TIME SERIES FORECASTING OF CARDIOPULMONARY SIGNALS DURING EXERCISE

A Thesis submitted to Gujarat Technological University

for the Award of

Doctor of Philosophy

in

Instrumentation & Control Engineering

By

Mitulkumar Baldevdas Patel
Enrollment No. 139997117001

under supervision of

Dr. Vipul A. Shah

GUJARAT TECHNOLOGICAL UNIVERSITY
AHMEDABAD

August–2018

TIME SERIES FORECASTING OF CARDIOPULMONARY SIGNALS DURING EXERCISE

A Thesis submitted to Gujarat Technological University

for the Award of

Doctor of Philosophy

in

Instrumentation & Control Engineering

By

Mitulkumar Baldevdas Patel
Enrollment No. 139997117001

under supervision of

Dr. Vipul A. Shah

GUJARAT TECHNOLOGICAL UNIVERSITY
AHMEDABAD

August– 2018

DECLARATION

I declare that the thesis entitled "**Time Series Forecasting Of Cardiopulmonary Signals During Exercise"** submitted by me for the degree of Doctor of Philosophy is the record of research work carried out by me during the period from JANUARY 2014 to MARCH 2018 under the supervision of **Dr. Vipul A. Shah** and this has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other University or other institution of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis. I shall be solely responsible for any plagiarism or other irregularities, if noticed in the thesis.

Signature of the Research Scholar: ……………………… Date……………

Name of Research Scholar: **Mitulkumar Baldevdas Patel**

Place: Gandhinagar

CERTIFICATE

I certify that the work incorporated in the thesis "**Time Series Forecasting Of Cardiopulmonary Signals During Exercise**" submitted by **Shri Mitulkumar Baldevdas Patel** was carried out by the candidate under my supervision/guidance. To the best of my knowledge: (i) the candidate has not submitted the same research work to any other institution for any degree/diploma, Associateship, Fellowship or other similar titles (ii) the thesis submitted is a record of original research work done by the Research Scholar during the period of study under my supervision, and (iii) the thesis represents independent research work on the part of the Research Scholar.

Signature of Supervisor: ………………………………Date: …/……/2018

Name of Supervisor: **Dr. Vipul A. Shah**

Place: Nadiad

Course-work Completion Certificate

This is to certify that Mr. Mitulkumar Baldevdas Patel enrolment no. 139997117001 is a PhD scholar enrolled for PhD program in the branch Instrumentation & Control of Gujarat Technological University, Ahmedabad.

(Please tick the relevant option(s))

☐ He/She has been exempted from the course-work (successfully completed during M.Phil Course)

☐ He/She has been exempted from Research Methodology Course only (successfully completed during M.Phil Course)

☐ He/She has successfully completed the PhD course work for the partial requirement for the award of PhD Degree. His/ Her performance in the course work is as follows-

Grade Obtained in Research Methodology (PH001)	Grade Obtained in Self Study Course (Core Subject) (PH002)
BB	BB

Supervisor's Sign

(Name of Supervisor)

Originality Report Certificate

It is certified that Ph.D. Thesis titled **"Time Series Forecasting Of Cardiopulmonary Signals During Exercise"** by **Mitulkumar Baldevdas Patel** has been examined by us.

We undertake the following:

a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.
b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.
c. There is no fabrication of data or results which have been compiled / analysed.
d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.
e. The thesis has been checked using **Turnitin software** (copy of originality report attached) and found within limits as per GTU Plagiarism Policy and instructions issued from time to time (i.e. permitted similarity index <=25%).

Signature of the Research Scholar: ……………………………Date: …./……/2018…

Name of Research Scholar: **Mitulkumar Baldevdas Patel**

Place: Gandhinagar

Signature of Supervisor: ……………………………… Date: …/……/2018…

Name of Supervisor: **Dr. Vipul A. Shah**

Place: Nadiad

Document Viewer

Turnitin Originality Report

Processed on: 02-May-2018 08:46 IST

ID: 957381376

Word Count: 19414

Submitted: 1

Time Series Forecasting of cardiopulmonary re... By GEC Gandhinagar 013

Similarity Index	Similarity by Source	
3%	Internet Sources:	2%
	Publications:	3%
	Student Papers:	0%

1% match (publications)
"Filtering for Removal of Artifacts", Biomedical Signal Analysis, 2015.

1% match (Internet from 01-May-2016)
http://www.cardiology.org

1% match (publications)
Elena Baralis, Tania Cerquitelli, Silvia Chiusano, Vincenzo D'elia, Riccardo Molinari, Davide Susta. "Early prediction of the highest workload in incremental cardiopulmonary tests", ACM Transactions on Intelligent Systems and Technology, 2013

1% match (publications)
Kenneth C. Beck, Idelle M. Weisman. "Methods for Cardiopulmonary Exercise Testing", S. Karger AG, 2002

1% match (Internet from 27-Oct-2015)
http://www.ijafrse.org

PhD THESIS Non-Exclusive License to
GUJARAT TECHNOLOGICAL UNIVERSITY

In consideration of being a Ph.D. Research Scholar at GTU and in the interests of the facilitation of research at GTU and elsewhere, I, **Mitulkumar Baldevdas Patel** (Full Name of the Research Scholar) having Enrollment No. **139997117001** hereby grant a non-exclusive, royalty free and perpetual license to GTU on the following terms:

a) GTU is permitted to archive, reproduce and distribute my thesis, in whole or in part, and/or my abstract, in whole or in part (referred to collectively as the “Work”) anywhere in the world, for non-commercial purposes, in all forms of media;

b) GTU is permitted to authorize, sub-lease, sub-contract or procure any of the acts mentioned in paragraph (a);

c) GTU is authorized to submit the Work at any National / International Library, under the authority of their “Thesis Non-Exclusive License”;

d) The Universal Copyright Notice (©) shall appear on all copies made under the authority of this license;

e) I undertake to submit my thesis, through my University, to any Library and Archives. Any abstract submitted with the thesis will be considered to form part of the thesis.

f) I represent that my thesis is my original work, does not infringe any rights of others, including privacy rights, and that I have the right to make the grant conferred by this non-exclusive license.

g) If third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required, I have obtained such permission from the copyright owners to do the acts mentioned in paragraph (a) above for the full term of copyright protection.

h) I retain copyright ownership and moral rights in my thesis and may deal with the copyright in my thesis, in any way consistent with rights granted by me to my University in this non-exclusive license.

i) I further promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to my University in this non-exclusive license.

j) I am aware of and agree to accept the conditions and regulations of Ph.D. including all policy matters related to authorship and plagiarism.

Signature of the Research Scholar: ______________________________

Name of Research Scholar: **Mitulkumar Baldevdas Patel**

Date:_____/______/2018_ Place: Gandhinagar

Signature of Supervisor: _____________________________________

Name of Supervisor: **Dr. Vipul A. Shah**

Date:_____/_______/2018__ Place: Nadiad

Seal:

Thesis Approval Form

The viva-voce of the Ph.D. Thesis submitted by **Shri Mitulkumar Baldevdas Patel** (Enrollment N o . **139997117001**) entitled "**Time Series Forecasting Of Cardiopulmonary Signals During Exercise**" was conducted on ………………………..………… (day and date) at Gujarat Technological University.

(Please tick any one of the following option)

☐ The performance of the candidate was satisfactory. We recommend that he/she be awarded the Ph.D. degree.

☐ Any further modifications in research work recommended by the panel after 3 months from the date of first viva-voce upon request of the Supervisor or request of Independent Research Scholar after which viva-voce can be re-conducted by the same panel again.

(Briefly specify the modifications suggested by the panel)

☐ The performance of the candidate was unsatisfactory. We recommend that he/she should not be awarded the Ph.D. degree.

(The panel must give justifications for rejecting the research work)

Name and Signature of Supervisor with Seal

----
1) (External Examiner 1) Name and Signature

----
2) (External Examiner 2) Name and Signature

----
3) (External Examiner 3) Name and Signature

ABSTRACT

Exercise-based evaluation of the cardiopulmonary system is always advisable over other invasive and costly diagnostics procedures. It has enough potential to judge the fitness of healthy persons as well as to detect symptoms of abnormality earlier. In the incremental exercise, amount of workload increases at every few minutes interval, so after initial stages, the amount of workload to the cardiac system is above average so sometimes it may overload the cardiopulmonary system and this is not tolerable every time especially when the subject is a cardiac patient. This research work presents a new approach for forecasting of cardiopulmonary signals after the premature end of the test. Here, three models are implemented which are the Adaptive filter, the Autoregressive moving average with exogenous terms (ARMAX) and the nonlinear ARX (NLARX), for time series forecasting of signals like Instant heart rate (HR) and respiration rate (RR). The models are implemented such that it utilizes subject's own past and current responses to forecast future response, there is no need of another database for training or any other purpose. Performance of these models are tested on normal as well as abnormal cardiac subjects. The normal database is collected with the help of young subjects facing physical stress level of exercise protocol BRUCE on Treadmill machine and abnormal cardiac patients database are collected from the online physiological database of PhysioNet. After validation of results and statistical analysis of forecasting errors, the NLARX model is found to be more accurate and reliable for forecasting of cardiopulmonary signals like Instant HR and Instant RR for both kinds of subjects, Normal as well as Abnormal. The performances of these three models ranked in descending order are: NLARX, ARMAX, and Adaptive filter. Additionally, the prediction algorithm is implemented for single step and multi-step-ahead (five step ahead) prediction of Instant RR with the help of artificial neural network structure- nonlinear autoregressive network (NARNET), which may help to avoid any critical situation by estimating the future trend of the signal.

Acknowledgement

First and foremost, I thank the superpower God for showering me with courage, confidence and support to complete this research work.

At first, my deepest gratitude to my supervisor, **Dr. Vipul A. Shah**, Professor & Head, Instrumentation & Control Engineering Department, Dharmsinh Desai University, Nadiad, for his consistent support, supervision, guidance and inspiration during my doctoral program. His invaluable suggestions and constructive criticisms from time to time enabled me to complete my work successfully.

The completion of this work would not have been possible without, the Doctorate Progress Committee (DPC) members: **Dr. Utpal T. Pandya**, Professor, Instrumentation & Control Engineering Department, Sarvajanik college of Engineering & Technology, Surat and **Dr. Saurin Shah**, Principal, Silver Oak College of Engineering & Technology Ahmedabad. I am really thankful for their rigorous examinations and precious suggestions during my research.

I would like to express my gratitude towards **my loving Parents and Sister** for all their efforts and encouragement throughout my life. **Pooja,** My wife, who has always stood by me, for everything I need and helped me to manage our social responsibility. Special love to my loving sons: **Akshat and Yug**, for all the maturity they demonstrated through the tenure.

I want to extend my special thanks to **Prof Piyush Patel, Prof. Ghanshyam Parmar** and **Prof. Maulindu Patel** for their assist whenever required**.** Lastly, I wish to thank one and all who helped me in finishing and furnishing my research work.

Mitulkumar Baldevdas Patel
Research Scholar, Gujarat Technological University

Table of Content

List of Abbreviation

Sr. No.	Abbreviation	Full form
1	HR	Heart Rate
2	RR	Respiration Rate
3	ARMAX	Autoregressive Moving Average With Exogenous Terms
4	NLARX	Nonlinear ARX
5	ATP	Adenosine Triphosphate
6	TMT	Treadmill Test
7	CPET	Cardiopulmonary Exercise Testing
8	ANN	Artificial Neural Network
9	K-NN	K Nearest Neighbor
10	ECG	Electrocardiogram
11	BP	Blood Pressure
12	ARMA	Autoregressive Moving Average
13	DWT	Discrete Wavelet Transform
14	CPE	Cardiopulmonary Efficiency
15	MSE	Mean Squared Error
16	MET	Metabolic Equivalent Of Task
17	WFI	Wellness-Fitness Initiative
18	USAFSAM	United States Air Force School Of Aerospace Medicine
19	CHF	Congestive Heart Failure
20	CMRR	Common Mode Rejection Ratio
21	NTC	Negative Temperature Coefficient
22	IIR	Infinite Impulse Response
23	LMS	Least Mean Square
24	NARNET	Nonlinear Autoregressive Neural Network
25	RMSE	Root Mean Squared Error
26	MAE	Mean Absolute Error
27	MAPE	Mean Absolute Percentage Error
28	MPH	Miles per hour

List of Symbols

Sr. No.	Symbol	Description
1	σ	Standard Deviation
2	^{0}C	Degree Celsius
3	$\pm$	Positive or Negative
4	min	Minute
5	sec	Second
6	ml	milliliter
7	HRmax	Maximum Heart rate
8	KΩ	Kilo-ohm
9	GΩ	Giga-ohm
10	pF	pico-Farad
11	MS/s	Mega samples per second

List of Figures

List of Tables

CHAPTER-1

Introduction

1.1 Overview

Exercise physiology highly depends on cardiopulmonary system's capacity. It is considered as a noninvasive way of diagnosis and advisable over other invasive and highly priced clinical options. It is also true that we can better evaluate the cardiopulmonary system under exercise conditions. As an effect of the exercise test, human body's resting metabolic rate may increase up to 20 times and cardiac output may increase as much as six times [1]. But finally its changes depend upon age, gender, body size, type of exercise, fitness level and the disease conditions. Cardiopulmonary response of exercise are limited by the two major factors, which are cardiac output that describes the capacity of the heart to function as a pump and arteriovenous oxygen difference that describe the capacity of the lung to oxygenate the blood delivered to it, as well as the capacity of the working muscle to extract this oxygen from the blood.

The energy for doing that mechanical work is derived from the common chemical intermediate adenosine triphosphate (ATP). Skeletal muscle stores very little ATP and therefore sustained exercise requires that ATP is replenished rapidly through metabolism of fats and carbohydrates. Only 20-25% of nutrients energy converted into the muscular work at maximum efficiency [2] and the remaining energy is converted in a non-usable form as heat energy, which increases the body temperature. To maintain the body temperature, cardiovascular system increases blood supply to the skin, Heat loss through sweat and lung expires air during ventilation is also increases. In short, as the amount of exercise increases, Heart and lung activity is also going to increases. A level of exercise, at which heart and/or lungs is unable to provide the necessary requirement of the human body (especially muscular) is generally assumed as the maximum workload that can be sustained by the individual.

Exercise based evaluation of the cardiopulmonary function is a noninvasive way of diagnostic and it links performance of physiological parameters to the underlying metabolic process (aerobic power). It can be used to measure the fitness of the individual or compare with others, to evaluate continuous improvement in health, for training prescription and in the medical diagnostic field. Generally, instruments like treadmill machine or Cycle ergometer are used for exercised based evaluation with incremental exercise protocol. Dynamic parameters related to the cardiopulmonary system are non-linearly varied with the amount of physical stress and it can be evaluated in terms of absolute value or percentage relative to pretest conditions. The selection of right exercise protocol is equally important and it is done on the basis of application like sports, kind of diagnosis, age etc. But the incremental exercise test is time-consuming and causes cardiopulmonary overload at later stages of incremental exercise.

Generally, in medical practice, doctor is suggests for the cardiopulmonary exercise testing (CPET) on the basis of a patient’s health condition, and the medical history. Sometimes in-charge medical exercise expert or doctor himself, go through a short questionnaire with patients to better understand the daily activity. A short physical examination of the patient is necessary to identify contraindications to performing CPET such as a joint related issue that can limit patient’s performance and doctor’s expectation level [3]. In such a case it may be possible that exercise session becomes short which includes practicing cycling on the bike, or walking on the treadmill while wearing electrodes and sensors, etc.

Progressively increasing workloads to the maximum are the most common in use for adult or children. The study requires cooperation from the patient, a real maximal study is difficult to obtain, either on the treadmill or the cycle. In fact, only about 20–30% of children actually achieve a true plateau of oxygen uptake [3], same scenario for another age group also. When the exercise test is dealing with patients, especially cardiac abnormality, than the situation becomes complicated and risky after initial few stages of incremental exercise. Sometimes for safety purpose, 90% of age-predicted Maximum achievable heart rate: (220-age), is assumed as exercise termination indication [3]. Many researchers have taken initiative to develop a system that can predict the future trend of the physiological signals especially related to cardiopulmonary system or done efforts to predict the maximum workload that can be sustained during exercise. For that, they have utilized linear and nonlinear models like ANN structure, K-NN network, ARMA models, ARX model, fuzzy model and many more.

Again, processing of the physiological signal acquired during exercise is a challenging task because the quality of low amplitude physiological signals is highly affected due to motion artifacts, baseline wondering and the addition of other signals which have same frequency components. Similarly, recording/ analyzing of ventilator parameters also create discomfort to patients.

1.2 Outline of the thesis

The rest of the thesis is organized into chapters as follows:

Chapter 2 provides overall background information on the research topics. It covers current scenario on exercise based physiological information gathering and analysis for diagnostic interpretation. It has also covered morphological changes in physiological signals due to exercise, the methodology used for interpretation and limitations of existing works. Based on the concluded survey, problem definition and objective behind this work are clearly mentioned.

Chapter 3 presents the actual way of exercise test conduction. It represents types of exercise equipment used and their pros and cons for selecting a proper equipment to justify particular application in a meaningful way. It has also covered types of exercise protocols and their selection criteria with many examples of existing exercise protocols for different clinical applications.

Chapter 4 covers the detail discussion on the methodology of the proposed system. It gives a clear description on four stages: Data acquisitions, Data processing, System modelling and Validation, of the proposed methodology in detail. In short, this chapter has covered types of database selected, details of the sensors, type of exercise, various stages of signal processing to get information in required forms, details of models/algorithm and their validation process.

Chapter 5 highlights the results of all models and their comparison for finding a model that best fits our application. All models are tested on both types of databases: the normal as well as the abnormal. It has also covered the facts derived from those results and their discussion.

Finally, ***Chapter 6*** covers the derived Conclusion, future scope of the study and major contribution of the thesis

At last list of publications in international journals and conferences, and references are given.

CHAPTER-2

Background, Literature Review And Motivation

2.1 Background

Despite advancement in technology in the clinical diagnosis and cure of cardiovascular diseases, the exercise test has its own importance compared to more expensive and invasive treatments. Excellent guidelines for conducting exercise test and selection of right exercise level have been suggested by researchers over the last few years. By doing that, researchers have contributed a lot to the understanding of more uniform applications of the exercise test. The risk to the patient associated with exercise testing cannot be neglected even with its excellent safety record. Proper Guidelines must be followed during exercise testing in the presence of medical experts. A right selection of exercise protocols is equally important when you are dealing with patients and the test should be terminated before reaching the maximum stress condition of individuals. ECG, pulse waveform, ventilation parameters etc. are easily obtainable physiological signals that can be used to analyze physical condition in a noninvasive way. Electrocardiogram is routinely used to assess the electrical and muscular function of the heart. Changes in amplitudes of ECG segment, heart rate, ST deviation, Athens QRS score etc. individually or in combination act as key indexes for exercise based evaluation of cardiovascular functions. The relative change in features of physiological signals due to physical stress is the major key index for interpretation of physical condition. Majority of human systems' behavior is nonlinear and their functions are interconnected with each other, so obviously, it can be represented more precisely with nonlinear models compared to linear models. The behavior of many physiological signals linearly vary with the intensity of exercise like HR, RR, BP, etc. Still, the physiological signals can be best represented mathematically through both linear and nonlinear components.

2.2 State of the art

For many researchers, prediction of future behavior of signals is an area of interest as it has enough potential for the diagnostic field. Many models/frameworks are suggested for time series forecasting of biomedical signals. It helps to estimate the future trend of signals from the past behavior of signals. In [4], authors utilized the concept of temporal fuzzy clustering for time series prediction. They have applied a temporal fuzzy clustering to estimate the parameters of a train set and the training period was used to extract the sub-sources parameters. The prediction algorithm applied to an RR- interval time series of rats with hyperbaric oxygen-induced generalized seizures. The initial 18 minutes of the RR series was defined for the training phase, same training model was used to predict RR interval for 18-25 minutes. In [5], authors have proposed a time-varying nonlinear prediction of biomedical signals like heart rate variability (HRV) time series and Electroencephalogram (EEG) signals that have significant nonlinearity and time-varying complexity. The concept is based on the identification of time-varying autoregressive (TV-AR) models through expansion of the TV coefficients onto a set of radial basis functions and on a K nearest neighbor (K-NN) local linear approximation for prediction. Simulation results show that the small neighborhood sizes can better estimate complexity in nonlinear time series. In [6], authors have proposed a nonlinear prediction method based on artificial neural network (ANN) structure to predict RR interval in ECG signal and compared the results with linear Autoregressive Moving average (ARMA) model. The nonlinear method used a concept of discrete wavelet transform (DWT) and implemented on ANN structure because the wavelet transform provides multi-resolution analysis and allows for time-frequency localization. The proposed approach uses low-frequency component of DWT because these components represent the long-term trend of RR intervals. Comparisons of results show that ARMA model has better accurate prediction compare to ANN model. All above-discussed works for time series prediction are not tested under exercised conditions.

In clinical field, the capacity of the individual to sustain exercise or physical stress is considered an indication of good physical fitness and can be used to compare fitness among healthy persons. It is more dependent on individual’s aerobic power (i.e., the ability to supply oxygen and remove carbon dioxide). The highest workload sustaining capacity in the incremental exercise test helps to analyze cardiac condition and decide further training session but it may cause cardiopulmonary overload in the later stages of the test [7]. At many places works related to the prediction of the maximum workload from the immature ending of exercise test or from factual data of individual are found. In [8] [7], authors have proposed a new approach that allows an early prediction of maximum workload that will be obtained during the cardiopulmonary test without execution of the complete test. That helps to save the time of test with less physical stress. Authors have defined a new index "The Cardiopulmonary efficiency index (CPE)" that integrally measures the overall cardiopulmonary response during initial stages of the incremental test. For the estimation of maximum workload, two approaches: the ANN and K-NN classifiers are implemented. The ANN is trained with training sets that contain the factual data and CPE sequence as inputs and maximum workload of individuals as the target of the neural network. The weights in the network are modified with the help of backpropagation to minimize the mean squared error (MSE) of prediction. The K-NN based classifier utilizes a reference knowledge base for which maximum workload value is known. This algorithm is searching similarity between new test and available reference knowledge database with the help of the factual and dynamic response of initial stages. The algorithm is trying to identify the most similar existing test and assuming its maximum workload as a maximum achievable workload of the new test person. Both, ANN and K-NN approaches are effective for the different working condition. ANN algorithm is more effective if training database contains a sufficient number of sets of varieties, initially it takes longer training time but after training, it is very efficient. While K-NN approaches are preferable if a number of training/reference databases are less. The K-NN algorithm analyzes the whole training set to perform each prediction hence classification typically requires a longer time. The k-NN algorithm execution time grows with the size of the training time. In [9], the author has derived generalized Equations for predicting maximum workload for young men and women. These Equations are derived by the cycle ergometer test performed on the thirty men and thirty women, validated with fifteen men and fifteen women response during the test. The maximum workload can be calculated from age and weight of subject with the help of the derived equations. The equation derived for alien population did not show fully satisfactory results for the local population. Along with maximum workload of the individual, the maximum oxygen uptake is also considered as a standard parameter for aerobic capacity estimation. In [10], authors suggested a new approach to predict maximum workload and maximum oxygen uptake with the help of a short cycle ergometer test. The predictions are on the basis of assumption: the Heart rate, workload, and oxygen uptake have a linear relationship and the maximum achievable heart rate for any age is (220 ─ age). The author has concluded that the estimation of maximum workload is more accurate than the estimation of oxygen uptake.

Today various physiological parameters and their changes during exercise play an important role in evaluating the physical condition and it is the current scenario of growing diagnostic field. The works related to morphological changes in physiological signals due to different exercise conditions and their interpretation for classification of diseases, fitness evaluation or improvement are found in the literature of past decade. In [11], authors have compared various clinical indexes based on ECG morphological changes in response to predefined exercise test and interpreted those indexes for the classification of subjects in two groups: Healthy and Ischemic. ECG based indexes like ST segment deviations and QRS waves amplitude changes with the concept of multivariate discriminant analysis were used to classify patients in the different groups. Results showed that a combination of QRS amplitude changes may define the closer interpretation of the heart response to exercise than ST change. A close study of the time constant of heart rate decay also reflects cardiovascular fitness. In [12], authors have proposed a relatively easy method to compare the cardiovascular fitness of individuals. A short duration and low-intensity exercise test was chosen to identify the cutoff point of the time constant of HR recovery and on that bases subjects were classified in fit and unfit groups. A cutoff point of the exercise varies with exercise intensity. This approach is a very good alternative of complex and costly cardiovascular fitness interpretation procedures. In [13], authors studied the physiological signals, before and after 12 minutes running exercise in nine weeks with a healthy person. Close observation of the collected data shows that there is an improvement in the physical condition of health. This interpretation was done on the basis of observation like HR of the later week is reduced compared to initial weeks for the same exercise, R–wave amplitude is increased which shows that cardiac reserve capacity has been improved and many more. While dealing with the morphological analysis of physiological signals, the major problems faced is the high noise content of the signals during stress due to small amplitude of signals, baseline wondering, muscular noise and respiration modulation [11].

Other exercise related work is found but it is not directly relevant to our research area. Like The reduction in heart rate in the first minute of recovery after a maximal exercise stress test is a predictor of cardiovascular mortality [14]. ST segment change of ECG represents the abnormal cardiac status and it can be used to identify cardiac ischemia conditions [15]. In

[16], [2] [17], authors have tried to provide a guideline for exercise testing of human body systems and behavior of physiological systems in response to physical stress. They have also discussed various standards of exercise, testing procedures and equipment for exercise based testing. At many places, it is recommended that relation between maximum HR and age is approximately "HRmax=220 ─ age" [18] [19]. In [3], authors have suggested 90% of age-predicted HRmax can be considered as an indication of exercise termination for safety purpose.

Conclusion survey:

The past discussed works here related to time series forecasting are for the clinical application but not tested under exercise conditions. The exercise condition has its own set of challenges. A huge amount of research works related to the prediction of the future trend of physiological signals with the help of linear and nonlinear models/algorithm are found. Sufficient amount of works have been done for the classification of abnormal conditions (for diagnostic applications) on the basis of physiological signals' features changes during exercise.

After the study of existing works led us to the following conclusion that there has been no scientific work found for time series prediction of cardiopulmonary signals like Instant HR and Instant RR under exercise condition. The major problems faced by the researcher while dealing with models is Training of the model. Performance of models highly depend on the variety of the training datasets like age, sex, BMI, a database of the various Geographical population. To avoid such an issue, in approach toward forecasting of future trends of the signals from available past response is highly demanded.

2.3 Definition of the Problem

Cardiopulmonary exercise testing has potential to measure the integrative functional capacity of pulmonary, cardiovascular and skeletal muscle system. It helps to take final diagnostic decisions for patients and for fitness evaluation of the normal subjects [18]. In the incremental exercise, amount of workload increases at every few minutes' intervals, so after initial stages, the amount of workload on the cardiac system is above average and this is not tolerable every time. Especially when the subject is a cardiac patients due to which active involvement of medical expert is mandatory.

So a novel approach towards forecasting future response of cardiopulmonary signals from available past values is appreciable. This work will help to avoid cardiac overload condition, reduce testing time and minimize risk factor for patients.

2.4 Objective of work

Exercise based evaluation of the cardiopulmonary system is a non-invasive method but it is time-consuming and the subject becomes overloaded at later stages. It is not always safe to test a patient at high workloads like cardiac patients and asthma patients. In such a case, instead of taking the risk, it is advisable to stop exercise test prematurely and then apply modelling concepts to forecast future response of cardiopulmonary signals.

- To implement the concept of time series forecasting with the help of linear and Nonlinear Models/ Algorithm. More precisely use of different models to predict future response of Instant HR and Instant RR on the basis of past observed response during incremental exercise protocol. Under the assumption of age-predicted maximum HR is an upper limit for individuals to sustain exercise.
- Identify the model that best fit to forecasting application of cardiopulmonary signals during incremental exercise test for both normal and abnormal (cardiac patients) subjects.
- My proposed system provides additional prediction facilities to existing exercise based evaluation system. That is able to predict future trend of cardiopulmonary signals after premature end of exercise test. Hence it can reduce physical stress level of subjects and reduce exercise test time.

CHAPTER-3

Exercise Test Modalities

The purpose of the exercise based study is to evaluate cardiopulmonary response by applying variable physical stress. The dynamic kind of exercises are used for clinical testing and it is more appropriate to check balance between Cardiac output, Blood supply and Aerobic capacity. The level of dynamic exercise should be increased progressively so the physiological system gets sufficient time to adjust with a progressive increase in workload. Many kinds of equipment with a predefined set of exercise protocols are available for different clinical and other applications.

3.1 Exercise Equipment

Instruments like cycle ergometer and treadmill are used for assessment of the cardiopulmonary system.

Cycle ergometer:

Working of cycle ergometer is applying a variable resistance to the pedaling speed, it may be a manual or an electrical mechanism. Individuals who are fatigued will reduce the speed of pedaling as resistance increases. Instruments may be calibrated in watt or kiloponds. The exercise on a cycle ergometer is non weight bearing kind, so kiloponds or watts can be converted to oxygen uptake in milliliters per minute. METs (the metabolic equivalent of task) can be obtained by dividing VO_2 in milliliters per minute by the body weight (in kg) and multiplying by 3.5 [16]. METs can be defined as the ratio of metabolic rate (the rate of energy consumption) during a specific physical activity to a reference metabolic rate.

$$1\ \text{W} = 6\ \text{kiloponds meter / min}$$

$$1\ \text{METs} = 3.5\ \text{ml}\ O_2/\ \text{kg}/\ \text{min}$$

Cycle ergometer is less expensive, occupies less space and is less noisy compared to the treadmill. The major advantage is less motion of the upper body, so it is convenient to collect other body parameters of interest like BP, ECG or ventilatory parameters with minimum motion artifacts. A major limitation with cycle ergometer is fatigue of the quadriceps muscles and overall physiological stress applied to the body is less compared to the treadmill that is why cycle ergometer based exercise produce a lower peak of VO_2. Sometimes arm ergometer may also be preferred for lower-limb disability patients.

(a)

(b)

FIGURE 3.1 (a) Sample picture of the treadmill (Captured at BM Department, G.E.C. Gandhinagar, India) and (b) Sample picture of the cycle ergometer (collected from ref. [18]) for exercise testing.

Treadmill:

Treadmill exercise is more natural and reflects greater overall muscle use. The work performed on a treadmill depends on two variables: the speed and grade of the treadmill. The major problem with treadmill testing is motion artifacts in signal acquisition and also discomfort during the recording of ventilatory parameters. But on a treadmill, there is a greater involvement of overall muscles so it produces a high value of VO_2 Peak compared to the cycle ergometer test. Subjects should not hold front or side rail of Treadmill machine tightly because it will reduce the amount of workload to the particular stage.

The comparison of cycle ergometer with Treadmill is shown in Table 3.1. It may help to select an exercise equipment during the exercise test.

METs can be estimated from the treadmill's speed and grade with the help of following Equation (3.1) [1].

$$\text{METs} = [(\text{mph} \times 26.8)(0.1 + \text{grade} \times .018) + 3.5]/3.5 \quad \ldots\ldots\ldots\ldots\ldots \quad (3.1)$$

METs can be estimated from cycle ergometer workload with the help of following Equation (3.2) [1].

$$\text{METs} = [(\text{Watts} \times 12/\text{Body weight in Kg}) + 7]/3.5 \ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \quad (3.2)$$

TABLE 3.1 Comparison of exercise equipment: cycle ergometer v/s treadmill [3] [18]

	Cycle Ergometer	**Treadmill**
Vo_{2max}	Lower	Higher
Leg Muscle Fatigue	Often Limits Performance	Less Often Limits Performance
Work Rate Quantification	Yes	Estimation Only
Blood Gas Collection	Easier	More Difficult
Instrumentation Noise And Artifacts	Less	More
Safety	Safer	Less Safer
Weight Bearing In Obese	Less	More
More Appropriate For	Patients	Active Normal

3.2 Exercise Protocol

There are several protocols that can be used with either a cycle ergometer or a treadmill. The type of workload depends on the manner in which work is applied:

- Ramp protocol
- Multistage exercise protocol (increasing intensity every few minutes)
- Constant work rate for given time period.

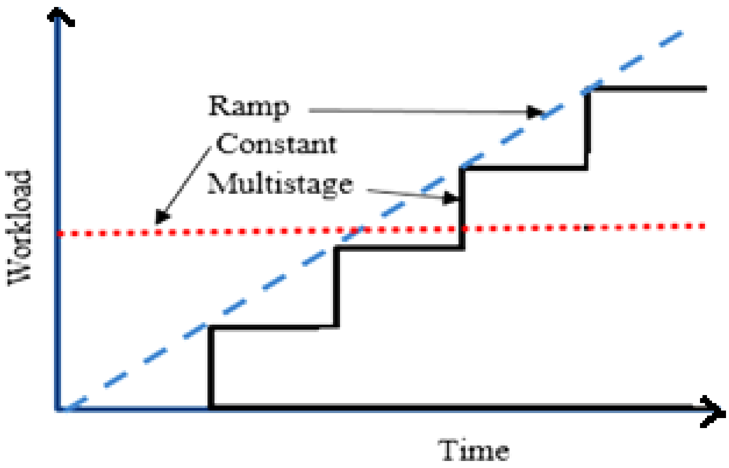


FIGURE 3.2 Comparison of Ramp, Multistage and Constant workload exercise protocol

Design of an incremental protocol has two major elements: the size of the increment in workload and the time period for each stage. In continuous, ramp protocol amount of workload applied to the individual is increased in a ramp fashion. Initially, the test is started with lower treadmill speed and gradually it is going to increase with inclination. The ramp angle of incline is progressively increased at fixed intervals (i.e. 10 to 60 seconds) starting at 0 grade [16]. In this type of protocol there is no steady state phase, so a precise estimation of exercise level is necessary otherwise premature termination of the test by subjects, leads to failure of the test. In multistage exercise protocol amount of workload increase at regular interval with larger change and workload will remain steady for that interval. It is proven that the difference in Peak exercise response, between the ramp and 1 min step cycle exercise protocols in exertional breathlessness kind of patients were small (<10 % limits of agreement) for physiological variables like oxygen uptake, minute ventilation and heart rate, representing that there is no major difference between these two kinds of exercise protocols [20]. In some protocols, either speed or elevation can be a constant and other variable varies with time like, The Balke protocol maintains a constant speed as elevation is increased 1%

every minute. While in constant workload protocol, amount of workload remains same for the whole test and that's why after initial few minutes dynamic responses remains same. A constant work protocol is preferable for findings of pulmonary gas exchange analysis, where respiratory air/ blood gas values at rest are used for comparison with exercise conditions. The basic schematic for comparison between three kinds of exercise protocol is shown in Fig. 3.2.

Normally clinical exercise duration is 8 to 12 minutes [17]. If the test duration is shorter than normal, the amount of work in multistage exercise protocol is less. Conversely if the test duration is greater than average then subjects may terminate exercise because of specific muscle fatigue. Longer test duration may cause an increase in core temperature, dehydration or discomfort. A key point is that exercise intensity should be between 8 to 12 minutes regardless of baseline fitness level. Therefore the selection of the modality and intensity of exercise is particularly important in cardiopulmonary exercise testing. Sometimes questionnaire related to their daily activity, energy consumptions and habits are discussed before exercise to decide exercise protocols.

Exercise testing is generally considered as safe procedures but still, sometimes cardiac overload results into the death of subject or medical emergency. Several symptoms during exercise testing give an indication of premature ending of the test, it may cover chest pain, fall in systolic blood pressure from resting value, a sign of respiratory failure, ST-segment elevation, muscle fatigue etc. Although exercised testing is considered as a safe and noninvasive way, still it should be performed under observation of medical staff which has knowledge of exercise physiology, which can identify abnormal response during exercise and able to provide some immediate treatment if necessary. That is why a good physician-patient communication about testing is advisable, and written informed consent should be obtained before the test.

Many different exercise protocols exist for the clinical and other fitness related interpretation purpose.

- The classical Balke treadmill protocol, developed for fitness assessment in military personnel. The Bruce treadmill protocol, developed for testing cardiopulmonary fitness in adults with the cardiac disease. The modified the Bruce protocol for testing children with cardiac disease [21].

- To improve firefighters' function, on duty effectiveness, the overall quality of life and assessment of firefighters' cardiopulmonary capacity, it is mandated to achieve WFI protocol (wellness-fitness Initiative) maximal exercise treadmill test [22].
- The modified Naughton protocol is recommended for cardiac patients for treadmill exercise testing [18].
- Protocol such as the modified Balke or the United States Air Force School of Aerospace Medicine (USAFSAM) is appropriate when testing patients with suspected coronary disease [1].
- But a recent survey confirms that BRUCE protocol is the most commonly used protocol for routine clinical testing [1].

Detail of BRUCE protocol is given in Table 3.2. A normal young person can run up to 3 to 4 stage of protocol but athlete or sports person can go beyond it.

TABLE 3.2 Details of Bruce protocol [39]

Stage	min	% grade	MPH	min/mile	km/h	min/km	METS
1	3	10	1.7	35:18:00	2.7	22:13	4
2	3	12	2.5	24:00:00	4	15:00	7
3	3	14	3.4	17:39	5.5	10:55	10
4	3	16	4.2	14:17	6.8	08:49	13
5	3	18	5	12:00	8	07:30	15
6	3	20	5.5	10:55	8.9	06:44	18
7	3	22	6	10:00	9.7	06:11	21

The overall comparison between commonly used exercises protocols are shown in Table 3.3. The protocols that are used for cardiac patients like the modified Naughton for Congestive Heart failure (CHF), has a low level of exercise intensity and METs compare to the Bruce protocol, which is used for healthy subjects or at the primary diagnostic stage.

TABLE 3.3 Comparisons between Commonly used protocols for graded, progressive exercise testing [1].

Functional class	Clinical status	O_2 cost ml/kg/min	METS	Bicycle ergometer	Treadmill protocols: Bruce		Balke-Ware	USAFSAM		"Slow" USAFSAM		McHenry		Stanford		ACIP		CHF		METS
				1 watt = 6.1 Kpm/min; For 70 kg body weight Kpm/min	3 min stages		% grade at 3.3 MPH 1 min stages							% grade at 3 MPH	% grade at 2 MPH					
					MPH	%GR														
					5.5	20														
Normal and I	Healthy, dependent on age, activity	56.0	16		5.0	18	26	MPH	%GR							MPH	%GR			16
		52.5	15				25, 24	3.3	25			MPH	%GR			3.4	24.0			15
		49.0	14	1500			23, 22					3.3	21			3.1	24.0			14
		45.5	13		4.2	16	21, 20	3.3	20					22.5						13
		42.0	12	1350			19, 18					3.3	18	20.0		3.0	21.0	MPH	%GR	12
	Sedentary health	38.5	11	1200	3.4	14	17			MPH	%GR			17.5		3.0	17.5	3.4	14.0	11
		35.0	10	1050			16, 15, 14	3.3	15	2	25	3.3	15	15.0				3.0	15.0	10
		31.5	9	900			13, 12					3.3	12	12.5		3.0	14.0	3.0	12.5	9
	Limited	28.0	8	750			11, 10	3.3	10	2	20	3.3	9	10.0	17.5	3.0	10.5	3.0	10.0	8
		24.5	7	600	2.5	12	9, 8			2	15			7.5	14.0			3.0	7.5	7
II	Symptomatic	21.0	6				7, 6	3.3	5			3.3	6	5.0	10.5	3.0	7.0	2.0	10.5	6
		17.5	5	450	1.7	10	5, 4			2	10			2.5	7.0			2.0	7.0	5
III		14.0	4	300	1.7	5	3, 2	3.3	0	2	5					3.0	3.0	2.0	3.5	4
		10.5	3	150			1	2.0	0	2	0	2.0	3	0	3.5	2.5	2.0	1.5	0.0	3
		7.0	2		1.7	0										2.0	0.0	1.0	0.0	2
IV		3.5	1																	1

USAFSAM = United States Air Force School of Aerospace Medicine
ACIP = asymptomatic cardiac ischemia pilot
CHF = congestive heart failure (modified Naughton)
Kpm/min = Kilopond meters/minute
%GR = percent grade
MPH = miles per hour

CHAPTER-4

Proposed System

The basic methodology of the proposed system for time series forecasting of instant HR and instant RR is broadly organized into four stages, which are Data acquisitions, Data processing, System modeling, and Validation, as shown in the Fig. 4.1 and discussed below.

Data acquisition stage has covered various technical aspects of signal recording. It has covered exercise equipment and level of physical stress (exercise protocol), sensors' descriptions with specifications, and placement of electrodes. Data processing stage has covered the details of all necessary signal conditioning from which signals have passed. System modelling stage has discussed the concept of model selection, training procedures and their parameters optimization. And last validation stage has discussed validation of models' forecasted response and calculation of errors.

FIGURE 4.1 Basic stages of the proposed system

4.1 Data Acquisitions

Here I have covered two major applications domain of cardiopulmonary exercise testing. First is the fitness relevant interpretation of normal subjects and other is for diagnostics clinical application relevant to cardiopulmonary patients. The detailed schematic for Data acquisitions and processing steps are shown in the Fig. 4.2.

For performance testing of implemented three models, two types of databases were used. The first database contains ECG and breathing signal of normal subjects, whose age are in between 18 to 32 years. Their database is collected with the help of DAQ NI ELVIS II at sampling frequency 200 Hz (sampling time 0.005 sec), it is suggested by PAN TOMPKINS for QRS detection [23].

Major Specifications of NI ELVIS-II Board:

- Number of channels: 8 differential or 16 single-ended channels
- ADC resolution: 16 bits
- Sample Rate Maximum: 1.25 MS/s single channel, 1.00 MS/s multi-channel
- Maximum working voltage for analog inputs (signal + common mode): ±11 V of Analog Input GND (AIGND)
- CMRR (DC to 60 Hz): 90 dB
- Input Impedance Device on AI+ or AI– to AIGND: >10 GΩ || 100 pF
- Accuracy: 1%

Because of global acceptance of the Bruce protocol, here all normal subjects have passed through treadmill exercise test with exercise intensity level of the Bruce protocol. And all young and normal subjects have completed at least 3 stages (9 min) of the Bruce protocol. For the uniformity in data analysis, I have processed with 9 min data of all normal subjects. During that, the ECG and breathing signals are recorded. ECG is recorded using surface electrodes and breathing signal is recorded using negative temperature coefficient (NTC) thermistor placed near to nasal track with the help of mask as shown in Fig. 4.3.

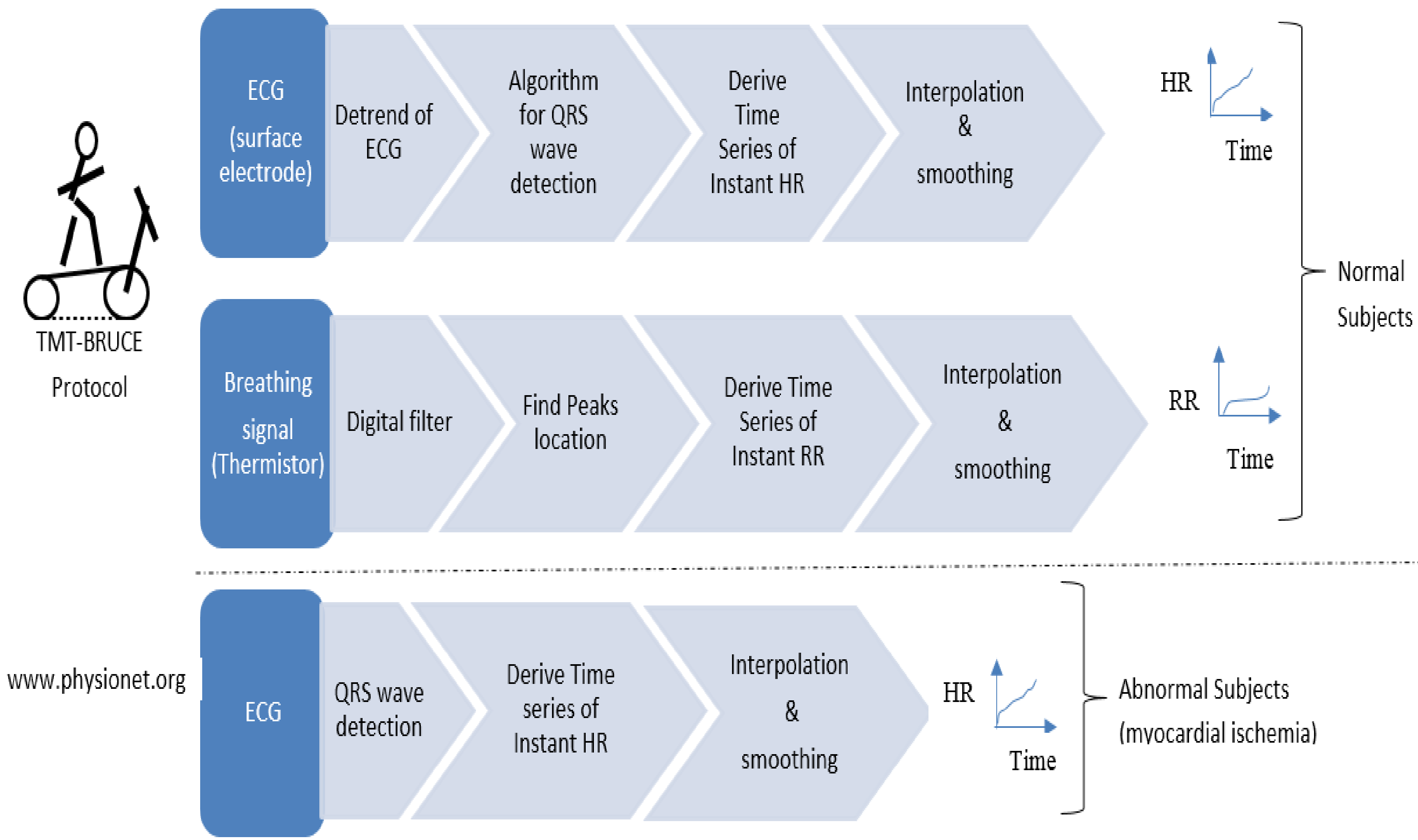


FIGURE 4.2 Basic schematic of data acquisitions and processing cardiopulmonary signals for Normal and Abnormal subjects.

FIGURE 4.3 Sample picture of data acquisition system

FIGURE 4.4 A thermal sensor for the recording of breathing pattern.

The selected surface temperature sensor is designed for human respiration studies or skin temperature measurements with extremely rapid response time shown in the Fig. 4.4. It is a type of stainless steel temperature sensor. It is 20KΩ NTC thermistor, whose resistance decreases as the temperature increases. As per the specifications given in datasheet, -25 to 125 ^{0}C is the temperature range, 150 ^{0}C is the maximum temperature range that sensor can tolerate without any damage and Accuracy $\pm$ 0.2 ^{0}C at 0 ^{0}C and $\pm$0.5 ^{0}C at 100 ^{0}C. The interface measures the resistance value R at a particular temperature and converts back to temperature using below given Steinhart-Hart Equation (4.1).

$$T = [k_0 + k_1 (\ln 1000R) + k_1 (\ln 1000R)^3]^{-1} - 273.15 \qquad (4.1)$$

Where,

T is temperature in ^{0}C

R is measured resistance in KΩ,

$K_0 = 1.02119 \text{ X } 10^{-3}$

$K_1 = 2.22468 \text{ X } 10^{-4}$

$K_2 = 1.33342 \text{ X } 10^{-7}$

FIGURE 4.5 Electrocardiogram (ECG) sensor

FIGURE 4.6 Surface metallic contact type clamp electrodes for the recording of cardiac electric potential

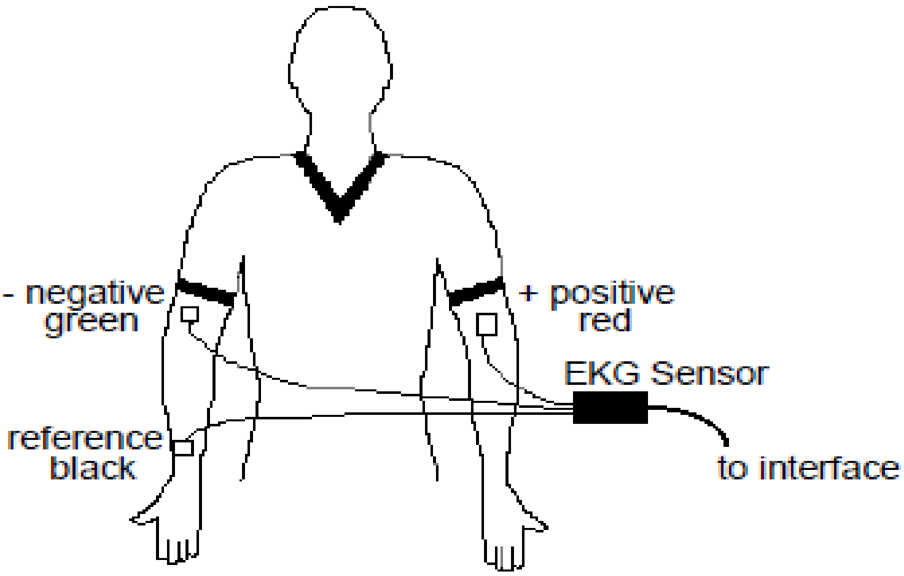


FIGURE 4.7 Placement of electrodes for the recording of cardiac electric potential [24]

For Acquisitions of cardiac electrical potential waveforms here 3 lead Vernier EKG (electrocardiogram) sensor is used. The major specifications of ECG sensors are offset ~1.00 volt ($\pm$0.3 v) and Gain of 1000 (1 V sensor output/1 mv body potential).

For the recording of cardiac electrical activity from the surface, a metallic contact type clamp electrodes are used as shown in Fig. 4.6. The structure of electrodes is such that it can maintain a good skin contact with minimum motion artifacts. While placing the electrodes on the body surface, it is important to make good contact with skin with the help of electrolyte gel for removing air medium and reduce mismatch of impedance between skin and electrode. The electrical signal produced by the heart and collected at the surface is small, so to maintain good contact between electrode and skin, scrub the area of skin with a paper towel to remove dead skin and oil. During the exercise condition over treadmill machine, the lower body has higher amount of motion activity compared to the upper body. That is why to reduce motion artifacts the reference electrode is placed on the right-hand wrist as suggested by the manufacturer of ECG sensor [24]. Remaining two are on the elbow of both hands in lead 1 configuration as shown in Fig. 4.7.

Another set of the database is collected from www.physionet.org “The MIT-BIH ST Change database” [25]. The PhysioNet offers free web access to large collections of recorded physiologic signals (PhysioBank) and related open source software (Physio Toolkit). The MIT-ST change database includes ECG recordings of varying lengths, most of which were recorded during exercise stress tests and which exhibit transient ST depression or elevation. This database consists ambulatory ECG recordings from subjects who have chances of myocardial ischemia recorded with sampling frequency 360Hz. On website, information regarding types of exercise is not given and different subjects have done exercise for different time length.

The MIT-BIH ST Change Database

DOI for
The MIT-BIH ST
Change Database:

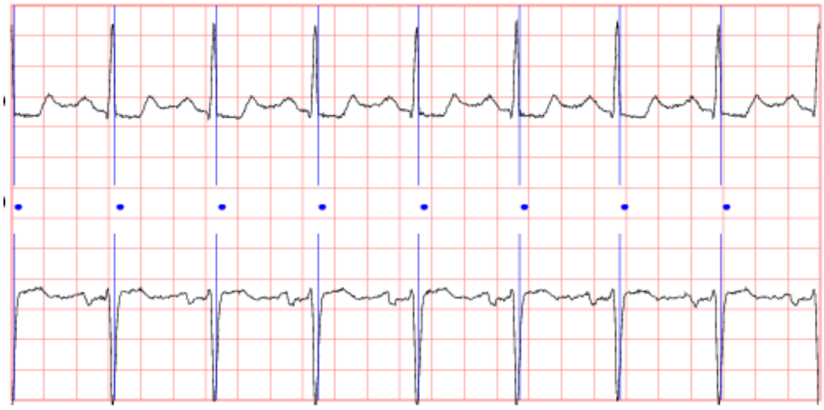
doi:10.13026/C2ZW2H

This database is described in

> Albrecht P. S-T segment characterization for long-term automated ECG analysis. M.S. thesis, MIT Dept. of Electrical Engineering and Computer Science, 1983.

Please cite this publication when referencing this material, and also include the standard citation for PhysioNet:

> Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/content/101/23/e215.full [http://circ.ahajournals.org/content/101/23/e215.full]]; 2000 (June 13).



This database includes 28 ECG recordings of varying lengths, most of which were recorded during exercise stress tests and which exhibit transient ST depression. The last five records (323 through 327) are excerpts of long-term ECG recordings and exhibit ST elevation.

Note that the annotation files contain only beat labels; they do not include ST change annotations, as in the European ST-T Database [../edb/].

FIGURE 4.8 Screenshot of "The MIT-BIH ST Change Database" available on web open resources www.physionet.org

4.2 Data Processing

The Basic steps of data processing for the normal and abnormal subjects are represented in Fig. 4.2. Signals recorded from Normal/Abnormal subjects are passed through a multi-stage processing sequence to present information in required necessary forms. Recorded or collected signals are either Raw ECG signal or breathing signal, to convert that information into a vector of heart rate and respiration rate following operations are performed on it.

4.2.1 For the Normal subjects database

Electrocardiogram processing:

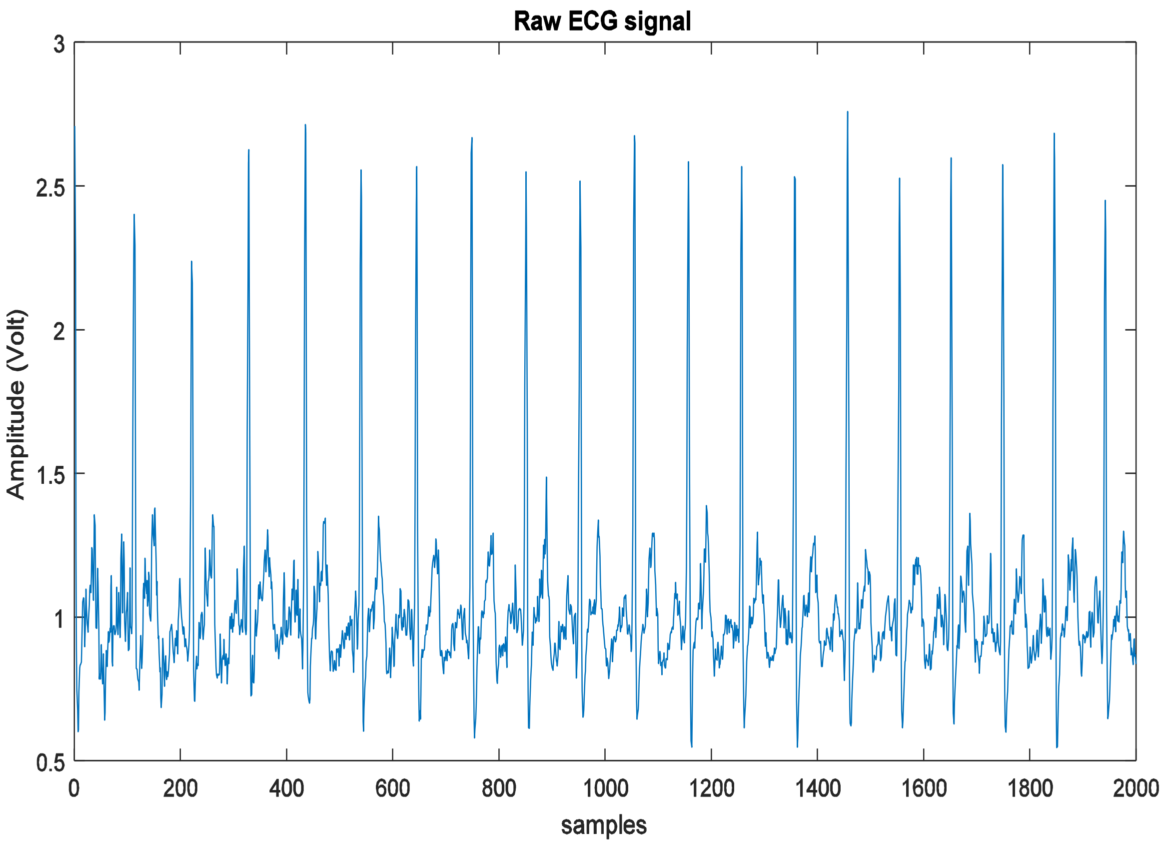


FIGURE 4.9 A sample raw ECG signal (database 6)

Collected Raw ECG signals showed baseline drift (nonlinear trend of the signal + offset) that is why it becomes difficult to deal with QRS complex. To remove that nonlinear trend here, first of all, the 10-degree polynomial curve is fitted to the signal that represents trend structure of ECG signal and that curve is subtracted from the original signal. The general

Equation for a polynomial is given in Equation (4.2). The effect of polynomial filter on sample ECG is shown in Fig. 4.10.

$$P(x) = p_1x^n + p_2x^{n-1}+\ldots+p_nx^1+p_{n+1} \tag{4.2}$$

Where

p -coefficients for a polynomial

n -Polynomial degree

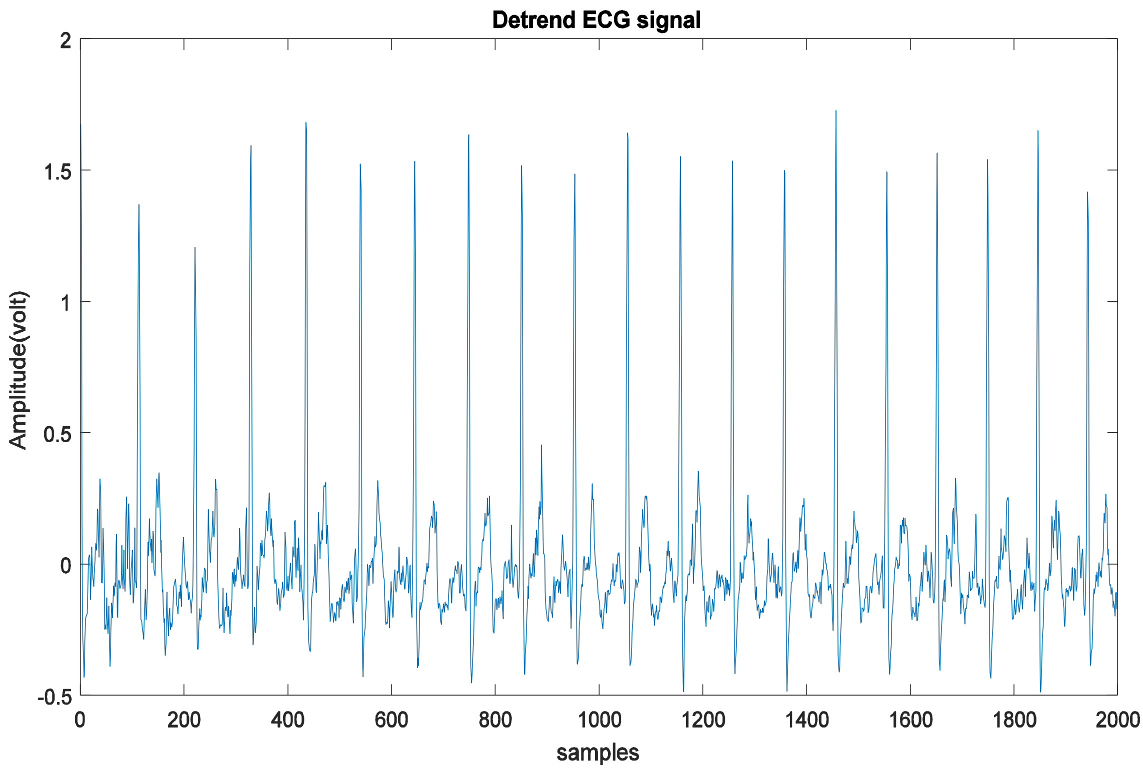


FIGURE 4.10 A sample Detrend ECG signal

To derive a time series of instant HR, it is necessary to find out the location of each R wave peak. Many algorithms are available for R wave location detection. Among them PAM-TOMPKINS is the most commonly used and effective algorithm [26]. Here I have followed PAM-TOMPKINS process for R wave detection. So after detrending that ECG signal, it is passed through various stages like Differentiator, squaring function and moving the window of size 20 integral process [23] [27].

- In the derivative stage, signal is differentiated with a five-point derivative function to extract QRS slope related information. This derivative is implemented with the following difference Equation (4.3).

$$y(nT) = \frac{2x(nT) + x(nT - T) - x(nT - 3T) - 2x(nT - 4T)}{8} \tag{4.3}$$

After signal passes through the Differentiator stage, the low-frequency component of P and T waves are further attenuated and high-frequency component due to the slope of QRS wave is further enhanced. A sample result of Differentiator stage is shown in Fig 4.11.

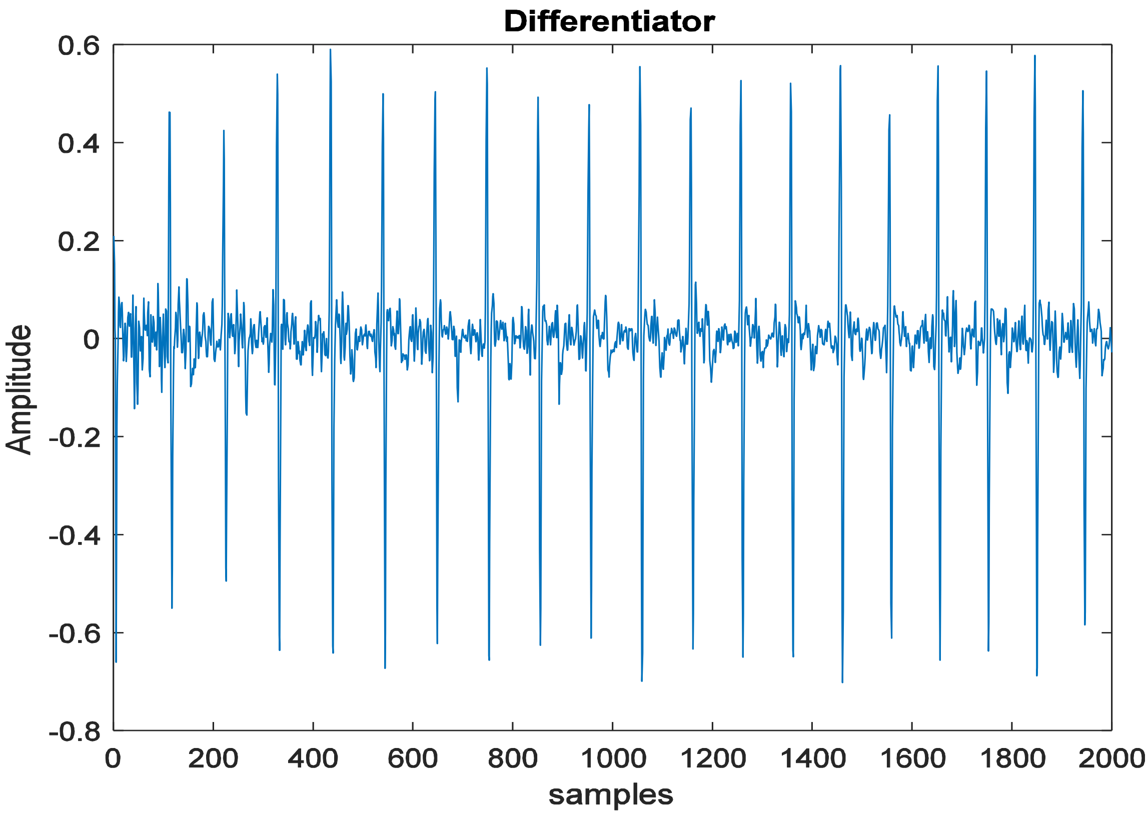


FIGURE 4.11 A sample ECG signal after Differentiator stage

- In Squaring stage, signal is further processed through following nonlinear function

$$y(nT) = (\,[x(nT)]^2) \tag{4.4}$$

The squaring operation makes all data samples of signal positive and it amplifies the output of derivative process nonlinearly. It emphasizes the QRS complex frequencies and suppresses the small difference arising from P and T waves. A sample result of the Squaring function is shown in Fig. 4.12.

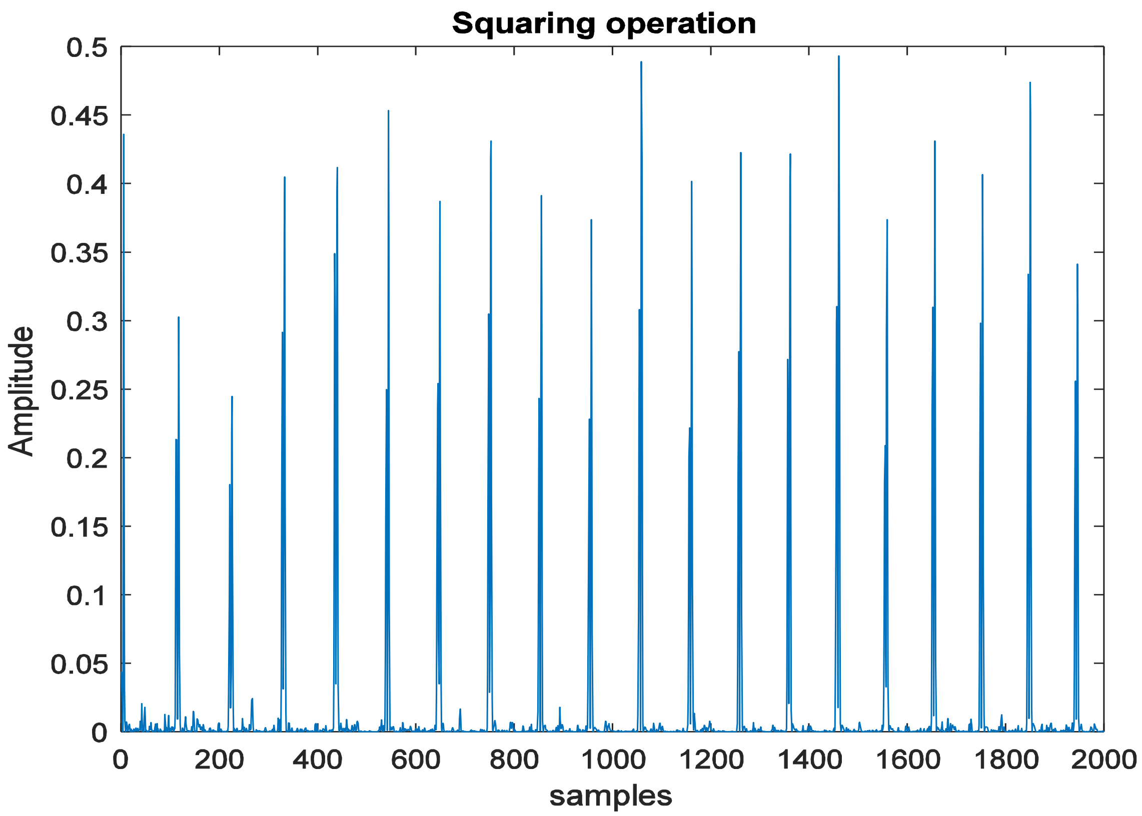


FIGURE 4.12 A sample ECG signal after squaring function

- Moving window integration is also called smoothing filter. There are multiple peak within the duration of the QRS complexes. The PAN-TOMPKINS algorithm performs smoothing with the help of following difference Equation (4.5).

$$y(nT) = \frac{1}{N}[\,x(nT-(N-1)T) + x(nT-(N-2)T) + \cdots + x(nT)] \qquad (4.5)$$

Where N is the number of samples in the width of the moving window. It is advisable to select this parameter N very carefully. Here I have selected 20 window size. A large value of window size may merge QRS and T wave and small value may reflect in output as multiple peaks for the single QRS complex. A sample result of the moving window integration function is shown in Fig. 4.13.

It shows a peak corresponds to QRS wave, it makes the task easy for finding the location of QRS wave.

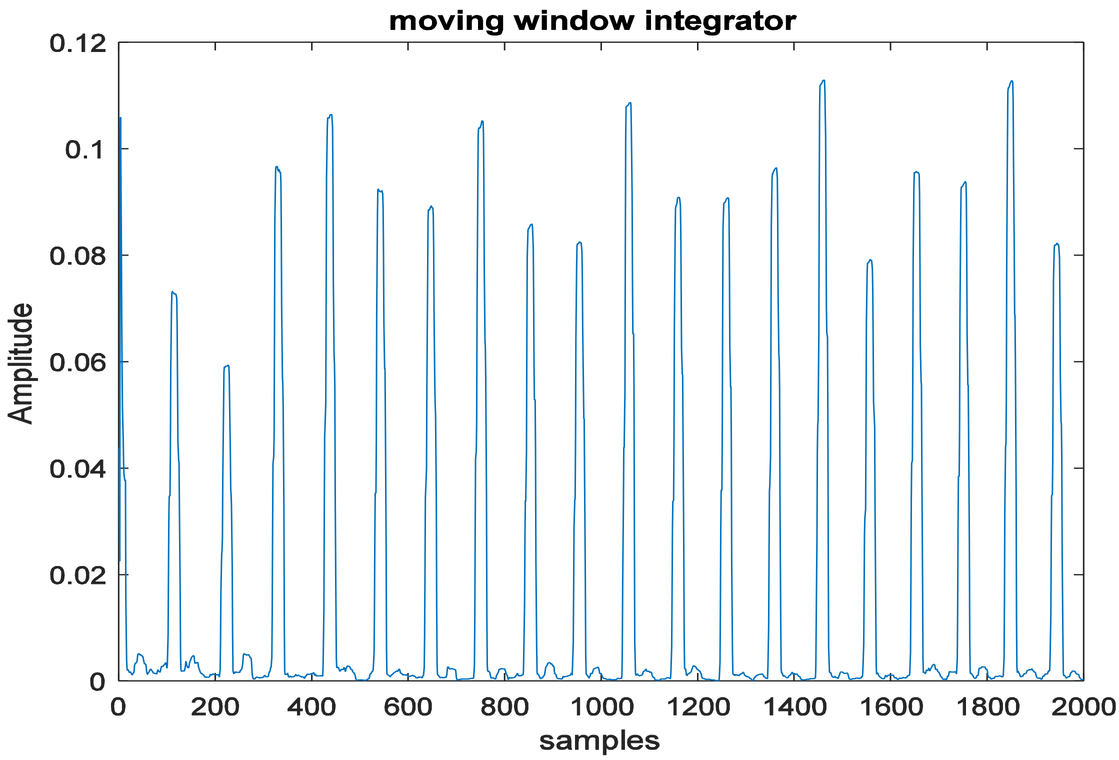


FIGURE 4.13 A sample ECG signal after moving window integrator stage

After getting the location of each R wave peak on the basis of amplitude threshold. It is very simple to derive time series of instant HR by finding time distance between two consecutive R wave peaks. It can be calculated by Equation (4.6) and shown in Fig. 4.14.

$$\text{Instant HR} = \frac{60}{\text{the time distance between to consecutive R wave peaks}} \tag{4.6}$$

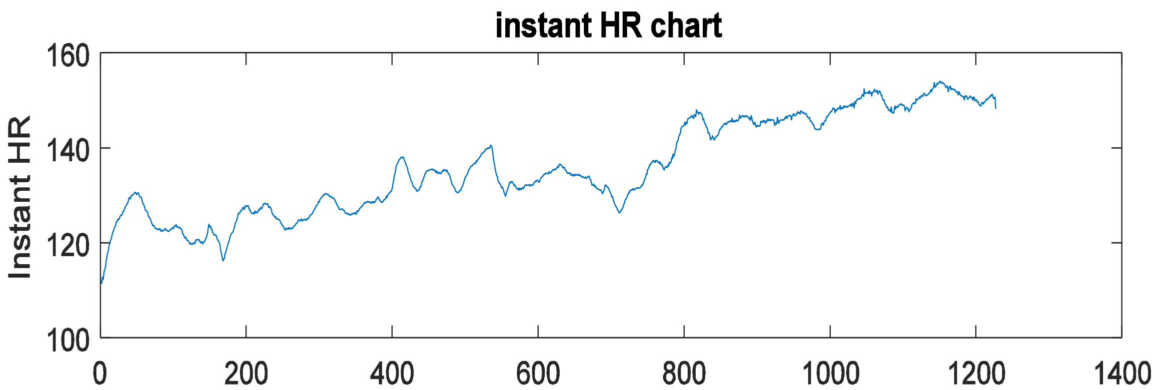


FIGURE 4.14 Instant HR chart calculated from time distance between two consecutive R wave locations

It is very difficult to detect all QRS waves in an ECG signal [26]. But sometimes few QRS complexes have very less amplitude so chances are that they may not detect by PAM TOMPKINS algorithms. Or it may happen that a short peak of noise is detected as a QRS signal. So in these both cases distance between two consecutive R waves is either increased or decreased by almost twice and it is reflected as a peak (sudden change in HR) in time series of instant HR. So to attenuate those peaks here, the concept of interpolation technique is applied. In this techniques, if an absolute change of two consecutive instant HR is more than 10 HR than next HR is replaced with an average of both one. Detail function is elaborated in Equation (4.7).

$$if \mid insatntHR(n+1) - instantHR(n) \mid > 10\,, \qquad then$$

$$instantHR(n+1) = \frac{instantHR(n) + insatntHR(n+1)}{2} \tag{4.7}$$

If there is no such peak of a sudden change in HR then signal information remains as it is before and after the interpolation operation.

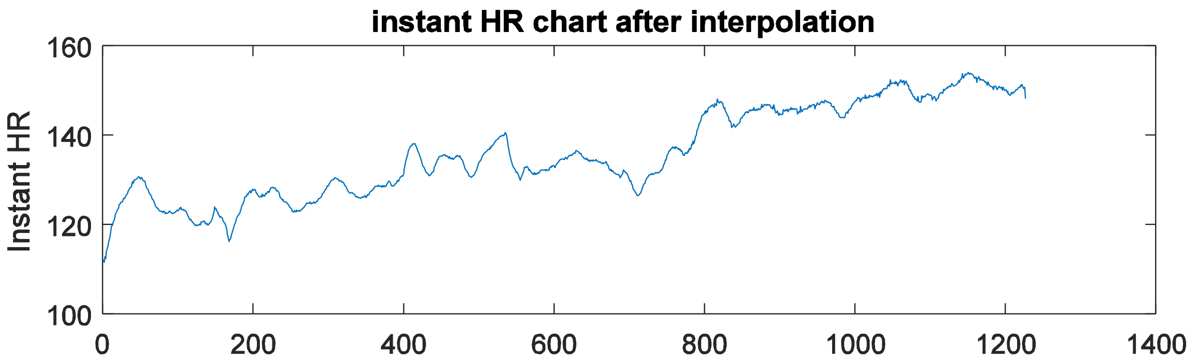


FIGURE 4.15 A sample result of Interpolation applied on vector Instant HR is shown in Fig. 4.14

At last again final signal of instant HR is smoothed (moving average integral) with the help of window size 30 as per equation (4.5). A sample result of smoothing operation is shown in Fig. 4.16.

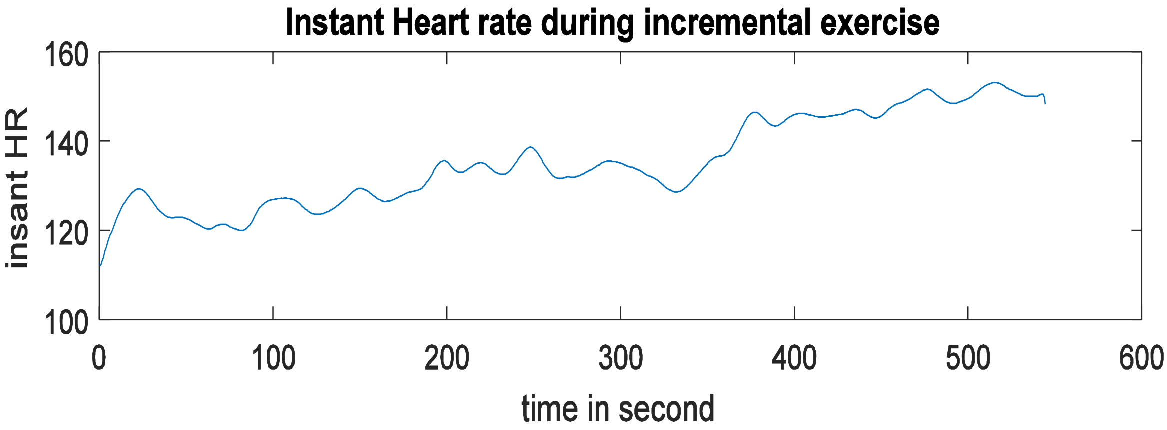


FIGURE 4.16 A Sample result of Smoothing operation applied on the signal of fig 4.15

Breathing signal processing:

As discussed above thermistor sensor is used for the recording of breathing cycle. A thermistor is placed near to the nasal passage for detecting a change in temperature of inhaling and exhaling air. The pattern of recorded breathing signal with the help of thermistor is like a sine wave. And the whole signal is shifted upward because breathing air temperature range is around the body and environment temperature. A sample result of original/ raw breathing signal is shown in Fig. 4.17.

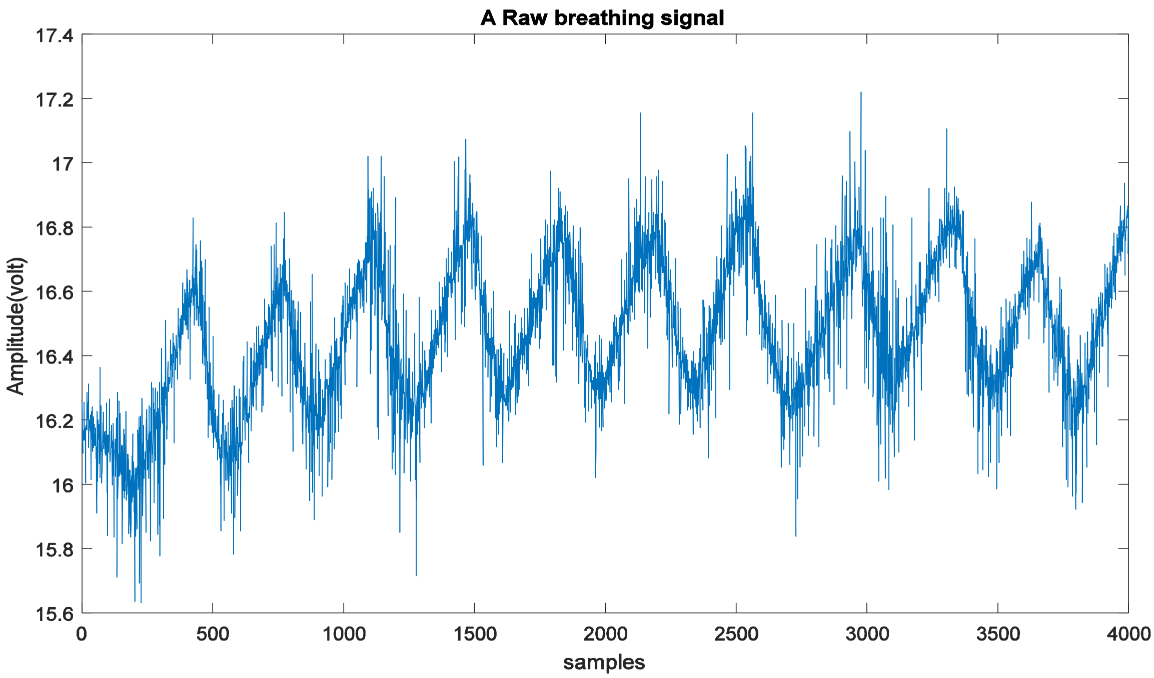


FIGURE 4.17 A sample of the original breathing signal

For better interpretation signal should center at zero level with inhale and exhale pattern on either side of zero level. To remove that positive drift (a signal with frequency 0 Hz), a digital filter whose transfer function represented in the Equation (4.8) is applied.

$$H(Z) = \frac{1}{T}\frac{[Z-1]}{[Z-0.995]} \tag{4.8}$$

In above Equation location of the pole on the positive real axis is at distance 0.995 from center and location of zero at unity distance from the center of Z plane. Positive real axis on Z plane corresponds to zero frequency. This digital filter is kind of IIR filter design with the help of pole-zero method. This filter's z-plane plot and amplitude-frequency responses are shown in Fig. 4.18 and Fig. 4.19. In general, the magnitude transfer function of a system for a particular value of z is given by the product of the distances from the corresponding point in the z-plane to all the zeros of the system's transfer function, divided by the product of the distances to its poles. The phase response is given by the sum of the angles of the vectors joining the point to all the zeros, minus the sum of the angles to the poles. It is obvious that the magnitude response of the filter is zero at z = 1, due to the presence of a zero at that point. Furthermore, for values of z away from z = 1, the distances to the zero at z = 1 and the pole at z = 0.995 will be almost equal; therefore, the gain of the filter will be close to unity for frequencies above 0 Hz.

A sample result of Digital filter is shown in Fig. 4.20 corresponds to breathing signal shown in Fig. 4.17.

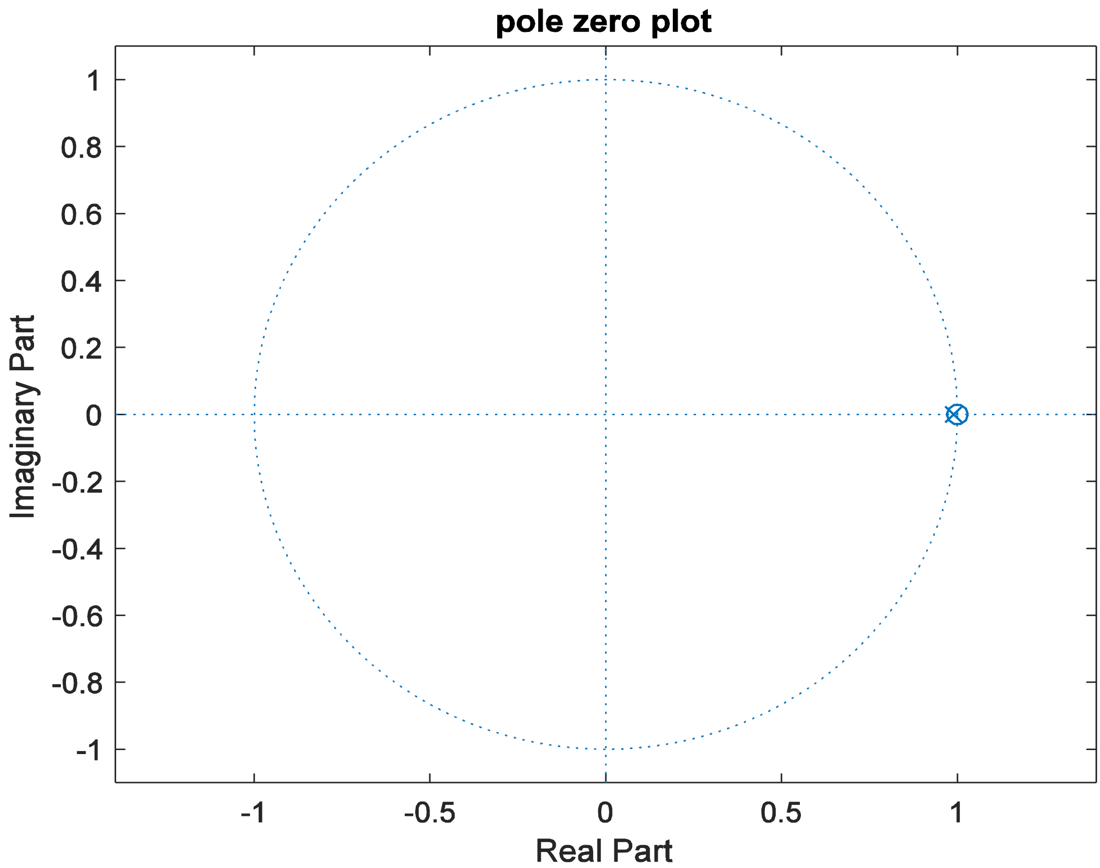


FIGURE 4.18 Z plane plot of the digital filter (Equation 4.8)

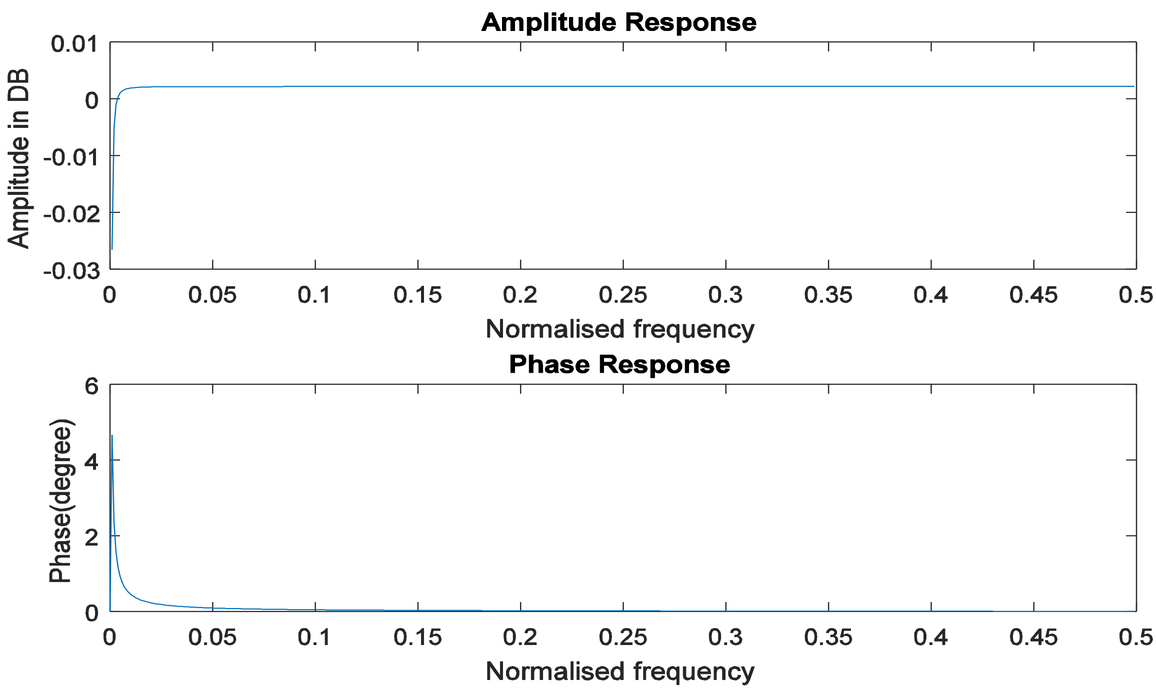


FIGURE 4.19 Amplitude and Phase Response of Digital IIR filter (Equation 4.8)

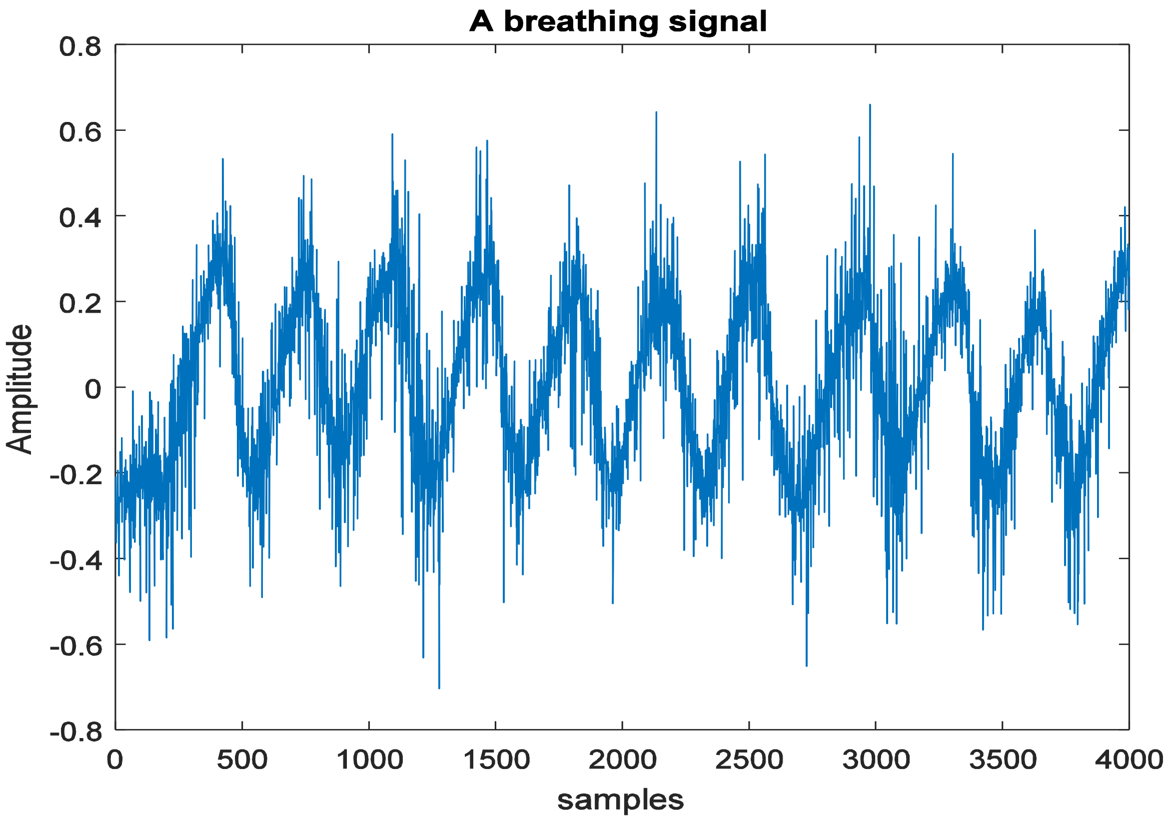


FIGURE 4.20 A breathing signal after digital filtering process (offset is removed)

After the filtration of the offset, breathing signal is symmetrical around baseline and inhale –exhale air temperatures are separated by zero level. Later the smoothing operation of window size 30 is applied and the result shown in Fig. 4.21.

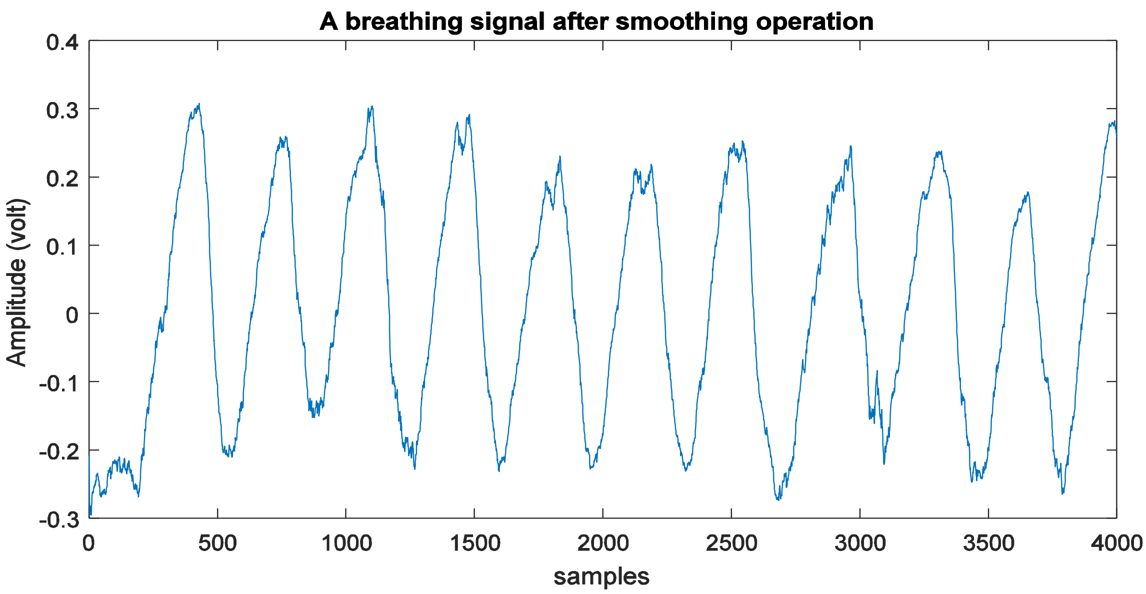


FIGURE 4.21 A sample breathing signal after smoothing operation

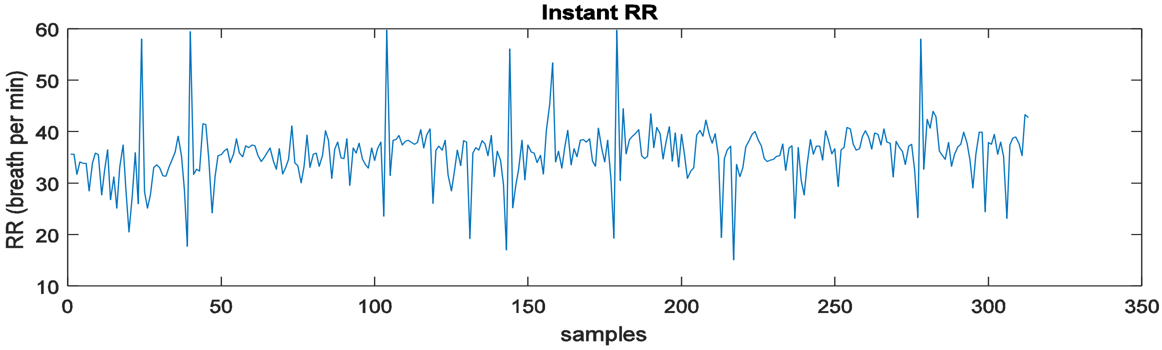


FIGURE 4.22 A sample result of Instant RR

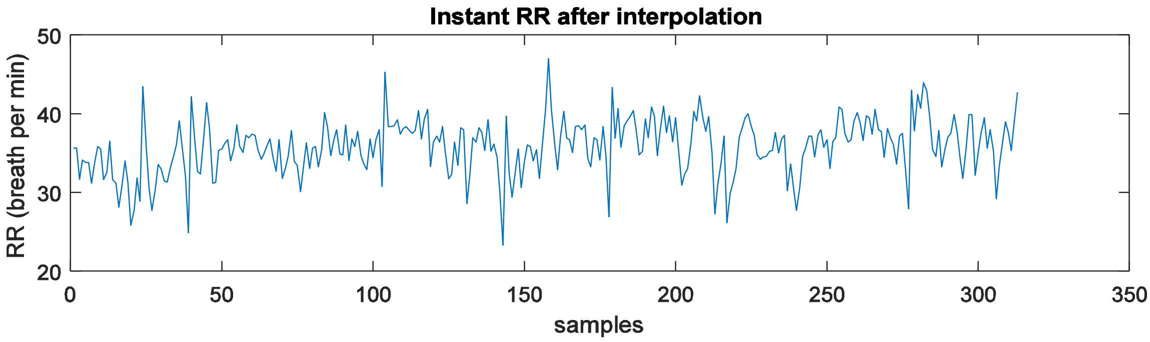


FIGURE 4.23 Instant RR after interpolation

Now to identify cyclic repetition time, it is necessary to find out the location of Peak, for that the concept of amplitude thresholding is used. After getting series of time distance between two consecutive peaks, it is converted to time series of instant RR using formula (60/ time between two consecutive peaks) as shown in Fig. 4.22. Due to chances of missing any peak or consider any noise as signal peak may generate a sharp peak in instant RR series. This can be removed with the help of interpolation of signals. In this technique, if an absolute change of two consecutive samples of instant RR is more than 5 RR than next one RR is replaced with an average of both which is shown in Fig. 4.23. And at last smoothing operation performed on the processed signals of window size 20, it is shown in Fig. 4.24.

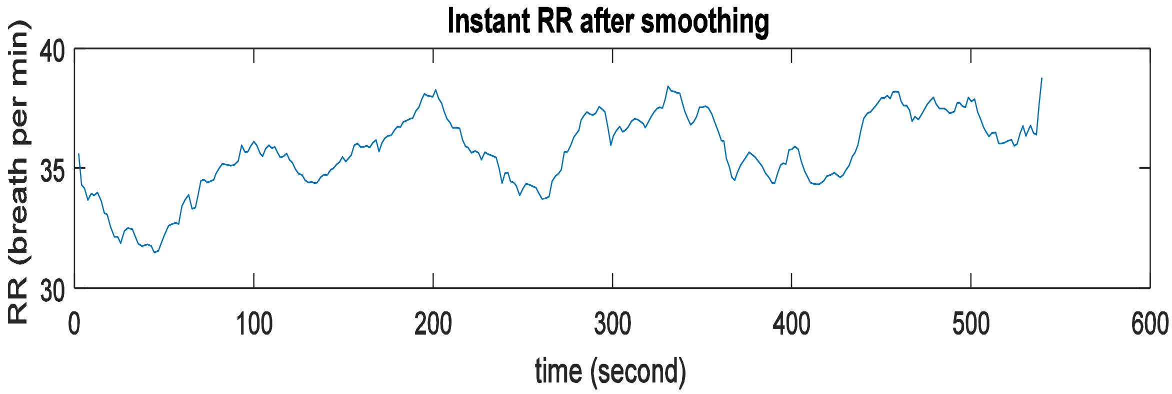


FIGURE 4.24 Instant RR after smoothing operation

4.2.2 For the Abnormal database

The abnormal database "The MIT-BIH ST Change Database" is collected from the online physiological database www.physionet.org. This database consists ambulatory ECG recordings from subjects who have chances of myocardial ischemia. The quality of ECG signal is good enough so that there is no requirement of the filtering process. The remaining processes are same as the processes of the normal database that is discussed previously and mentioned in Fig. 4.2. I have performed few tasks like finding peak location on the basis of amplitude threshold, Calculation of instant HR from the location of two consecutive R peak, Interpolation of derived instant HR signal and at last 30 point smoothing of instant HR signal.

4.3 System Modelling

The core concept behind system modelling is the forecasting of future response of cardiopulmonary signals from available past response during physical exercise on a treadmill machine. Here three algorithms are implemented for the time series forecasting of instant HR and instant RR, which are an Adaptive filter, autoregressive-moving average with exogenous input (ARMAX), and nonlinear ARX (NLARX) models. With the help of these models, I have tried to forecast cardiopulmonary response of later stages in incremental exercise. Additionally, I have done a small work for single step and multi-step-ahead prediction with the help of artificial neural network concept. Performance of these models is evaluated by comparing predicted data with original data.

4.3.1 Implemented Models for Time series forecasting

Adaptive filter

An adaptive filter has an adaptation algorithm that is meant to monitor the environment and varies the filter transfer function accordingly. Adaptive filters have been successfully applied in a variety of fields, often quite different in nature. Filtering can be defined in the simplest terms that it is a process of noise removal from a measured process in order to reveal or enhance information about any quantity of interest, more specifically for prediction, involves forecasting information sometime into the future given the current and past data. The adaptive filters are also well known for their applications like System identification, Channel Identification, Plant identification, Echo cancellation for long-distance transmission, Adaptive noise canceling and Inverse plant modeling.

The purpose of adaptive filtering algorithms is to adjust the tap-weight vector of the adaptive FIR filter to minimize the prediction MSE. The aim of the filter optimization procedure would be to reach the bottom of the bowl-like function. For that solution can be achieved by the well-known optimization method called steepest descent, which uses the gradient vector to gradually descend step by step to the minimum error.

First of all, for better training, the signal vector should be normalized with the help of min-max theorem. Its basic formula is given in Equation (4.9).

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4.9}$$

Where x is the time series vector of the signal like Instant HR and Instant RR.
X_{min} and X_{max} are the minimum and maximum value of the signal.

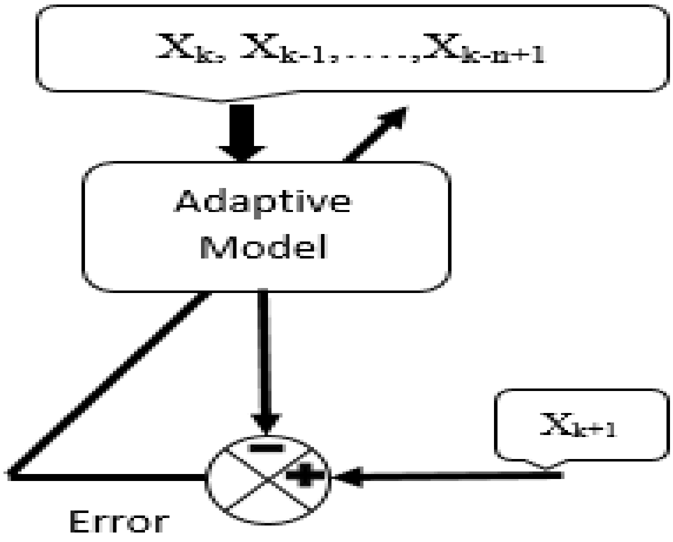


FIGURE 4.25 Adaptive filter training concepts

The general schematic of adaptive filter training concepts is shown in Fig. 4.25 and detail training concept of the Adaptive model is shown in Fig. 4.26.

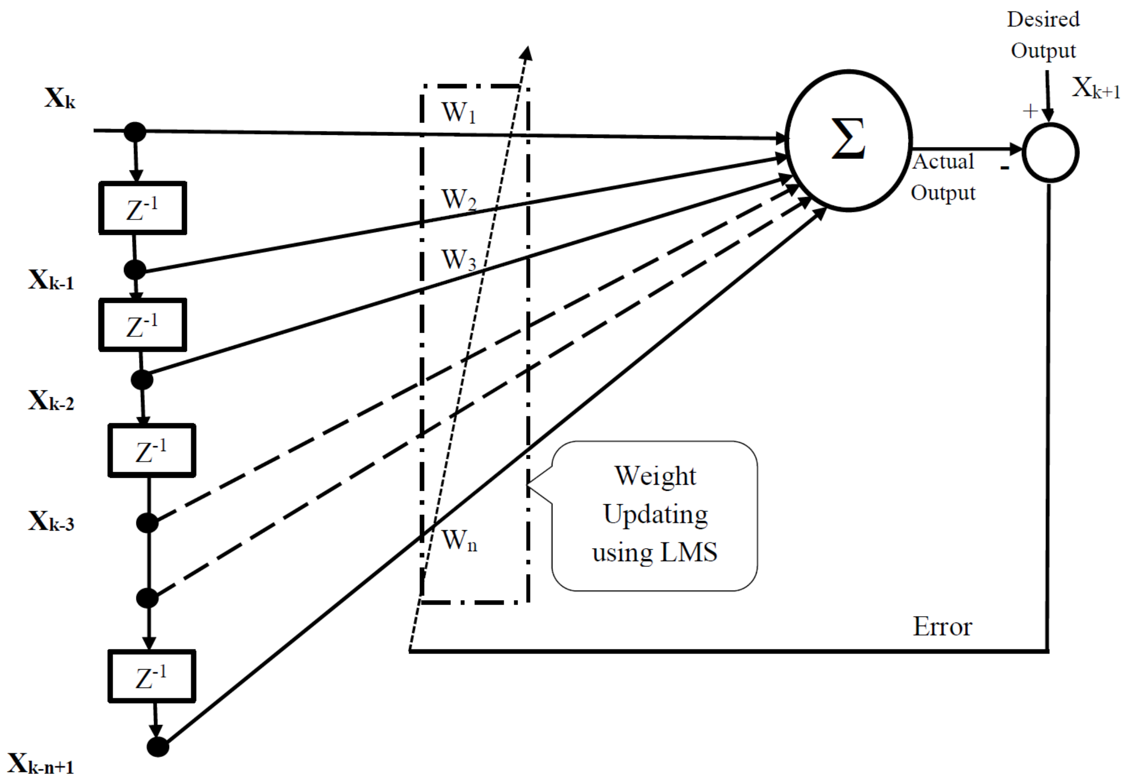


FIGURE 4.26 Detail training concepts of Adaptive FIR filter

The adaptive model is nothing but an FIR filter with initially random tap weights. Also, divide the available signal vector into training and validation sets. The training set contains Ntr samples from an original database of N samples. For better understanding assume that order of Adaptive FIR filter is n then the input with sample X_k, X_{k-1}… X_{k-n+1}, n is number of inputs or order of FIR filter, where k=n….Ntr-1and reference output as X_{k+1}. The actual output of the adaptive filter is weighted sum of input vectors that is why it is a kind of linear predictive model. Initially, weights are random so the output of the adaptive model is not matched with reference output (X_{k+1}), hence generated an error is quite large and same error is used to adjust the tap weights. The purpose of adaptive filtering algorithms is to adjust the tap-weight vector of the adaptive filter and minimize the error between desired and actual prediction. Least Mean Square (LMS) algorithm is used to adjust tap weight vector. The LMS algorithm is based on the method of steepest descent, where the new tap-weight vector w(n+1) is given by the present tap-weight vector w(n) plus a correction proportional to the error e(n) and current input x(n) [28] [29].

$$W\ (n+\ 1) =\ w\ (n) +\ 2\ \mu\ e\ (n) x\ (n) \tag{4.10}$$

Where

$W = [W_1\ W_2\ W_3 \ldots\ldots\ldots W_n]$

e (n) is scalar error (desired minus the actual output)

x (n) is input tap weight vector at instant

μ is learning rate (0 to 1)

This expression is also known as the Widrow-Hoff LMS algorithm. The lower value of μ takes more time for simulation but gives better performance and large value of μ takes less simulation time but gives a large error of prediction compared to the small value of μ. For our system, the value of μ at 0.01 is selected.

In our time series forecasting application, the available past time series signal vector is used to train the adaptive model in sets according to the order of the filter. The implemented adaptive filter is also tested for different FIR order to right order selection. After successful training, the same adaptive model is used to forecast time series of future values.

Performance of adaptive model can be evaluated by comparing forecasted response with the actually available response of the same subject.

ARMAX model

According to statistical analysis of time series the ARMAX (autoregressive-moving average with exogenous terms) model provides a description of a stationary signal in terms of three polynomials, one for the auto-regression (AR), the second for the moving average (MA) and the third for exogenous inputs terms.

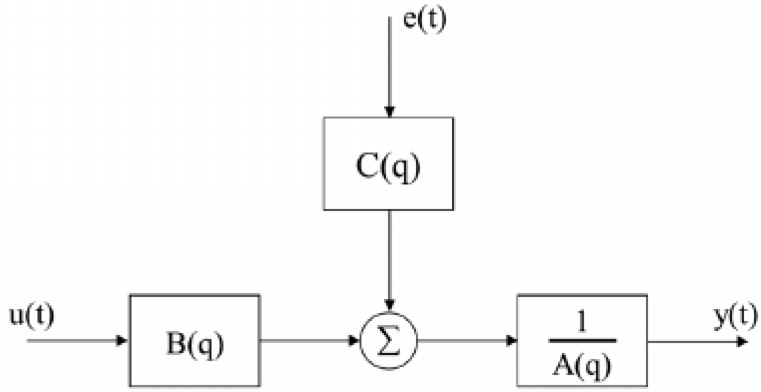


FIGURE 4.27 ARMAX model block diagram

Same can be represented with the help of Equation (4.11).

$$y(t) + a_1 y(t-1) + \cdots + a_{na} y(t-na) = b_1\, u(t-nk) + \cdots + b_{nb}\, u(t-nk-nb+1)$$
$$+c_1 e(t-1) + \cdots + c_{nc} e(t-nc) + e(t) \qquad (4.11)$$

Where

y(t)— Output at time t.

n_a — Number of poles.

n_b — Number of zeroes plus 1.

n_c — Number of C coefficients.

n_k — Number of input samples that occur before the input affects the output, also called the dead time in the system.

y(t−1)…y(t− n_a) — Previous outputs on which the current output depends.

$u(t-n_k)\dots u(t-n_k-n_b+1)$ — Previous and delayed inputs on which the current output depends.

$e(t-1)\dots e(t-n_c)$ — error term.

A more compact way to write the difference Equation is:

$$A(q)y(t) = B(q)u(t-nk) + C(q)e(t)$$

Where

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}$$

$$B(q) = b_1 + b_2 q^{-1} + \cdots + b_{n_b} q^{-n_b+1}$$

$$C(q) = 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}$$

q - The delay operator

The parameters n_a, n_b, and n_c are the orders of the ARMAX model, and nk is the delay.

In the case of time series data, it has no input channels and only one output channel, then ARMAX behaves like an ARMA model for the time series [30] and it has only autoregressive and moving average term. So its order is [n_a n_c]. Now the Equation for ARMA is

$$\mathrm{A(q)y(t) = C(q)e(t)} \tag{4.12}$$

The current output is predicted as a weighted sum of past output values and weighted of past error. The detail structure of ARMA model is depicted in Fig. 4.28.

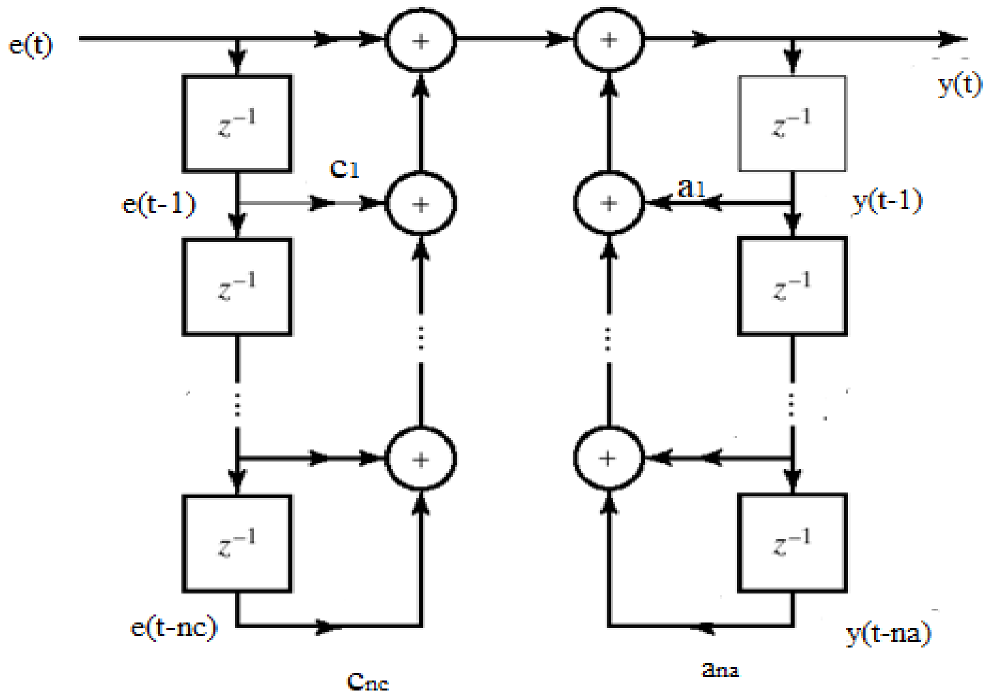


FIGURE 4.28 The detailed structure of ARMA model

For our application, I have tested this model by varying AR order and found an order for which prediction error is minimum. Order of MA is selected as 01. Performance of forecasted response can be validated by comparing with actual response of the same subject.

NLARX model

As a name suggests, NLARX is a nonlinear ARX model, it can be considered an extension of linear ARX models. The nonlinear ARX model has a flexible nonlinear mapping function [30].

$$y(t) = f(y(t-1), y(t-2), \dots \dots u(t), u(t-1), u(t-2) \dots \dots) \qquad (4.13)$$

The input to the function F is the regressors. For time series forecasting application the current output depends on only pasts time-lagged variables so it has an only Autoregressive component and there is no input to the model. The nonlinear model can better represent nonlinear behavior of system compared to a linear models. The nonlinearity estimator comprises both linear and nonlinear functions that act on the model regressors. The structure block diagram of a nonlinear ARX model for time series forecasting is shown in Fig. 4.29.

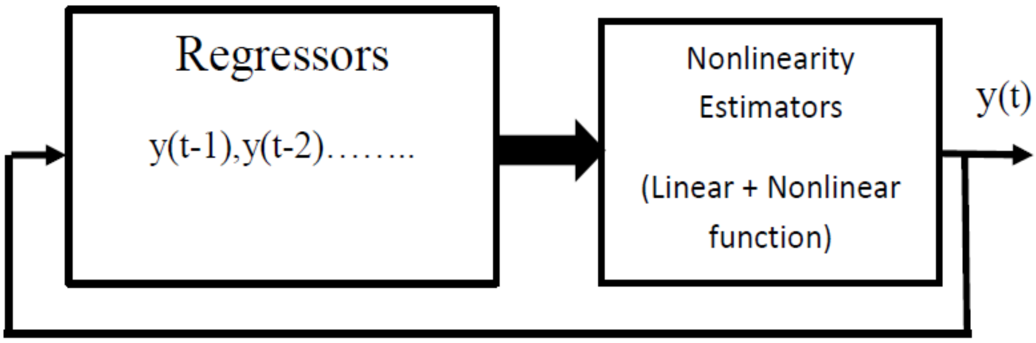


FIGURE 4.29 The structure of nonlinear ARX model for time series forecasting applications

For best representation of a system as an NLARX model, it is necessary to find out a number of regressor/orders. All the regressors are the inputs to both the linear and the nonlinear function blocks of the nonlinearity estimator. The function for nonlinearity estimator can be written in general form as Equation (4.14) [30].

$$F(x) = L^T(x-r) + d + g\big(Q(x-r)\big) \tag{4.14}$$

Where

x is a vector of the Regressors,
r is the mean of the regressors x,
d is a scalar offset,
$(L^T(x-r)+d)$ is the output of the linear function,
$g\big(Q(x-r)\big)$ is the output of the nonlinear function.

Nonlinearity estimators may be tree-partition networks, wavelet networks, and multilayer neural networks. Here wavelet network is used as a nonlinearity estimator, and it is prescribed in Equation (4.15).

$$g(x) = \sum_{k=1}^{n} a_k K\big(\beta_k (x - \gamma_k)\big) \tag{4.15}$$

Where

K is the wavelet function

γ_k , β_k are translation and dilution vector

a_k is the weight coefficient vector

This model can be tested by varying AR order (no. of past Regressors). Performance of forecasted response can be validated by comparing with actual response of the same subject.

4.3.2 Single and Multi-step ahead prediction

All systems of a human body are nonlinear and interconnected with each other. An artificial neural network is capable to behave like nonlinear system after training. Here Concept of the artificial neural network is used to predict single and multi-step-ahead prediction of respiration rate. Single step ahead prediction can be described by nonlinear function F as per Equation (4.16).

$$y(n+1) = F\big(y(n), y(n-1), y(n-2), y(n-3) \ldots\ldots y(n-p)\big) + e(n) \tag{4.16}$$

Where future value y (n+1) can be predicted from past p samples, where P represent the Autoregressive order. e (n) is residual. Similarly, the function of m steps ahead prediction can be described as Equation (4.17).

$$y(n+m) = F\big(y(n), y(n-1), y(n-2), y(n-3) \ldots\ldots y(n-p)\big) + e(n) \tag{4.17}$$

The neural network is a collection of interconnected elements that are first trained to perform a predefined task and later it can be tested to unknown inputs. Performance of neural network depends on a number of hidden layers, a number of neurons in each layer, the structure of network, training datasets, the algorithm used to adjust weights and activation function of neurons. In general time series prediction kind of application only past response is available, the current and past values of the signal used to predict future signal. In such a case nonlinear autoregressive neural network (NARNET) structure is best fit for time series forecasting [31]. A general schematic of NARNET is shown in Fig. 4.30.

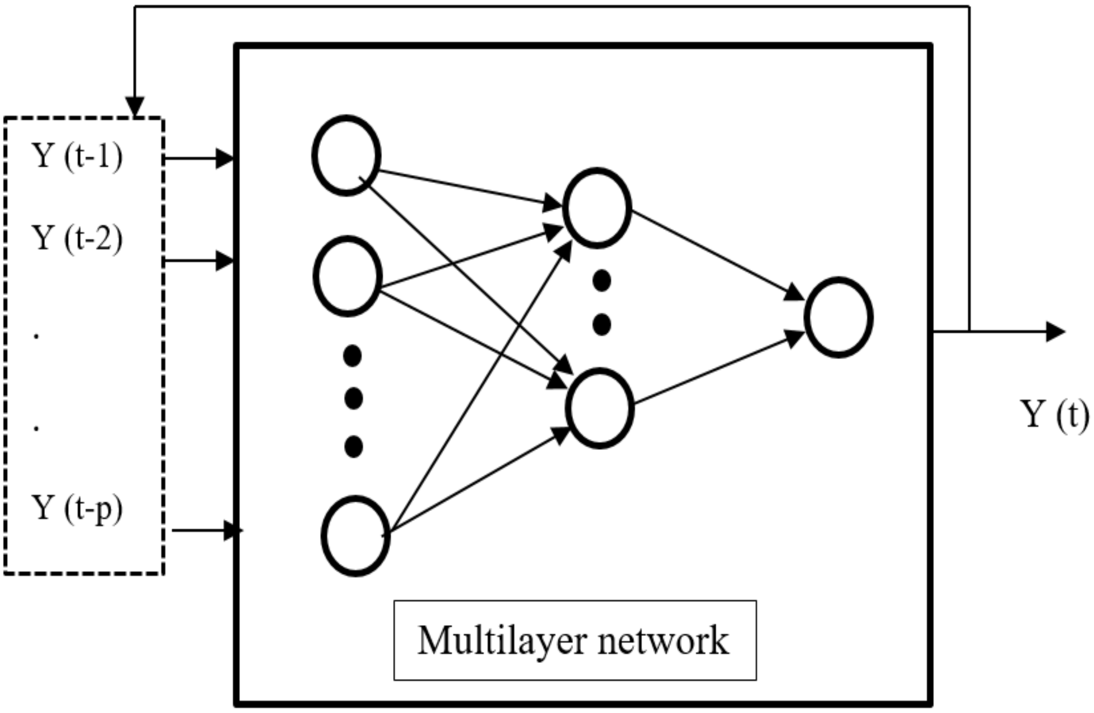


FIGURE 4.30 General schematic of a nonlinear autoregressive network

NARNET with a two hidden layers which has 10 and 5 neurons in two layers, and a number of inputs (previous values) 25 is selected for successful prediction of instant RR in the case of single and multi (5 step) ahead prediction, it is shown in Fig. 4.31. The selected activation function for neuron of hidden layers is tan-sigmoid and for output layer, it is a linear function.

Tan-sigmoid function:

$$y = tansig(x) = \frac{2}{(1 + e^{-2x})} - 1 \tag{4.18}$$

Linear function: $y(x) = x$ (4.19)

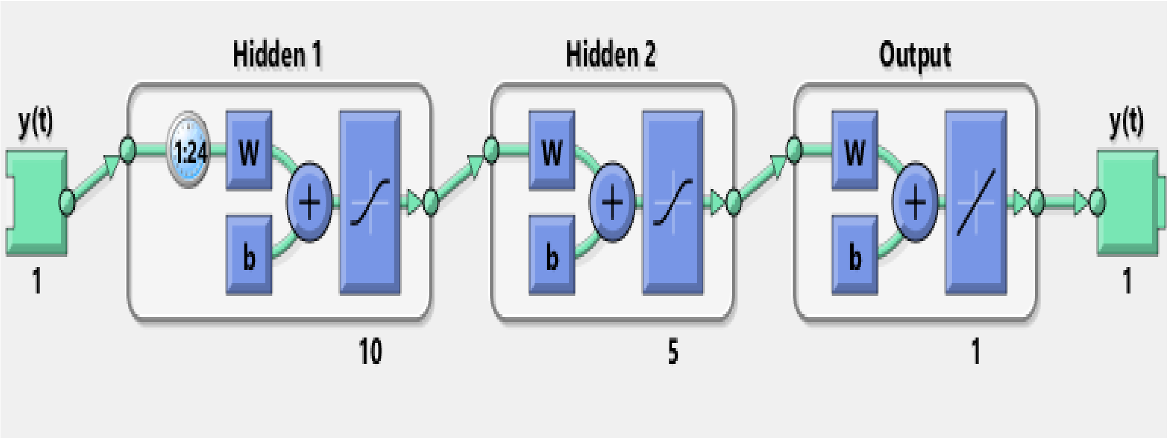


FIGURE 4.31 Implemented Structure of NARNET with 2 hidden layers

For single or multi-step ahead prediction, all subjects selected for the test are normal and young (Normal database), they run up to 3 stages (each stage of 3 minutes) of BRUCE exercise protocol on a treadmill machine. Out of these 9 min, 7 min dataset were for the training of neural network and last 2 min were for validation purpose. For training purpose here levenberg-Marquardt backpropagation algorithm was used that optimized weight and bias values. During training, a dataset of instant RR up to 7 min are divided differently for single and multi-step ahead prediction. For example for 5 step-ahead prediction y(n+5) is the reference output and y (n)…..y (n-24) past output is feedback to the input of the network, in between bias and weight are adjusted with the help of optimization algorithm.

4.4 Validation

To evaluate the performance of forecasting models and to choose best forecasting models for cardiopulmonary response during incremental exercise, the actual forecasted response of models compared with observed values. Statistics like the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE) were applied for meaningful interpretation of deviation between actual and observed values. These types of errors are advisable for time series forecasting in the case of the same scale and the data processing procedures were performed [32].

Mathematical formulas of forecasting errors:

The Root mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2} \tag{4.20}$$

The mean absolute error

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\ (y_i - y_i')\ | \tag{4.21}$$

The mean absolute percentage error

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} |\frac{(y_i - y_i')}{y_i}| \tag{4.22}$$

Where y_i - The observed actual value and

y_i' - The model's forecasted future response.

RMSE indicates–how close the observed data values are to the models predicted values. It has the property of being in the same units as the response signal. The lower value of RMSE indicates the better fit of model's response with observed value. As here the opposite signed errors do not offset one another, RMSE gives an overall idea of the error occurred during forecasting. The MAE is the average of the absolute value of the difference between the forecasted value and the actual value. For a good forecast, the obtained MAE should be as small as possible. The MAE uses the same scale as the signal being measured. The MAE and RMSE are scale-dependent accuracy measures and therefore cannot be used to make comparisons between signals series using different scales. MAPE measures prediction of the accuracy of a forecasting method in statistics and usually expressed as a percentage. It is independent of the scale of measurement and opposite signed errors do not offset each other.

All three kinds of errors is linear to each other with some scaling factor so anyone can utilized any error for model comparison purpose.

In case of normal healthy databases, I had recorded ECG and breathing signal under exercise conditions as mention above for 540 sec (9 min) from that almost 420 seconds data (7 min or 77.77% of total span) was utilized for training of models and after successful training, same trained model was used to forecast future response for the duration of 421 to 540 seconds (8^{th} and 9^{th} min, later 22.23 % of total span). The forecasted response is being compared with the actual response to evaluate the performance of individual models and in between models comparison purpose.

And in case of the Abnormal database: “The MIT-BIH ST Change Database” contains only variable length ECG signals of cardiac abnormal subjects. So for the performance evaluation of models, Initial 77.77% of the total span is used for training of models and the remaining for validation of models. Same way performance of models is also tested by changing forecasted duration by selecting initial 66.66% of total span for training and remaining for validation of models.

The order selection of models is again a challenging task, the performance of three models is tested for a set of order range. For the Adaptive model it is tested by varying FIR order, ARMAX and NLARX model are tested for AR order range. For Instant HR, order range is 1 to 100 and for instant RR, it is 1 to 50 for all models. All models are tested on 20 normal subjects’ database recorded in the laboratory as well as 20 abnormal databases collected from the physionet.org online source.

For the case of single and multi-step ahead prediction, NARNET model is validated with the Normal database. The network was trained with past 7 min database of instant RR, the same model is used for single step and 5 step ahead prediction in the forecasted duration (later 2 min, 8^{th}-9^{th} min). To check the performance of the network, predicted response is compared with an actual response (8^{th} and 9^{th} min), and above discussed statistical errors are calculated.

CHAPTER-5

Results And Discussions

This section presents the forecasting results of three models and their forecasted error comparisons for both kinds of databases, Normal database as well as abnormal cardiac database. Models' parameter estimation only depends on their own past data (the recorded data of the initial period of the test) so there is no need for any similar database for reference or training purpose. This is the biggest advantage of the proposed system.

5.1 Time Series Forecasting

Forecasting can be understood as selecting models (or Parameters of Model) that best fit to historical past data and using them to predict future observations. The model's performance for time series forecasting is determined by its performance at predicting the future.

5.1.1 Results of the Normal database

All normal subjects had run on Treadmill machine for 9 min, during that ECG and breathing signals are recorded, from that the vector of instant HR and Instant RR are derived through above discussed signal processing steps. Out of 9 min (540 sec), signal's initial 420 sec (77.77% of total span) data of both vectors were utilized for training of models and remaining 421^{st} to 540^{th} sec data were used to validate the performance of Models. After training the models have forecasted future behavior of signals for 421^{st} sec to 540^{th} sec. The performance of all three models: Adaptive filter, ARMAX and NLARX are tested on both signals. And for comparison of models performance, the statistical errors: RMSE, MAE and MAPE, are calculated and compared with others. For the first kind of normal databases, experimental results of 20 subjects are discussed in table 5.1 and table 5.2. In table 5.1, results of three models for normal subject's Instant HR are presented. In table 5.2, results of three models for normal subject's Instant RR are presented. Order mentioned in both tables are corresponding to the minimum forecasting error. In case of Adaptive filter it is FIR filter order and for ARMAX and NLARX models it is Autoregressive order. Orders are selected after testing all models for range of orders like for Instant HR, models are tested for 1 to 100 and for Instant RR models are tested for 1 to 50.

Adaptive Model is used in such a way that parameters of FIR filter can be estimated from available past time series of Instant HR & RR. Here, a sample result of database 06 is represented in Fig. 5.1 and 5.2. In Fig. 5.1, the forecasting results of adaptive filter are depicted. On the top graph, RMSE corresponding to FIR order 1 to 100 is shown. From that order 02 has minimum error that is why instant HR signal is estimated for forecasted duration 421st sec to 540th sec with FIR order 02. And calculated RMSE for forecasted duration is 1.4634. Similarly, for instant RR, minimum RMSE found for the forecasted duration is 1.2356 (Order: 50) as shown in Fig. 5.2.

A sample result of database 06 for ARMAX model is discussed here. For database 6, Minimum RMSE error found for the forecasted duration is 1.1487 (Order: 18) for instant HR shown in Fig. 5.3 and similarly, for instant RR, minimum RMSE found for the forecasted duration is 1.0993 (Order: 16) as shown in Fig. 5.4.

Similarly, the performance of NLARX model tested on database 06 of the normal database is discussed here. For database 06 of normal subjects, Minimum RMSE error found for the forecasted duration is 1.1359 (Order: 73) in time series forecasting of instant HR shown in Fig. 5.5. And similarly, for instant RR, minimum RMSE found for the forecasted duration is 0.6731 (Order: 07) as shown in Fig. 5.6.

In Table 5.1, the models' forecasting performance of instant HR of 20 Normal subjects are summarized. In Table 5.2, the models' forecasting performance of instant RR of 20 Normal subjects are summarized.

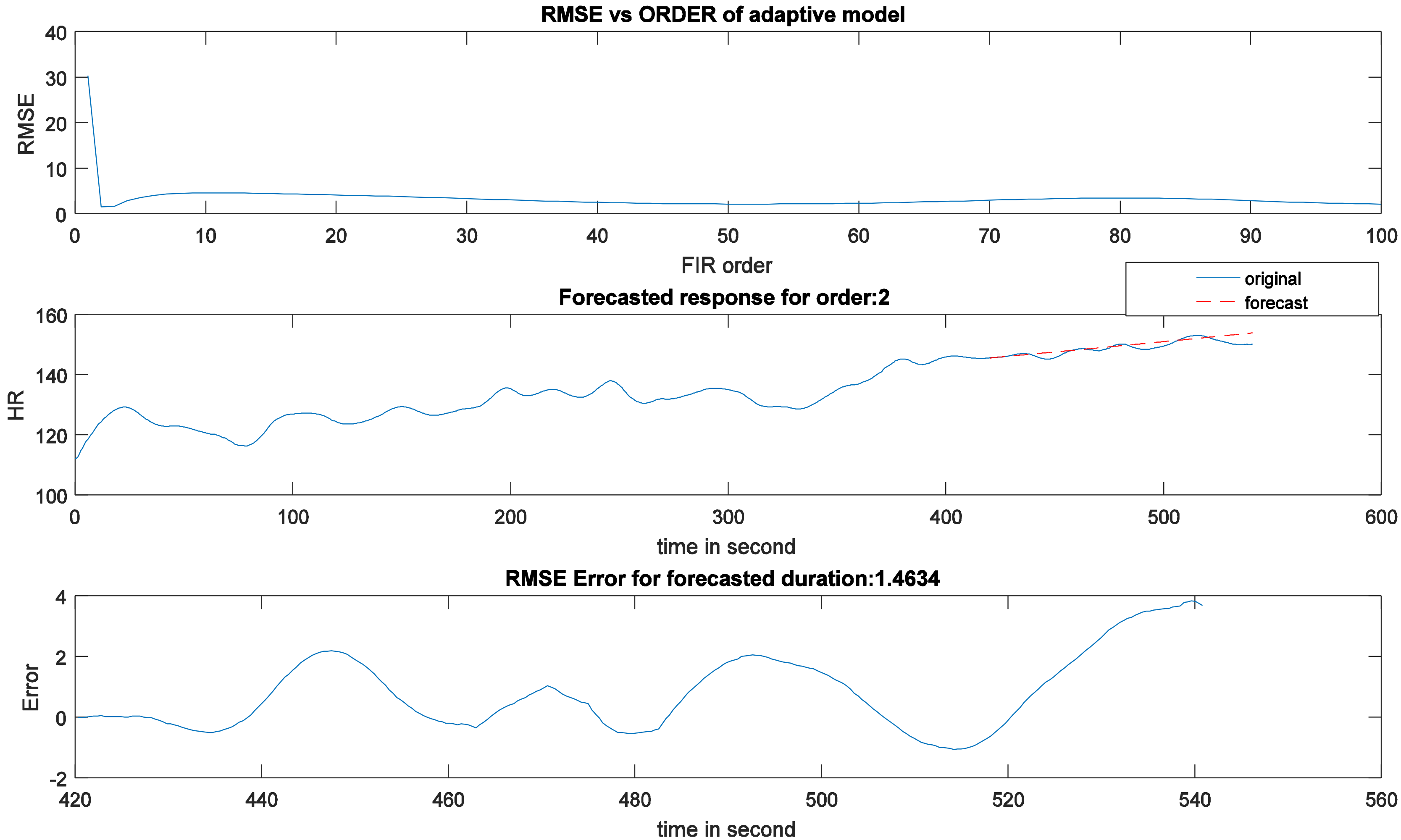


FIGURE 5.1 The forecasting result of Adaptive filter for the instant HR of normal database 06. RMSE vs. FIR order (Top), Time series forecasting of instant HR for duration 421st to 540th sec, order 02 (Middle), Error plot: RMSE 1.4634 (Bottom)

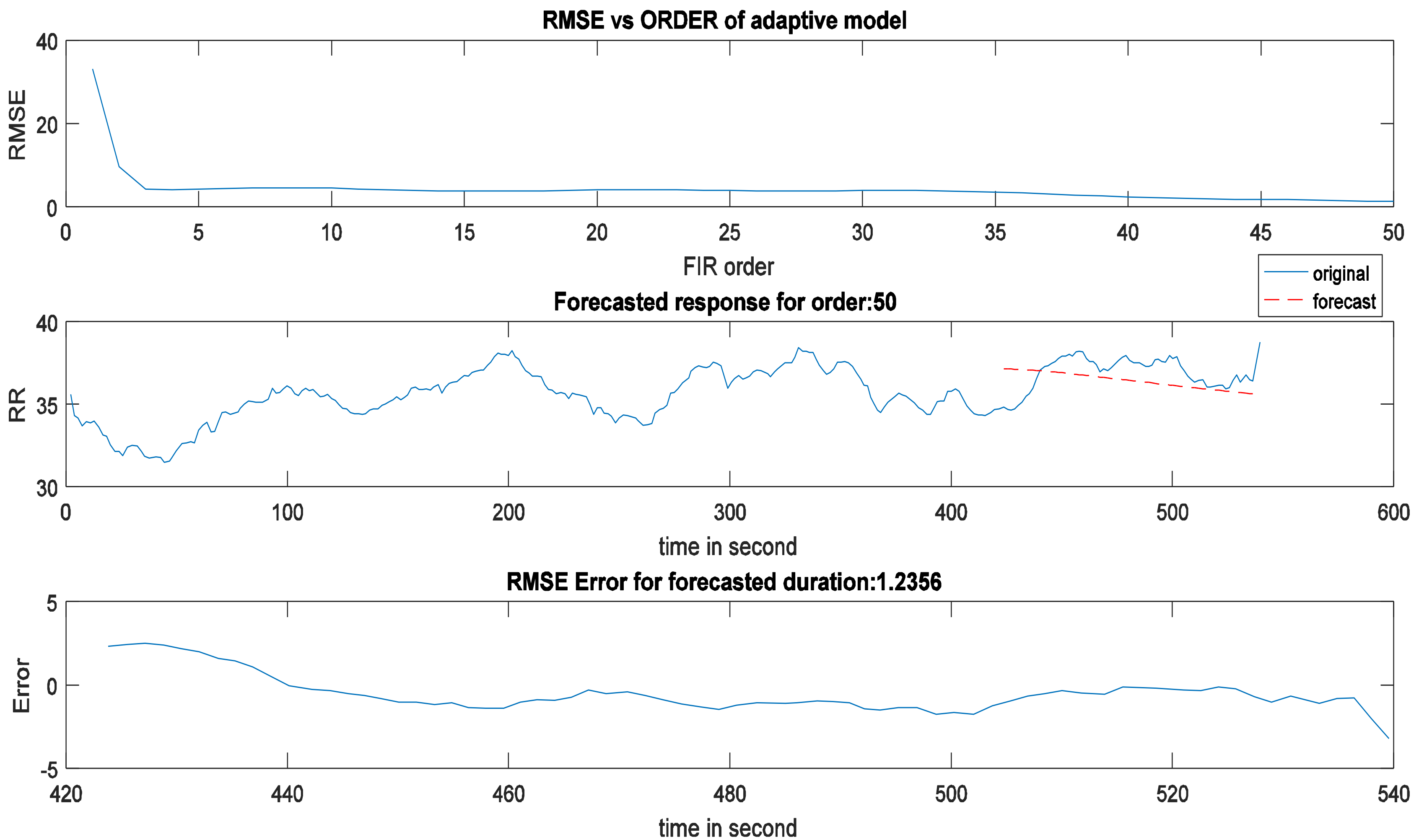


FIGURE 5.2 the forecasting result of Adaptive filter for the instant RR of normal database 06. RMSE vs. FIR order (Top), Time series forecasting of instant RR for duration 421st to 540th sec, order 50 (Middle), Error plot: RMSE 1.2356 (Bottom)

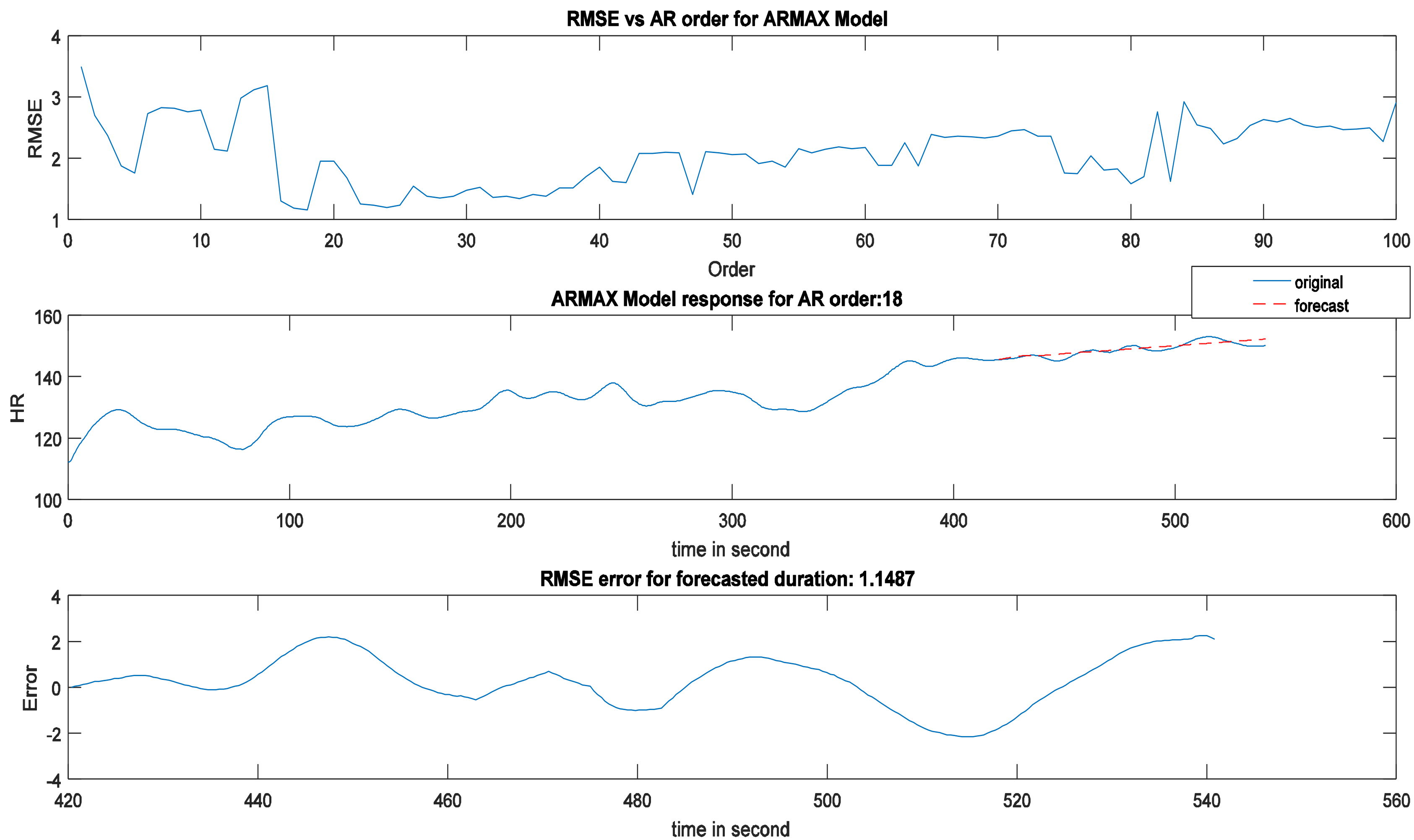


FIGURE 5.3 the forecasting result of ARMAX model for the instant HR of normal database 06. RMSE vs. AR order (Top), Time series forecasting of instant HR for duration 421st to 540th sec, order 18 (Middle), Error plot: RMSE 1.1487 (Bottom)

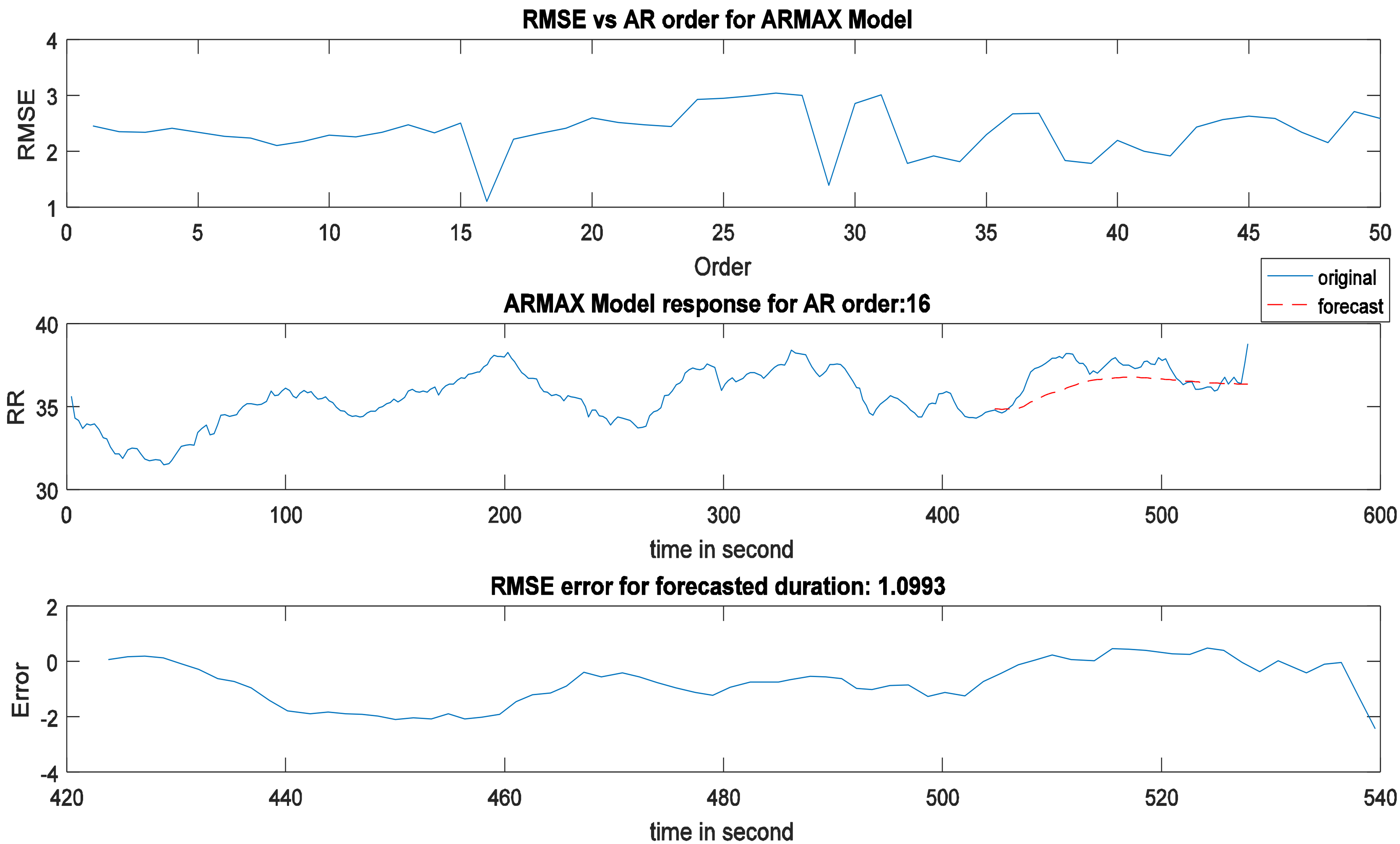


FIGURE 5.4 the forecasting result of ARMAX model for the instant RR of normal database 06. RMSE vs. AR order (Top), Time series forecasting of instant RR for duration 421st to 540th sec, order 16 (Middle), Error plot: RMSE 1.0993 (Bottom)

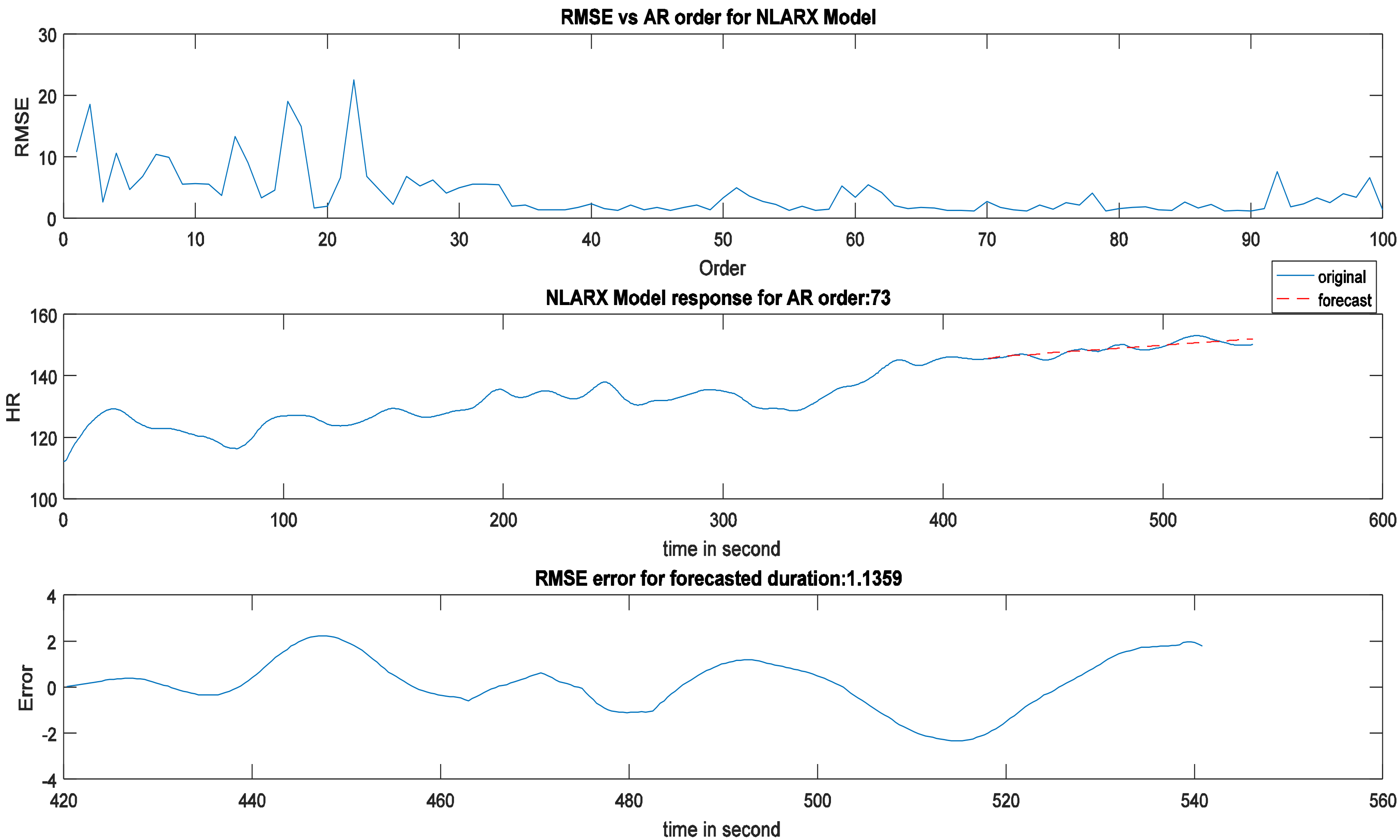


FIGURE 5.5 the forecasting result of NLARX model for the instant HR of normal database 06. RMSE vs. AR order (Top), Time series forecasting of instant HR for duration 421st to 540th sec, order 73 (Middle), Error plot: RMSE 1.1359 (Bottom)

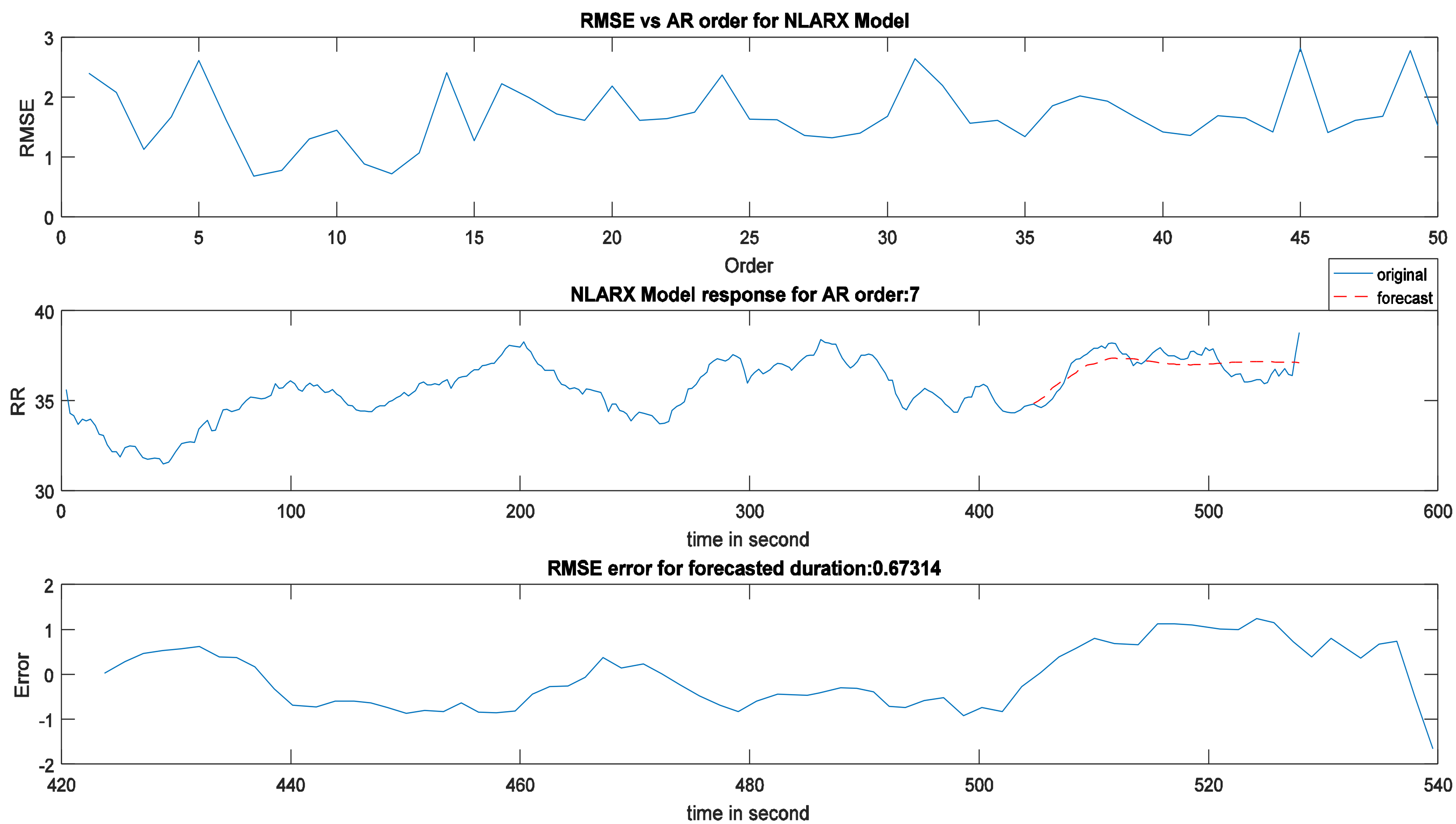


FIGURE 5.6 the forecasting result of NLARX model for the instant RR of normal database 06. RMSE vs. AR order (Top), Time series forecasting of instant RR for duration 421st to 540th sec, order 7 (Middle), Error plot: RMSE 0.67314 (Bottom)

TABLE 5.1 Summary of the models' forecasting performance for **instant HR of 20 Normal subjects**. 421st to 540th sec is the forecasted time horizon.

Database No	M/F	Age	ADAPTIVE HR				ARMAX HR				NLARX HR			
			RMSE	MAE	MAPE	Order	RMSE	MAE	MAPE	Order	RMSE	MAE	MAPE	Order
1	M	22	4.0561	2.4696	0.0164	13	4.0946	2.5650	0.0170	5	4.0203	2.3706	0.0158	24
2	M	32	4.1188	3.4694	0.0208	5	5.2245	4.4805	0.0269	1	4.5433	3.5578	0.0213	71
3	M	29	8.6854	7.1868	0.0480	2	5.0950	4.4853	0.0304	4	4.1992	3.3409	0.0230	3
4	M	22	21.0052	18.9124	0.1111	1	3.0104	2.1445	0.0124	100	2.7108	2.1231	0.0123	12
5	M	22	5.1349	3.9043	0.0247	6	4.5094	3.4744	0.0220	11	5.4004	4.3260	0.0274	7
6	M	22	1.4634	1.0720	0.0072	2	1.1487	0.9101	0.0061	18	1.1359	0.8938	0.0060	73
7	F	20	11.0783	8.6421	0.0437	1	10.3601	7.9951	0.0404	67	4.8783	3.7686	0.0193	23
8	M	21	10.3314	5.9953	0.0287	46	10.1109	5.9160	0.0284	4	6.6782	4.6252	0.0226	14
9	F	19	9.3139	7.5198	0.0532	3	9.7449	7.9931	0.0579	2	14.0284	11.7342	0.0869	1
10	M	21	14.5270	12.5281	0.0841	2	7.4814	6.4153	0.0429	2	3.0266	2.3108	0.0157	2
11	F	24	3.8951	3.0475	0.0170	25	6.9537	5.0707	0.0279	44	3.8708	3.0787	0.0172	36
12	F	18	15.7193	13.6067	0.0749	1	6.2117	4.8231	0.0262	22	5.6148	4.5306	0.0244	14
13	M	24	35.1256	32.4337	0.2365	2	8.0850	7.1398	0.0521	23	5.5951	4.7900	0.0353	85
14	F	23	2.3678	1.4971	0.0151	4	2.3757	1.4711	0.0148	82	1.6882	1.0562	0.0106	18
15	F	22	13.6536	11.8997	0.0643	1	5.2399	4.2035	0.0225	60	5.3060	4.1752	0.0223	81
16	F	22	3.0704	2.8140	0.0149	1	1.7617	1.5907	0.0084	1	1.7606	1.5662	0.0083	60
17	M	24	1.0556	0.8997	0.0052	1	4.2965	3.1668	0.0183	91	1.1994	0.9924	0.0057	17
18	M	25	13.2033	9.8836	0.0638	97	13.2033	9.4838	0.0612	99	4.0183	2.4545	0.0155	22
19	M	24	4.3865	3.4033	0.0187	4	4.6192	3.7983	0.0209	11	3.6509	2.9882	0.0164	61
20	F	20	6.7792	5.0633	0.0310	18	8.0962	5.9083	0.0357	31	6.6941	4.8224	0.0294	96
			ADAPTIVE				ARMAX				NLARX			
MEAN			9.4485				6.0811				4.5010			
STD. DEVI (σ)			8.1386				3.1607				2.8009			

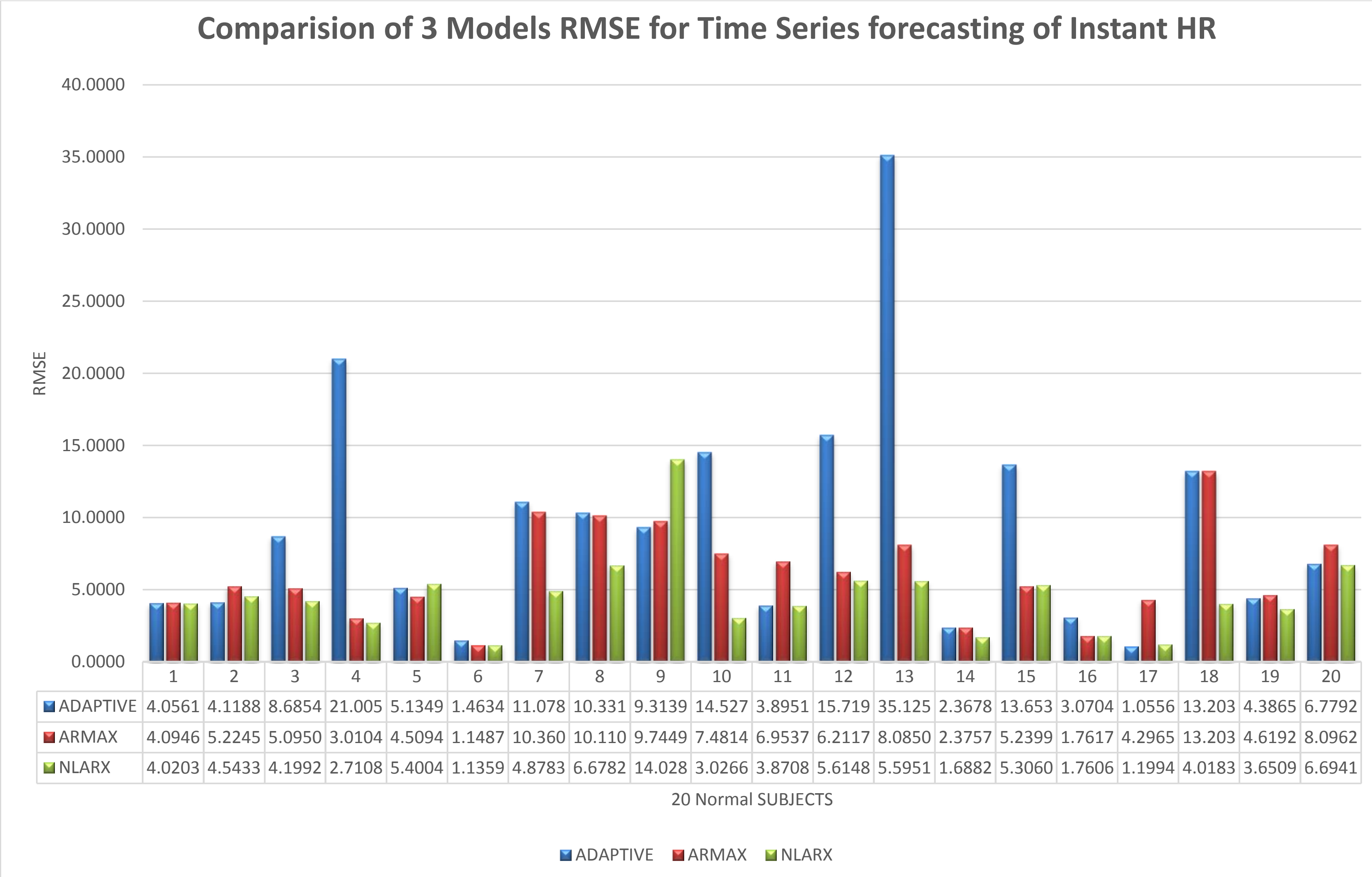


FIGURE 5.7 Comparison of the model's forecasting RMSE for Instant HR of 20 Normal subjects (Reference to Table 5.1)

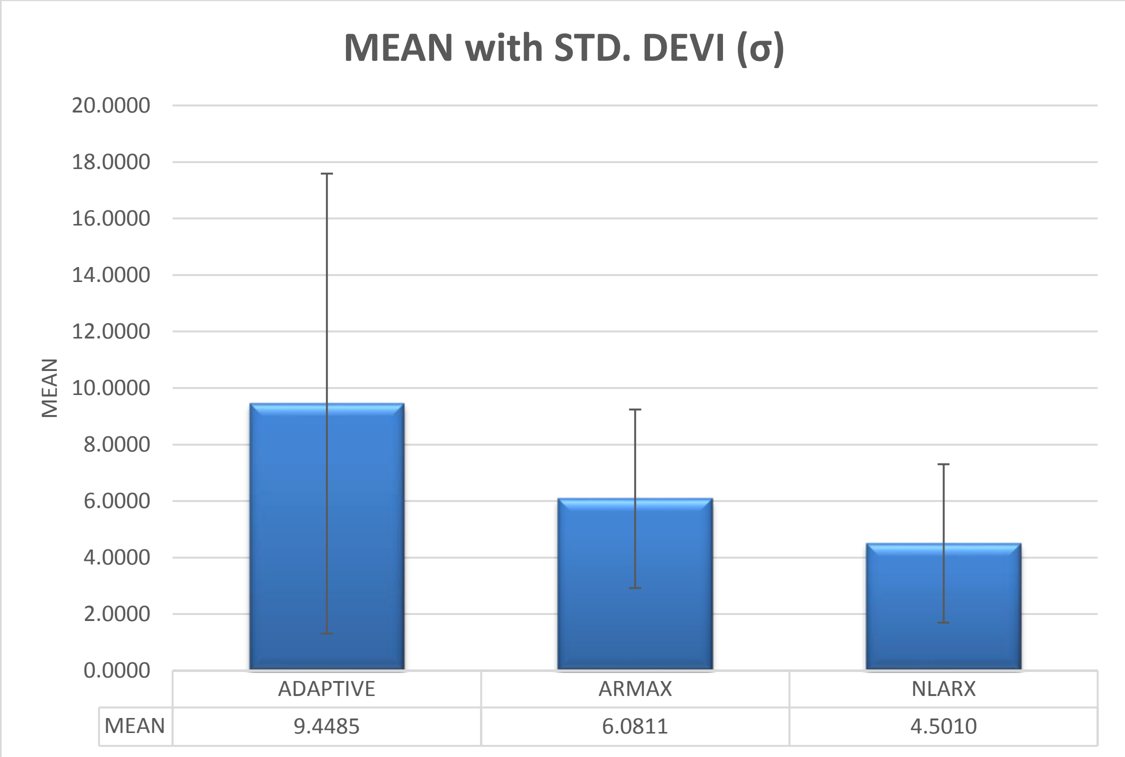
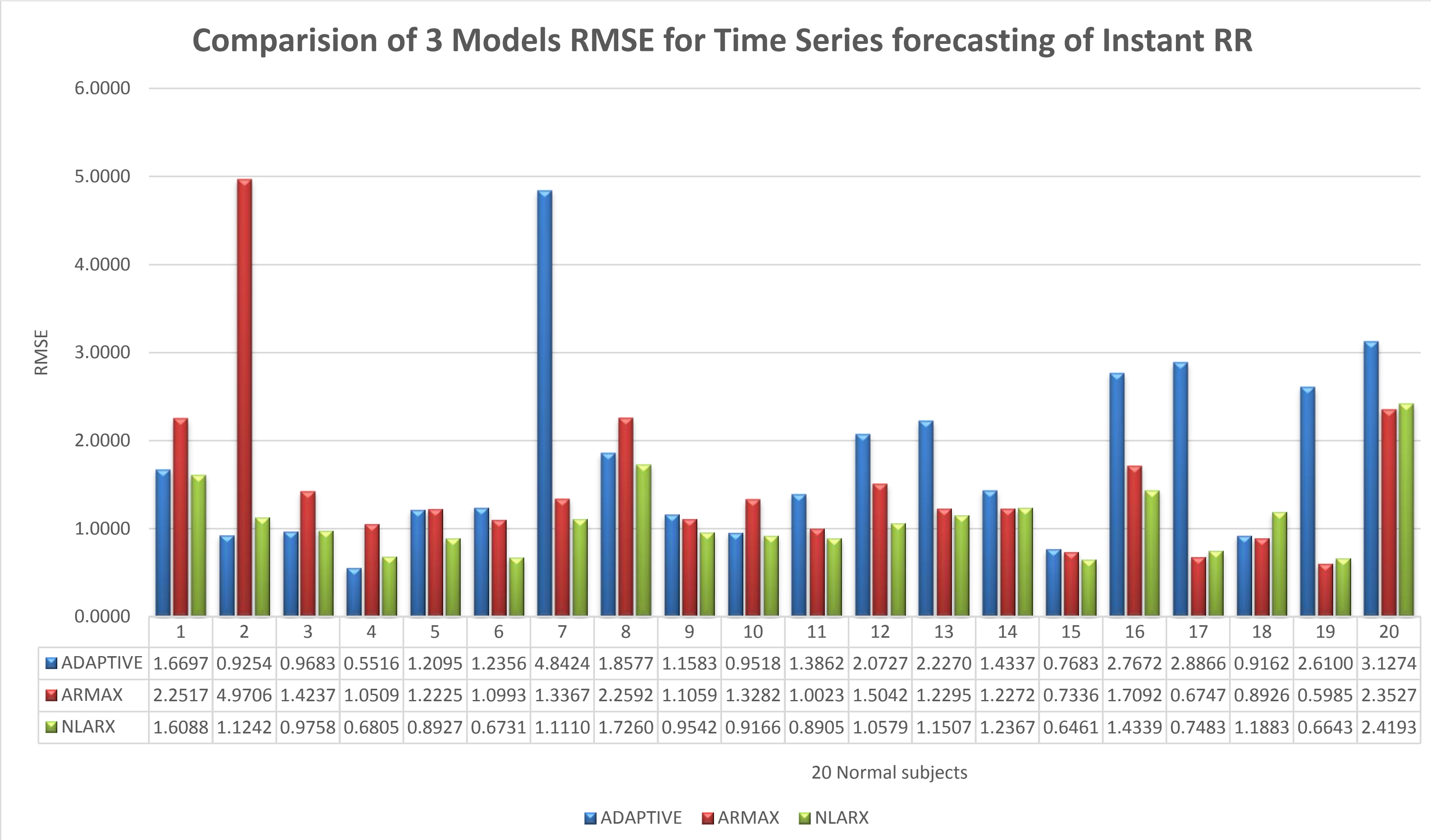


FIGURE 5.8 Mean and Standard deviation of three models RMSE for Instant HR of 20 Normal subjects (reference to Table 5.1)

In Table 5.1, Summary of all three models' forecasting performance for instant HR of 20 Normal subjects is shown. And for visual comparison of the same results, it is shown in column chart format in Fig. 5.7. All three errors: RMSE, MAE, MAPE are linearly proportional to each other with some scaling factor, so for models' comparison purpose we can choose any one error. In a column chart, I have considered only RMSE. In the majority of cases, NLARX model has less or near to less error compared to other two models. So to prove the reliability of models in time series forecasting of instant HR during exercise condition, I have calculated MEAN and standard deviation (σ) and depicted in Fig. 5.8. Standard deviation indicates how values are spread around the mean of the dataset. Low standard deviation shows that the data is clustered closely around the mean hence it is more reliable. Here the NLARX model is found to have a low value of MEAN and standard deviation, hence it is more reliable and it has low average forecasting error compared to Adaptive and ARMAX model. The adaptive model has the worst performance for time series forecasting of instant HR during exercise condition.

TABLE 5.2 Summary of the models' forecasting performance for **instant RR** of 20 Normal subjects. 421st sec to 540th sec is the forecasted time horizon

Database No	Male/ Female	Age	ADAPTIVE RR				ARMAX RR				NLARX RR			
			MSE	MAE	MAPE	Order	RMSE	MAE	MAPE	Order	RMSE	MAE	MAPE	Order
1	M	22	1.6697	1.4596	0.0380	34	2.2517	1.6921	0.0428	19	1.6088	1.3422	0.0350	19
2	M	32	0.9254	0.8357	0.0152	46	4.9706	4.4668	0.0805	9	1.1242	0.8199	0.0148	25
3	M	29	0.9683	0.8132	0.0264	7	1.4237	1.1126	0.0365	6	0.9758	0.7954	0.0259	23
4	M	22	0.5516	0.4606	0.0119	47	1.0509	0.9087	0.0234	10	0.6805	0.5775	0.0149	26
5	M	22	1.2095	0.9431	0.0283	5	1.2225	0.9758	0.0292	1	0.8927	0.6948	0.0208	1
6	M	22	1.2356	1.0529	0.0284	50	1.0993	0.8727	0.0239	16	0.6731	0.5969	0.0162	7
7	F	20	4.8424	4.4541	0.0940	3	1.3367	1.0681	0.0224	50	1.1110	0.9000	0.0188	21
8	M	21	1.8577	1.6088	0.0315	1	2.2592	1.6965	0.0322	1	1.7260	1.4492	0.0280	2
9	F	19	1.1583	0.9768	0.0278	17	1.1059	0.8835	0.0253	15	0.9542	0.8345	0.0236	9
10	M	21	0.9518	0.6700	0.0182	39	1.3282	0.9936	0.0266	8	0.9166	0.6710	0.0183	1
11	F	24	1.3862	1.2437	0.0297	25	1.0023	0.8050	0.0192	1	0.8905	0.7608	0.0184	9
12	F	18	2.0727	1.6440	0.0383	3	1.5042	1.2323	0.0280	1	1.0579	0.8745	0.0200	5
13	M	24	2.2270	1.9712	0.0543	6	1.2295	1.0903	0.0299	32	1.1507	0.8260	0.0223	21
14	F	23	1.4337	1.1866	0.0432	13	1.2272	1.0596	0.0384	45	1.2367	1.0552	0.0385	42
15	F	22	0.7683	0.6568	0.0168	3	0.7336	0.6264	0.0160	6	0.6461	0.5606	0.0144	36
16	F	22	2.7672	2.6167	0.0757	6	1.7092	1.5571	0.0450	37	1.4339	1.2979	0.0374	14
17	M	24	2.8866	2.5780	0.0780	5	0.6747	0.5312	0.0159	6	0.7483	0.6092	0.0181	25
18	M	25	0.9162	0.7665	0.0245	27	0.8926	0.7416	0.0238	10	1.1883	0.7697	0.0254	19
19	M	24	2.6100	2.0670	0.0554	32	0.5985	0.4829	0.0127	22	0.6643	0.5498	0.0144	33
20	F	20	3.1274	2.7062	0.0624	3	2.3527	1.8193	0.0413	25	2.4193	1.8992	0.0436	26
			ADAPTIVE				**ARMAX**				**NLARX**			
MEAN			1.7783				1.4987				1.1049			
STD. DEVI (σ)			1.0519				0.9586				0.4334			



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ADAPTIVE	1.6697	0.9254	0.9683	0.5516	1.2095	1.2356	4.8424	1.8577	1.1583	0.9518	1.3862	2.0727	2.2270	1.4337	0.7683	2.7672	2.8866	0.9162	2.6100	3.1274
ARMAX	2.2517	4.9706	1.4237	1.0509	1.2225	1.0993	1.3367	2.2592	1.1059	1.3282	1.0023	1.5042	1.2295	1.2272	0.7336	1.7092	0.6747	0.8926	0.5985	2.3527
NLARX	1.6088	1.1242	0.9758	0.6805	0.8927	0.6731	1.1110	1.7260	0.9542	0.9166	0.8905	1.0579	1.1507	1.2367	0.6461	1.4339	0.7483	1.1883	0.6643	2.4193

FIGURE 5.9 Comparison of the model's forecasting RMSE for Instant RR of 20 Normal subjects (Reference to Table 5.2)

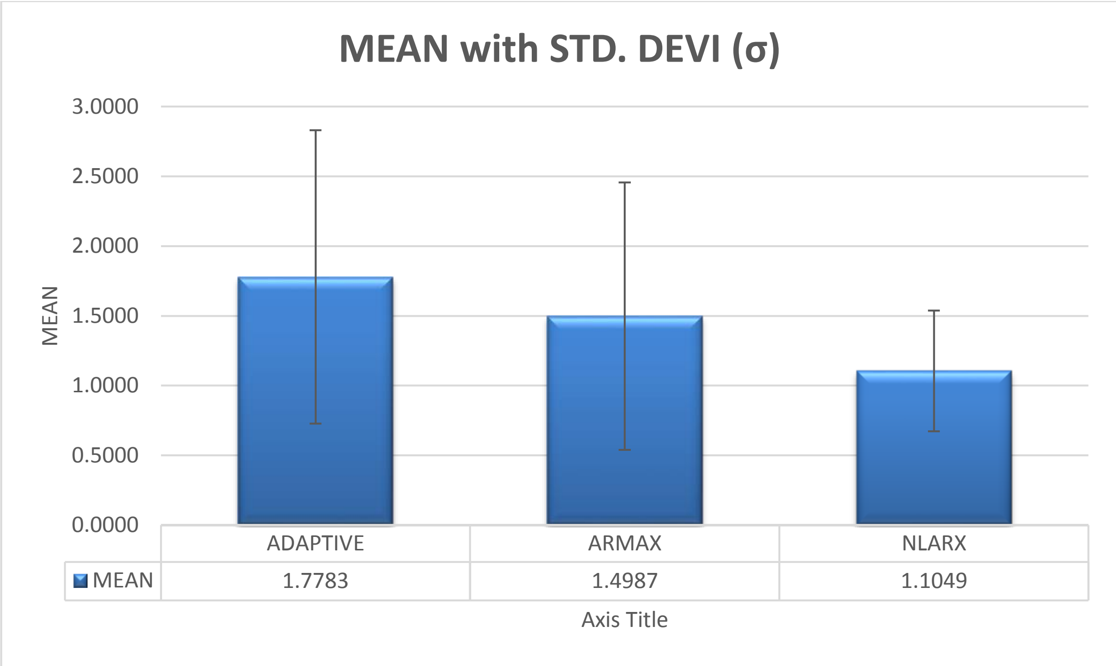


FIGURE 5.10 Mean and Standard deviation of models RMSE for Instant RR of 20 Normal subjects (reference to Table 5.2)

In Table 5.2, Summary of all three models' forecasting performance for instant RR of same 20 Normal subjects is shown. And for visual comparison of the same results, it is shown in column chart format in Fig. 5.9. In a column chart, I have considered only RMSE. In the majority of case, NLARX model has less or near to less error compared to other models. So to prove the reliability of models in time series forecasting of instant RR during exercise condition, I have calculated MEAN and standard deviation (σ) and depicted in Fig. 5.10. Standard deviation indicates how values are spread around the mean of the dataset. Low standard deviation shows that the data is clustered closely around the mean hence it is more reliable. Here the NLARX model is found have a low value of MEAN and standard deviation, hence it is more reliable and it has low average forecasting error compare to other Adaptive and ARMAX model. The adaptive model has the worst performance for time series forecasting of instant RR during exercise condition.

5.1.2 Results of the abnormal database

Cardiopulmonary exercise testing has a variety of applications in sports and medical diagnostic field (more precisely cardiac abnormality). To test the model's performance over cardiac patients, I have used "The MIT-BIH ST Change Database" [33]. This database includes ECG recordings of variable time length, recorded during exercise stress tests and exhibit transient ST change. The ST segment abnormality (elevation or depression) is an indication of myocardial ischemia or infarction. Here a detailed description of incremental exercise condition is not specified. The performance of three models for time series forecasting of instant HR derived from abnormal database: “The MIT-BIH ST Change Database” are shown in Table 5.3 and Table 5.4 for different forecasting duration. These databases are of variable lengths so their starting and ending time of exercise mentioned in both tables are derived from Instant HR plot on the assumption that HR increases with incremental exercise.

In selected abnormal databases, Patients have done exercise for different exercise duration, hence for the uniformity, signal’s initial 77.77% of net exercise duration is utilized for training of models and remaining for the validation of models. And the detailed simulation results of abnormal databases are shown in Table 5.3. As a sample to discuss results, the database 308 is used, this subject had done exercised from 352nd sec to 950th sec (so the net length of exercise is 598 seconds), out of that signal of 352nd sec to 816th sec (77.77% of total span) was used for training of model and signal of 817th sec to 950th sec was reserved for validation of models. Results of three models: Adaptive filter, ARMAX and NLARX models are displayed in sequence in the Fig. 5.11, 5.12 & 5.13 for the same database.

In Fig. 5.11, the forecasting results of adaptive filter are depicted. On the top graph, RMSE corresponding to FIR order 1 to 100 is shown. From that order 04 has minimum error that is why instant HR signal is estimated for forecasted duration 817th sec to 950th sec with FIR order 04. And calculated RMSE for forecasted duration is 5.7074. Same way Fig. 5.12 and 5.13 can be interpreted for ARMAX and NLARX models.

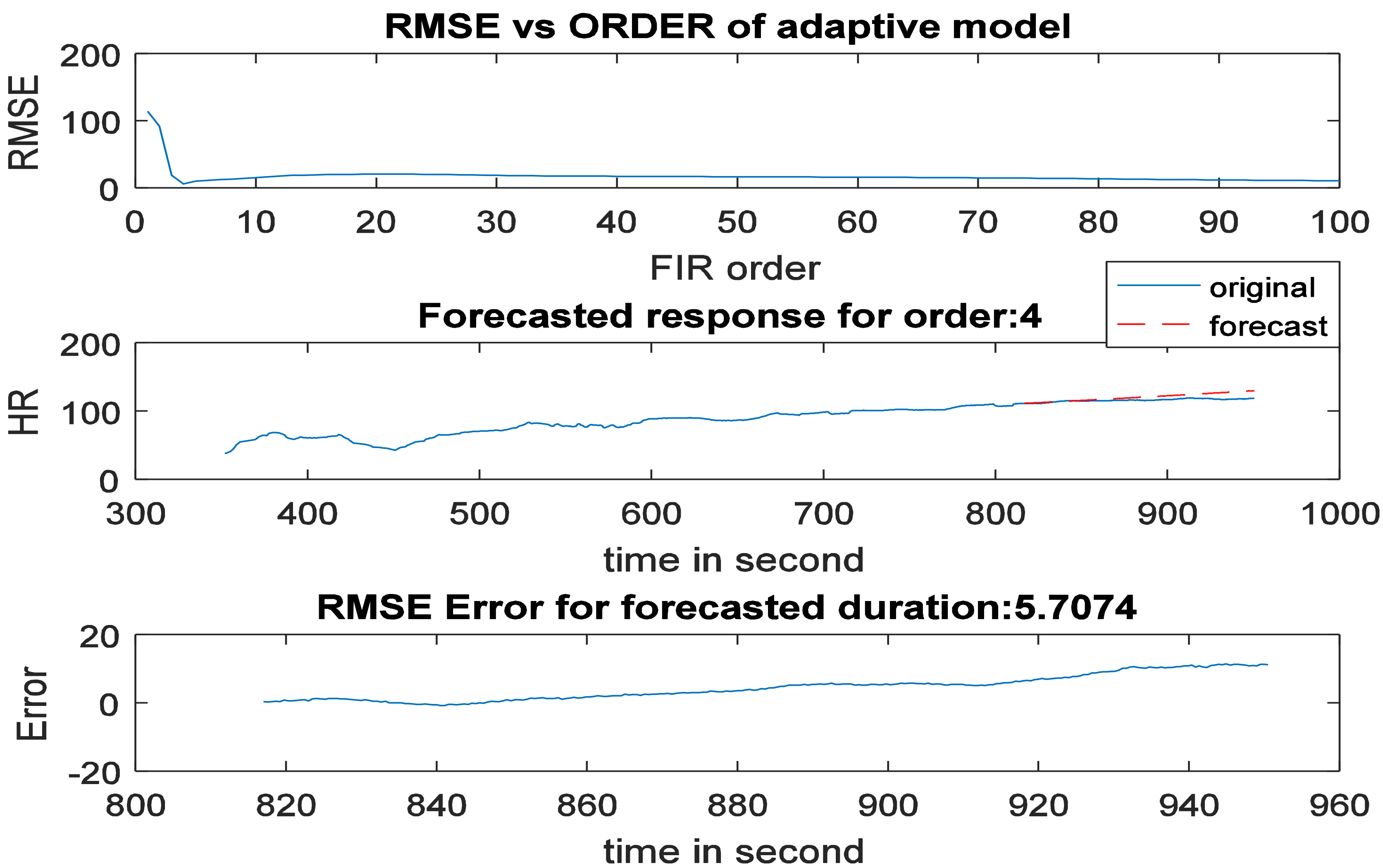


FIGURE 5.11 the forecasting result of an Adaptive model for the instant HR of abnormal database 308. RMSE vs. FIR order (Top), Time series forecasting of instant HR for duration 817th to 950th sec, order 04 (Middle), Error plot: RMSE 5.7074 (Bottom)

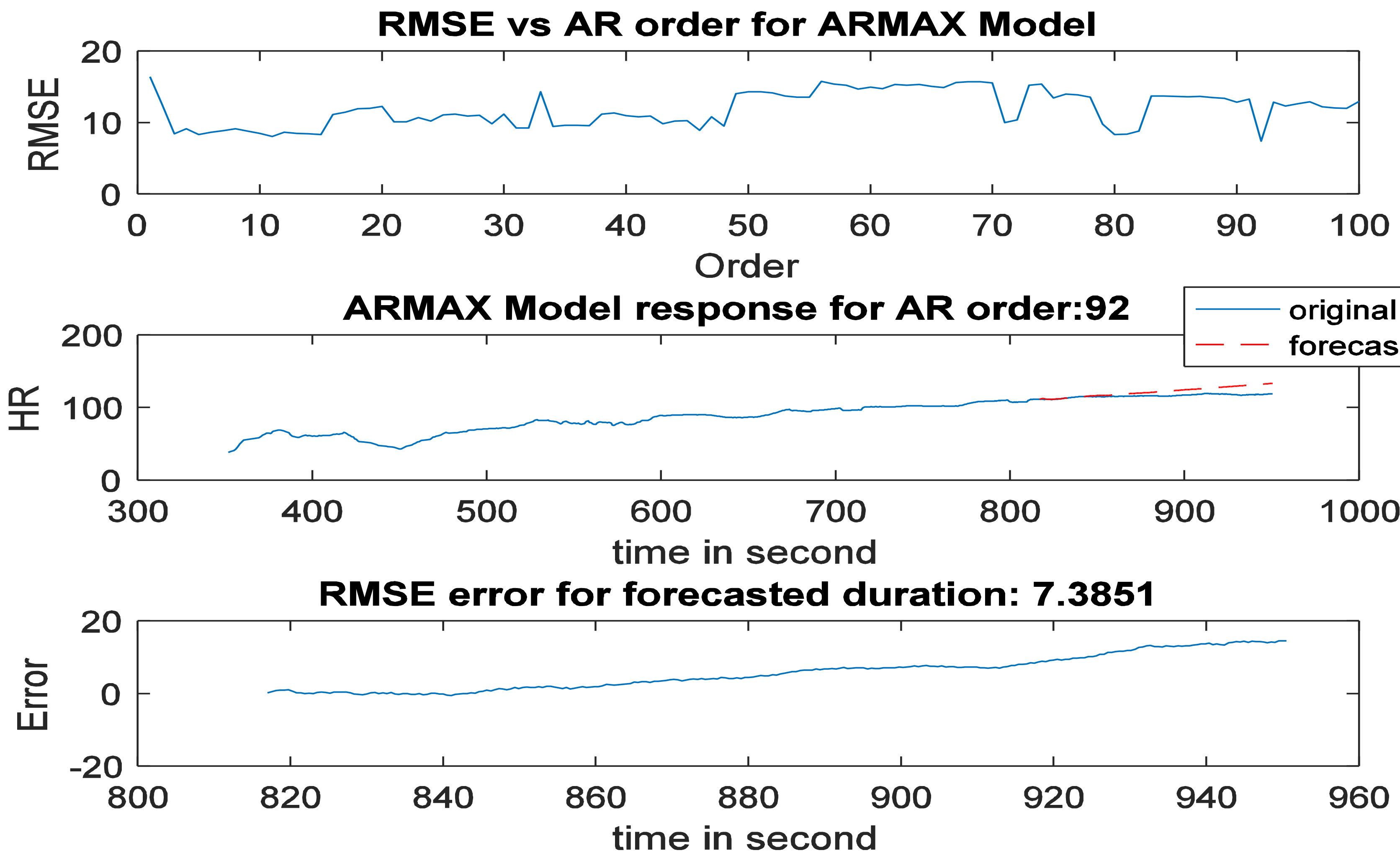


FIGURE 5.12 the forecasting result of ARMAX model for the instant HR of abnormal database 308. RMSE vs. AR order (Top), Time series forecasting of instant HR for duration 817th to 950th sec, order 92 (Middle), Error plot: RMSE 7.3851 (Bottom)

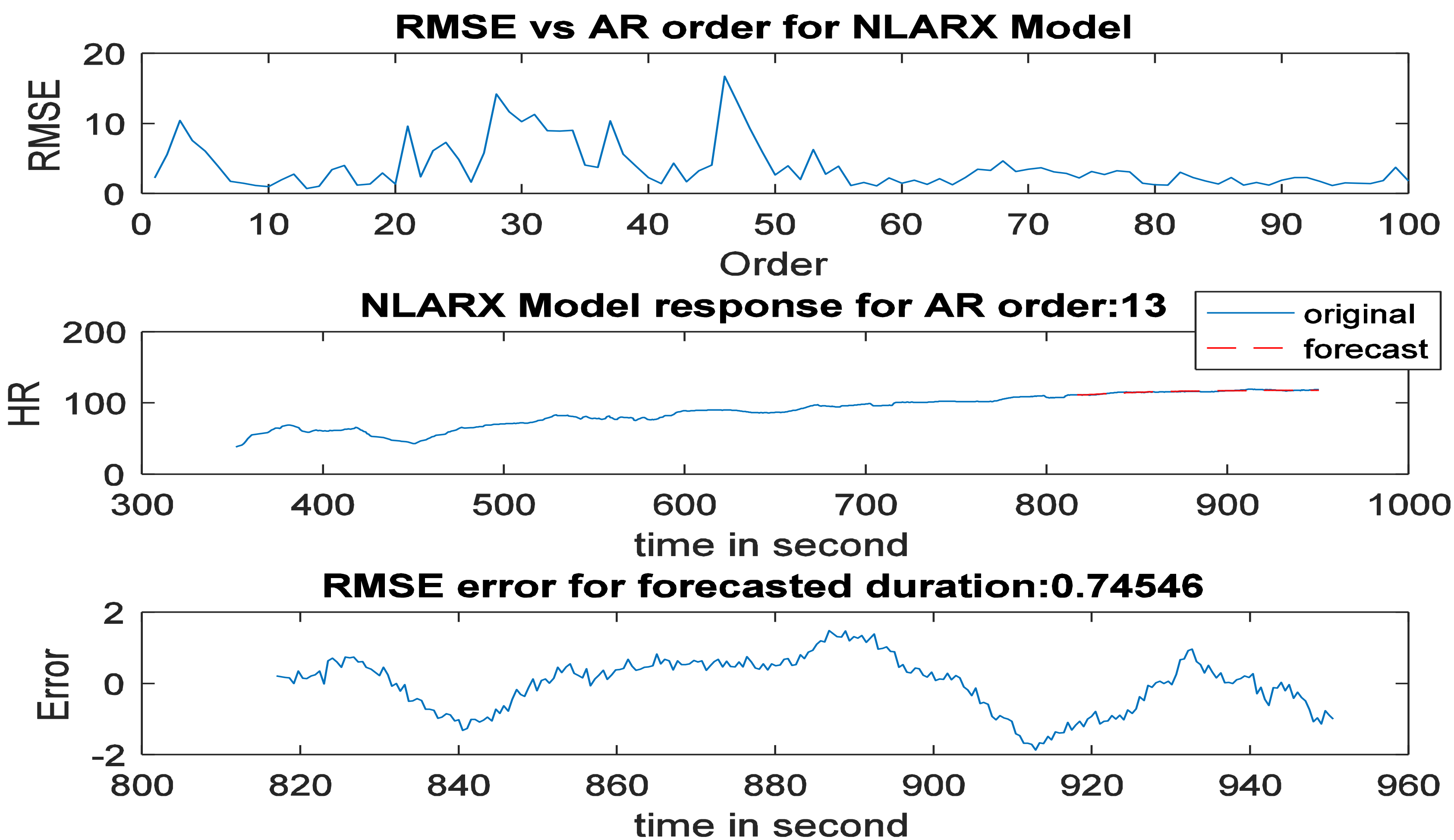
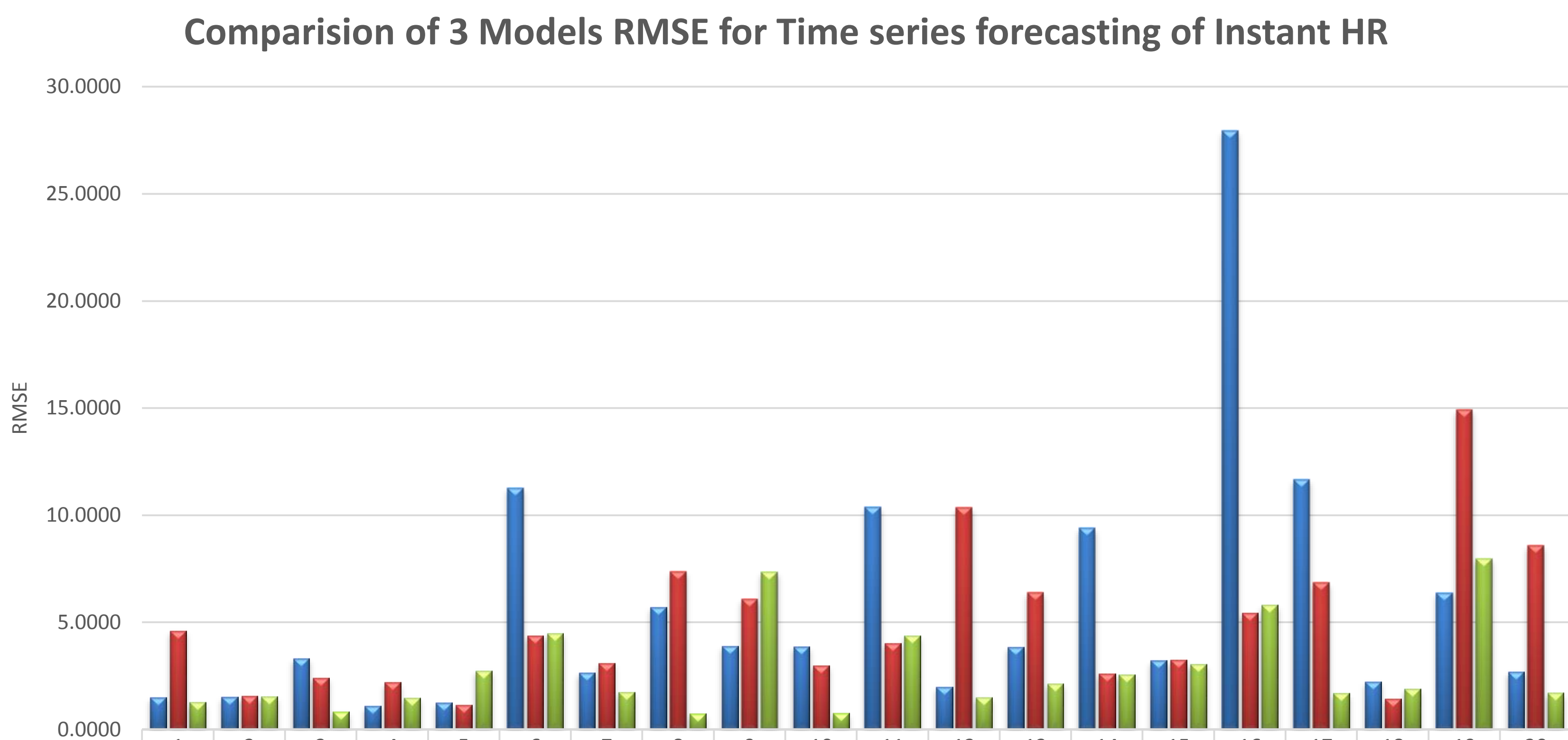


FIGURE 5.13 the forecasting result of NLARX model for the instant HR of abnormal database 308. RMSE vs. AR order (Top), Time series forecasting of instant HR for duration 817th to 950th sec, order 13 (Middle), Error plot: RMSE 0.74546 (Bottom)

TABLE 5.3 Summary of the models' forecasting performance for Instant HR of 20 abnormal subjects. Signal of Initial 77.77% of net exercised duration used for modelling and remaining for validation.

SR. NO.	Data-base	ADAPTIVE				ARMAX				NLARX				Exercise Duration (sec)		Net Exercise Duration (sec)
		HR				HR				HR						
		RMSE	MAE	MAPE	order	RMSE	MAE	MAPE	order	RMSE	MAE	MAPE	order	END	START	
1	300	1.4980	1.2420	0.0111	2	4.6010	4.1163	0.0367	1	1.2851	1.0814	0.0096	75	620	0	620
2	301	1.5178	1.2246	0.0111	83	1.5647	1.2938	0.0119	1	1.5451	1.2392	0.0114	98	1100	0	1100
3	302	3.3126	2.6434	0.0232	3	2.4008	1.8968	0.0167	32	0.8280	0.6802	0.0059	24	1025	0	1025
4	303	1.0938	0.9544	0.0098	2	2.1948	1.9221	0.0200	9	1.4704	1.2106	0.0125	21	1560	0	1560
5	304	1.2524	1.0500	0.0130	12	1.1501	0.9345	0.0116	23	2.7300	2.1837	0.0274	4	1280	640	640
6	306	11.2700	9.4328	0.0814	3	4.3844	4.0243	0.0332	23	4.4912	4.1147	0.0340	99	2775	2050	725
7	307	2.6486	2.3101	0.0243	6	3.0836	2.7474	0.0294	18	1.7547	1.2381	0.0127	29	1450	750	700
8	308	5.7074	4.4955	0.0392	4	7.3851	5.7754	0.0504	92	0.7455	0.6130	0.0053	13	950	352	598
9	309	3.8954	2.9693	0.0231	8	6.0911	5.0635	0.0396	1	7.3538	6.7203	0.0518	17	1350	0	1350
10	310	3.8799	3.4179	0.0258	3	2.9865	2.4145	0.0182	15	0.7753	0.5529	0.0041	29	440	0	440
11	311	10.3893	6.9515	0.0613	4	4.0271	3.4927	0.0291	6	4.3850	3.8152	0.0311	57	1220	550	670
12	312	1.9801	1.6232	0.0140	6	10.3503	7.6331	0.0749	91	1.5133	1.2809	0.0115	33	730	0	730
13	313	3.8489	3.0789	0.0233	69	6.3962	5.1357	0.0391	12	2.1323	1.7042	0.0129	55	600	0	600
14	314	9.3994	8.3619	0.0776	4	2.5971	2.2604	0.0207	59	2.5731	2.2264	0.0205	62	560	0	560
15	316	3.2207	2.2278	0.0177	4	3.2385	2.2372	0.0178	10	3.0442	2.2405	0.0176	6	560	0	560
16	317	27.9520	25.4015	0.2269	3	5.4300	4.5006	0.0397	94	5.8137	4.9049	0.0434	100	970	0	970
17	319	11.6687	10.7941	0.0789	4	6.8477	5.8038	0.0424	79	1.6853	1.3647	0.0099	29	260	0	260
18	320	2.2366	1.8942	0.0147	38	1.4399	1.1695	0.0092	41	1.8976	1.4800	0.0117	26	1350	850	500
19	321	6.3735	5.0795	0.0435	16	14.9256	14.2323	0.1248	96	7.9692	6.5883	0.0567	85	950	0	950
20	322	2.6845	2.4806	0.0184	41	8.5904	6.8005	0.0506	4	1.7160	1.3011	0.0097	4	250	0	250
		ADAPTIVE				ARMAX				NLARX						
MEAN		5.7915				4.9842				2.7854						
STD. DEVI (σ)		6.2407				3.4372				2.1312						



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ADAPTIVE	1.4980	1.5178	3.3126	1.0938	1.2524	11.270	2.6486	5.7074	3.8954	3.8799	10.389	1.9801	3.8489	9.3994	3.2207	27.952	11.668	2.2366	6.3735	2.6845
ARMAX	4.6010	1.5647	2.4008	2.1948	1.1501	4.3844	3.0836	7.3851	6.0911	2.9865	4.0271	10.350	6.3962	2.5971	3.2385	5.4300	6.8477	1.4399	14.925	8.5904
NLARX	1.2851	1.5451	0.8280	1.4704	2.7300	4.4912	1.7547	0.7455	7.3538	0.7753	4.3850	1.5133	2.1323	2.5731	3.0442	5.8137	1.6853	1.8976	7.9692	1.7160

20 Abnormal Subjects

ADAPTIVE ARMAX NLARX

FIGURE 5.14 Comparison of the models RMSE for Instant HR of 20 abnormal subjects (Reference to Table 5.3)

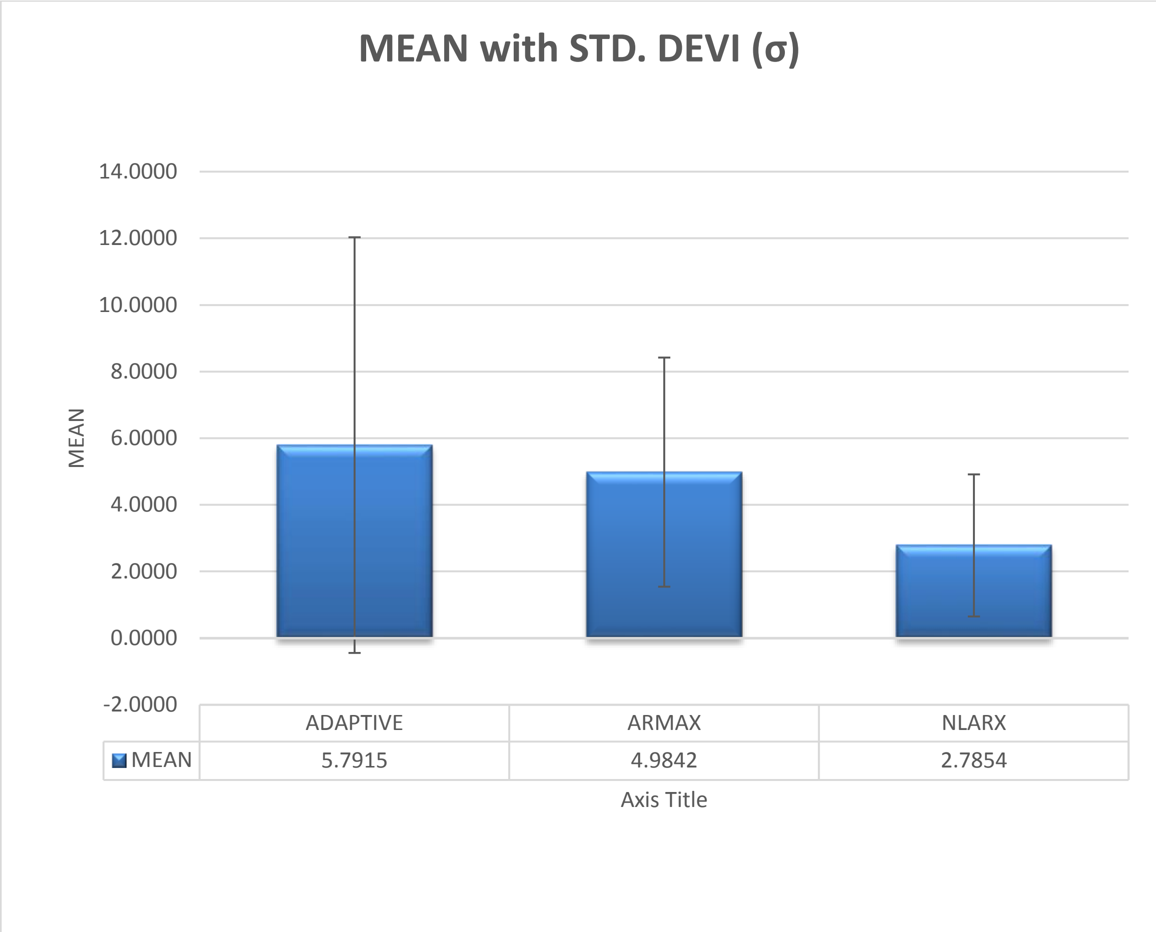


FIGURE 5.15 Mean and Standard deviation of models RMSE for Instant HR of 20 abnormal subjects (reference to Table 5.3)

In Table 5.3, Summary of all three models' forecasting performance for instant HR of the 20 abnormal subjects is shown, where signal's initial 77.77% of net exercised duration are used for training of models and remaining for model's validation. And for visual comparison of the same results, it is shown in column chart format in Fig. 5.14. In a column chart, I have considered only RMSE. In the majority of case, NLARX model has less or near to less error compared to other models. So to prove the reliability of models in time series forecasting of instant HR during exercise condition, I have calculated MEAN and standard deviation (σ), it is depicted in Fig. 5.15. Standard deviation indicates how values are spread around the mean of the dataset. Low standard deviation shows that the data is clustered closely around the mean hence it is more reliable. Here the NLARX model is found to have a low value of MEAN and standard deviation, hence it is more reliable and it has low average forecasting error compare to Adaptive and ARMAX model, in the case of abnormal subjects also. The adaptive model has the worst performance for time series forecasting of instant HR during exercise condition for abnormal subjects.

I have also done experiments on same “The MIT-BIH ST Change Database” by increasing forecasted duration and decreasing length of instant HR signal used for models training purpose. Here signal’s initial 66.66% of net exercise duration is utilized for training of models and remaining for validation of models. As a sample, to discuss results, the database 308 is used, this subject had done exercised from 352^{nd} sec to 950^{th} sec (so the net length of exercise is 598 seconds), out of that, signal of 352^{nd} to 750^{th} sec (initial 66.66 % of total span) was used for training of model and signal of 751^{st} sec to 950^{th} sec was reserved for validation of models. Results of three models: Adaptive filter, ARMAX and NLARX models are displayed in sequence in the Fig. 5.16, 5.17 and 5.18 for the same database.

In Fig. 5.16, the forecasting results of adaptive filter are depicted. On the top graph, RMSE corresponding to FIR order 1 to 100 is shown. From that order 07 has minimum error that is why instant HR signal is estimated for forecasted duration 751^{st} sec to 950^{th} sec with FIR order 07 and calculated RMSE for forecasted duration is 2.6968. Same way Fig. 5.17 and 5.18 can be interpreted for ARMAX and NLARX models.

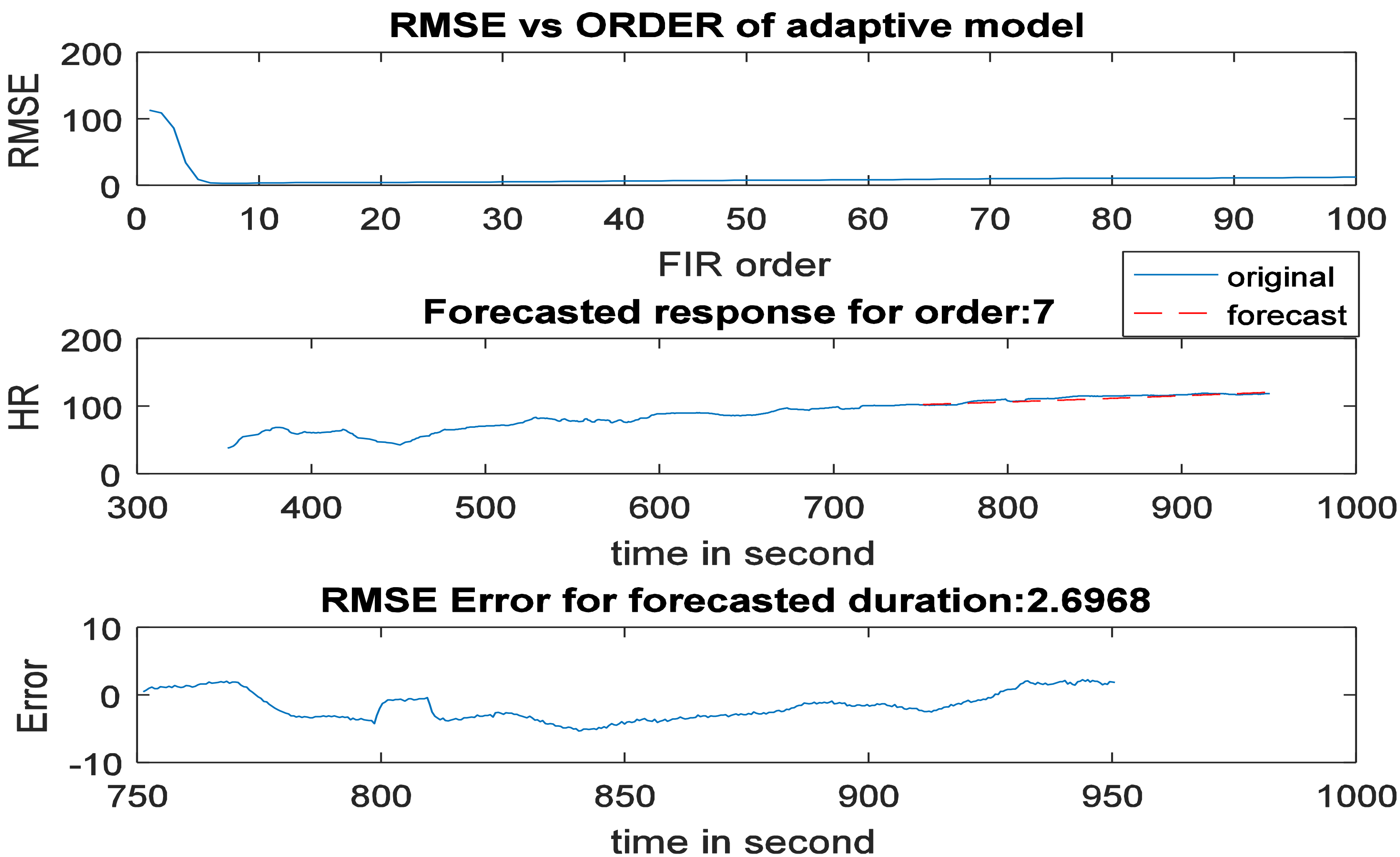


FIGURE 5.16 the forecasting result of an Adaptive model for the instant HR of abnormal database 308. RMSE vs. FIR order (Top), Time series forecasting of instant HR for duration 751st sec to 950th sec, order 07 (Middle), Error plot: RMSE 2.6968 (Bottom)

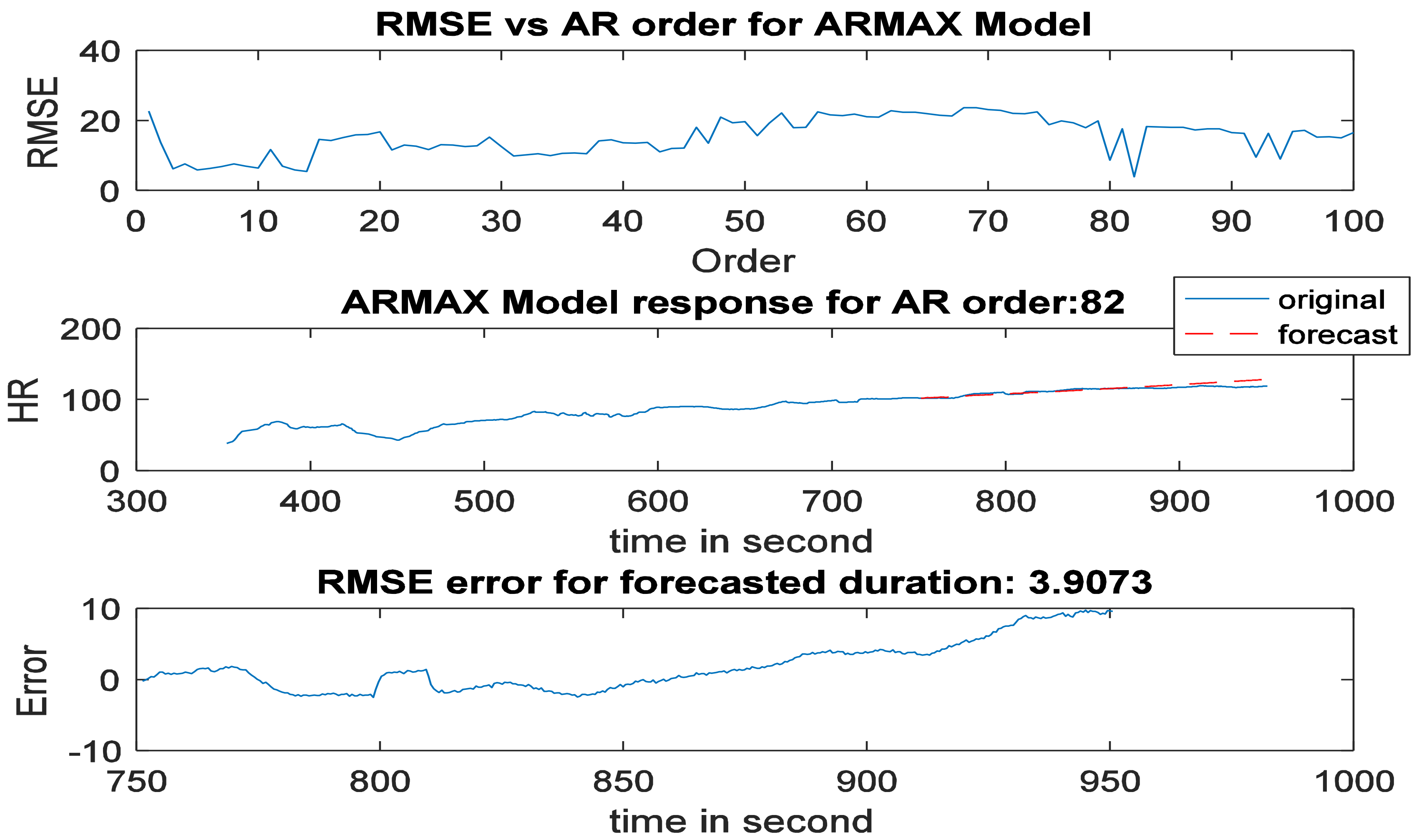


FIGURE 5.17 the forecasting result of ARMAX model for the instant HR of abnormal database 308. RMSE vs. AR order (Top), Time series forecasting of instant HR for duration 751st sec to 950th sec, order 82 (Middle), Error plot: RMSE 3.9073 (Bottom)

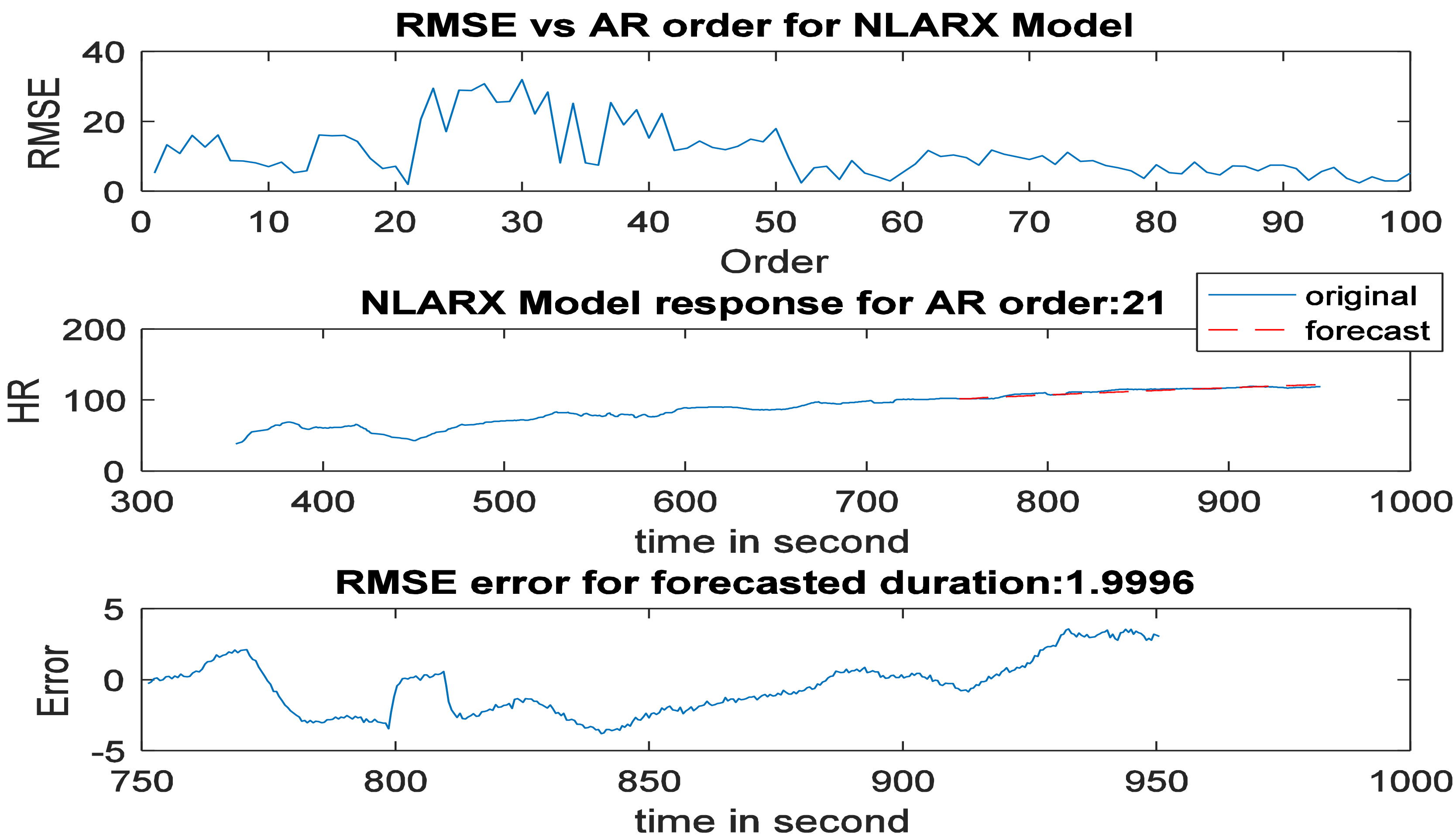
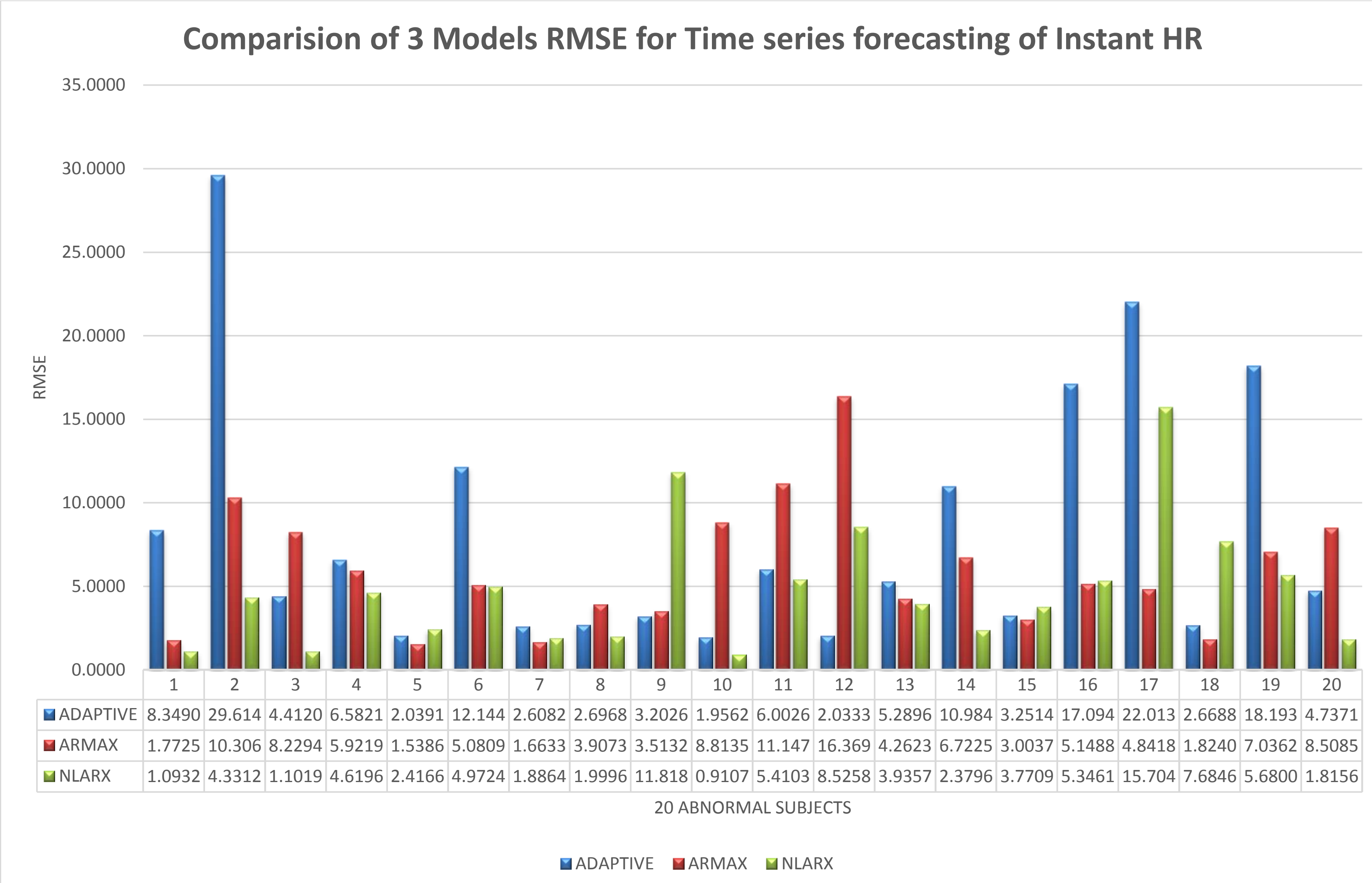


FIGURE 5.18 the forecasting result of NLARX model for the instant HR of abnormal database 308. RMSE vs. AR order (Top), Time series forecasting of instant HR for duration 751st sec to 950th sec, order 21 (Middle), Error plot: RMSE 1.9996 (Bottom)

TABLE 5.4 Summary of the models' forecasting performance for Instant HR of 20 abnormal subjects. Signal's initial 66.66 % of net exercised duration used for modelling and remaining for validation

SR. NO.	Data base	ADAPTIVE				ARMAX				NLARX				Exercise Duration (sec)		Net Exercise Duration (sec)
		HR				HR				HR						
		RMSE	MAE	MAPE	order	RMSE	MAE	MAPE	order	RMSE	MAE	MAPE	order	END	START	
1	300	8.3490	7.3962	0.0666	3	1.7725	1.4219	0.0128	57	1.0932	0.9337	0.0084	38	620	0	620
2	301	29.6146	26.6156	0.2744	6	10.3062	9.5992	0.0976	64	4.3312	3.5485	0.0351	96	1100	0	1100
3	302	4.4120	3.7738	0.0346	3	8.2294	7.1929	0.0657	2	1.1019	0.8845	0.0077	20	1025	0	1025
4	303	6.5821	5.2682	0.0562	3	5.9219	5.4916	0.0582	7	4.6196	4.0392	0.0429	20	1560	0	1560
5	304	2.0391	1.5770	0.0193	8	1.5386	1.1968	0.0148	97	2.4166	2.0622	0.0261	19	1280	640	640
6	306	12.1447	11.2113	0.0954	3	5.0809	3.7434	0.0305	96	4.9724	4.0817	0.0338	34	2775	2050	725
7	307	2.6082	2.1992	0.0247	8	1.6633	1.2113	0.0131	25	1.8864	1.3663	0.0147	26	1450	750	700
8	308	2.6968	2.4027	0.0213	7	3.9073	2.8736	0.0260	82	1.9996	1.6553	0.0147	21	950	352	598
9	309	3.2026	2.5662	0.0202	3	3.5132	2.8529	0.0222	99	11.8189	9.2392	0.0777	1	1350	0	1350
10	310	1.9562	1.6451	0.0125	3	8.8135	7.2212	0.0553	12	0.9107	0.6559	0.0049	13	440	0	440
11	311	6.0026	3.7403	0.0337	3	11.1478	10.1404	0.0849	2	5.4103	4.7873	0.0402	32	1220	550	670
12	312	2.0333	1.7034	0.0166	11	16.3697	12.1113	0.1296	91	8.5258	7.2944	0.0745	72	730	0	730
13	313	5.2896	4.4290	0.0345	25	4.2623	3.3613	0.0264	59	3.9357	3.1870	0.0248	18	600	0	600
14	314	10.9842	9.3690	0.0864	4	6.7225	5.6301	0.0517	72	2.3796	1.9546	0.0180	100	560	0	560
15	316	3.2514	2.6263	0.0207	5	3.0037	2.1759	0.0175	15	3.7709	3.1991	0.0251	1	560	0	560
16	317	17.0949	13.8068	0.1312	3	5.1488	4.0622	0.0362	90	5.3461	4.3307	0.0388	100	970	0	970
17	319	22.0130	17.0388	0.1321	6	4.8418	3.9775	0.0305	8	15.7040	12.2826	0.0930	50	260	0	260
18	320	2.6688	2.2109	0.0177	6	1.8240	1.5464	0.0124	2	7.6846	6.2283	0.0512	12	1350	850	500
19	321	18.1931	15.9279	0.1527	5	7.0362	6.0766	0.0555	97	5.6800	4.5949	0.0397	64	950	0	950
20	322	4.7371	3.1133	0.0236	5	8.5085	6.1874	0.0471	9	1.8156	1.4588	0.0110	62	250	0	250
		ADAPTIVE				ARMAX				NLARX						
MEAN		8.2937				5.9806				4.7702						
STD. DEVI (σ)		7.7811				3.7895				3.7920						



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ADAPTIVE	8.3490	29.614	4.4120	6.5821	2.0391	12.144	2.6082	2.6968	3.2026	1.9562	6.0026	2.0333	5.2896	10.984	3.2514	17.094	22.013	2.6688	18.193	4.7371
ARMAX	1.7725	10.306	8.2294	5.9219	1.5386	5.0809	1.6633	3.9073	3.5132	8.8135	11.147	16.369	4.2623	6.7225	3.0037	5.1488	4.8418	1.8240	7.0362	8.5085
NLARX	1.0932	4.3312	1.1019	4.6196	2.4166	4.9724	1.8864	1.9996	11.818	0.9107	5.4103	8.5258	3.9357	2.3796	3.7709	5.3461	15.704	7.6846	5.6800	1.8156

FIGURE 5.19 Comparison of models performance for Instant HR forecasting (reference to Table 5.4)

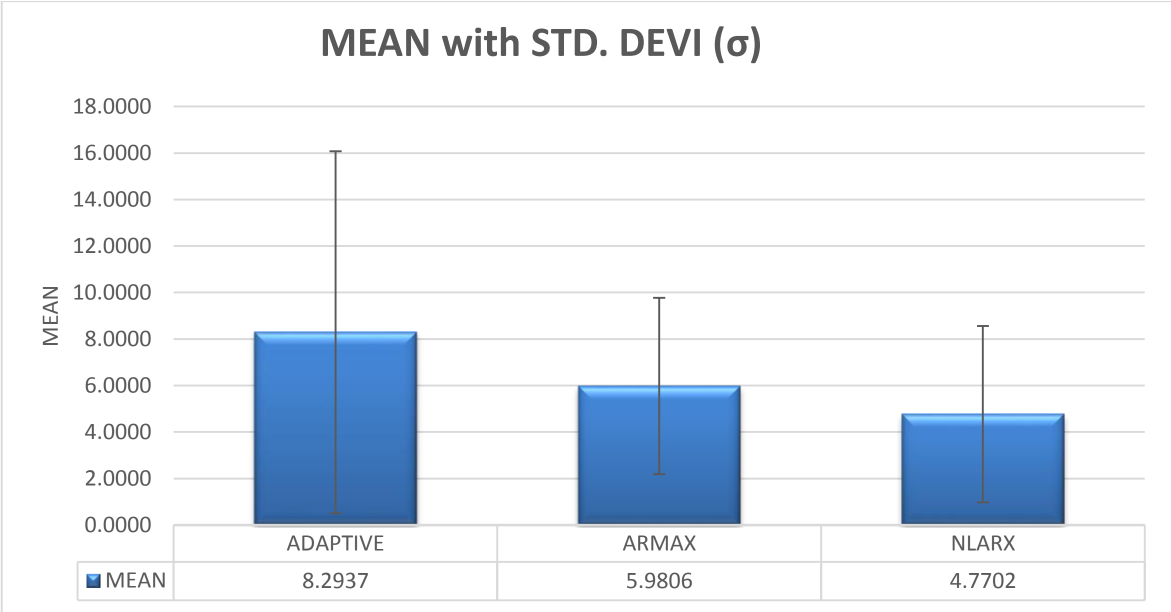


FIGURE 5.20 Mean and standard deviation of three models in Instant HR prediction of 20 Abnormal subjects (reference to table 5.4) where initial 66.66 % of total span data used for training

In Table 5.4, Summary of all three models' forecasting performance for instant HR of the 20 abnormal subjects are shown, where signal's initial 66.66% of net exercised duration is used for training of models. And for visual comparison of the same results, it is shown in column chart format in Fig. 5.19. In column chart, I have considered only RMSE. In majority of case NLARX model have less error or near to less compared to other models. So to prove the reliability of models in time series forecasting of instant HR during exercise condition, I have calculated MEAN and standard deviation (σ), it is depicted in Fig. 5.20. Here the NLARX model is found to have low value of MEAN, so it indicates overall less time series forecasting error compared to other models. But standard deviation of ARMAX and NLARX models are almost similar, which indicates that spreading of errors over the MEAN is almost equal. The adaptive model has worst performance for time series forecasting of instant HR during exercise condition of abnormal subjects.

The abnormal databases selected for Table 5.3 and Table 5.4 are same, the only difference is their training and forecasting durations. Table 5.3 represents the results of abnormal database for signal's initial 77.77% of net exercise duration is used for training and remaining for validation, while in Table 5.4, signal's initial 66.66% of net exercise duration is used for training and remaining for validation. In comparison of both tables, it has come to notice that as forecasting horizon increases, estimation errors of all three models also increase.

5.2 Single and Multi-step ahead prediction

Here, the Concept of the artificial neural network is used to predict single and multi-step-ahead prediction of respiration rate. The neural network is a collection of interconnected elements that are first trained to perform a predefined task and later it can be tested to unknown inputs. The structure NARNET is used with two hidden layers and number of neurons in each layer are [10 5] with 25 inputs (past outputs) feedback from output with the help of 24 delay blocks. This section represents results of NARNET where subjects are young whose age is in between 18 to 32 years, it is summarized in Table 5.5. Model's performance is evaluated with the RMSE, MAE and MAPE error calculation. As a sample result of database 01 from Table 5.5 is depicted in Fig. 5.21 and 5.22 for single step and five step ahead prediction of Instant RR.

Out of 540 sec of recorded database, initial 420 sec signal is used for training of NARNET. For training, the levenberg-Marquardt backpropagation algorithm is used that optimizes weight and bias values. During training, a dataset of instant RR up to 7 min (420 sec) is divided differently for single and five-step ahead prediction. For example, for 5 step-ahead prediction y (n+5) is the reference output and y (n)…..y (n-24) past output is feedback to the input of the network, in between bias and weight are adjusted with the help of optimization algorithm. After successful training of NARNET structure, the same network is used to predict the single or five-step ahead for the duration of 421st sec to 540th sec. At last the predicted response of network is compared with observed values to validate models.

To interpret the results of table 5.5, forecast errors, their MEAN and standard deviation are calculated. It has come to notice that all errors, their MEAN and spreadness of error around MEAN (Standard deviation) are large in Five step ahead prediction compared to single step ahead prediction because as prediction horizon increases, the prediction error also increases. It is also true that performance of NARNET changes due to change in number of layers, neurons in each layer, activation function, training method etc. Unfortunately, the only way to determine the number of layers in a network is by trial and error [34]. The NARNET models are not tested on Instant HR signals because of its long training time and our main objective is to forecast future behavior of signals from their own past response. So every time long training time for each subject is not appropriate.

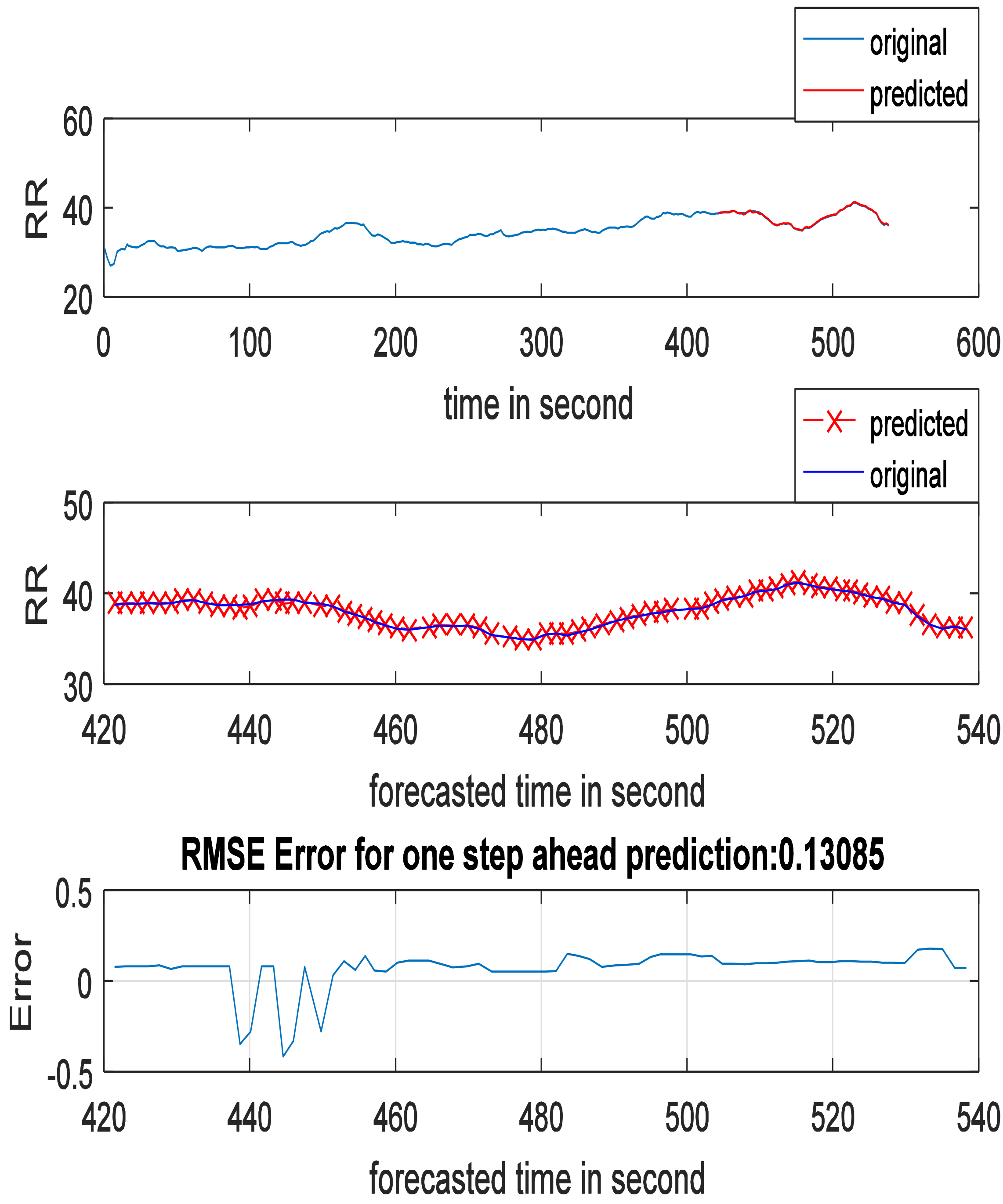


FIGURE 5.21 the results of a Single step ahead prediction for instant RR signal of database 01 (Table 5.5), observed instant RR signal with predicted response(Top), Observed and predicted signals for prediction time horizon 421st sec to 540th sec (Middle). Error plot (Bottom)

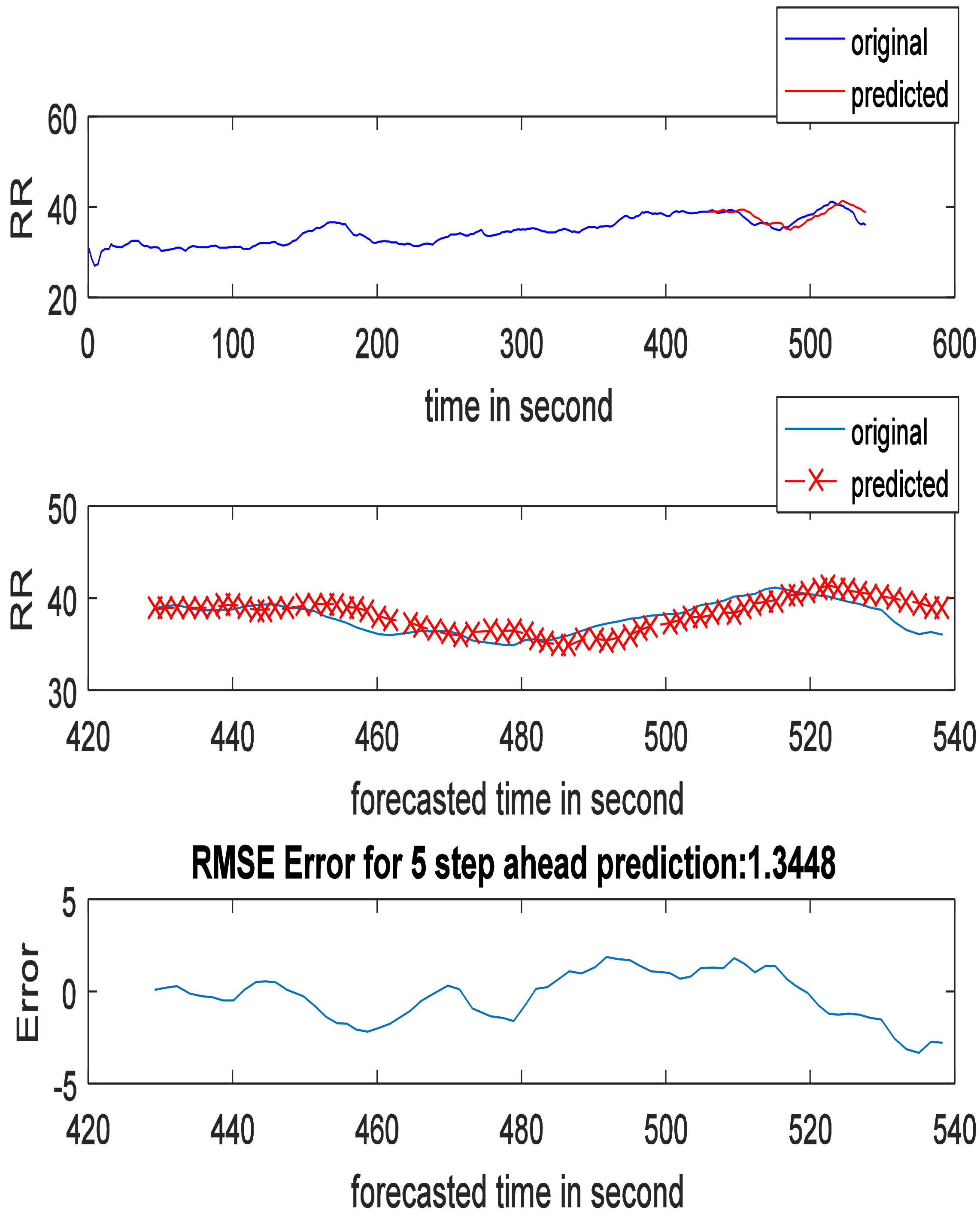


FIGURE 5.22 the results of five-step ahead prediction for instant RR signal of database 01 (Table 5.5), observed instant RR signal with predicted response(Top), Observed and predicted signals for prediction time horizon 421st sec to 540th sec (Middle). Error plot (Bottom)

TABLE 5.5 The summary of single and five-step ahead prediction of Instant RR signal using NARNET for the Normal database

Data-base	Male/ Female	Age	Instant RR					
			One step ahead prediction			Five step ahead prediction		
			RMSE	MAE	MAPE	RMSE	MAE	MAPE
1	M	22	0.1309	0.1119	0.0029	1.3448	1.1013	0.0287
2	M	32	0.263	0.2415	0.0044	0.9255	0.747	0.0135
3	M	29	0.5792	0.5753	0.0184	0.8263	0.6362	0.0203
4	M	22	0.195	0.1594	0.0042	0.5168	0.4198	0.0108
5	M	22	0.6879	0.6806	0.0202	0.9631	0.7714	0.023
6	M	22	0.1596	0.1588	0.0043	1.0884	0.7115	0.0194
7	F	20	0.1913	0.1865	0.0039	1.047	0.9331	0.0193
8	M	21	0.1785	0.1462	0.0028	1.5277	1.2888	0.0248
9	F	19	0.3644	0.321	0.009	0.8473	0.7279	0.0206
10	M	21	0.1432	0.1424	0.0039	0.8391	0.6522	0.0178
11	F	24	0.0596	0.0516	0.0012	0.7192	0.5755	0.0137
12	F	18	0.24	0.2196	0.005	0.7803	0.6653	0.0151
13	M	24	0.1366	0.1297	0.0035	1.1426	0.7979	0.0218
14	F	23	0.1301	0.1234	0.0045	1.126	0.9632	0.0345
15	F	22	0.0684	0.059	0.0015	0.6704	0.4824	0.0123
16	F	22	0.0795	0.0701	0.002	0.7315	0.5978	0.017
17	M	24	0.7226	0.7224	0.0216	1.3793	1.1569	0.0342
18	M	25	0.1399	0.1358	0.0043	0.74	0.5603	0.018
19	M	24	0.6992	0.6989	0.0183	0.5646	0.4399	0.0115
20	F	20	0.2287	0.1802	0.0041	2.4582	1.5251	0.0351
MEAN			0.2699			1.0119		
STD. DEVI (σ)			0.2193			0.4350		

CHAPTER-6

Conclusion, Major Contributions And Scope Of Further Work

6.1 Conclusion

- Prediction of future is always a challenging task and area of interest in many fields. Cardiopulmonary exercise testing is a completely non-invasive technique and it has a wide range of application like fitness evaluation, sports, clinical diagnostic, etc. But there is a chances of overload to cardiopulmonary system. Hence, to reduce testing time and testing a patient at lower physical stress for avoiding any serious incidents, a novel approach is successfully presented here. The core concept is time series forecasting of future behaviour of cardiopulmonary signal after immature ending of the exercise test. This work presents three models an Adaptive, ARMAX and NLARX for time series forecasting of Instant HR and RR signals during incremental exercise.

- All these models are implemented such that it forecast future behavior of signals from available past signal values of the same subject, recorded before actual end of exercise test.

- From statistical analysis of results, it is proven that NLARX model is the most suitable for time series forecasting of cardiopulmonary signals during incremental exercise test for normal as well as abnormal subjects. The performances of three models ranked in descending sequence are: NLARX, ARMAX and Adaptive model.

- For all three models, it is also true that, as the forecasting time horizon increases, forecasting error also proportionally increases.

- Additionally single step and multi-step ahead prediction of the signals may give an idea about the future condition of patient health and helps to avoid any critical situation.

6.2 Major contributions

The presented research work addresses the major safety issues in the exercise-based evaluation of the cardiopulmonary system due to cardiac system overload. Due to which continuous involvement of medical experts is required.

The novel contribution of the presented work is:

- Implementation of linear and nonlinear models to forecast future values of signals.

- Comparative analysis of results to identify the best suitable model for time series forecasting of cardiopulmonary signals during incremental exercise.

- Performance of model is tested on both kinds of the database: the normal as well as the abnormal cardiac subjects.

- The implemented models are using only current and past signal values of the same subject to forecast future values of signal, without the need of other similar databases. Implemented models can be applied to a database of any age group and sex. The implemented model's core concept is time series forecasting so it can be used on other physiological signals with little modification.

- Implemented single and multi-step ahead prediction of physiological signal helps to avoid any critical situation by giving estimated values of the future.

6.3 Scope of Further work

One can extend this work by selecting other physiological parameters under different physical stress protocol or one can go for disease-specific study.

Also, the performance of forecasting can be improved with the help of other models or optimization algorithms.

Further utilization of this work is possible for prediction of maximum workload that anyone can sustain during incremental exercise test. After immature end of test in the safe limit, the concept of time series forecasting for HR can be applied to forecast up to the maximum achievable HR. it can be age dependent or any other way. Time at which signal reaches to maximum HR can be used to identify maximum achievable stage of exercise test and workload.

Someone can also think in the direction of defining a new fitness index from physiological signal's features recorded during exercise test to define or compare fitness level.

List of Publications

- M. B. Patel and V. A. Shah, “Single and multi-step ahead prediction of instant Respiration rate in incremental exercised condition with neural network,” *International Journal of Advance Engineering and Research Development,,* vol. 5, no. 1, pp. 483-488, January 2018.

- M. B. Patel and V. A. Shah, “Comparison of models performance for forecasting of Instant Heart Rate and Respiration Rate during incremental exercise,” *International Journal of Engineering Science Invention*, vol. 6, no. 8, pp. 12-19, August 2017.

- M. B. Patel and V. A. Shah, “A Review on Technical Aspects of Cardiopulmonary Exercise Testing,” *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 6, no. 8, pp. 6157-6162, August 2017.

- M. B. Patel and V. A. Shah, “Methods for Prediction of Instantaneous Heart rate During Exercise,” in ISTE Gujarat Section, 1^{st} International Conference, Bangkok, 10-11^{th} June 2016.

- M. B. Patel and V. A. Shah, “Exercised based evaluation of Cardiopulmonary system,” in International Conference on Engineering: Issues, opportunities and Challenges for Development, S.N. Patel Institute of Technology & Research Centre, Umrakh, Bardoli, 11^{th} April 2015.

List of References

[1] V. F. Froelicher and J. Myers, Manual of Exercise Testing, Third Edition ed., Philadelphia: Mosby Inc., 2007.

[2] D. A. Burton, K. Stokes and G. M. Hall, “Physiological effects of exercise,” *Continuing Education in Anesthesia, Critical Care & Pain,* vol. 04, no. 06, pp. 185-188, 2004.

[3] I. M. Weisman and R. J. Zeballos, Clinical exercise testing, vol. 32, C. T. Bolliger, Ed., Basel: Karger, 2002.

[4] S. Policker and A. B. Geva, “A new algorithm for time series Prediction by Temporal fuzzy clustering,” in *Proceedings of 15th international conference on pattern Recognition*, barcelona, spain, 2000.

[5] L. Faes, K. H. Chon and G. Nollo, “A method for time-varying Nonlinear Prediction of complex Nonstationary Biomedical Signals,” *IEEE transaction on Biomedical Engineering,* vol. 56, pp. 205-209, February 2009.

[6] Z. German-sallo and C. Ciufudean, “RR interval Prediction in ECG signals,” in *IEEE international conference on Electrical and Power Engineering (EPE)*, Calin Ciufudean, 2014.

[7] E. Baralis, T. Cerquitelli, S. Chiusano and V. D'elia, “Early Prediction of the Highest Workload in Incremental Cardiopulmonary Tests,” *ACM Transactions on Intelligent Systems and Technology,* vol. 4, pp. article 70:1-20, Sept 2013.

[8] E. Baralis, T. Cerquitelli, S. Chiusano, V. D’Elia, R. Molinari and D. Susta, “Predicting the highest workload in cardiopulmonary test,” in *IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS)*, Perth, WA, Australia, 2010.

[9] F. Nogueira and F. Pompeu, “Maximal Workload Prediction Models in the Clinical Cardio-pulmonary Effort Test,” *Arquivos Brasileiros de Cardiologia,* pp. 137-145, Aug 2006.

[10] F. J. P. Arts, H. Kuipers, A. E. Jeukendrup and W. H. M. Saris, “A short cycle Ergometer test to predict Maximum workload and Maximum oxygen Uptake,” *International Journals of Sports Med.,* vol. 14, pp. 460-464, 1993.

[11] J. Garcia, P. Serrano, R. Bailbn, E. Gutierrez, A. d. Rio, J. Casasnovas, I. Ferreira and P. Laguna, “Comparison of ECG Based Clinical Indexes during Exercise Test,” in *Proceedings of IEEE computing in cardiology conference*, Cambridge,MA,USA, 2000.

[12] L. Wang, S. W. Su and B. Celler, “Time constant of heart rate recovery after low level exercise as a useful measure of cardiovascular fitness,” in *28th IEEE EMBS Annual international conference*, New York city, USA, 2006.

[13] L. Xu, Y. Zhong, S. Yin, Y. Zhang, Y. Shen, D. Hao, Y. Hu and R. Zhang, “ECG and Pulse Variability Analysis for Exercise Evaluation,” in *Proceeding of IEEE, Automation and Logistics (ICAL)*, Chongqing, China, 2011.

[14] L. R. Davrath, I. Pinhas, A. Beck, M. Scheinowitz, D. Elian and S. Akselrod, “Heart rate recovery after exercise: a study by wavelet analysis,” in *Computers in Cardiology*, Thessaloniki Chalkidiki, Greece, 2003.

[15] A. Amal and G. Reshmi, "Cardiac Ischemia Diagnosis Using Stress ECG Analysis," in *Third International Conference on Advances in Computing and Communications (ICACC),*, Cochin, India, 2013.

[16] G. F. Fletcher, G. . J. Balady, E. A. Amsterdam, B. Chaitman, R. Eckel, J. Fleg, V. . F. Froelicher, A. . S. Leon, I. L. Pina, R. Rodney, D. G. Simons-Morton, M. A. Williams and T. Bazzarre, "Exercise Standards for Testing and Training-A Statement for Healthcare Professionals From the American Heart Association," *American Heart Association Scientific Statement,* pp. 1694-1740, June 2001.

[17] D. E. Forman, J. Myers, C. J. Lavie, M. Guazzi, B. Celli and R. Arena, "Cardiopulmonary Exercise Testing: Relevant but Underused," *Postgraduate Medicine,* vol. 122, no. 6, pp. 68-86, November 2010.

[18] k. Albouaini, M. Egred, A. Alahmar and D. J. Wright, "Cardiopulmonary exercise testing and its application," *Postgraduate Medical Journal,* p. 675–682, 2007.

[19] R. A. Robergs and R. Landwehr, "The Surprising History Of The "HRmax=220-Age" Equation," *Journal of Exercise Physiology online,* vol. 5, pp. 1-10, May 2002.

[20] S. M. Revill, K. E. Beck and M. D. L. Morgan, "Comparison of the Peak Exercise Response Measured by the Ramp and 1-min Step Cycle Exercise Protocols in Patients With Exertional Dyspnea," *chest journal online,* vol. 121, pp. 1099-1105, April 2002.

[21] S. A. Ward and P. Palange, Clinical Exercise testing, vol. 12, UK: European Respiratory Society, 2007.

[22] D. C. Drew-Nord, J. Myers, S. R. Nord, R. K. Oka, O. Hong and E. S. Froelicher, "Accuracy of peak VO2 assessments in career firefighters," *Journal of Occupational Medicine and Toxicology,* p. 6:25, 2011.

[23] R. M. Rangayyan, Biomedical Signal Analysis A case study approach, New York: John Wiley & Sons, Inc., 2002.

[24] "EKG Sensor," Vernier Software & Technology, [Online]. Available: https://www.vernier.com/products/sensors/ekg-sensors/ekg-bta/.

[25] "The MIT-BIH ST Change Database," 28 October 2016 . [Online]. Available: https://physionet.org.

[26] R. A. Alvarez, A. J. Mendez Penin and X. A. V. Sobrino, "A comparision of three QRS detection algorithms over a public database," *Procedia Technology,* vol. 9, pp. 1159-1165, 2013.

[27] W. J. Tompkins, Biomedical Digital Signal Processing, Englewood Cliffs,New jersey: Prentice Hall, 1993.

[28] B. Widrow and S. D. Steams, Adaptive signal processing, New Jersey: Printice-hall Inc., 1985.

[29] S. Haykin, Adaptive filter Theory, New Jersey: Prentice-hall Inc, 1996.

[30] Mathworks, "MATLAB System Identification Toolbox User's Guide," 2016. [Online].

[31] L. G. B. Ruiz, M. P. Cuéllar, M. D. Calvo-Flores and . M. D. C. Pegalajar Jiménez, "An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings," *Energies* , vol. 9, no. 9, pp. 684-705, 2016.

[32] M. V. Shcherbakov, A. Brebels, S. N. Lvovna, A. P. Tyukov, T. A. Janovsky and V. A. Kamaev, "A Survey of Forecast Error Measures," *World Applied Sciences Journal 24 (Information Technologies in Modern Industry, Education & Society),* pp. 171-176, 2013.

[33] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C.-K. Peng and H. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals Circulation 101(23):e215-e220 [Circulation Electronic Pages;

http://circ.ahajournals.org/content/101/23/e215.full],” 13 June 2000. [Online]. Available: doi:10.13026/C2ZW2H.

[34] S . I. Gallant, Neural Network Learning and Expert Systems, Cambridge, Massachusetts, London: The MIT Press, 1995.

[35] S. Paul Linder, S. Wendelken and J. Clayman, “Detecting exercise induced stress using the Photoplethysmogram,” in *28th IEEE EMBS Annual international conference,* New York city, USA, 2006.

[36] Z. Meng, J. Bali, S. Cui and S. Wang, “Exercise testing in hypertension patients,” in *23rd IEEE Annual EMBS international conference*, Istanbul, Turkey, 2001.

[37] X. Li and J. Bai, “Assessing Cardiovascular function status with non-linear indexes from HRV of exercise testing,” in *IEEE Engineering in Medicine and Biology 27th annual conference*, Shanghai, china, 2005.

[38] E. Baralis, T. Cerquitelli, A. Giordano, A. Mezzani, D. Susta and X. Xiao, “Predicting Cardiopulmonary Response to Incremental Exercise Test,” in *IEEE 28th International Symposium on Computer-Based Medical Systems (CBMS)*, Sao Carlos, Brazil, 2015.

[39] “Bruce protocol,” 29 December 2017. [Online]. Available: http://en.wikipedia.org/w/index.php?title=Bruce_protocol&oldid=817558929.

[40] N. Uth, H. Sorensen, K. Overgaard and P. K. Pedersen, “Estimation of VO2max from the ratio between HRmax and HRrest--the Heart Rate Ratio Method,” *European Journal of Applied Physiology,* pp. 111-115, 2004.

[41] R. Johnston, “PFT BLOG,” 24 May 2015. [Online]. Available: https://www.pftforum.com/blog/treadmill-protocols/.

[42] A. Cabasson, O. Meste, G. Blain and S. Bermon, “Quantifying the PR interval pattern during dynamic exercise and recovery,” *IEEE Transactions on Biomedical Engineering,* vol. 56, no. 11, pp. 2675-2683, November 2009.