Analysis of Heart Rate Dynamics using Linear and Nonlinear Models

by
Md. Sanaullah Hakim

A Thesis
Submitted to the Department of
Electrical and Electronic Engineering, BUET,
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Electrical and Electronic Engineering

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

2002

The thesis titled **Analysis of Heart Rate Dynamics using Linear and Nonlinear Models** Submitted by **Md. Sanaullah Hakim** Roll No: **9506201 P** Session **1994-95-96** has been accepted as satisfactory in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronic Engineering** on **June 19, 2002.**

BOARD OF EXAMINERS

19 June 2002

1.	Dr. Md. Aynal Haque Associate Professor Dept. of EEE, BUET, Dhaka	Chairman
2.	Dr. Md. Kamrul Hasan Associate Professor Dept. of EEE, BUET, Dhaka	Member
3.	Dr. M. M. Shahidul Hassan Professor & Head Dept. of EEE, BUET, Dhaka	Member (Ex-officio)
4.	Dr. Abdul Wadud Chowdhury Assistant Professor of Cardiology National Institute of Cardiovascular Diseases & Hospital, Dhaka	Member (External)

CANDIDATE'S DECLARATION

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Signature of the Candidate

19.06.02

Md. Sanaullah Hakim

ACKNOWLEDGEMENTS

The author would like to express his indebtedness and gratitude to his supervisor Dr. Md. Aynal Haque for his endless patience, friendly supervision and invaluable assistance in making a difficult task a pleasant one.

The author wishes to express his thanks and regards to the Head of the Department of Electrical and Electronic Engineering, BUET, Dhaka, for his support during the work.

Sincerest thanks to friends and colleagues for their constant support and valuable suggestions of the work.

CONTENTS

ABSTRACT

Modeling of time series data plays an important role in determining the underlying physical mechanism generating and controlling the data. Linear and nonlinear parametric and nonparametric modeling techniques can be used in data prediction in many diversified fields. In this thesis, two parametric methods, linear autoregressive (AR) and nonlinear neural network (NN) models, are used to predict heart rate. Nonparametric chaotic modeling technique is used to differentiate abnormal cardiac rhythm from the normal one.

The instantaneous heart rate (IHR) is calculated from the measured ECG of 5 adult human subjects and 5 IHR time series data sets are constructed, each for more than half an hour. The AR model parameters and order are determined by Burg method and Akaike information criterion (AIC) statistics, respectively. The results of one-step ahead prediction of 1000 IHR data for each data set show that it can predict precisely IHR of 3 data sets. On the other hand, radial basis function NN, trained by K-means clustering and least square algorithms can accurately predict IHR of all the 5 data sets. This indicates that nonlinearity may be present in IHR time series.

Central tendency measure (CTM), a characteristic of chaotic modeling is applied to the IHR of all 5 data sets. In attempt to verify whether CTM can differentiate normal and abnormal heart rate dynamics, random noise is added to normal IHR and CTM is measured for this noisy IHR. Analysis of variance shows that the CTM of normal IHR (*0.971±0.016*) is significantly different from that of noisy IHR (*0.829±0.034*). The CTM of apnea IHR is also calculated and the results show that this CTM (*0.223±0.038*, *N=9*) is also significantly different from that of normal IHR.

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Cardiac activity

The heart is a muscular organ located in the thoracic cavity of the body. It is the physiological center of the vascular system. The heart is divided longitudinally into left and right halves, each consisting of two chambers, an atrium and a ventricle. The cavities of the atrium and ventricle on each side of the heart are directly connected, but the right half is not directly connected with the left half. Thus, right and left atria and right and left ventricles are functionally distinct.

The heart pumps blood through out the body. The timing of the pumping activity is determined by electrical events. Contraction of atrial muscle due to electric pulses forces blood from the left atrium to the left ventricle through the mitral valve, and from the right atrium to the right ventricle through the tricuspid valve. Contraction of the ventricles results in ejection of blood from the left ventricle through the aortic valve into the systemic circulation, and from the right ventricle through the pulmonary valve into the pulmonary circulation. Fig. 1.1 shows the front view of a mammalian heart and the blood circulation in it.

Malfunction of the pumping action of the heart leads to morbidity and death, as does the malfunction of the system generating and conducting the electrical impulses in the heart even if the heart muscle is physiologically normal. Following a heart attack (coronary occlusion), where regions of heart muscle have been injured because of an interruption in blood supply to the heart's tissues (ischemia), the affected muscle generally loses some or all of its ability to contract. At the same time, severe disturbances in the electric conduction system may occur. Depending on the extent of the injury, the mechanical impairment of the localized parts of the heart often does not profoundly compromise its ability to pump blood. On the other hand, malfunction of the electric system will result in improper contraction of the otherwise viable muscle because of a lack of coordinating trigger. Indeed, many people die within minutes or hours of a coronary occlusion, which is termed as sudden cardiac death. The pumping phenomenon in a normal heart is periodic. A normal heart beats spontaneously and rhythmically. Any abnormality in the conduction system or other regions of the heart can result in a large variety of disturbances of the rhythm of the heart: conduction may be slowed in a region or completely blocked, spontaneous beats may arise at one or more abnormal sites, heart rate may be slowed down (bradycardia), or increased (tachycardia). The occurrence of some of the above mentioned events for a prolonged period and/or above certain normal levels may cause sudden cardiac death. Approximately, half of all cardiovascular deaths are sudden.

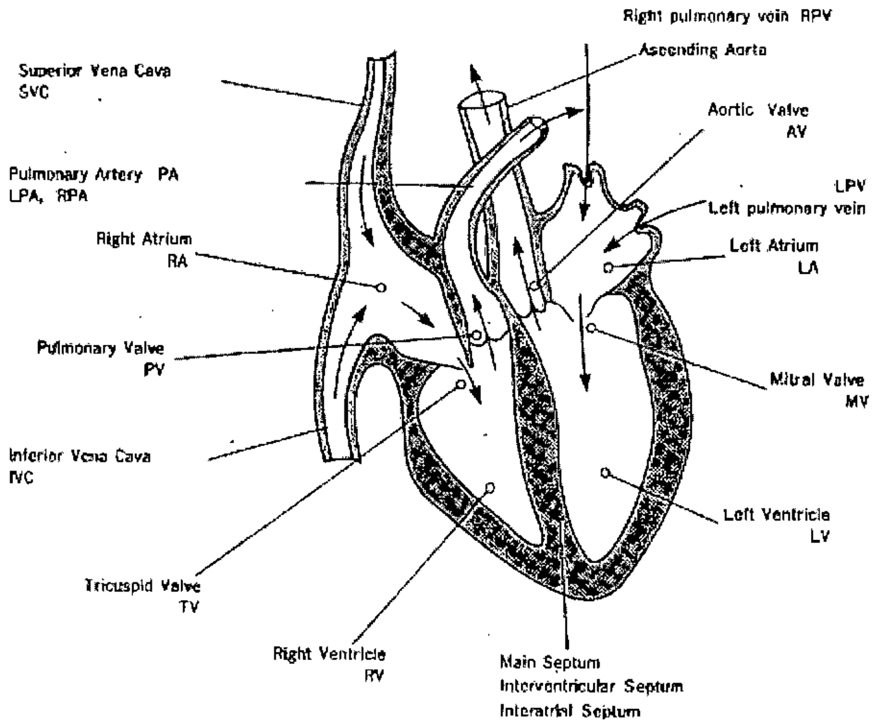

Fig. 1.1 Heart front view and blood circulation

Artificial cardiac pacemakers and defibrillators are designed to generate electric pulses in order to correct certain rhythmic disturbances. Yet, the death due to cardiac arrest is still a potential problem. Better understanding of the underlying developmental and functional characteristics of normal and abnormal hearts can help to prevent some of these occurrences.

1.2 ECG and Heart Rate

Cell to cell communication in the heart is electrical. Currents arising from depolarization of a cell are generally large enough to depolarize the neighboring cells. Adjacent cells of the heart are joined end to end at structures containing gap junctions which allow action potentials to be transmitted from one cell to another. Hence, activity initiated in a region will spread throughout the heart. A heart beat originates from the activity of specialized muscle cells known as pacemaker located in the sinoatrial (SA) node in the right atrium. A number of factors control the beat rate of the pacemaker cells, notably chemicals released by activity of sympathetic and parasympathetic nerve fibers. In addition, control can be affected by the atrioventricular (AV) node, located in the right ventricle. The impulse generated in the SA node is conducted through the atrial muscle to the AV node. Normally the region between atria and ventricles does not conduct the cardiac impulses except through the AV node. Conduction velocity in the AV node is relatively slow, resulting in a delay between activation of the atria and the ventricles. The cardiac impulses emerge from the AV node into the bundle of His which then divides into the left and right bundle. The left bundle further divides into fascicles and both bundles then ramify into a mesh-like network, the Purkinje fibers, which cover a wide area of the ventricles and interventricular septum, generally on the inner surface. Conduction through this specialized conduction system is relatively fast. The system serves to bring excitation rapidly to large areas of the ventricles from which it then spreads more slowly through the muscle mass of the ventricles. Heart cells exhibit refractoriness, i. e., once they fire a certain period of time must elapse before they can once again be excited. This characteristic leads to an orderly spread of activity under normal circumstances. Excitation of each cell is followed by a recovery period. Fig. 1.2 shows a schematic block diagram of the electric conduction system of the heart. For a normal sinus beat, the excitation proceeds from left to right as indicated. Retrograde conduction may occur under certain circumstances, as well as the development of some accessory pathways from atrial to ventricular muscles, in addition to the normal conduction system, as shown by the dotted line in Fig. 1.2. The electric potential thus developed in the heart propagates through the surrounding cells and muscles and constitutes a measurable potential over the body surface.

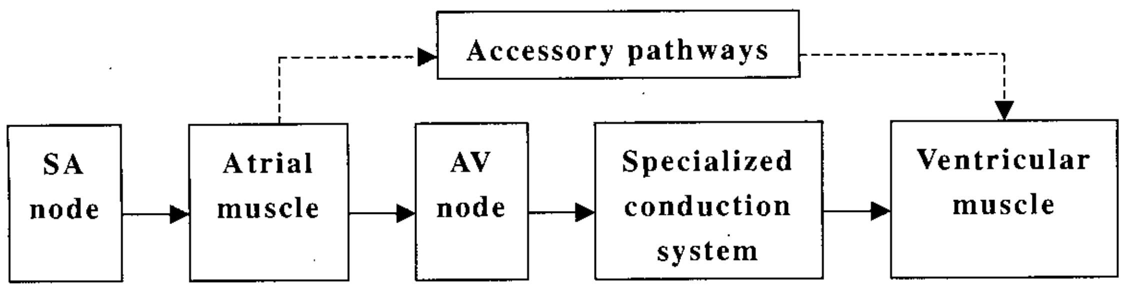

Fig. 1.2 Block diagram of the electrical conduction system of heart

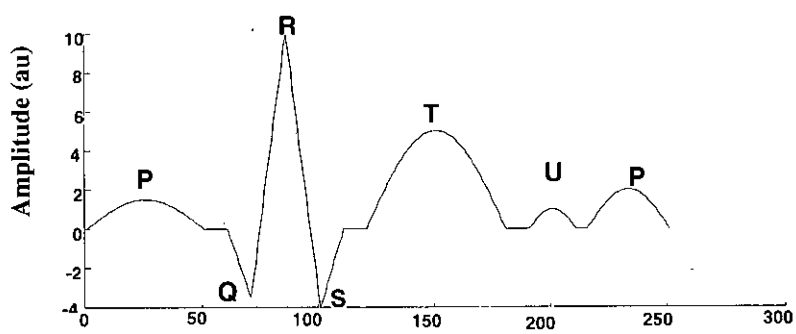

Fig. 1.3 A typical ECG pattern presented for one cardiac cycle

The electrocardiogram (ECG or EKG) is a graphical recording or display of the time-variant voltages produced by the myocardium of the cardiac system. Figure 1.3 shows a diagrammatical waveform of an ECG for one cardiac cycle. There are three distinct waves in each cycle of ECG. The first, designated as P wave, results from the spread of excitation through the atria. The second, designated as QRS complex, results from the spread of excitation through the ventricles. The third is the T wave which results from the recovery of excitations of cells in the ventricles. In addition, a fourth deflection, a U wave generated due to the relaxation of the ventricular muscles is sometimes seen to follow the T wave.

Heart rate (HR) is calculated from the measured ECG. Two types of HR are normally calculated: mean heart rate (MHR) is from the ECG of a fixed duration, normally one minute; and instantaneous heart rate (IHR) is from the successive peaks. (R-wave). Together with the ECG, MHR is often used to assess cardiac activities. The heart rate can vary widely even in a short period of time due to natural variability. MHR can not explain the issues of arrhythmias which are the major causes of sudden cardiac death. A parameter other than MHR may illuminate more light on the better understanding of the cardiac functions and other phenomena associated with it. The beat to beat analysis of cardiogenic signals, i.e., IHR is such a parameter. It has long been recognized that the HR, arterial blood pressure, and other hemodynamic parameters fluctuate on a beat-to-beat basis [1]. The beat-to-beat fluctuations in hemodynamic parameters reflect the dynamic response of the cardiovascular control systems to a host of naturally occurring physiological perturbations. The sympathetic and parasympathetic nervous systems are usually considered to be the principal systems involved in short-term cardiovascular control on the time scale of seconds to minutes. Heart rate fluctuations reflect the regulatory activity of the autonomic nervous system (ANS), which modulates the intrinsic sinus node firing rate. In fact, the instantaneous heart rate (IHR) can be thought as the output of a feedback control loop that is continuously regulated by the ANS. Lack of beat-to-beat heart rate variability is an important predictor of impending death.

1.3 Modeling

In recent years, a particular attention has arisen in noninvasive medical diagnostic procedures. Because biosignals recorded on the body surface reflect the internal behavior and status of the organism or its parts, they are ideally suited to provide essential information of these organs to the clinicians without any invasive measures. Here some questions arise. How are the recorded time courses of the signals to be interpreted with regard to a diagnostic decision? What are the essential features and in what code is the information hidden in the signals? These questions can be answered by the recognition of detailed signal patterns. Since signal processing is concerned with the mathematical representation of a signal and the algorithmic operation carried out to extract the information present in the signal, the systems generating and controlling the signal must be understood first.

Modeling is the mathematical description of a physical system. Once a model of a system is available, one can analyze its overall behavior and predict the response of its output to different inputs. The difficulty, however, lies in determining such a mathematical description. Indeed, although one can derive a model for some simple systems, such as electric motors, there are many systems whose outright complexity makes such a task impossible. In such cases, one has to resort to the numerical techniques of system identification [2]. In attempting to determine empirically the mathematical descriptions associated with systems, modeling techniques are used. The classical theory of time series analysis has been well developed over the past two decades, and excellent accounts of this theory are available [2]-[7]. An important assumption that is made in the classical theory is that the structure of the series can be described by a linear model.

Two broad categories of methods are used in the modeling of time series data, namely, parametric and nonparametric methods. Nonparametric methods make no assumption about how the data are generated, and mainly use Fourier based approach and periodogram. Nonparametric methods are relatively simple, well understood and easy to compute using the FFT algorithm. However, these methods require the availability of long data records in order to obtain necessary frequency resolution required in many applications. Furthermore, these methods suffer from the leakage effects due to windowing that are inherent in finite-length data records. Often, the leakage masks weak signals that are present in the data.

On the other hand, parametric methods of data modeling eliminate the need of window functions. As a consequence, parametric or model based methods avoid the problem of leakage and provide better resolution than do the FFT based nonparametric methods. This is specially true in applications where short data records are available due to time-variant of transient phenomena. Both linear and nonlinear parametric time-series models are used for the prediction of data. Linear models based on filter theory have a long and rich history of development over the last six decades. The historical account of the first three decades is available in [8]. The pioneer work of Weiner [9] on optimum linear filtering for statistically stationary signals is specially significant. The generalization of the Weiner filter theory to dynamical systems with random inputs was developed in early sixties [10], [11]. There are numerous references on linear prediction and lattice filters. Tutorial treatments on these subjects are found in [12], [13]. The books by Haykin [14], Markel and Gray [15], and Tretter [16] provide comprehensive treatments of these subjects. Applications of linear prediction to spectral analysis are found in the books by Kay [17] and Marple [18], to geophysics [19] and to adaptive filtering [14].

1.4 Aim of the Thesis

HR is a major indicator of cardiac activity. Moreover, it also reflects the conditions of other functional systems, such as the autonomic nervous and respiratory systems. Biological data are modeled with an ultimate aim of differentiating the data generated by normal subject with that by a patient. The objective of this research work is to apply linear and nonlinear modeling techniques in the prediction of heart rate time series data. First the test of linearity of heart rate is attempted applying linear autoregressive and nonlinear neural network model to predict heart rate. It is assumed that the better predictability of any model will give indications about the nature of the data, i.e., deterministic, random or otherwise. Based on the results obtained from linearity test, other methods of heart rate modeling will be applied to investigate the possibility of categorical division of diseases.

1.5 Organization of the Dissertation

Three techniques of heart rate modeling are considered. Chapter 2 and 3 describe the application of autoregressive and neural network models, respectively, to predict instantaneous heart rate. The results show that the heart rate may be nonlinear. Based on this, another nonlinear model, chaos theory is considered to differentiate between normal heart rate with abnormal one. Chaotic modeling is presented in chapter 4. Chapter 5 represents the conclusions of the findings.

CHAPTER 2

AUTOREGRESSIVE MODELING OF HEART RATE

2.1 Introduction

Linear prediction is an important topic in digital signal processing with many practical applications. Two of the most important applications of linear estimation are power spectral estimation and linear filtering. Many of the phenomena that occur in nature are best characterized statistically in terms of averages. Most signals in nature are characterized as random processes. For example, meteorological phenomena such as the fluctuations in air temperature and pressure are best characterized statistically as random processes. Thermal noise voltages generated in resistors and electronic devices are additional examples of physical signals that are well modeled as random processes. Due to the random fluctuations in such signals, one must adopt a statistical viewpoint, which deals with the average characteristics of random signals. In particular, the autocorrelation function of a random process is the appropriate statistical average that is used for characterizing random signals in the time domain, and the Fourier transform of the autocorrelation function provides the transform from time domain to frequency domain.

Autoregressive (AR) and autoregressive moving average (ARMA) are the mostly used linear parametric models in the prediction of a time series in a wide variety of applications [20]. For a linear time invariant system, the general system transfer function $H(z)$ is given by [21],

$$H(Z) = \frac{B(Z)}{A(Z)} = \frac{\sum_{k=0}^{Q} b_k z^{-k}}{\sum_{k=1}^{P} a_k z^{-k} + 1} \qquad (2.1)$$

where b_k and a_k are the system coefficients that determine the location of the zeros and poles of $H(z)$, respectively. The parameters P and Q are the system orders. For the linear system with the rational system function $H(z)$ given by eqn. (2.1), the output $x(n)$ is related to the input $u(n)$ by the difference equation,

$$x(n) + \sum_{k=1}^{P} a_k x(n-k) = \sum_{k=0}^{Q} b_k u(n-k) \tag{2.2}$$

Depending on the system parameters, the parametric models are of 3 major types. If the system has both finite poles and zeros and the corresponding difference equation is given by eqn. (2.2), and the model is termed as ARMA model. If $b_0 = 1$, $b_k = 0$, $k>0$, the linear system has $H(z) = 1/A(z)$, and the system is an all-pole system. This model is known as AR model. In this case, the difference equation for the input-output relationship is,

$$x(n) + \sum_{k=1}^{P} a_k x(n-k) = u(n) \tag{2.3}$$

If $a_k = 0$, $k \geq 1$, the linear system has $H(z) = B(z)$, and the system is an all-zero or moving average (MA). In this case, the difference equation for the input-output relationship is,

$$x(n) = \sum_{k=0}^{Q} b_k u(n-k) \tag{2.4}$$

Out of the above-mentioned models, the most popular of the time series modeling approaches to prediction is AR estimator. This is because accurate estimates of the AR parameters can be found by solving a set of linear equations. For accurate estimation of the parameters of MA or ARMA processes, we need to solve a set of highly nonlinear equations. When the AR modeling assumption is valid, estimators are obtained which are less biased and have a lower variability than conventional Fourier based estimators. Other names by which the AR estimator is known are the maximum entropy estimator and the linear prediction estimator. Although the theoretical foundations for the latter two estimators differ from those of the AR estimators, in practice all the approaches are identical. The difficulty with adopting either the maximum entropy or linear prediction philosophies is that the all-pole filter assumption implicit in both approaches is not highlighted. As a consequence, application to non-AR time series usually results in poor quality estimates with no clues provided as to the reasons why. The AR modeling is the vehicle used to describe this class of high resolution estimators. In the prediction of time series data, we have a single function. Normally in eqn. (2.3) the input $u(n)$ is immeasurable and assumed to be independent and identically (i. i. d.) white noise process. In the prediction of actual noise free data, IHR in our case, let us assume that $u(n) = 0$ for all n of interest. Let us assume that we are provided with data as $x[m]$, $m= 0, 1, 2,\ldots, n-1$ and we need to predict $x[n]$. In that case the AR model can be written as,

$$x[n] = -\sum_{i=1}^{P} a_P[i]x[n-i] \tag{2.5}$$

where a current data is a linear function of P (known as model order) former data multiplied by the AR parameters $a_P[i]$. To design an effective AR model for prediction, we need to calculate model order P and AR parameters $a_P[i]$, $i= 1,2,3,\ldots\ldots\ldots,P$. There are various techniques to determine the values of $a_P[i]$ and P.

2.2 Determination of AR Parameters

The methods normally used to estimate AR parameters are the Yule-Walker method, the Burg method, unconstrained least-squares methods and sequential estimation method. In the Yule-Walker and unconstrained least-squares methods, we simply calculate the autocorrelation from the data and use it to solve for the AR parameters [21]. Sequential estimation of AR parameters is used in situations where data are available on a continuous basis and the estimates can be updated as new data points become available [21]. In contrast to the other three methods which estimate the AR parameters directly, the Burg method [17] estimates the reflection coefficients from the data and uses Levinson recursion [22]. Due to some advantages, the Burg method has been used in this study for the determination of AR parameters. It's major advantages are:

(a) It results in high frequency resolution.

(b) It yields a stable AR model.

(c) It is computationally efficient.

The Burg method for estimating the AR parameters may be viewed as an order recursive least square lattice method based on the minimization of the forward and backward errors in linear predictors, with the constraint that the AR parameters satisfy the Levinson recursion. The reflection coefficients used in determining AR parameters are obtained by minimizing prediction error power for different order predictors in a recursive manner. Specially, based on the Levinson recursion algorithm, if estimates of the reflection coefficients $k_1, k_2, k_3, \ldots, k_P$ are available, the AR parameters may be obtained as follows:

The autocorrelation of N data at 0 lag is,

$$r_{xx}[0] = \frac{1}{N}\sum_{n=0}^{N-1}|x[n]|^2 \tag{2.6}$$

The first AR parameter is,

$$a_P[1] = k_1 \tag{2.7}$$

For $k=2, 3, \ldots, P$

$$a_k[i]=a_{k-1}[i]+ k_k a_{k-1}[k-i], \qquad i=1, 2, \ldots, k-1 \tag{2.8}$$

and $a_k[i]=k_k$ for $i=k$. The AR filter parameters are $a_P[1], a_P[2], \ldots, a_P[P]$. It remains only to calculate the reflection coefficients. In deriving the kth reflection coefficient, Burg assumed that the $(k-1)$st order prediction error filter coefficients had already been calculated as $a_{k-1}[1], a_{k-1}[2], \ldots, a_{k-1}[k-1]$, having been obtained by minimizing $(k-1)$st order prediction error power. Burg proposed to calculate k_k by minimizing average of the estimates of the forward and backward prediction error powers. Defining forward and backward prediction error powers as ρ^f_k and ρ^b_k, respectively, we need to minimize the average error power ρ_k, where ρ_k is,

$$\rho_k = \frac{\rho_k^f + \rho_k^b}{2} \tag{2.9}$$

Let us assume that we have an AR model with model order k and $\hat{x}$[n] and $\hat{x}$[n-k] be the forward and backward predicted values of the data x[n] and x[n-k], respectively. Then we may write,

$$\hat{x}[n] = -\sum_{i=1}^{k} a_k[i]x[n-i] \tag{2.10}$$

and,

$$\hat{x}[n-k] = -\sum_{i=1}^{k} a_k[i]x[n-k+i] \tag{2.11}$$

Defining forward and backward errors of prediction as e^f_k and e^b_k, respectively, we can write,

$$e_k^f[n] = x[n] - \hat{x}[n] = x[n] + \sum_{i=1}^{k} a_k[i]x[n-i] \tag{2.12}$$

and,

$$e_k^b[n] = x[n-k] - \hat{x}[n-k] = x[n-k] + \sum_{i=1}^{k} a_k[i]x[n-k+i] \tag{2.13}$$

The forward and backward prediction error powers ρ^f_k and ρ^b_k then become,

$$\rho_k^f = \frac{1}{N-k}\sum_{n=k}^{N-1}\left|e_k^f[n]\right|^2 \tag{2.14}$$

$$\rho_k^b = \frac{1}{N-k}\sum_{n=k}^{N-1}\left|e_k^b[n]\right|^2 \tag{2.15}$$

The lattice filter relations which describe the model order update of the forward and backward prediction error time series can be obtained from Levinson recursion algorithm [23]. These are,

$$e_k^f[n] = e_{k-1}^f[n] + k_k e_{k-1}^b[n-1] \tag{2.16}$$

and,

$$e_k^b[n] = e_{k-1}^b[n-1] + k_k e_{k-1}^f[n] \tag{2.17}$$

where,

$$e_0^f[n] = e_0^b[n] = x[n]$$

When the relations of eqns. (2.16) and (2.17) are substituted into eqns. (2.14) and (2.15) and then into (2.9), we obtain,

$$\rho_k = \frac{1}{2(N-k)} \sum_{n=k}^{N-1} [\left| e_{k-1}^f[n] + k_k e_{k-1}^b[n-1] \right|^2 + \left| e_{k-1}^b[n-1] + k_k e_{k-1}^f[n] \right|^2] \qquad (2.18)$$

Differentiating ρ_k with respect to k_k and solving for k_k, we obtain,

$$k_k = -2\sum_{n=k}^{N-1} e_{k-1}^f[n] e_{k-1}^b[n-1] / \sum_{n=k}^{N-1} (\left| e_{k-1}^f[n] \right|^2 + \left| e_{k-1}^b[n-1] \right|^2) \qquad (2.19)$$

Equation (2.19) is used for calculating reflection coefficients k_k and the AR parameters $a_k[i]$ are determined by equation (2.8).

2.3 Determination of AR Model Order

One of the most important aspects of the use of the AR model is the selection of the order P. As a general rule, if we select a model with too low an order, we obtain a highly smoothed data fit. On the other hand, if P is selected too high, we run the risk of introducing spurious low-level peaks in the data variation. One indication of the performance of the AR model is the mean square value of the residual error, which, in general, is different for each of the estimators described above. The characteristics of this residual error are that it decreases as the order of the AR model is increased. We can monitor the rate of decrease and decide to terminate the process when the rate of decrease becomes relatively slow. It is apparent, however, that this approach may be imprecise and ill-defined, and other methods should be investigated. Much works have been done by the various researchers on this problem and many experimental results have been given in the literature.

Two of the better known criteria for selecting the model order have been proposed by Akaike [24] - [25]. With the first, called the final prediction error (FPE) criterion [24], the order is selected to minimize the performance index,

$$FPE(p) = \sigma^2 \left(\frac{N+p-1}{N-p+1} \right) \qquad (2.20)$$

Where σ^2 is the estimated variance of the prediction error. This performance index is based in minimizing the mean-square error for a one-step predictor.

The second criterion proposed by Akaike [25] is called the Akaike information criterion (AIC) statistics which is based on selecting the order that minimizes,

$$AIC(P) = \ln \sigma^2 + \frac{2P}{N} \qquad (2.21)$$

Note that the term σ^2 decreases and therefore $ln(\sigma^2)$ also decreases as the order of the AR model is increased. However, $2P/N$ increases with an increase in P. Hence a minimum value is obtained for some P.

2.4 Statistical Analysis of Data

When a time-series data is predicted by any method, it is imperative to test the validity of that model by comparing predicted data with the corresponding original one. Statistical analysis of hypothesis testing at a given level of significance is very often used for this purpose. Analysis of variance (ANOVA) [26] is the technique for testing the equality of several means simultaneously. ANOVA splits the total variation of data into meaningful components that measure different sources of variation. If the observations are classified according to the single criterion, such as IHR, the method is called one-way classification of ANOVA. The detailed method of hypothesis testing using one-way classification of ANOVA is provided in appendix (Appendix A).

2.5 Calculation of Heart Rate

ECG was measured from 5 healthy human subjects aged 22-30 years. The ECG was stored in a microcomputer through a 12-bit A/D converter. The digitized signals were temporarily saved on the hard-disk driver and then transferred to a file server. The frequency at which the signals are to be sampled may cause errors in the calculation of IHR [27]. Due to the limited capacity of the computer memory, the signals were sampled at a low frequency of 50 Hz. On the other hand, in order to calculate the IHR with an inaccuracy of less than one beat/min (bpm), the sampling frequency (f) which satisfies the following relation is needed,

$$f = \frac{2(60 - \Delta t)}{(\Delta t)^2} \tag{2:22}$$

where $\Delta t = 60/(IHR)$. For example, for a maximum IHR value of 120 bpm, the minimum sampling frequency is 476 Hz. The signals sampled at 50 Hz were restored by the *sinc* function,

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin[2\pi \ f_M \ (t - nT)]}{2\pi \ f_M \ (t - nT)} \tag{2.23}$$

where T is sampling interval (i.e., *20* msec) and f_M is Nyquist's reflection frequency (25 Hz). In order to minimize time of calculation, the signal restoration was made only for the period of 40 msec when three points sampled were found to constitute a peak of ECG (R wave) corresponding to each heartbeat. After restoration of the R wave, the time interval between two successive peaks was determined and the IHR was then calculated from the time interval. The number of sampled points used for restoration by eqn. (2.23) was determined to be 401 (i.e., n = 200) in a preliminary experiment. In the preliminary experiment, the ECGs were sampled at a frequency of 500 Hz and the IHR was calculated from the time intervals of two successive peaks (referred to as HR_{500}). Then, single values of every 80 sampling points (which correspond to the signals sampled at 50 Hz) were extracted and the ECG waves were restored by eqn. (2.23) changing n-value from 100 with a step of 5. The IHR was calculated for each wave restored with different values of n. It was found that the IHR so calculated was consistent with HR_{500} when n was more than 200 in all data sets. IHR of a person was stored for more than half an hour and in this way 5 data sets were constructed. The calculated IHR is provided in figure 2.1, where the data is presented for 30 minutes for each person. In each panel, the letter and the number indicates the subject's identification and age (in years), respectively. Table 2.1 shows the statistical characteristics of the data sets. From figure 2.1, it is seen that the data sets have different types of IHR variations. Since the variation of IHR is not well understood from the data of 30 minutes, IHR for each data set is presented in figure 2.2 for first 3 minutes for more close-up view of the data sets.

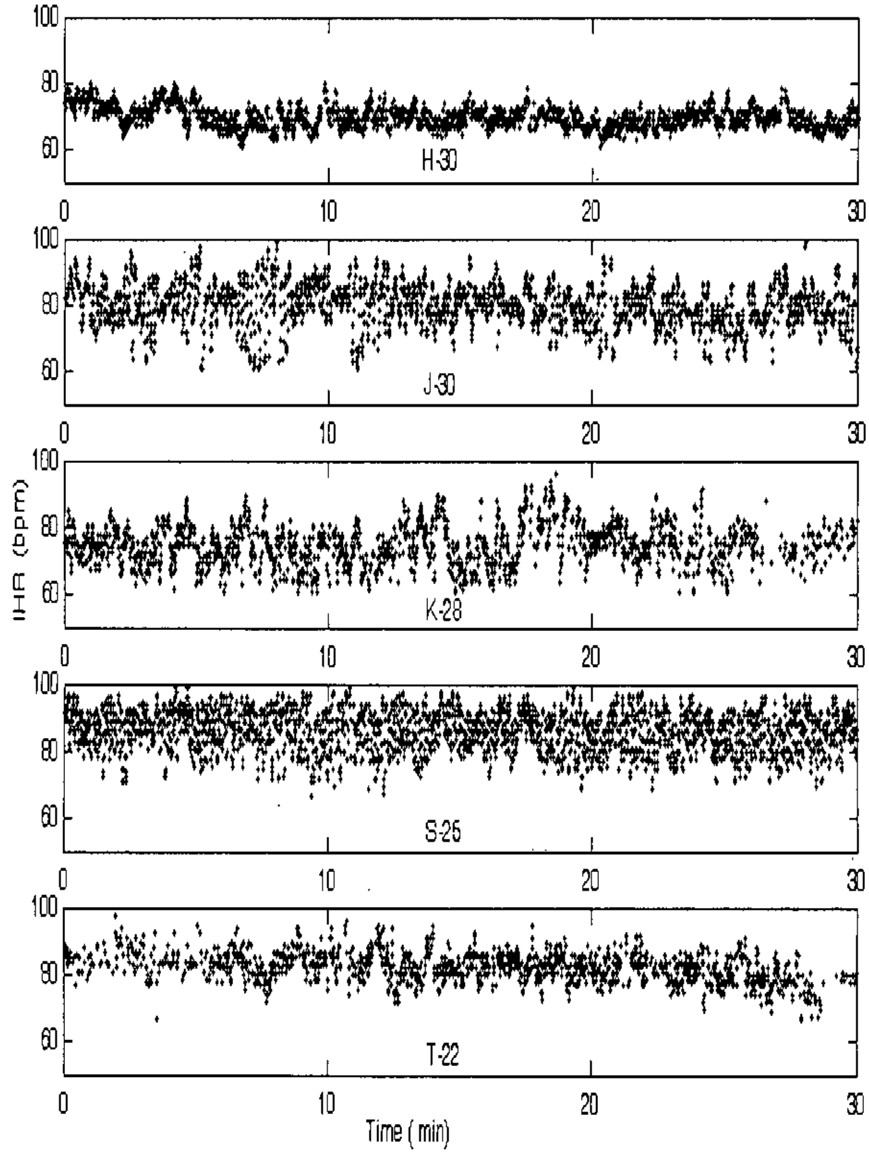

Fig. 2.1 Examples of IHRs measured from adult healthy subjects presented for 30 minutes.

Table 2.1 Mean, SD and N of measured IHR

Data set	Mean (bpm)	SD (bpm)	N
H-30	70.1	3.2	2401
J-30	80.6	6.4	2362
K-28	75.2	6.2	2112
S-25	86.7	6.1	2482
T-22	82.4	4.6	2301

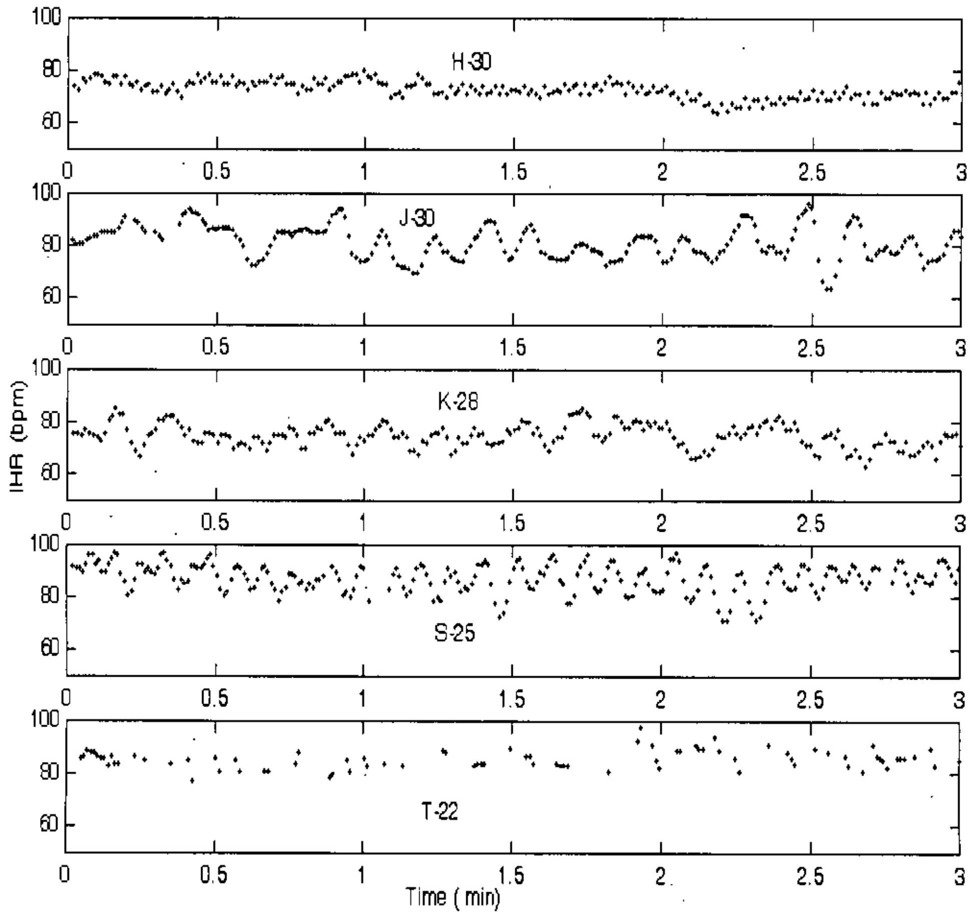

Fig. 2.2 Examples of IHRs measured from adult healthy subjects presented for 3 minutes.

2.6 Results

The AR model as described in this chapter was applied to predict the IHR. The variable $x(n)$ is each data point of IHR, i.e., $x(n) = IHR(n)$, where IHR is presented as a discrete-time data instead of a continuous one. Total number of IHR data points are taken as 2000 ($n = 0, 1, \ldots, 1999$). The AR parameters were determined by Burg method and the model order by AIC. For each data set, first 1000 IHR was provided as the input of the AR model. All the parameters of the model were then determined and with these parameters the next 1000 IHR were predicted. The original data and the error of prediction (difference between original and corresponding predicted data) are presented in Fig. 2.3, where, the upper plot shows the original data and the lower one depicts the error. Only 20 data and the respective error are presented for better visual inspection. ANOVA test as described in appendix A was applied to test the statistical significance of the differences between predicted and original IHR. There is no difference between original and predicted IHR in 3 data sets, but for the rest 2 data sets , the predicted IHR is significantly different from the original one. The mean and standard deviation (SD) (mean ± SD) of original as well as predicted IHR of data points with the model order (MO) are depicted in Table 2.2. The f-values were calculated applying one-way ANOVA technique between the predicted and original IHR.

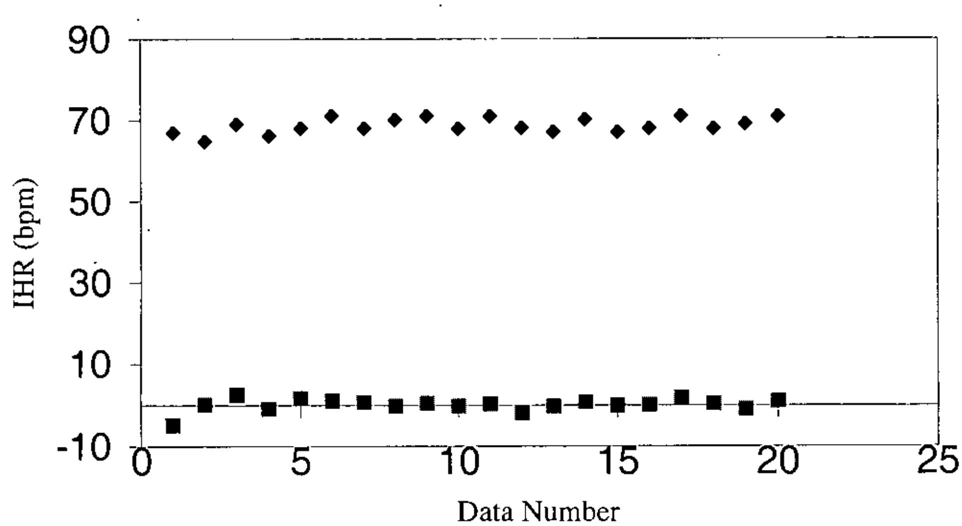

Fig. 2.3 Example of autoregressive model based prediction of IHR

Table 2.2 Comparison of original and AR predicted IHR

Data set	Original IHR (bpm)	Model order	Predicted IHR (bpm)	f
H-30	69.4 ± 2.7	4	72.5 ± 2.6	512.06
J-30	79.5 ± 5.6	6	80.2 ± 5.1	1.20
K-28	75.7 ± 6.6	6	76.2 ± 5.9	0.80
S-25	86.4 ± 5.9	4	87.8 ± 4.8	17.3
T-22	81.4 ± 4.8	4	82.5 ± 3.5	1.96

2.7 Discussion

The study presented in this chapter was performed with an interest to know the applicability of AR model in predicting HR. AR model is widely used for linear prediction of time series data. It has a highly flexible structure that can be used to parameterize the dynamics of a system. It is also a very useful tool in the spectral analysis of time series data.

The Burg method has been extensively used in calculating AR parameters due to its high frequency resolution, stability and computational efficiency. However, it has several disadvantages as well. First, it exhibits line splitting at high signal-to-noise ratios. By line splitting, we mean that the $x(n)$ may have a single sharp peak, but the Burg method many result in two or more closely spaced peaks in specially the power spectrum of the signal. For high order models, the method also introduces spurious peaks. Furthermore, for sinusoidal signal in noise, Burg method exhibits a sensitivity to the initial phase of a sinusoid, especially in short data records. This sensitivity is manifested as a frequency shift from the true frequency resulting in an frequency bias that is phase dependent frequency bias. Several modifications have been proposed to overcome some of the more important limitations of the Burg method namely the line splitting, spurious peaks, and frequency bias. Basically, the modifications involve the introduction of a weighting (window) sequence of the squared forward and backward errors. Despite all the limitations, Burg method is widely used in AR parameter determination.

The optimal selection of model order is necessary to avoid the over-flatening and spurious peaks in the power spectra of the data and hence to have an optimum model fit. Although, various techniques can be used for model order selection depending upon the data length, AIC performs similarly or better irrespective of data length [9]. Despite its bias to the data length, AIC statistics is mostly used method for model order selection.

The results show that AR model can not precisely predict time series data for all IHR data examined in this study. There is no significant difference between the original and predicted IHR in three data sets, namely, J-30, K-24 and T-22 in 95% confidence level (Critical value of F is 3.84). But, this model fails to predict IHR in rest two data sets (H-30 and J-25). Since for these 2 data sets the model order is only 4, we have tried to predict IHR by increasing P manually. No improvement of the predictive ability of AR model was found for $P<8$. However, in both cases, no significant difference between original and predicted IHR was found when P is greater than or equal to 8.

Although linear AR model is used to predict IHR, the assumption of linearity is itself a very dubious one. The theory of Volterra [28] and Wiener [29] on functional series representation has provided great stimulus to the development of non-linear models, but unfortunately Wiener's representation is too general and the statistical estimation of the Wiener kernels is unwieldy.

Linear models have many drawbacks. The application of these models inherently assumes that the modeled HR time series is linear. While linear system identification methods provide a comprehensive view of characterizing fluctuations in HR, such techniques cannot determine whether nonlinear coupling is also present. Experimental evidences indicate the presence of nonlinear interactions between the sympathetic and parasympathetic nervous system in the control of heart rate [30].

Based on the experimental findings, recent studies suggest that the HR dynamics may be nonlinear or even chaotic [31]. In the presence of nonlinearity or chaos in the HR dynamics, linear models may be inadequate. Consequently, nonlinear model structures other than the linear ones should provide a more complete description of the HR system dynamics. The nonlinear extension of ARMA model, the bilinear model, was shown to be more adequate in the modeling of HR dynamics [32] which proves that the HR dynamics is controlled by bilinear operation of the ANS. This indicates the presence of nonlinearity in the HR dynamics. If a nonlinear technique can model HR more accurately than a linear one, we can assume that there is more influence of nonlinear control on HR. Neural network (NN) modeling is such a nonlinear technique. The application of NN modeling technique is provided in next chapter.

CHAPTER 3

NEURAL NETWORK MODELING OF HEART RATE

3.1 Introduction

Artificial neural network (ANN) models have been studied for many years in the hope of achieving human-like performance in the fields of different types of signal processing. These models are composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. Computational elements or nodes are connected via weights that are typically adapted during use to improve performance. There has been a recent resurgence in the field of ANN caused by new topologies and algorithms, analog VLSI implementation techniques, and the belief that massive parallelism is essential for high performance pattern recognition. One of the most important applications of ANN is its ability to approximate any continuous function by adaptation in response to examples of the mapping action [33]. Particularly, ANN is found to be useful in predicting nonlinear and chaotic time series [34]. Typically, a set of examples of P digitized signals $x(1), x(2), x(3), \ldots, x(P)$ are presented for the prediction of $x(P+1)$. The requirement is to design a neural network that implements an approximation to the signals with a suitable function. The prediction can be performed with many types of neural networks. Two of the most extensively used neural nets are multilayer perceptron (MLP) and radial basis function (RBF) networks.

The building block of MLP is the perceptron which maps the original input vector to form a weighted sum and adds a bias value. The function by which input is mapped varies, but the mostly used function is a sigmoid nonlinear one. The capabilities of single perceptron are limited to linear decision boundaries and simple logic functions. The implementation of complex decision boundaries and arbitrary

Boolean expressions can be performed by cascading perceptrons in layers, and the resultant network is an MLP net. The individual perceptrons in the network are called neurons or nodes. MLP are feedforward nets with one or more layers of nodes between the input and output nodes. The capabilities of MLP stem from the nonlinearities used within nodes. If nodes were linear elements, then a single layer net with appropriately chosen weights could exactly duplicate those calculations performed by any multilayer net.

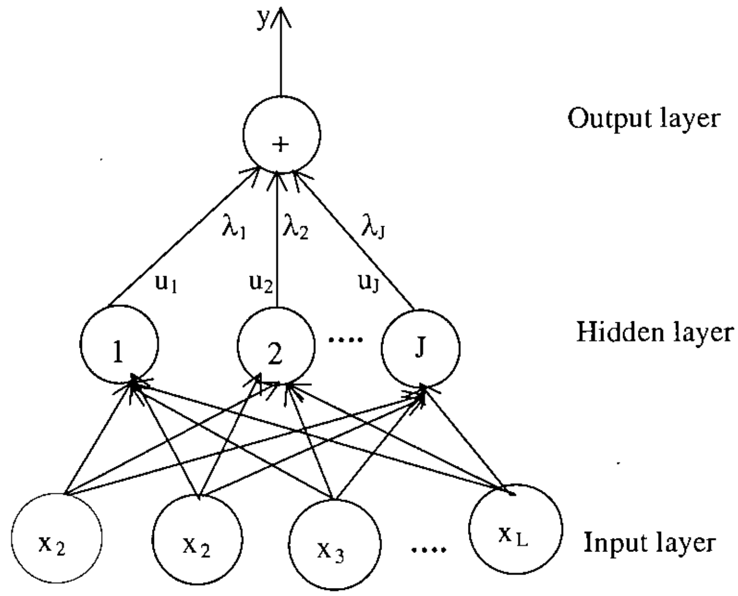

Fig. 3.1 Structure of a radial basis function neural network

An RBF network is a two-layer network as shown in figure 3.1. It has the on-line learning ability [35] which is very important to real time signal processing applications. The input layer is the source nodes that provide physical access points to the input signals. The hidden and output layers consist of computation nodes. The output nodes form a linear combination of the basis functions computed by the hidden layer nodes. For the sole purpose of data prediction, the output layer has only one node. The basis functions in the hidden layer produce a localized response to input signals. That is, they produce a significant nonzero response only when the input falls within a small localized region of the input space. The synaptic connections from the input layer to the hidden layer are nonlinear, whereas those from the hidden layer to the output layer are linear.

In this study, the RBF classifier is used to predict IHR. For comparison, the results of prediction with NN is compared to that with AR model.

3.2 Model Description

Although total number of data points provided in the input layer is L, the model takes only P (model order) points for prediction of a data. The basis functions (outputs of hidden layer) are of different types and that used in this work is most commonly used Gaussian kernel of the form:

$$u_j = \exp\left[-\frac{(x-w_j)^2}{2\delta_j^2}\right], \qquad \mathrm{j} = 1, 2, \ldots, \mathrm{J} \tag{3.1}$$

where x is the input pattern; u_j, w_j, δ_j^2 are the output, weight (center of the Gaussian) and normalization parameters (width of the Gaussian) respectively associated with the jth node in the hidden layer; and J is the number of nodes in the hidden layer. Let us assume that the output of the network is y, which is in this case x_{L+1}. Then y is given by:

$$y = \sum_{j=1}^{J} \lambda_j u_j \tag{3.2}$$

where λ is the weight vector for output node.

A number of parameters are to be determined in the training process of the network: model order (P), the center (w_j) and width (δ_j^2) of the Gaussian kernel, total number of nodes needed in the hidden layer (J) and the connectionist weights from the hidden layer to the output node (λ_j).

There are a variety of approaches to learning in the RBF network [36], [37]. Most of them start by breaking the problem into two stages: learning in the hidden layer (determination of w_j, δ_j^2 and J), followed by learning in the output layer (λ_j). The center and width can be chosen randomly or by some self-organizing method such as a clustering algorithm (unsupervised method). The learning in the output layer is supervised. Once an initial solution is found, a supervised learning algorithm is sometimes applied to both layers simultaneously to fine-tune the parameters of the network. There are numerous clustering algorithms that can be used for the determination of centers of the nodes in the hidden layer. A popular choice is the K-means clustering algorithm [38], [39]. This algorithm is perhaps the most widely known clustering algorithm because of its simplicity and good performance [40], [41]. The K-means clustering algorithm is summarized in Table 3.1.

Table 3.1 K-means clustering algorithm

Procedure K_MEANS

Initialize the cluster centers $w_j, j = 1, 2, \ldots\ldots\ldots, J$;

/* typically these are set to the first J training samples */

repeat

/* Group all patterns with the closest cluster center */

for all x_i **do**

Assign x_i **to** θ_j, **where** $w_j = min \, / \, x_i - w_j \, /$;

endloop

/* Compute the sample means */

for all w_j **do**

$$w_j = \frac{1}{M_j} \sum_{x_i \in j} x_j \quad ;$$

endloop

until there is no change in cluster assignments

from one iteration to the next ;

end ;

The width of the hidden layer nodes (δ_j^2), which represents a measure of the spread of the data associated with each node is determined once the clustering algorithm is complete. Although they can be determined in a variety of ways, the most common is to make them equal to the average distance between the cluster centers and the training pattern, that is:

$$\delta_j^2 = \frac{1}{M_j} \sum_{x \in \theta_j} (x - w_j)^2 \qquad (3.3)$$

where θ_j is the set of training patterns grouped with cluster center w_j and M_j is the number of patterns in θ_j. The learning in the output layer is performed after the parameters of the basis functions have been determined; that is, after learning in the first layer is complete for a fixed number of hidden layer nodes.

Table 3.2 Least square (LS) algorithm

```
Procedure LS
Initialize the weights, λj, to small random values for all j;
        repeat
                Choose next training pair (u, y)
                 /* Compute outputs */
                y = λ^T u ;
                /* Compute mean square errors */
                e = (y-d)^2 ; /* d is actual output */
                /* Update weights */
                for all j do
                    λj = λj + μeuj ;   /* μ is a constant in the range
                                        0≤ μ ≤ 1; μ decreases as the
                                        iteration proceeds */
                endloop
        until LS is found ;
end ;
```

The output layer is typically trained using the least square (LS) algorithm [29], [30]. The training set consists of input/output pairs (u, y) as before, but now the input patterns are processed by the hidden layer before being presented to the output layer for use in the training algorithm. Table 3.2 summarizes the LS algorithm. The number of nodes in the hidden layer (J) can be determined by a trial and error method or by the use of AIC as described in eqn. (2.21) of chapter 2. The model order can also be determined by AIC statistics.

3.3 Results

The same 5 data sets described in chapter 2 are subjected to prediction. For each data set, first 1000 IHR was provided as the input of the NN model and all the parameters of the model were determined by the training of the network. In the training phase, the clustering algorithm takes data in groups, i.e., if we provide model order as P, first P points are taken in the first iteration. In the next iteration, the first data is discarded and 2nd to $P+1$ data points are taken. In this way, the algorithm takes all the 1000 IHR data for the determination of the NN parameters. Although AIC statistics can be used for determining model order, we used the same model orders as found by AR modeling for all the 5 data sets studied. In the prediction phase, although 1000 IHR points are provided as input, the model takes only P number of data for prediction. For the prediction of 1001st IHR, the input are the IHR of points $1000\text{-}P+1$ to 1000, and so on. In general, if we want to predict $IHR(1000+n)$, the input data are $IHR(1000\text{-}P+n)$ to $IHR(1000+n\text{-}1)$, where $n = 1, 2, \ldots, 1000$. In this way, all the data in the range of $1001\text{-}2000$ points were predicted. Figure 3.2 depicts a typical example of neural network based prediction result where the upper plot shows the original heart rate while the corresponding prediction error (original HR-predicted HR) is shown in lower plot. ANOVA technique was applied to test the statistical significance of the differences between predicted and original IHR. The original and predicted data characteristics of the 5 data sets are presented in Table 3.3 with the number of hidden layer nodes. It is observed that there is no difference between original and predicted IHR in all the 5 data sets ($F_{critical} = 3.84$).

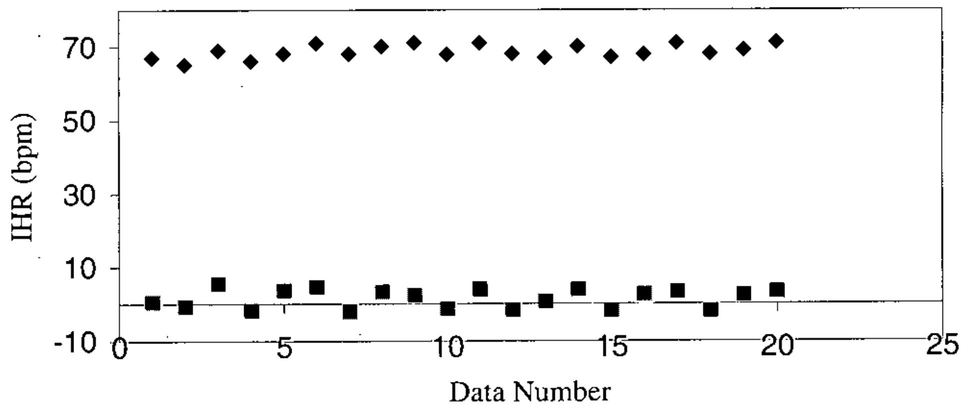

Fig. 3.2 Example of neural network model based prediction of IHR

Table 3.3 Comparison of original and NN predicted IHR

Data set	Original IHR (bpm)	No. of nodes in hidden layer	Predicted IHR (bpm)	*F*
H-30	69.4 ± 2.7	3	70.4 ± 2.60	2.16
J-30	79.5 ± 5.6	2	80.4 ± 5.40	1.18
K-28	75.7 ± 6.6	3	76.3 ± 6.70	0.92
S-25	86.4 ± 5.9	3	86.9 ± 5.80	0.63
T-22	81.4 ± 4.8	3	82.5 ± 4.70	1.68

Table 3.4 Comparison of AR and NN predicted IHR

Data set	AR predicted IHR (bpm)	NN predicted IHR (bpm)	*f*	Better Model
H-30	72.5 ± 2.6	70.4 ± 2.60	340.60	NN
J-30	80.2 ± 5.1	80.4 ± 5.40	0.90	
K-28	76.2 ± 5.9	76.3 ± 6.70	0.10	
S-25	87.8 ± 4.8	86.9 ± 5.80	14.30	NN
T-22	82.5 ± 3.5	82.5 ± 4.70	0.10	

The comparison of the predicted data with NN and AR models are provided in table 3.4. The ANOVA was applied to the predicted data with the two models. The results indicate that NN model performs better in 2 data sets, namely H-30 and S-25, while statistically similar performance was obtained with the both techniques in the rest 3 data sets (J-30, K-24 and T-22).

Since the HR was calculated using a man-made nonlinear equation, nonlinearity may prevail due to this. To verify whether nonlinear control of ANS is present in the HR dynamics, the same techniques of prediction were applied to the beat-to-beat time difference (Δt) of each data set. Same types of results were obtained, that is, NN performed better in the prediction of beat-to-beat interval for the same 2 data sets (H-30 and S-25) while both the methods performed statistically similarly for the rest 3 data sets.

3.4 Discussion

The present study was performed with an interest to know nature of the IHR. Since ANS modulates IHR, the predictive modeling also provides a way to understand the controlling mechanism of ANS. Due to its robustness in the predicting ability neural network has been used to differentiate between random and deterministic processes [42]-[44]. Also, it is able to detect signals corrupted with unknown noise [45]. This study shows that nonlinear NN model works better in the prediction of IHR. Only the predictive ability of NN model does not establish that the control of ANS on the HR is nonlinear, but it gives an indication that the HR dynamics may contain nonlinearity. Although the number of data sets analyzed in this study is very few (only 5), it gives an indication that future studies of HR dynamics analysis should focus on the nonlinear phenomenon of HR controlling mechanisms.

Since the results found in this study indicate the presence of nonlinearity in HR dynamics, it is imperative that other nonlinear techniques such as fractal and chaotic models should provide the capability to distinguish between normal and pathologic HR dynamics. The application of chaos theory is depicted in the next chapter.

CHAPTER 4

CHAOTIC MODELING OF HEART RATE

4.1 Introduction

Recently, chaos theory is having its popularity to many researchers who work for the analysis of nonlinear data for which most mathematical models produce intractable solutions. The basic idea behind the chaos theory is that any iterative function does not suddenly become chaotic. Rather, it goes from the stage of convergence to the stage of bifurcations and then finally to chaotic. First introduced by May [46], considerable works have been done on the theoretical aspects of chaos [47], [48]. Its application area ranges from metrology [46] to medicine and biology [49]. HR dynamics may exhibit chaos [31], [49]. Some works on the modeling of HR dynamics using chaos have been recently reported [50]-[52].

Chaotic systems are by nature variable, and this variability serves as an important mechanism for adaptability and flexibility. Such plasticity is essential for coping with the exigencies of an unpredictable and changing environment. The absence of a characteristic time scale for heart rate fluctuations, for example, broadens the frequency response of the cardiovascular system and prevents the system from being locked into one frequency response. Recent evidence from a number of laboratories suggests that the normal dynamics of other physiological systems are also chaotic [53]. Evidence for the existence of chaos in the central nervous system is also supported by dimensional analysis of electroencephalogram (EEG) waveforms of healthy individuals [54]. Chaos in neuroendocrine function is suggested by time series analysis of serum hormone levels in healthy human subjects [48]. Such dynamics do not display the regularity expected by a classical homeostatic system. The typical homeostatic systems have the characteristics such as,

1. System will settle down to a constant steady state if unperturbed.
2. Fluctuations result from external influences.

3. Destabilizing factors such as disease or aging are anticipated to increase order (increase chaos).

On the other hand, for a chaotic system,

1. System does not settle down to a constant steady state.
2. Fluctuations arise from the internal feedback and do not require external perturbation.
3. Destabilizing factors such as disease or aging usually decrease the degree of complex variability (reduce chaos).

Based on the assumptions that the HR is nonlinear and complex, an approach has been taken to quantify the variability of the HR dynamics. Two methods have been applied to achieve this objective. The first one involves the graphical representation of the HR time series data with the help of phase space portrait while the other is the measurement of central tendency that quantifies the degree of variability.

4.2 Phase Space Portrait

The first step in the description of nonlinear processes is usually an attempt to reconstruct its phase space portrait. The general technique of this approach is to generate several different scalar signals $x_k(n)$ from the original time series $x(n)$ in such a way as to reconstruct an m-dimensional space, where, under some conditions, we can obtain a good representation of the attractor of the dynamical system which can help to visualize the possibility of the chaos. The simplest and most popular way to do this is to use time delays. We can write,

$$x_k(n) = x(n+(k-1)D),\ \ k = 1,2,\ldots,m \tag{4.1}$$

where, D is the time delay. In this manner, an m-dimensional signal can be generated.

The most crucial part in this approach is to determine the value of D. If D is too small, then $x(n)$ and $x(n+D)$ are too close to serve as independent coordinates. On the other hand, if D is too large, chaos may make $x(n)$ and $x(n+D)$ disconnected. So, optimum choice of D is necessary. Normally three methods are used to determine D.

They are [33]:

1. Optimal filling of the phase-space portrait method,
2. Autocorrelation function (ACF) method and
3. Mutual information method.

Out of the above mentioned methods, the autocorrelation method is extensively used for its simplicity. A suitable D is chosen as the first zero-crossing point of the autocorrelation function of the original time series. If the ACF does not cross zero, then the first local minima of it serves to the determination of D. This method was applied in determining D for each of 5 data sets of IHR time series described in chapter 2. In each case 2000 data points are taken. The obtained values of D for H-30, J-30, K-28, S-25 and T-22 data sets are 4, 6, 5, 5 and 4, respectively. A typical phase-space portrait with 2000 point IHR (of data set H-30) is provided in Fig. 4.1.

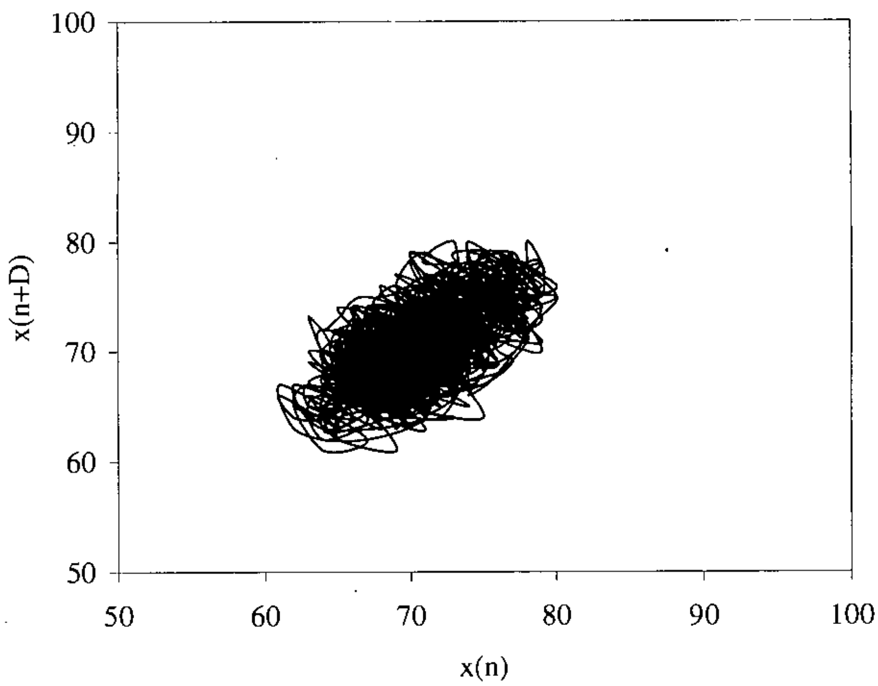

Fig. 4.1 Phase space portrait of IHR

4.3 Central Tendency Measure

Since chaos manifests itself by the irregularity or the deviation of data with respect to its average value, a quantitative measure of the relative deviation is helpful in analyzing chaotic nature of the signal. The use of central tendency measure (CTM) is a measurement of the deviation of data from the center of the image of data. This image is a two-dimensional graph of the time series data, IHR in this case, plotted as a second-order difference plot [50]. If the discrete time data is $x(n)$, the abscissa and the ordinate of the plot is $x(n+1)-x(n)$ and $x(n+2)-x(n+1)$, respectively. CTM is calculated by selecting a circular region of radius r around the origin of the difference plot, counting the number of data that fall within the circle and dividing by the total number of data. If N is the number of data, then,

$$CTM = \frac{1}{N-2}\sum_{n=0}^{N-3} C_n \tag{4.1}$$

where, C_n is the count of data points within r, and is defined as,

$$C_n = 1 \quad \text{if } \sqrt{[x(n+2)-x(n+1)]^2 + [x(n+1)-x(n)]^2} < r \tag{4.2}$$

$$= 0 \quad \text{otherwise}$$

The value of r is chosen depending on the character of the data. In this study the statistical ANOVA technique was applied. For all the 5 normal IHR data sets, r was increased by 1 from 1. Then f-values were calculated between the CTM for a particular value of r with that for $r+1$. When the CTM falls below the critical value (=5.32), that particular r is chosen as the radius of the circle. It is found that the optimum value of r is 8 for the normal data sets studied here.

The 2000 IHR points, corresponding to about 30 minutes, were subjected to the computation of CTM with the value of r taken as 8. The values of CTM for 5 IHR data sets are represented in Table 4.1. In an interest to know how the deviation of IHR from its center occurs when the data is corrupted with noise, random white noise was added with IHR. The noise was generated using the "random" function as built in the C programming language, with mean taken as 0 and variance as that of the original IHR. This noise incorporated data is termed as "noisy IHR". CTM was also calculated for all noisy IHR data sets and the results are provided in Table 4.1. The mean values of CTM for 5 data sets of each type are also provided in the table.

Table 4.1 CTM of normal and noisy IHR

Data set	CTM for normal IHR ($r = 8$)	CTM for noisy IHR ($r = 8$)
H-30	0.998	0.868
J-30	0.963	0.845
K-28	0.972	0.837
S-25	0.959	0.780
T-22	0.965	0.813
Mean	**0.971**	**0.829**

4.4 Discussion

In the last three decades, we have seen the emergence of a new descriptive scheme that can be used successfully to deal with many problems of the complexity of natural world. The widespread applicability of chaos makes it a tool for almost all science and other fields as well. It has already been proved that the application of chaos covered a large area of the systems with inherent nonlinear mechanisms. Chaos theory brought together biomedical, physical and mathematical sciences for work on the problems that were formerly thought to be outside some of the artificially set of ranges in each field.

It has been shown that the analysis of HR, particularly HR variability, by nonlinear dynamics can significantly improve the identification of the causes of sudden cardiac death, in comparison with the conventional linear analysis in the time or frequency domain [51]. Uncorrelated stochastic signals fill the phase space evenly, but regular deterministic signals create clear periodic trajectories. Several kinds of correlated stochastic and chaotic signals are more likely to fill the typical selected regions of the phase space than others. Irregular time series of chaotic processes can be identified advantageously by their phase space plots which show their well organized strange attractors.

The results presented in Table 4.1 show the nature of CTM for normal as well as noisy HR. The CTM with both kinds of HR was applied to the ANOVA technique which calculates the f-value to test the difference of CTM of normal HR with noisy one. The CTM of normal IHR (0.971 ± 0.016) has statistically significant difference ($f = 74.55$) from that of noisy IHR (0.829 ± 0.034), where $F_{critical} = 5.32$.

Table 4.2 Data characteristics and CTM of apnea IHR

Data set	Duration (sec)	Mean of IHR (bpm)	SD of IHR (bpm)	IHR length	CTM ($r = 8$)
1	151	104.8	11.8	261	0.178
2	139	113.1	21.2	247	0.269
3	119	109.7	9.6	216	0.196
4	108	126.9	14.2	225	0.292
5	144	132.9	17.1	313	0.215
6	122	108.3	10.1	218	0.236
7	105	116.8	15.1	201	0.216
8	106	126.7	19.6	218	0.185
9	103	114.0	18.7	191	0.222

Since the addition of only random noise with IHR may not indicate abnormality or disease, we tried to apply the method of CTM to truly abnormal heart rate. To collect abnormal ECG, the help of the web available ECG data base of the Physionet [55] was considered. Out of the huge data base, only partial data of apnea ECG of 9 data sets could be down loaded. The duration of each data set varies. From the ECG, IHR was calculated applying the same methodology as described in chapter 2. The duration, mean, SD, number of IHR data points (N) of each data set are provided in Table 4.2. The calculated CTM for $r = 8$ for each data set is also provided in Table 4.2. For these apnea data sets, the mean and SD of CTM is found to be *0.223* and *0.038*, respectively. Application of ANOVA between CTM of normal IHR (5 data) and of apnea IHR (9 data) shows that the two CTM values differ significantly ($f = 1751.0$, $F_{critical} = 4.75$).

Although phase space portrait and second-order difference plot were used by many researchers to determine the chaotic nature of ECG and HR dynamics, Cohen et. al. [50] first applied CTM to differentiate between normal and pathologic cardiac signals. In this work, the parameter used as cardiac signal is beat-to-beat interval (BBI), the time difference (Δt) between successive R-waves of ECG. The authors took r as *0.1* in the calculation of CTM of BBI, and CTM of normal and congestive heart failure (CHF) subjects were calculated, although the method of determining r is not presented. Since the optimum selection of r is very much crucial to use CTM as a diagnostic tool, we used the ANOVA technique to calculate it. Using ANOVA to the BBI of 5 normal data sets, we found that the optimum value of r is *0.07*. In an attempt to compare our results of CTM with that of Cohen et. al. [50], we measured CTM of normal and apnea BBI taking r as *0.1* and *0.07*. The results are summarized in Table 4.3.

Some interesting points are to be noted from Table 4.3. First, the trend of variation of CTM for these two r is same for normal BBI, but very much uncorrelated for apnea BBI. This is due to the fact that the variance of IHR as well as BBI for normal subjects is less. On the other hand, the IHR and BBI variation for apnea subjects are high, a slight variation of r may result significant different in CTM. Secondly, the mean CTM of normal BBI is almost same as that of normal IHR. This can be explained in term of the coefficient of variation (CV), defined as the ratio of SD to mean. The average CV of IHR is 0.0315 and that of BBI is 0.0321. Since the variations of two parameters are very close, their CTMs must also be almost equal. For $r = 0.1$, the CTM of normal BBI is found to be 0.975 ± 0.018 $(N=5)$ and that of apnea BBI is 0.591 ± 0.048 $(N=9)$, the two CTM values are significantly different

($f=283.36$). For the case of $r = 0.07$, the CTMs of normal and apnea BBI are 0.902 ± 0.040 and 0.355 ± 0.060, respectively, and f is found to be 331.59. Both the r can differentiate apnea BBI from the normal one. Since the value of f is greater for $r=0.07$, we can conclude that it has greater ability of distinguishing between normal and abnormal cardiac signals.

Table 4.3 CTM of normal and apnea BBI

Normal BBI			Apnea BBI		
Data set	CTM ($r = 0.1$)	CTM ($r = 0.07$)	Data set	CTM ($r = 0.1$)	CTM ($r = 0.07$)
H-30	0.998	0.953	1	0.548	0.378
J-30	0.964	0.919	2	0.514	0.425
K-28	0.952	0.845	3	0.622	0.290
S-25	0.975	0.891	4	0.655	0.453
K-22	0.984	0.904	5	0.624	0.376
Mean	**0.975**	**0.902**	6	0.583	0.292
			7	0.558	0.312
			8	0.648	0.370
			9	0.571	0.302
			Mean	**0.581**	**0.355**

Since both the normal as well as diseased heart rate can be chaotic [31], [50], it is still difficult to find the condition of heart from the phase space portrait and the dimension of the manifold associated with this. On the other hand, the quantitative description of CTM may provide a better way to distinguish between different cardiac abnormalities. To provide a strong base to this seemingly powerful tool, CTM should be applied to HR as well as ECG of the persons with different varieties of cardiac malfunctions.

CHAPTER 5

CONCLUSIONS

5.1 Discussions

This work describes the prediction of heart rate with linear and nonlinear methods. Every method has advantages and limitations in terms of predicting ability, computational complexity, learning process, speed, etc. Out of the many linear methods of prediction, AR model is the most widely used one. Although ARMA model can provide superior prediction in comparison to AR model, the computational complexity of finding ARMA parameters makes it time consuming and less suitable for on-line prediction of data. There are many methods to determine the AR parameters. Out of them the Burg method is widely used due to its computational simplicity, since it is based on the recurrence of autocorrelation function. The Burg method uses the minimization technique of average error due to forward and backward prediction, and involves the solution of a set of linear difference equations. The major advantages of the Burg Method for estimating the parameters of the AR model are:

- It results in high frequency resolution.
- It yields a stable AR model.
- It is computationally efficient.

Out of the many methods of AR model order determination, AIC statistics is widely used due to its superior performance over other ones. All the other methods have the drawback that they are very much sensitive to data length. On the other hand, AIC statistics performs better or at least similarly irrespective of data length.

RBF neural network is widely used for the following advantages:

- It has on-line learning ability

- It can be trained by linear adaptive algorithms
- Its regularization ability makes it a robust technique in all ill-posed problem such as model fitting from data

Although linear models are widely used for the prediction of HR, presently there is indication that heart rate may be nonlinear or even chaotic. But still, linear models provide a comprehensive view of the mechanism generating and controlling ECG and subsequently HR. AR model has also the ability to correlate pathophysiological mechanisms to HR dynamics. The predictive modeling of the 5 normal IHR data sets proves that nonlinear NN model performs similarly or better than linear AR model. This may indicate the presence of nonlinearity in the HR dynamics.

The second attempt of this study is to apply chaos theory to differentiate normal HR from abnormal one. The measure of variability, i.e., CTM is calculated here. The results show that CTM can distinguish abnormal HR from the normal one. Here, the optimum selection of r is necessary. Statistical analysis as presented in this study or other such methods can be employed for this purpose. Despite the noted ability of CTM, it is still premature to use it as a diagnostic tool since higher HR variability can be seen in a normal subject too due to the natural causes.

5.2 Future Perspectives

The HR can very widely even for a short period of time, and even for the case of healthy subjects. Although HR is a major indicator of the cardiac condition, it does not provide any indication about the relative change in magnitude and/or duration of each wave of ECG. The waves and complexes of ECG should be considered before taking any decision regarding the condition of the heart. The pattern recognition techniques can play an important role in analyzing rhythm disturbances.

Appendix A

Analysis of Variance

Random samples of size n are selected from each of k populations. It is assumed that the k populations are independent and normally distributed with means μ_1, μ_2, ..., μ_k and a common variance σ^2. We wish to derive appropriate methods for testing the following hypotheses:

Null hypothesis $\quad H_0: \mu_1 = \mu_2 = ... = \mu_k$,

Alternate hypothesis $\quad H_1$: At least two of the means are not equal.

Let x_{ij} denotes the jth observation from the ith population as arranged in Table A1. Here $T_{i.}$ and $\bar{x}_{i.}$ are the total and mean, respectively of all observations in the sample from the ith population, $T_{..}$ and $\bar{x}_{..}$ are the total and mean, respectively of all nk observations.

Table A1: Random samples

	Population				
	1	2	i	k	
	x_{11}	x_{21}	x_{i1}	x_{k1}	
	x_{12}	x_{22}	x_{i2}	x_{k2}	
	.	.	.	.	
	.	.	.	.	
	x_{1n}	x_{2n}	x_{in}	x_{kn}	
Total	$T_{1.}$	$T_{2.}$	$T_{i.}$	$T_{k.}$	$T_{..}$
Mean	$\bar{x}_{1.}$	$\bar{x}_{2.}$	$\bar{x}_{i.}$	$\bar{x}_{k.}$	$\bar{x}_{..}$

From the data we compute the following:

Total sum of squares (*SST*),

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 - \frac{T_{..}^2}{nk} \tag{A1}$$

Sum of the squares for column means (*SSC*),

$$SSC = \frac{\sum_{i=1}^{k} T_{i.}^2}{n} - \frac{T_{..}^2}{nk} \tag{A2}$$

Sum of the squares for error, $SSE = SST - SSC$

$$s_1^2 = \frac{SSC}{k-1} \tag{A3}$$

If H_0 is true, s_1^2 is an unbiased estimate of σ^2. However, if H_1 is true, *SSC* will have a larger numerical value and s_1^2 overestimates σ^2. A second and independent estimate of σ^2, based on $k(n-1)$ degrees of freedom is given by,

$$s_2^2 = \frac{SSE}{k(n-1)} \tag{A4}$$

The estimate s_2^2 is unbiased regardless of the truth or falsity of the null hypothesis. When null hypothesis is true, the ratio

$$f = \frac{s_1^2}{s_2^2} \tag{A5}$$

is a value of the random variable F having the F distribution with $k-1$ and $k(n-1)$ degrees of freedom. The value of $F_{critical}$ at the α level of significance can be represented by this equation.

$$F_{critical} = f_{\alpha}[k-1, k(n-1)] \tag{A6}$$

$F_{critical}$ can be found by using F distribution table [26]. If the computed value of f is less than or equal to $F_{critical}$, the null hypothesis is accepted. If $f > F_{critical}$, the null hypothesis is rejected.

BIBLIOGRAPHY

[1] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Barger and R. J. Cohen, "Power spectrum analysis of heart rate fluctuation: a quantitative probe to beat-to-beat cardiovascular control," *Science*, v+ol. 213, pp. 220-222, 1981.

[2] L. Ljung, *System Identification: Theory for the user*, Englewood cliffs, NJ: Prentice Hall, 1997.

[3] G. E. P. Box and G. M. Jenkins, *Time series, Forecasting and Control*, San Francisco: Holden-day, 1970.

[4] T. W. Anderson, *The Statistical Analysis of Time Series*, New York and London: Wiley, 1971.

[5] E. J. Hannan, *Time Series Analysis*, London: Methuen Monographs, 1962.

[6] E. J. Hannan, *Multiple Time Series*, New York and London: Wiley, 1970.

[7] T. S. Rao and M. M. Gabr, "A test for Linearity of Stationary Time Series," *J. Time Series Analysis*, vol. 1, no. 2, 1981.

[8] T. Kailath, "A view of three decades of linear filter theory," *IEEE Trans. Information Theory*, vol. 20, pp. 146-181, 1974.

[9] N. Weiner, *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, Wiley, New York, 1949.

[10] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, vol. 82D, pp. 35-45, 1960.

[11] R. E. Kalman and R. S. Bucy, "New results in linear filtering theory," *Trans. ASME, J. Basic Eng.*, vol. 83, pp. 95-108, 1961.

[12] J. Makhoul, "Linear prediction: a tutorial review," *Proc. IEEE*, vol. 63, pp. 561-580, 1975.

[13] B. Friedlander, "Lattice filters for adaptive processing," *Proc. IEEE*, vol. 70, pp. 829-867, 1982.

[14] S. Haykin, *Adaptive Filter Theory*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1991.

[15] J. D. Markel and A. H. Gray Jr., *Linear Prediction of Speech*, Springer-Verlag, New York, 1976.

[16] S. A. Tretter, *Introduction to Discrete-Time Signal Processing*, Wiley, New York, 1976.

[17] S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, Englewood Cliffs: NJ, 1988.

[18] S. L. Marple Jr., *Digital Spectral Analysis with Applications*, Prentice Hall, Englewood Cliffs, NJ, 1987.

[19] E. A. Robinson and S. Treitel, *Geophysical Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1980.

[20] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, New York, NY: McGraw-Hill, 1984.

[21] J. K. Proakis and D. G. Malonakis, *Digital Signal Processing: Principles, Algorithm and Applications*, Prentice-Hall (India), 1997.

[22] J. P. Burg, "A new analysis technique for time series data," *NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics*, pp. 17-23, August, 1968. Reprinted in *Modern Spectrum Analysis*, D. G. Childers ed., IEEE Press, New York.

[23] J. Durbin, "Efficient estimations of parameters in moving average models," *Bio-metrica*, vol. 46, pp. 306-316, 1959.

[24] H. Akaike, "Power spectrum model fitting through autoregressive model fitting," *Ann. Inst. Stat. Math.*, vol. 21, pp. 407-409, 1969.

[25] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. on Automatic Control*, vol. 19, pp. 716-723, 1974.

[26] R. E. Walpole, *Introduction to Statistics*, Collier Macmillan, 2nd Ed.

[27] M. Merri, D. C. Farden, J. G. Mottley and E. L. Titlebaum, "Sampling frequency of electrocardiogram for spectral analysis of the heart rate variability," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 99-106, 1990.

[28] V. Volterra, *Theory of Functionals and an Integral and Integrodifferential Equations*, London: Dover, 1930.

[29] N. Wiener, *Nonlinear Problems in Random Theory*, London: Wiley, 1958.

[30] M. N. Levy, "Sympathetic-parasympathetic interactions in the heart," *Circ. Res.*, vol. 29, pp. 437-445, 1971.

[31] A. L. Goldberger, "Is normal heart beat chaotic or homeostatic?," *News in Physiol. Sci.*, vol. 6, pp. 87-91, 1991.

[32] D. J. Christini, M. Bennett, K. R. Lutchen, H. M. Ahmed, J. M. Hausdorff and N. Oriol, "Application of linear and nonlinear time series modeling to heart rate dynamics analysis," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 411-415, 1995.

[33] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, vol. 78, pp. 1481-1497, 1990.

[34] D. S. Broomhead and D. Lowe, "Multivariate functional interpolation and adaptive networks," *Complex Systems*, vol. 2, pp. 321-355, 1988.

[35] J. Moody and C. J. Darken, "Fast learning in networks of locally tuned processing units," *Neural Computation*, vol. 1, pp. 281-294, 1989.

[36] S. Chen, C. F. N. Cowan and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Network*, vol. 2, pp. 302-309, 1991.

[37] M. T. Musavi, W. Ahmed, K. H. Chan, K. B. Faris and D. M. Hummels, "On the training of radial basis function classifiers," *Neural Networks*, vol. 5, pp. 595-603, 1992.

[38] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley: New York, 1973.

[39] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*, Addison Wesley, Reading, MA, 1974.

[40] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, vol. 4, pp. 4-22, 1987.

[41] D. R. Hush and B. G. Horne, "Progress in supervised neural network," *IEEE SP Mag.*, vol. 10, pp. 8-39, 1993.

[42] J. B. Elsner, "Predicting time series using a neural network as a method of distinguishing chaos from noise," *Science*, vol. 257, pp. 1230-1235, 1992.

[43] M Casdagli, "Nonlinear prediction of chaotic time series," *Physica D*, vol. 35, pp. 335-356, 1989.

[44] H. Leung and T. Lo, "Chaotic radar signal processing over the sea," *IEEE J. Ocean. Eng.*, vol. 18, pp. 287-295, 1993.

[45] D. M. Hummels, W. Ahmed and M. T. Musavi, "Adaptive detection of small sinusoidal signals in non-Gaussian noise using an RBF neural network," *IEEE Trans. Neural Network*, vol. 6, pp. 214-219, 1995.

[46] R. M. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, pp. 459-467, 1976.

[47] M. J. Feigenbaum, "Quantitative universality for a class of nonlinear transformations," *J. Stat. Physics*, vol. 19, no. 1, pp. 25-53, 1978.

[48] L. Glass and M. C. Mackey, *From Clocks to Chaos*, Princeton University Press, Princeton, 1988.

[49] A. L. Goldberger, D. R. Rigney and B. J. West, "Chaos and fractals in human physiology," *Scientific American*, vol. 262, no. 2, pp. 42-49, 1990.

[50] M. E. Cohen, D. L. Hudson and P. C. Deedwania, "Applying continuous chaotic modeling to cardiac signal analysis," *IEEE EMB Mag.*, vol. 15, pp. 97-102, 1996.

[51] D. Hoyer, K. Schmidt, R. Bauer, U. Zwiener, M. Kohler, B. Luthke and M. Eiselt, "Nonlinear analysis of heart rate and respiratory dynamics," *IEEE EMB Mag.*, vol. 16, pp. 31-39, 1997.

[52] O. Fujt and J. Holcik, "Applying nonlinear dynamics to ECG signal processing," *IEEE EMB Mag.*, vol. 17, pp. 96-101, 1998.

[53] L. Glass, "Nonlinear dynamics of physiological function and control," *Chaos,* vol. 1, no. 3, pp. 247-250, 4991

[54] R. Hornero, A. Alonso, N. Jimeno, A. Jimeno and M. Lopez, "Nonlinear analysis of time series generated by schizophrenic patients," *IEEE EMB Mag.*, vol. 18, pp. 84-90, 1999.

[55] CinC Challenge 2000 data sets: available via
http://www.physionet.org/physiobank/database/apnea-ecg/