior once the seizure stops. In general for ECGs, the nonstationarities introduced by the partial seizures may not follow, in synchrony, their duration or intensity. The changes in dynamics can be caused by changes in the rate and/or the regularity with which the heart beats and it may take a while for the heart to return to 'normal' behavior after the occurrence of a partial seizure. For the purposes of this research this relationship is adequate as the main aim is to determine if the index can pick up the differences in the dynamics.

The data used was taken from the Physionet archives [44; 45] from a study done on partial epilepsy. The data is taken from five patients that had multiple seizures [44]. Five minute sections of each ECG signal were selected that included the event of the partial seizure. The signals were sampled at a rate of 200 times a second resulting in $60,000$ samples to cover the five minute period. The five minute section was divided into 50 windows each covering a duration of six seconds, and consisting of 1200 samples.

3.4 Conclusion

This chapter presented the correlation integral (autocorrelation integral) and the cross correlation integral. The methods of nonstationarity detection (based on the correlation integral) reviewed in literature were presented. The developed index for the detection of nonstationarity was presented as well as the methodology used to implement it. It was shown that this index is uniquely used to map similarities and differences, between time windows, to a scale of $0 \rightarrow 1$ and multiplying this by 100 allows the index to be viewed as a percentage. It is argued that the index has the ability, not just to detect stationarity but also allows for the quantification thereof. The manner of quantification allows for the comparison of nonstationary behavior in signals in a way that is novel. Assumptions that are made regarding the test are stated. **Chapter 4:** presents the results of Stationarity Index tests on chaotic signals. The stationary behavior of the signals tested on is already known. The purpose is to assess the ability of the test to discriminate between nonstationary and stationary time series.

Chapter 4

Results I

4.1 Results of Nonstationarity Tests on Chaotic Signals

The stationarity index method was tested firstly on chaotic signals to test its ability to detect nonstationary behavior. The first two chaotic signals are known stationary signals. They were the Henon chaotic signal *Equation (4.1)*, *Figure 4.1* [7], and the Lorenz chaotic signal *Equation (4.2) Figure 4.4* [23; 24; 39]. The third chaotic signal was a modified version of *Equation (4.2)* in which the variables r, σ and β were incremented with each iteration to create a nonstationary signal *Figure 4.5*.

$$x_{n+2} = 1 - \alpha x2_{n+1} + \beta x_n \quad \text{with} \quad \alpha = 1.4 \quad \text{and} \quad \beta = 0.3 \tag{4.1}$$

$$\frac{dx}{dt} = \sigma(y - x), \quad \frac{dy}{dt} = rx - y - xz, \quad \frac{dz}{dt} = xy - \beta z \tag{4.2}$$

Here $\beta = \frac{8}{3}$, $\sigma = 10$ and $r = 30$.

The stationarity index plot, with respect to the first reference window, of the Henon signal had peaks of approximately 7% as shown in *Figure 4.2*, an indication of stationary behavior. The Stationarity Index plot with respect to all reference windows shows a global maximum of less than 12% as shown in *Figure 4.3*. This is an indication that every window of the signal is in the same region of phase space and therefore the signal is stationary. A Stationarity Index plot, with respect to the first reference window, is carried out for a stationary Lorenz signal *Figure 4.4*. It gave peaks of approximately 14.7% as shown in *Figure 4.5*, an indication of stationary behavior.

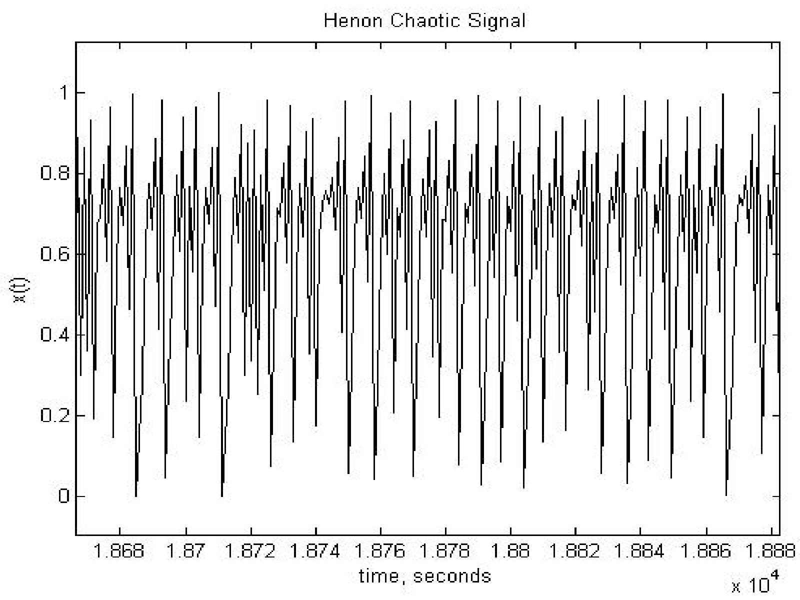


Figure 4.1: Henon signal

A SI plot, with respect to the first reference window, of the nonstationary Lorenz signal is shown *Figure 4.6*. This plot was distinctively different from the plots for the first two stationary signals . The plot rose sharply and maintained highs of above 70% often reaching 100%. This was a clear indication of nonstationary behavior. Close ups of this Lorenz signal give a visual idea of the nonstationary behavior, all the close ups are equal in length covering 2000 samples. *Figure 4.7* is a section of the Lorenz taken from the beginning of the signal. *Figure 4.8* is a similar section taken from the center and finally *Figure 4.9* is a section taken towards the end. It is easy to see the nonstationary behavior in the magnitude as well as the shape of the signal. Furthermore a Stationarity Index plot with respect to all reference windows, *Figure 4.10*, showed a generally high index everywhere except along the diagonal. Along the diagonal the 'ref window' is the same as the 'compared window'. Essentially each reference window is compared with itself. As each reference window is identical to itself the index will be zero.

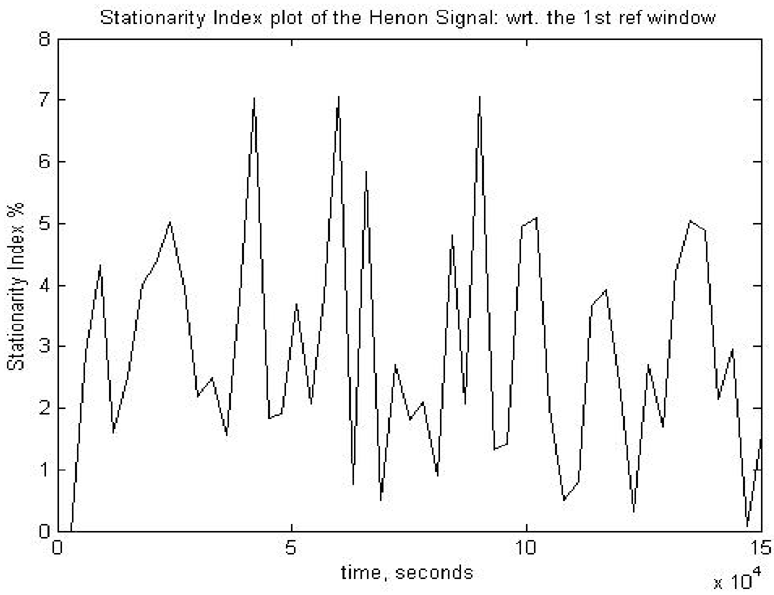


Figure 4.2: Stationarity Index plot of Henon signal: wrt. the 1st ref window

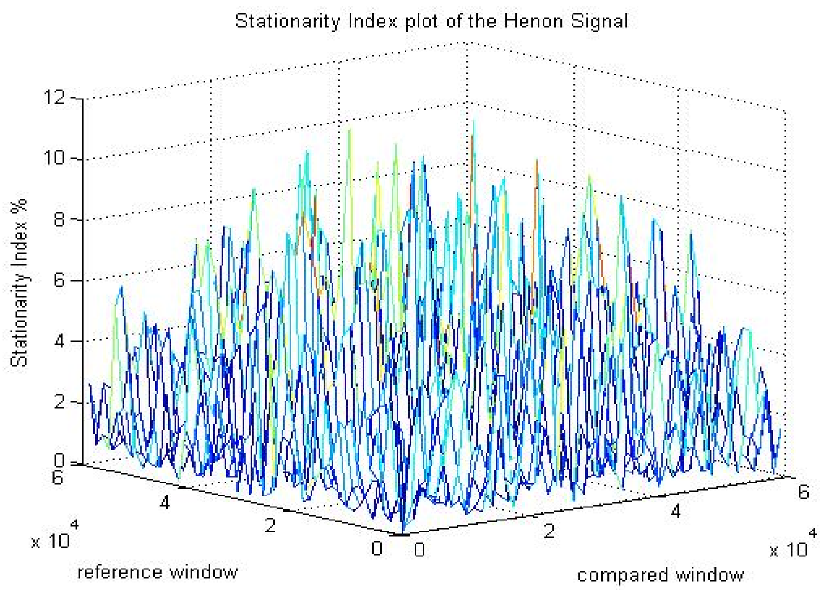


Figure 4.3: Stationarity Index plot of Henon signal 3D plot

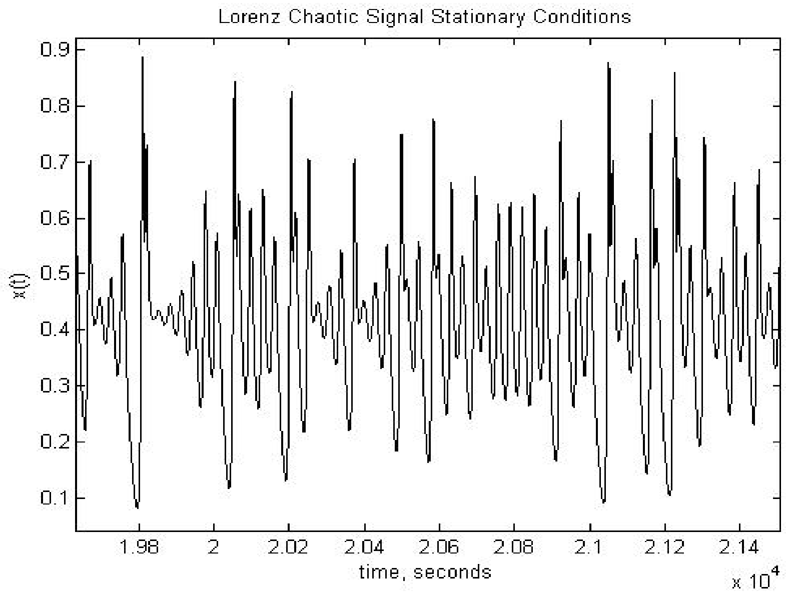


Figure 4.4: Lorenz signal Stationary Conditions

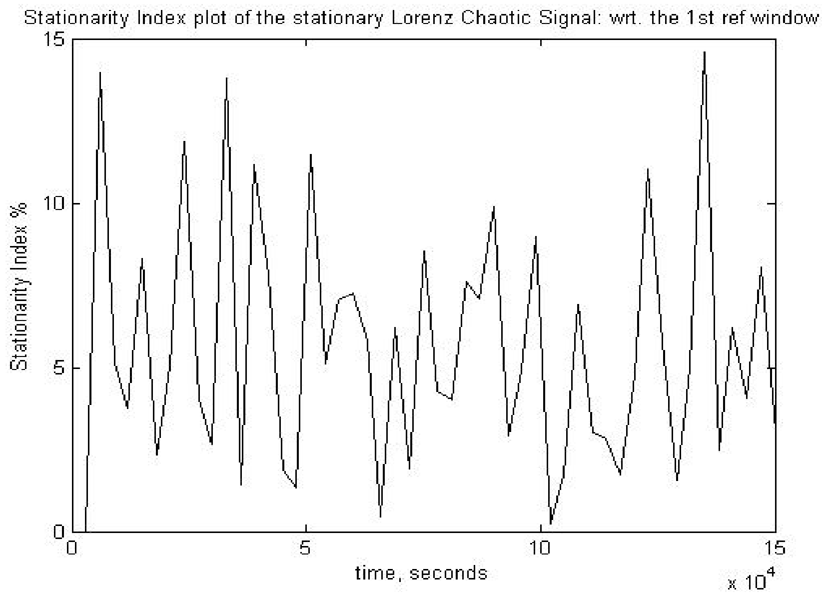


Figure 4.5: Stationarity Index plot of a Lorenz signal Stationary Conditions: wrt. the 1st ref window

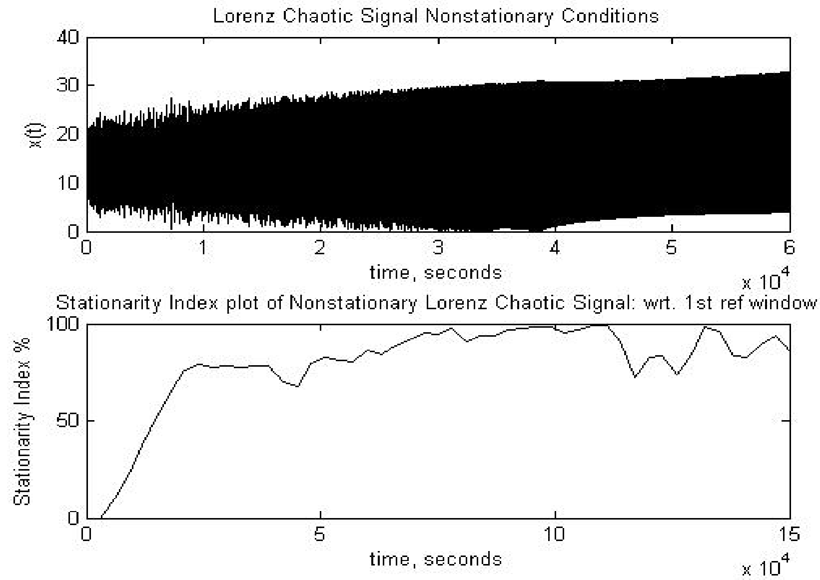


Figure 4.6: Stationarity Index plot of a Lorenz signal Nonstationary Conditions: wrt. the 1st ref window

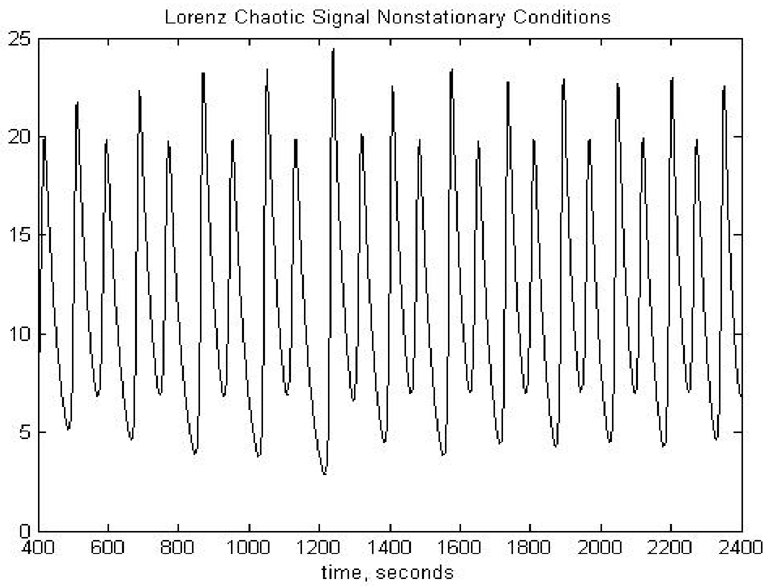


Figure 4.7: Nonstationary Lorenz Section 1

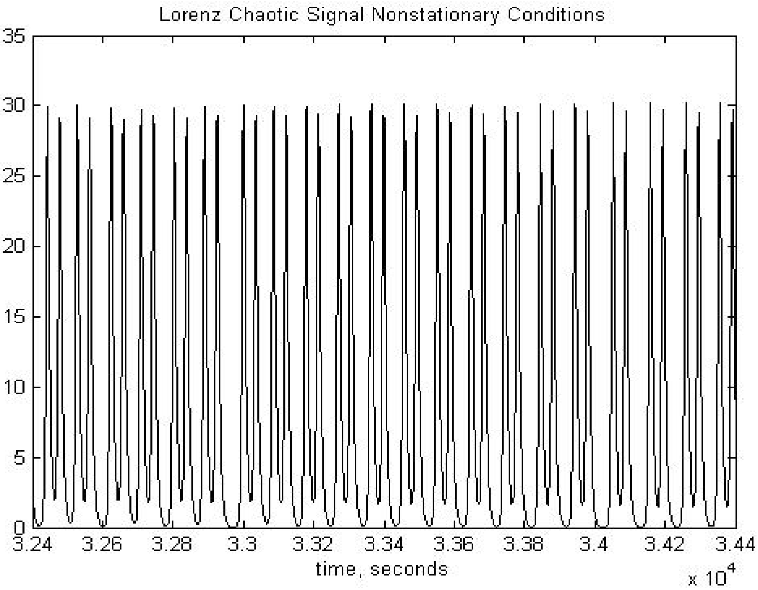


Figure 4.8: Nonstationary Lorenz Section 2

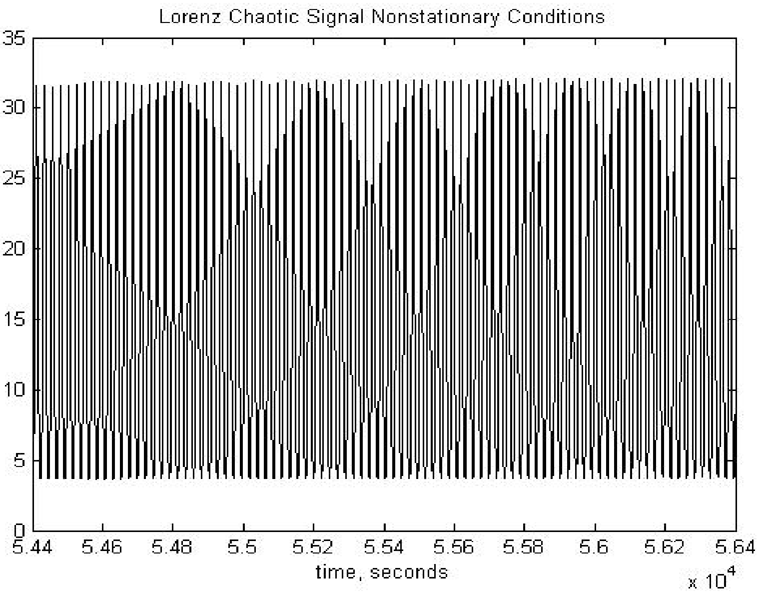


Figure 4.9: Nonstationary Lorenz Section 3

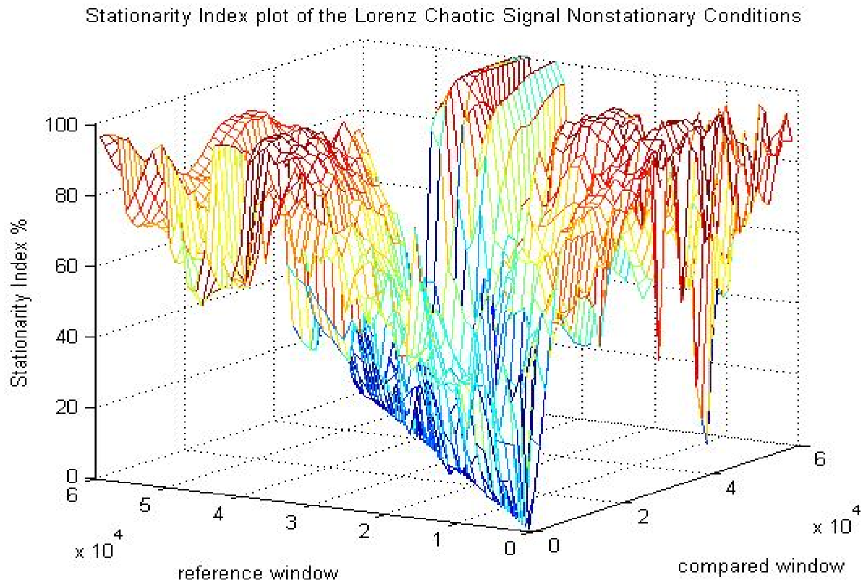


Figure 4.10: Stationarity Index plot of a Lorenz signal Nonstationary Conditions 3D plot

4.2 Conclusion

This chapter presented the results of Stationarity Index tests on chaotic signals. The purpose of the tests was to assess the sensitivity of the index to the stationarity/non-stationarity of the signals presented. It was found to be sensitive in discriminating between stationary and nonstationary behaviour in the signals present. **Chapter 5:** presents the results of Stationarity Index tests on electrocardiogram (ECG) and electroencephalogram (EEG) signals. These signals have disturbances in them caused by the onset of partial epileptic seizures. They are used to observe whether the index reflects the changes in dynamics that occur in the signals.

Chapter 5

Results II

5.1 Results of Nonstationarity Test on ECG and EEG Signals

For each reference window the Theiler window was calculated as well as the error tolerance ε. These were used to calculate the stationarity index between this window and every other window of the signal. An example of the resultant plot follows (*Figure 5.1*). The $Z - axis$ shows the stationarity index, as the index is calculated between the reference window $Y - axis$ and the compared windows $X - axis$. The time represented on the X and Y axes is the time at which the respective windows end. For example, if the sample started at a time of $3950secs$ then the time corresponding to the end of the first reference window, given that it covers a duration of $6secs$, is $3956secs$ as this is the time at which the window ends.

A clearer representation follows. *Figure 5.2* is the plot of the stationarity index with respect to the first reference window. The reference window is taken from a section of normal ECG behavior. The stationarity index goes high as regions of the ECG recorded during the partial seizure are approached. This is a clear indication of a change in the dynamics of the signal, the index reaches a high of 90% indicating that there is almost no phase space overlap between the two sections. Sections in which the index remains low are sections that exhibit the same kind of dynamics as the reference window.

(*Figure 5.3*) is a plot of the stationarity index with respect to the 28th reference window. This is a window taken out of a section of signal recorded during the partial seizure. The stationarity Index is moderately high in the beginning $\approx 40\%$, evidence

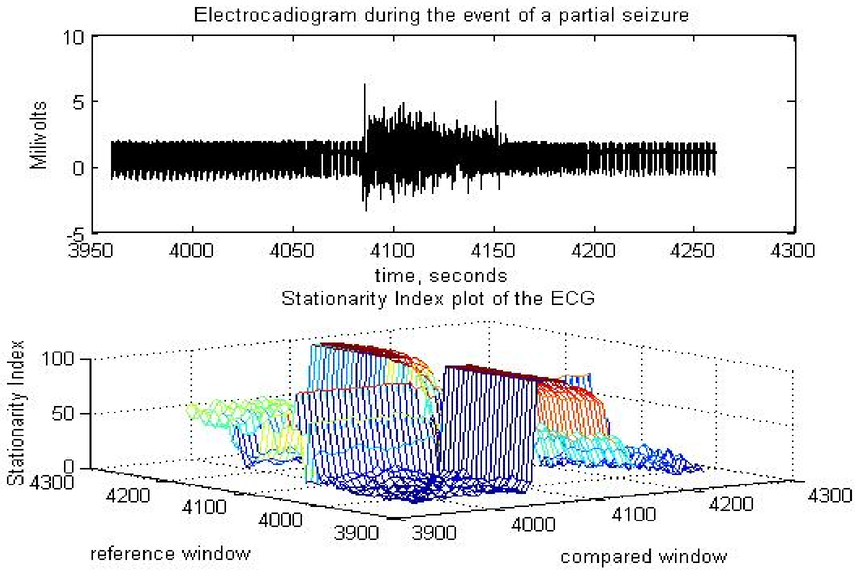


Figure 5.1: Stationarity Index plot of an ECG signal

that the dynamics displayed in the reference window are not exactly the same as those displayed by the windows exhibiting normal ECG behavior. As the stationarity index is calculated between this reference window and windows from areas displaying similar dynamics to its own, the index drops. This is a clear indication of similarities in the dynamics displayed in the two windows. A similar plot to *Figure 5.1* is obtained from an EEG signal (*Figure 5.4*). *Figure 5.5* also shows a plot of the SI with respect to the first reference window. The nonstationarities introduced by the occurrence of the partial seizure are clearly seen.

Figure 5.6 $\rightarrow$ *Figure 5.10* are the results of SI index plots carried out on several ECG signals. Again the major changes in the dynamics are as a result of the onset of partial seizures. All the SI plots are with respect to the first reference window of the signal. *Figure 5.6* is the SI plot of an ECG that records changes that are the result of a partial seizure that occurs from $876 - 972secs$. The index shoots up to about 60% at the time of the partial seizure. It can be seen, even visually, that the nonstationary behavior carries on after $972secs$, which is the time that the seizure ceases. *Figure 5.7* is the SI plot of an ECG that records changes that are the result of a partial seizure occurring from $10551 - 10576secs$. It can be seen, even visually, that the nonstationary behavior carries on after $10576secs$, which is the time that the seizure cease. *Figure 5.8* is the SI plot of an ECG that records changes that are the result of a partial seizure occurring from $5074 - 5182secs$. It can be seen, even visually, that the nonstationary behavior carries on after $5182secs$, which is

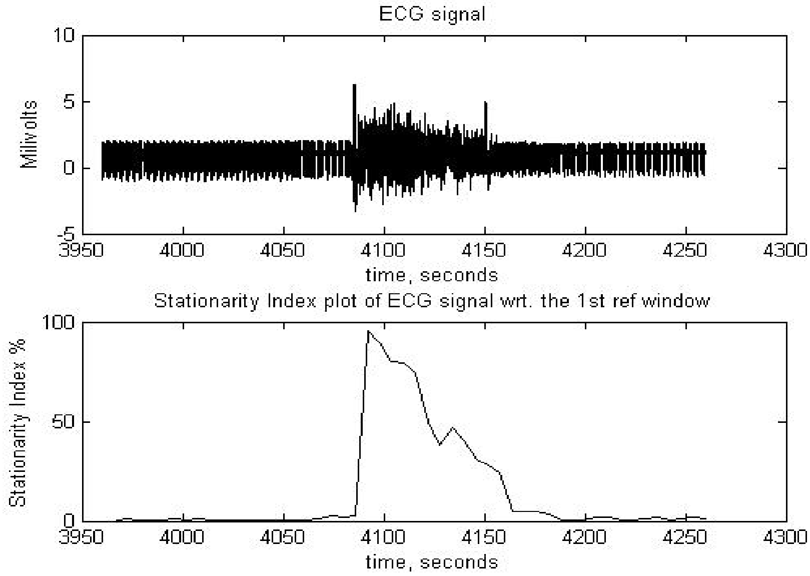


Figure 5.2: Stationarity Index plot of an ECG signal: wrt. the 1st ref window

the time that the seizure ceases. *Figure 5.9* is the SI plot of an ECG that records changes that are the result of a partial seizure occurring from $9267 - 9377secs$. It can be seen that the nonstationary behavior starts to reduce before $9377secs$, which is the time that the seizure ceases. *Figure 5.10* is the SI plot of an ECG that records changes that are the result of a partial seizure occurring from $1210 - 1315secs$. It can be seen that the nonstationary behavior carries on slightly after $1315secs$, which is the time that the seizure ceases. *Figure 5.11* is the SI plot of an ECG that records changes that are the result of a partial seizure occurring from $4082 - 4171secs$. It can be seen that the nonstationary behavior correlates well with the duration of the seizure. The relationship between the variance of the dynamics of the signal and the duration of the seizures is not the focus of the investigation; the focus is on the Stationarity Index's ability to pick out and quantify that variance of dynamics. It is interesting to note, however, that a study was carried out on the same signals [44] in which Power Spectral Density estimates were used to analyse the data. It was found that low-frequency heart rate oscillations were present after the seizure in some cases [44] i.e. changes in the normal behavior of the heart rates that persist after the seizure has finished.

In general the SI index picked out changes in the dynamics and these changes coincided with the onset of partial seizures. The changes in dynamics persisted in most cases after the seizures had ceased. As previously stated the relationship between the variance of the dynamics of the signal and the duration of the seizures is not

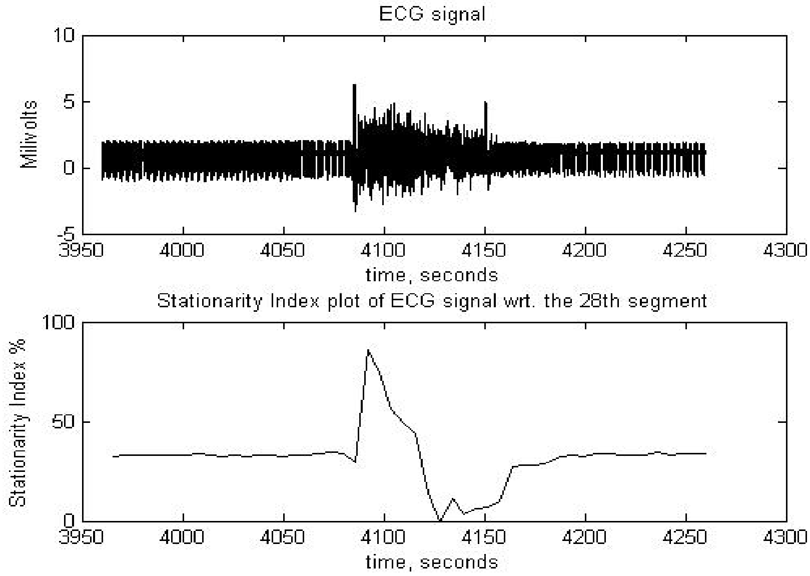


Figure 5.3: Stationarity Index plot of an ECG signal: wrt. the 28th ref window

the focus of the investigation. The SI demonstrates the ability to give a measure of changes in the dynamics. This measure varies directly with the extent to which the dynamics change.

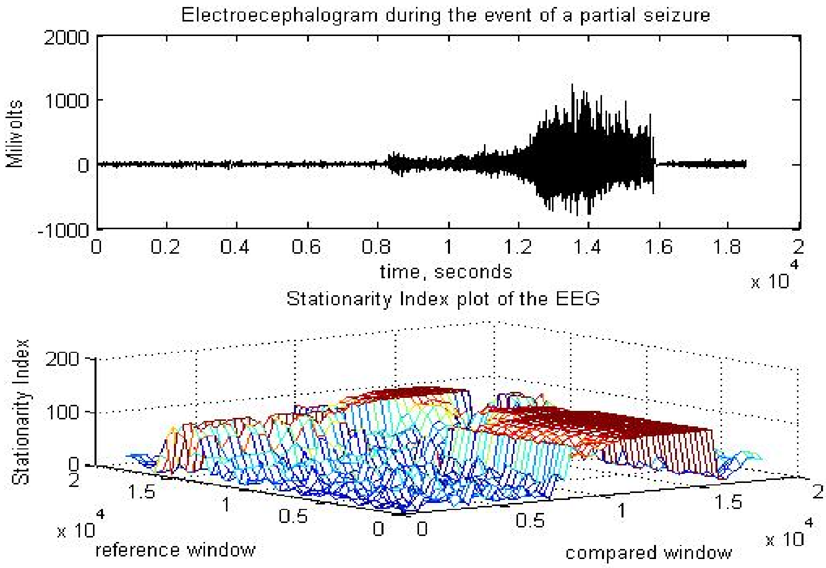


Figure 5.4: Stationarity Index plot of an EEG signal

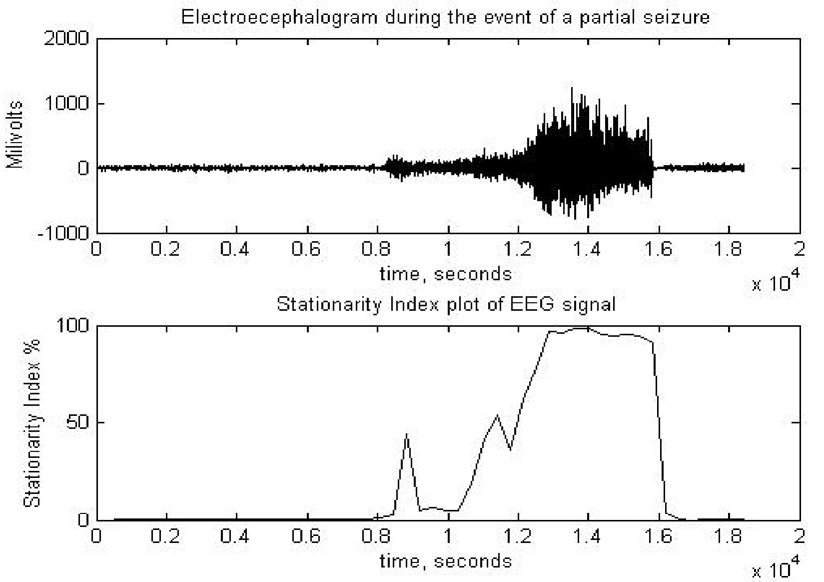


Figure 5.5: Stationarity Index plot of the above EEG signal wrt. the 1st ref window

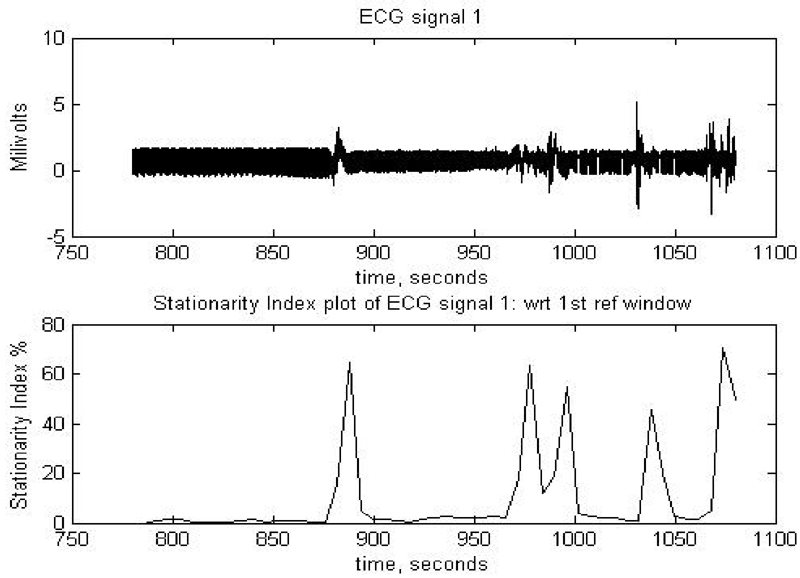


Figure 5.6: Stationarity Index plot of ECG signal 1: wrt. the 1st ref window

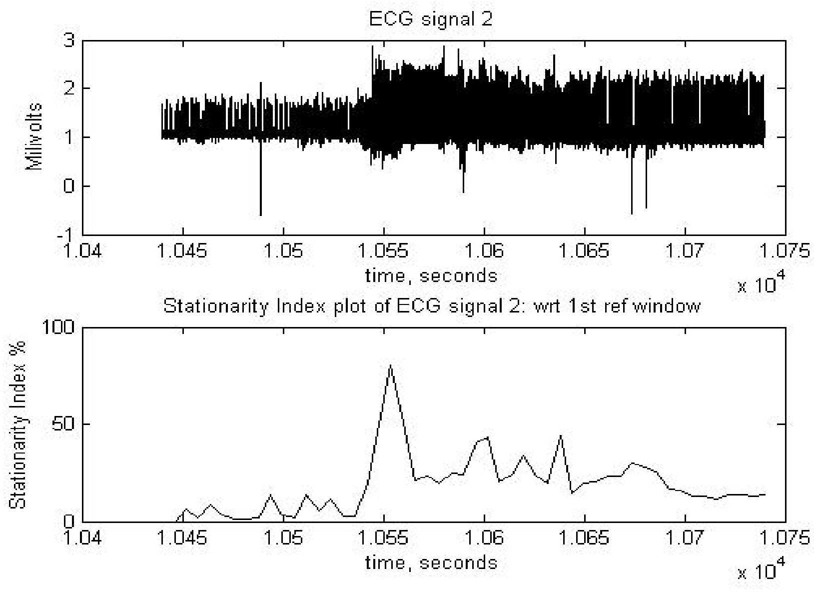


Figure 5.7: Stationarity Index plot of ECG signal 2: wrt. the 1st ref window

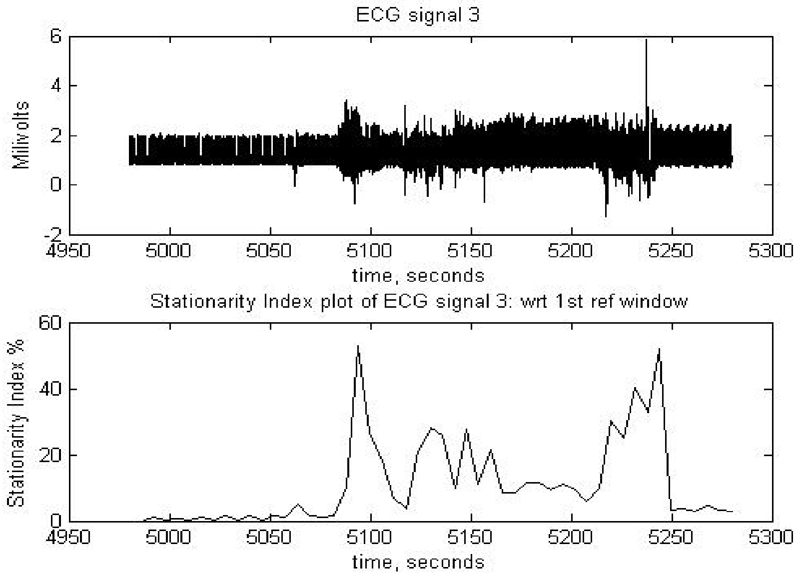


Figure 5.8: Stationarity Index plot of ECG signal 3: wrt. the 1st ref window

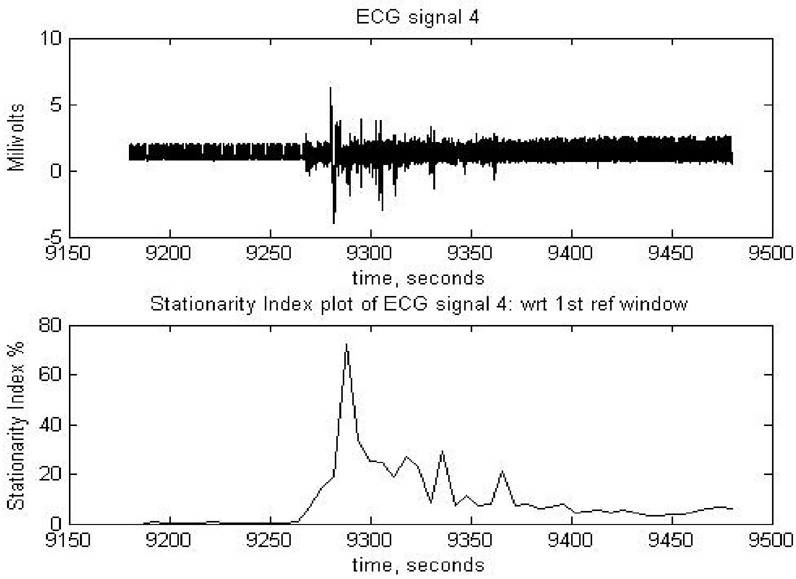


Figure 5.9: Stationarity Index plot of ECG signal 4: wrt. the 1st ref window

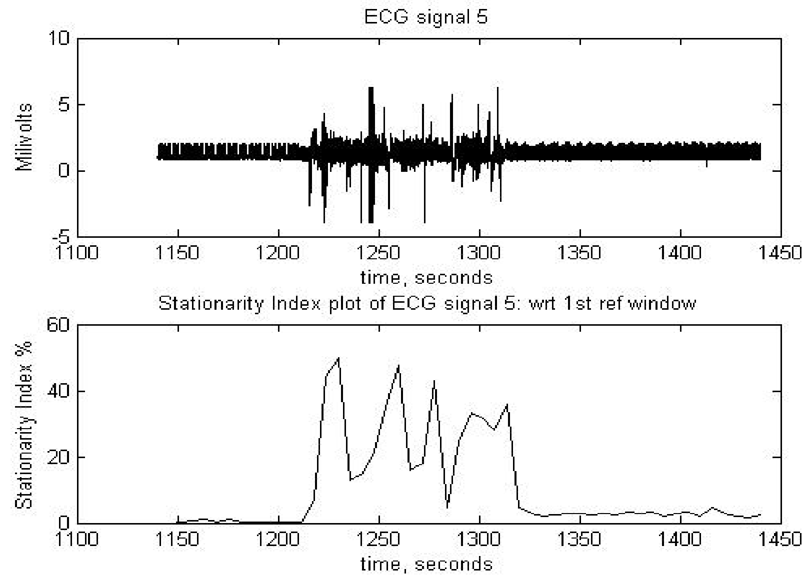


Figure 5.10: Stationarity Index plot of ECG signal 5: wrt. the 1st ref window

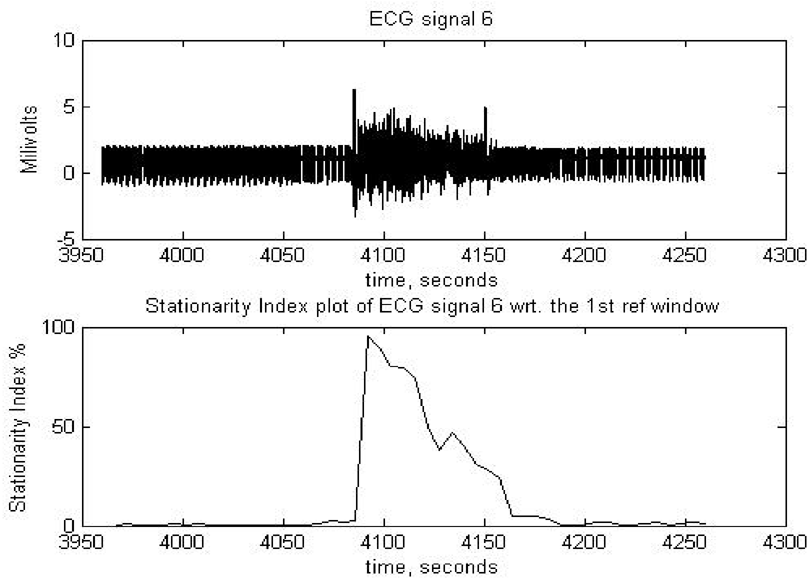


Figure 5.11: Stationarity Index plot of ECG signal 6: wrt. the 1st ref window

The method can be used in segmentation algorithms to determine whether a segment of a signal is stationary. Current methods include the analysis of the Power Spectral Density (PSD) [4]. Similar comparisons between windows are used and for each window the PSD is compared to that of the reference window. If the differences cross a predetermined threshold the compared window is taken to be the beginning of a new segment. A priori information is required about the signals being segmented. With the Stationarity Index method, no a priori information is required. Furthermore the occurrence of the partial seizures were clearly reflected in the SI plot. Therefore the SI shows potential in the analysis of the occurrence of events especially in cases in which the changes in the dynamics are not readily observable from visual examination. The SI can also be used as a means of identifying certain conditions by means of comparing the dynamics of the signals.

5.2 Conclusion

This chapter presented the results of Stationarity Index tests on electrocardiogram (ECG) and electroencephalogram (EEG) signals. The index is sensitive to the changes in dynamics. In most of the cases the changes in SI from low to high correspond exactly with the onset of the partial seizures. This demonstrates potential for this index to be used in the analysis of biomedical signals. **Chapter 6:** summarises the findings of the dissertation and identifies areas for further research.

Chapter 6

Conclusion

Following a review of nonstationarity detection techniques a unique index (the Stationarity Index) was developed. This index enables the detection of nonstationarity and allows for a measure of the extent to which the signal is nonstationary. The mapping of stationarity it provides allows for the comparison of the variation in dynamics of different time series. This is a unique concept. Upon testing with stationary and nonstationary chaotic signals it successfully discriminated between stationary and nonstationary behavior in the signals analysed. This test served to determine the index's ability to do this. The Stationarity Index was then tested on electrocardiogram (ECG) and electroencephalogram (EEG) signals. The signals had changes in their dynamics present as a result of the onset of partial seizures. The Stationarity Index method efficiently differentiated pairs of windows that exhibited similar dynamics from pairs that displayed different dynamics. This index proves useful in the analysis of biomedical signals as it can be used in segmentation to divide nonstationary signals into partially stationary segments. The index rose to a high of almost 100% in cases where there were hardly any similarities between sections of an ECG signal that exhibited different dynamics. Furthermore given that biological signals may exhibit different behavior under different conditions, e.g. different psychotropic drugs appear to have their own 'EEG portraits', there is the possibility of the use of the index to determine when these conditions are present. Such information could aid in the identification of conditions that are present in a patient. The Stationarity Index therefore proves a useful tool for the purposes of analysis and comparison of biomedical nonstationary signals. Further work includes testing to validate or disprove some of the assumptions that were made with regard to this analysis. These tests would include investigations into how much the autocorrelation of signals and different Theiler windows affect the Stationarity Index. This would include testing to determine whether two signals have to have the same autocorrelation α in order to be directly compared. They would also include testing to determine how the lengths of signals compared affects the Stationarity Index i.e. whether the dynamics of two signals, of different lengths yet with the same number of windows, can be compared directly via the Stationarity Index. Other testing would include investigation of the Stationarity Index's sensitivity to other types of nonstationarity. An example would be tests to determine whether it picks out changes in the dynamics of EEGs caused by the changes in sleep stages. And finally comparisons with other methods to evaluate the relative performance of the Stationarity Index method.

Appendix A

Papers

The following are papers that have been submitted based on the work carried out in this research:

- The use of the cross correlation integral in ECG and EEG profile analysis, Bunty B.E. Kiremire and Tshilidzi Marawala, Conference on Collaborative Research for Technological Development, December 2007, Kampala Uganda - Accepted.
- Nonstationarity detection: The use of the cross correlation integral in ECG, and EEG profile analysis, B.B.E. Kiremire and Tshilidzi Marawala, 2008 Congress on Image and Signal Processing, May 2008 Hainan China - Accepted.

References

[1] R. Manuca and R. Savit, "Stationarity and nonstationarity in time series analysis," *Physica D*, vol. 99, pp. 134 – 161, May 1996.

[2] J. Hu, J. B. Gao, and K. D. White, "Estimating measurement noise in a time series by exploiting nonstationarity," *Chaos Solitons and Fractals*, vol. 22, pp. 807–819, 2004.

[3] T. Schreiber, "Detecting and analysing nonstationarity in time series using nonlinear cross predictions," *Physical Review Letters*, vol. 78, pp. 843 – 846, February 1997.

[4] J. D. Bronzino, *The Biomedical Engineering Handbook.* CRC Press, 1995.

[5] A. A. Fingelkurts and A. A. Fingelkurts, "Problems and solutions in electroencephalogram neuropsychopharmacology," in *8th Multidisciplinary International Conference of Biological Psychiatry "Stress Behavior" At-Petersburg, Russia*, May 2004.

[6] O. E. Rössler, "An equation for continuous chaos," *Physics Letters*, vol. 57 A, pp. 397– 398, July 1976.

[7] D. Yu, W. Lu, and R. G. Harrison, "Space time-index plots for probing dynamical nonstationarity," *Physics Letters A*, vol. 250, pp. 323 – 327, December 1998.

[8] S. D. Mohanty, "Robust test for detecting nonstationarity in data from gravitational wave detectors," *Physical Review D*, vol. 61, May 2000.

[9] M. B. Kennel, "Statistical test for dynamical nonstationarity in observed time series data," *Physical Review E*, vol. 56, pp. 316 – 321, July 1997.

[10] K. Gröchenig and G. Zimmermann, "Hardy's theorem and the short-time fourier transform of schawtz functions," *J. London Math. Soc.*, vol. 63, pp. 205–214, 2001.

[11] W. Sun and X. Zhou, "Irregular wavelet/gabor frames," *Applied and Computational Harmonic Analysis*, vol. 13, pp. 63–76, July 2002.

[12] S. Qian and D. Chen, *Joint Time-Frequency Analysis - Methods and Applications*. Prentice Hall, 1996. Short-Time Fourier Transform Section available on the Internet via: http://zone.ni.com/devzone/cda/ph/p/id/235 - Last accessed 15th March 2007.

[13] J. V. Samar, "Wavelet analysis of neuroelectric waveforms: A conceptual tutorial," *Brain and Language*, vol. 66, pp. 7–60, 1999.

[14] J. R. Whitla, P. Embrechts, A. M. Herzberg, H. K. Kalbfleischand, and W. N. Traves, "An introduction to wavelets with applications to andrews' plots," *Journal of Computational and Applied Mathematics*, vol. 64, pp. 41–56, 1995.

[15] H. Burri, P. Chevalier, M. Arzi, P. Rubel, G. Kirkorian, and P. Touboul, "Wavelet transforms for analysis of heart rate variability preceeding ventricular arrhythmias in patients with ischemic heart disease," *International Journal of Cardiology*, vol. 109, pp. 101–107, 2006.

[16] G. Tzanetakis, G. Essl, and P. Cook, "Audio analysis using the discrete wavelet transform," *In Proceedings of WSES International Conference, Acoustics and Music: Theory and Applications (AMTA 2001)*, 2001.

[17] J. Wexler and S. Raz, "Discrete gabor expansions," *Signal Processing*, vol. 21, pp. 207–220, 1990.

[18] S. Joshi and J. M. Morris, "On a novel critically-sampled discrete-time real gabor transform," *Signal Processing*, vol. 61, pp. 9–22, 1997.

[19] M. J. Bastiaans and M. C. W. Geilen, "On the discrete gabor transform and the discrete zak transform," *Signal Processing*, vol. 49, no. 3, pp. 151–166, 1996.

[20] H. G. Feichtinger and K. Gröchenig, "Gabor frames and time-frequency analysis of distributions," *Journal of Functional Analysis*, vol. 146, pp. 464–495, 1997.

[21] J. M. Morris and Y. Lu, "Discrete gabor expansion of discrete-time signals in l2 (z) via frame theory," *Signal Processing*, vol. 40, pp. 155–181, 1994.

[22] V. Krüger and G. Sommer, "Gabor wavelet networks for effecient head pose estimation," *Image and Vision Computing*, vol. 20, pp. 665–672, 2002.

[23] J. B. Gao, "Detecting nonstationarity and state transitions in a time series," *Physical Review E*, vol. 63, May 2001.

[24] J. B. Gao, "Recurrence time statistics for chaotic systems and their applications," *Physical Review Letters*, vol. 83, pp. 3178 – 3181, October 1999.

[25] J. P. Zbilut, A. Giuliani, and C. L. Webber, "Recurrence quantification analysis and principle components in the detection of short complex signals," *Physics Letters A*, vol. 237, pp. 131 – 135, January 1998.

[26] C. Ahlstrom, P. Hult, and P. Ask, "Thresholding distance plots using true recurrence points," *Acoustics, Speech and Signal Processing - proceedings*, vol. 3, p. 3, 2006.

[27] C. Rieke, K. Sternickel, R. G. Andrzejak, C. E. Elger, P. David, and K. Lehnertz, "Measuring nonstationarity by analysing the loss of recurrence in dynamical systems," *Physical Review Letters*, vol. 88, June 2002.

[28] A. Provenzale, L. A. Smith, R. Vio, and G. Murante, "Distinguishing between low-dimensional dynamics and randomness in measured time series," *Physica D*, vol. 58, pp. 31 – 49, March 1992.

[29] S. Conforto and T. D'Alessio, "Spectral analysis for non-stationary signals from mechanical measurements: A parametric approach," *Mechanical Systems and Signal Processing*, vol. 13, no. 3, pp. 395–411, 1999.

[30] S. Mukhopadhyay and P. Sircar, "Parametric modelling of non-stationary signals: A unified approach," *Signal Processing*, vol. 60, pp. 135–152, 1997.

[31] P. Sircar and M. S. Syali, "Complex am signal model for non-stationary signals," *Signal Processing*, vol. 53, pp. 35–45, 1996.

[32] T. Tambouratzis and M. Antonopoulos-Domis, "A non-stationary signal correlator for on-line transit time estimation," *Annals of Nuclear Energy*, vol. 29, pp. 1299–1313, 2002.

[33] G. Andria, F. Attivissimo, and A. Lanzolla, "Interpolation techniques for non-stationary signal analysis," *Measurement*, vol. 21, no. 1, pp. 1–16, 1997.

[34] S. Nandi and D. Kundu, "Analysing non-stationary signals using generalized multiple fundamental frequency model," *Journal of Statistical Planning and Inference*, vol. 136, pp. 3871–3903, 2006.

[35] J. I. I. Hussain, H. Kanki, and N. W. Mureithi, "Chaos in the unbalance response of a rigid rotor in cavitated squeeze film dampers without cetering springs.," *Chaos, Solutions and Fractals*, vol. 13, pp. 929–945, 2001. Accepted February 2001.

[36] B. Pilgram, W. Schappacher, W. N. Löscher, and G. Pfurtscheller, "Application of the correlation integral to respiratory data of infants during rem sleep," *Biogical Cybernetics*, vol. 72, pp. 543–551, May 1995.

[37] R. Thieberger, "The use of luminosity effects to calculate scaling regimes in the galaxy distribution," *Astronomy and Astrophysics*, vol. 397, no. 2, pp. 401–403, 2003.

[38] A. Cecen and A. Ugur, "On testing for nonlinear dependance and chaos in financial time series data," in *Systems, Man and Cybernetics, 2005 IEEE International conference on*, vol. 1, pp. 203–208, IEEE, October 2005.

[39] I. V. Fel'dshtein, "The cross correlation integral: Its features and application to nonstationarity detection in time series," *Technical Physics*, vol. 45, no. 6, pp. 667 – 671, 2000.

[40] H. Kantz, "Quantifying the closeness of fractal measures," *Physical Review E*, vol. 49, pp. 5091 – 5097, June 1994.

[41] C. D. Cutler, "A theory of correlation dimension for stationary time series," *Philosophical Transactions: Physical Sciences and Engineering*, vol. 348, pp. 342–355, Sept 1994.

[42] J. Theiler, "Estimating fractal dimension," *Optical Society of America*, vol. 7, pp. 1055–1073, June 1990.

[43] J. Theiler, "Spurious dimension from correlation algorithms applied to limited time-series data," *Physical Review A*, vol. 34, September 1986.

[44] I. AlAweel, K. B. Krishnamurthy, J. M. Hausdorff, J. Mietus, J. R. Ives, A. Blum, D. L. Schomer, and A. L. Goldberger, "Post ictal heart rate oscillations in partial epilepsy," *Neurology*, vol. 53, pp. 1590 – 1592, October 1999.

[45] "The sleep heart health study polysomnography database." Available at: www.physionet.org last accessed 6th Dec 2007.