MODELLING THE DIFFERENTIAL DERMAL POTENTIALS: A DETAILED STUDY

A dissertation submitted to the
Jadavpur University, Kolkata
for the award of the degree of
Doctor of Philosophy
in
Engineering

Submitted by
Somali Nandy

Supervised by
Prof. Ratna Ghosh
&
Former Prof. Bhaswati Goswami

Department of Instrumentation and Electronics Engineering
Jadavpur University
West Bengal, India
December 4, 2023

JADAVPUR UNIVERSITY
KOLKATA - 700032, INDIA

INDEX NO. 278/16/E

1. Title of the Thesis:

MODELLING THE DIFFERENTIAL DERMAL POTENTIALS: A DETAILED STUDY

2. Name, Designation & Affiliation of the Supervisors

Supervisor 1: Dr. Ratna Ghosh

Professor
Instrumentation and Electronics Engineering Department
Jadavpur University
Salt Lake Campus
Kolkata - 700 106
West Bengal, India

Supervisor 2: Dr. Bhaswati Goswami

Former Professor
Instrumentation and Electronics Engineering Department
Jadavpur University
Salt Lake Campus
Kolkata - 700 106
West Bengal, India

3. List of Publications:

a. Journals:

S. Nandy, A. Sarkar, B. Goswami, and R. Ghosh, "Stochastic ACF-ZCL model of time-series signals with an application to restfulness assessment," *Measurement*, vol. 213, pp. 112641, 2023 May 31.

b. International conferences:

S. Nandy, B. Goswami, R. Ghosh, "Deriving a linear model for passively acquired bio-potentials using sample ACF," in: *2018 IEEE Applied Signal Processing Conference (ASPCON)*, 2018, pp. 64–68.

4. Patent:

R. Ghosh, B. Goswami, A. Bhattacharya, D. Mondal, S. Biswas, S. Nandy, and A. Sarkar, "An expert system to determine the human condition using bilateral endosomatic eda and the method thereof," Indian patent IN201931038304, Oct. 23, 2022, grant No. 402187. [Online]. Available: $https://patentscope.wipo.int/search/en/detail.jsf?docId = IN309729256\&_cid = P10 - L776DE - 43223 - 1$

5. List of Presentations in National/International/Conferences /Workshops:

S. Nandy, B. Goswami, R. Ghosh, "Deriving a linear model for passively acquired bio-potentials using sample ACF," in: *2018 IEEE Applied Signal Processing Conference (ASPCON)*, 2018, pp. 64–68.

PROFORMA - 1

"Statement of Originality"

I **Somali Nandy** (Index No. 278/16/E, date- 16/06/2016) do hereby declare that this thesis entitled " **MODELLING THE DIFFERENTIAL DERMAL POTENTIALS: A DETAILED STUDY**" contains literature survey and original research work done by the undersigned candidate as part of Doctoral studies.

All information in this thesis have been obtained and presented in accordance with existing academic rules and ethical conduct. I declare that, as required by these rules and conduct, I have fully cited and referred all materials and results that are not original to this work.

I also declare that I have checked this thesis as per the "Policy on Anti Plagiarism, Jadavpur University, 2019", and the level of similarity as checked by iThenticate software is 1%.

Somali Nandy
Signature of Candidate:

Date : 7/12/2023

Certified by Supervisor(s):

Ratna Ghosh 7/12/23
(Dr. Ratna Ghosh)
Professor
Dept. of Instrumentation &
Electronics Engineering
Jadavpur University
Kolkata 700106, India

Professor
Dept. of Instrumentation & Electronics Engg.
Jadavpur University
Salt Lake, 2nd Campus
Kolkata - 700106

Bhaswati Goswami 7/12/23
(Dr. Bhaswati Goswami)
Former Professor
Dept. of Instrumentation &
Electronics Engineering
Jadavpur University
Kolkata 700106, India

Former Professor
Dept. of Instrumentation & Electronics Engg.
Jadavpur University
Salt Lake, 2nd Campus
Kolkata - 700106

PROFORMA - 2

Certificate from the Supervisors

This is to certify that the thesis entitled **MODELLING THE DIFFERENTIAL DERMAL POTENTIALS: A DETAILED STUDY** submitted by **Somali Nandy**, who got her name registered for the award of Ph.D.(Engg.) degree of Jadavpur University (Index No. 278/16/E, date-16/06/2016), is absolutely based upon her own work done under the supervision of Prof. Ratna Ghosh and Former Prof. Bhaswati Goswami and that neither her thesis nor any part of the thesis has been submitted for any degree/ diploma or any other academic award anywhere before.

Ratna Ghosh 7/12/23

(Dr. Ratna Ghosh)
Professor
Dept. of Instrumentation &
Electronics Engineering
Jadavpur University
Kolkata 700106, India

Professor
Dept. of Instrumentation & Electronics Engg.
Jadavpur University
Salt Lake, 2nd Campus
Kolkata - 700106

Bhaswati Goswami 7/12/23

(Dr. Bhaswati Goswami)
Former Professor
Dept. of Instrumentation &
Electronics Engineering
Jadavpur University
Kolkata 700106, India

Former Professor
Dept. of Instrum... Electronics Engg.
Jad... University
Salt Lake, 2nd Campus
Kolkata - 700106

ACKNOWLEDGEMENT

Throughout this Ph.D work, I have received a great deal of support and assistance from many people. This work would not have been possible without their advice, guidance and encouragement.

I would like to acknowledge and give my gratitude to my supervisors, Prof. Ratna Ghosh and Former Prof. Bhaswati Goswami, Dept. of Instrumentation and Electronics Engineering, Jadavpur University for their unwavering guidance, patience, and invaluable mentor ship throughout my research. Their expertise, dynamism, vision, sincerity, motivation and support have deeply inspired me in shaping the direction and quality of this work.

I would like to acknowledge Visvesvaraya Ph.D scheme, for providing me the required financial support as fellowship (Visvesvaraya fellowship) during the entire period of this research work. I am grateful to Head of the Department and all the staff and faculty members of the Department of Instrumentation and Electronics Engineering, Jadavpur University for their direct and indirect cooperation and assistance towards completion of this work.

I also like to thank my lab mates, Dr. Arindam Sarkar, Dr. Aditi Bhattacharya who provided encouragement, cooperation and assistance when needed.

I would also like to thank my parents, my younger sister for their love, prayers, caring, support and sacrifices for educating and preparing me for my future. A special thanks goes to my husband, Debasish Chowdhury for unwavering support, inspiration, encouragement and belief in my abilities and also to my beloved son Baivonavo Chowdhury for his unconditional love, patience, and understanding throughout this journey. His cheerful smile and innocent laughter provided much-needed breaks, joy during long study hours.

I also owe it all to Almighty God for granting me the wisdom, health and strength to undertake this research task and enabling me to complete it.

December 4, 2023

Somali Nandy

(Somali Nandy)

Abstract

Bio-potentials and Electrodermal Activity (EDA) are useful in the field of psycho physiology since they both provide valuable information about physiological responses to various stimuli and situations, including emotional and psychological states. The differential dermal potentials (DDP) considered in this study are a special type of EDA signals that are acquired in differential mode from the limbs of a subject. These signals have been introduced very recently and quite a few successful applications using these signals have also been demonstrated. Hence it is useful to model these signals and study their characteristics in detail for a deeper understanding of human physiology and behavior.

The aim of this work is to develop various models of the differential dermal potentials and to determine the use and efficacy of the parameters derived from these models in studying and/or classifying human conditions.

One of the first steps for developing a model is to determine the signal order. This aspect has been addressed in this work from a new perspective based on the derivation of the sample autocorrelation function (ACF) of three types of signals. The signal order is determined in terms of its zero crossing lag (ZCL), which is the lag value at which the 1^{st} zero crossing of the sample ACF of the signal occurs. Closed form solutions for the ZCL are determined for these 3 types of signals: a) the deterministic i^{th} order power law series, denoted as y_{dk}; b) the stochastic i^{th} order power law series, denoted as y_{sk}; c) the deterministic polynomial series y_{pk}; and these are used to determine the ZCL of the acquired signal, denoted as y_k, which is modelled as a n^{th} order stochastic polynomial.

There exists a one-to-one map between the ZCL and its order i for the y_{dk} signal. This in conjunction with the shift in the ZCL within limits due to noise or polynomial coefficient interactions forms the basis for determining the order n of the time-series signal y_k for all the models proposed in this work.

Three different models have been proposed in this work which are denoted as i) the stochastic ACF-ZCL model, or simply ACF model; ii) the KF_{LPC} model; and iii) the KF_{TS} model.

In the proposed stochastic ACF-ZCL model, the deterministic trend or tonic component of these signals is the polynomial regression time series with the order n as determined using the ZCL while the resultant error time series represents the stochastic phasic component.

The Kalman filter (KF) is a popular technique used to clean these noises and obtain dynamic system models that respond to internal and external valid inputs. The popular autoregressive (AR) modelling approach provides the required linear prediction coefficients (LPC) of the state space model. The KF model developed in this work using this standard approach is termed the KF_{LPC} model. The proposed ACF model is used to derive ab-initio another KF model. y_{pk}, which is the polynomial time-series (TS) developed in the ACF model, is used to obtain the description of its state space model and so it is termed the KF_{TS} model. The KF_{LPC} and KF_{TS} models differ in their underlying state space model descriptions.

These three models have been determined for a total of 5706 DDP signals in order to compare them in terms of their parameters and statistical characteristics. The ACF model is also compared with corresponding models whose orders are determined using the standard partial ACF (PACF) method. The various parameters of the robustness and sensitivity metrics of the two KF models are also compared.

The results show that the deterministic tonic component of the ACF model typically has a lower order polyfit component than the corresponding PACF model and the ACF model residuals follow the beta distribution more consistently. The ACF model as well as both KF models provide consistent tonic and phasic components. Several system characteristics are quite similar for both KF models. However, the KF_{LPC} model estimates show a distinct bias and the nature of the metrics in the two KF models are distinctly separate.

The parameters obtained from each of the three models have been used individually to study the effect of long duration of rest on a subject. A total of 272 sets of independent signals obtained from subjects in restful supine condition for 10 minutes have been used for the study. For this, each of these signals have been segmented into 5 non-overlapping sets, or states, of 2-minutes signals.

The changes in the maximal entropy of the tonic component of the ACF model of the acquired signals with increasing rest support an existing research finding that the subject achieves a fully rested condition typically within 4-6 minutes of no-nap supine rest. Furthermore, the variations of the ZCL, system order, mean of the tonic component, SD as well as the spectral entropy of the residuals of this model support the emerging determinism of the system with rest.

The spectral entropy (SE) of the KF_{TS} model is an useful identifier of the tonic components, the phasic components and the noise residuals. The medians and ranges of the SE in these 3 cases indicate a sequential departure from regularity to irregularity as expected. No such distinguishing features are observed for the KF_{LPC} models.

Two other human conditions have been classified in this study using the random forest (RF) classifier in WEKA, version 3.9.4 classification software on these models. These are:
(i) hypertensive and normotensive classification or detecting the presence or absence of hypertension in subjects
and (ii) sitting and supine posture classification or detecting the change in posture of human subjects.

In case of the hypertensive classification, only 3, 3 and 1 parameters of the ACF, KF_{LPC} and KF_{TS} models respectively are identified to be most significant. In the posture change study, 5 parameters are selected for the ACF model while only 1 parameter is selected in case of both the KF models.

The ACF parameters for classifying hypertension are all normalized variances of the phasic components, while for the posture change the chosen parameters are 2 each from the phasic and the tonic components and 1 model coefficient. In case of both the KF models for both classifications, the chosen parameters are all related to the process noise covariance q_{com} for the compromise metric.

Three standard cross-validation (CV) methods have been applied on the selected model parameters of the 3 models in both the classification studies. These are 10 fold CV (10FCV), leave one out CV (LOOCV) and leave one subject out CV (LOSOCV). In all cases, it is observed that the results of 10FCV and LOOCV are either identical or very close. However, since the subject bias effect is expectedly removed in LOSOCV, the results for the method differ significantly from those of the other two methods.

The features selected for the ACF model provide the most reliable and best F1-score result in case of the hypertensive classification. The balanced and highest specificity establishes the utility of the KF_{TS} model also in this classification despite the overall drop in F1-scores.

In case of the sitting and supine posture classification, the KF_{LPC} and ACF models show high 10FCV accuracies and relatively low or medium effects of subjective bias for the classification of posture change. The KF_{TS} model with its single selected feature is not suitable for this differentiation.

These studies validate all three proposed models. Their utility in studying and/or classifying 3 different human conditions have also been ascertained in terms of the effect of these conditions on the individual model parameters and metrics.

List of Abbreviations

ACF	Auto-Correlation Function
PACF	Partial Auto-Correlation Function
Ag-AgCl	Silver-Silver Chloride
BP	Blood Pressure
bpm	Bits per minute
CDF	Cumulative Distribution Function
DAS	Data Acquisition System
DDP	Differential Dermal Potential
ECG, EKG	Eleccardiogram
ERG	Electroretinogram
EDA	Electro-Dermal Activity
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electroocoulogram
ERG	Electroretinogram
GSR	Galvanic skin response
SCR	Skin Conductance response
SCL	Skin Conductance Level
SPL	Skin Potential Level
EDL	Electrodermal Level
EDR	Electrodermal Response
DC	Direct Current
AC	Alternating Current
LH	Left Hand signal
LL	Left Leg signal
RH	Right Hand signal
RL	Right Leg signal
10FCV	10 fold cross validation
LOOCV	Leave One Out Cross Validation
LOSOCV	Leave One Subject Out Cross Validation
LPC	linear prediction coefficients
AR	Autoregressive
MA	Moving Average
ARX	Autoregressive with external input
ARMA	Autoregressive Moving average
ARIMA	Autoregressive Integrated Moving Average

HMM	Hidden Markov Models
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
AUC	Area under the Curve
SpO2	Saturation of Peripheral Oxygen
SSE	Sum of Square Error
TNR	True negative rate
TPR	True positive rate
FPR	False positive rate
TN	True negative
TP	True positive
FN	False negative
FP	False positive
ROC	Receiver Operating Characteristic
PRC	Precision Recall Curves
ZCL	Zero-Crossing Lag
ZCI	Zero-Crossing Instant
IQR	Inter-quartile range
BCI	Brain-Computer Interface
SVM	Support Vector Machines
PAR	Polynomial Autoregressive
BL	bilinear
KF	Kalman Filter
RF	Radio Frequency
SD	Standard Deviation
TS	Time Series
SE	Spectral Entropy
RF	Random Forest
RT	Random Tree
PPV	Positive Prediction Value
AI	Artificial Intelligence
BCI	Brain-Computer Interface
DWT	Discrete Wavelet Transform
KF	Kalman Filter
EKF	Extended Kalman Filter
EKS	Extended Kalman Smoother
UKF	Unscented Kalman Filter
ML	Machine learning
OS	Operation System
PC	Personal Computer
PR	Pulse Rate
RMS	Root Mean Square
RMSE	Root Mean Square Error

NN	Neural Networks
SC	Skin Conductance
fMRI	Functional Magnetic Resonance Imaging
PPG	Photoplethysmograph
SNS	Sympathetic Nervous System
WBAN	Wireless Body Area Network
SNR	Signal-to-noise ratio
Norvar	Normalized variance
Norkur	Normalized kurtosis
WHO	World Health Organization
IMU	Inertial Measurement Unit

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Human health condition monitoring involves the study of physiological processes and the assessment of various electrical signals produced by the human body [2]. Bio-potentials that are generated within the human beings can provide valuable information about the functioning and health of various physiological processes [3,4]. These are employed in polysomnography (sleep studies) to monitor various physiological parameters during sleep, helping diagnose sleep disorders like sleep apnea and insomnia [5]. These are also used to study cognitive processes, memory, attention, emotion, and other aspects of human cognition and behavior [6]. Recently, these signals are also being used in various artificial intelligence (AI) applications, communications, remote health monitoring and control, geriatrics and so on. Overall, bio-potentials have numerous applications in clinical, research as well as social scenarios, offering valuable insights into the functioning of the human body and the brain.

Bio-potentials arise naturally from the activities of various cells, tissues, and organs. Hence, some specific signals are popularly used in specific health monitoring applications. Electrocardiogram (ECG or EKG) is widely used to diagnose and monitor various heart-related conditions, such as arrhythmia, heart attacks, and other cardiac abnormalities [7]. In particular, cardiac pacemakers and defibrillators use the ECG to monitor the heart's electrical activity and deliver appropriate electrical impulses to regulate the heart rate or treat life-threatening arrhythmia [8,9]. Electroencephalogram (EEG) is used to study brain function, diagnose neurological disorders like epilepsy, and analyze sleep patterns [10]. Electromyogram (EMG) helps assess muscle health, detect neuromuscular disorders, and guide rehabilitation treatments [11]. Electro-oculogram (EOG) is used to study eye movements and diagnose certain eye disorders [12,13]. Electroretinogram (ERG) records the electrical responses of the retina to light stimulation, aiding in the diagnosis of retinal

diseases and visual disorders [14].

In recent studies, researchers have employed electrodermal activity (EDA) in various studies to investigate a wide range of generic issues. EDA is a common term used to represent all electrical phenomena manifesting at the skin surface of human beings, typically as transient or ever-changing time-domain signals [15] [16]. Some common types of EDA signals are the Galvanic Skin Response (GSR), Skin Conductance Level (SCL), Skin Conductance Response (SCR) and Skin Potential Level (SPL). These provide valuable insights into an individual's psycho-physiological responses and are commonly used in psychology [17], neuroscience [18], human-computer interaction [19], and other fields to study emotions [20] such as happiness, sadness, fear, stress, attention as well as various aspects of human physiology. The area of other applications are biometric authentication, clinical and psychiatric research, neuromarketing research, gaming and virtual reality and sleep research.

Recently, a particular type of EDA signals has been introduced by Bhattacharya et al. [21], [22], [23]. Based on their method of acquisition, these signals have thereafter been termed as Differential Dermal Potentials (DDP) in Sarkar et al. [24]. Certain successful human condition monitoring applications using these signals have also been demonstrated [25], [23], [24], [26]. Despite these applications, the only model of these signals is a quasi-linear stochastic one based on its slow changing time-series characteristics. Yet, even this simplistic model and the basic statistical features of this signal offer some interesting insights into the underlying physiological processes [24].

There is thus a scope to develop models of these signals that can be used to study the signal characteristics.

An intrinsic feature of an acquired time-series data is that adjacent observations are usually related. Based on these inter-relations, different time-series analysis techniques have been developed to derive models of such data and interpret them [27]. In particular, bio-potential modelling is an essential tool that helps in the understanding of the underlying biophysical processes. These can thus be used as diagnostic tools to assess different aspects of human health. When developing these representative mathematical, computational, or empirical models, it is necessary to consider the specific attributes and properties of the signals being modelled [28]. In several cases, the models incorporate physiological parameters and tissue properties to simulate the behaviour of bio-potentials in specific organs or tissues [29]. Many bio-potentials exhibit spatial distributions across the body, such as EEG and ECG [30,31]. However, in general, the bio-potentials show slow time-varying

patterns as an effect of the underlying dynamic physiological processes.

Some useful by-products of these modelling are the development of advanced signal processing algorithms for denoising [32], feature extraction [33] and pattern recognition [34] to name a few. Such models can be used to predict abnormal electrical patterns associated with various medical conditions and so, these can aid in early diagnosis and monitoring of diseases. As a specific case, the Brain-Computer Interface (BCI) systems [35] rely mostly on modelling EEG signals and is used to decode and control brain activity [36] and translate it into commands for external devices [37]. In another case, 1^{st} order model of DDP signals has been developed where signals are fited with straight line of suitable slope through the zero crossing instant (ZCI) of the mean subtracted deviation signals [23], [24].

Thus, it is useful to obtain proper models of the bio-potentials that can yield comparable variations in the model outputs and the actual signals. The causative physiological processes can then be studied and interpreted based on their effects on various model parameters.

1.1 Review of literature

In view of the observations stated earlier, a comprehensive literature survey is provided on the modelling of time-series signals; the EDA and its applications; the DDP signals and their study.

A survey of some specific applications is also stated followed by the details of four specific experiments that were conducted by Sarkar [1], [24] to acquire the DDP signals for these applications.

It is to be mentioned that the data acquired in these experiments have been used in the present work for all the model validation and application studies.

1.1.1 Modelling time-series signals

Time-series signals, which are typically acquired in equi-spaced time intervals, generally exhibit some trend or periodicity and/or some random characteristics [38–40]. Predicting, smoothing, and filtering are essential tasks in time-series analysis. Of these, prediction is often used for forecasting [39] future events as well as past events which are related with various geological events, such as earthquakes [41], volcanic eruptions [42], or other natural phenomena; smoothing techniques are used to remove noise or short-term fluctuations from time-series data; and filtering involves separating the signal

(useful information) from noise in time-series data. A good model captures the qualitative features of the dynamics of the system and aids in the understanding of all available observations of the system [43]. In particular, it is well known that several real-time signals like EDA have a slowly varying *tonic* component along with a fast changing *phasic* component [15,44–46].

In the present work, three different models of the time-series DDP signals have been developed, each with its specific deterministic tonic component and corresponding stochastic phasic component.

(A) Order

It is known that the goodness of any model depends upon the detection of the proper model order and its coefficients. One of the favoured techniques for model order determination is based on the Autocorrelation function (ACF). It provides a measure of the internal correlation within a time series data and in doing so, detects the system order also [47]. Depending upon the nature and condition of the input random variables, the sample ACF has been proposed in different ways for both stationary and non-stationary random processes [48]. In an extension to include ergodic processes, the ACF has been derived for continuous as well as discrete events. In all these cases, the variables have been defined as samples of a random process at two different time instants.

The model order of several signals of biological origin like ECG or speech signals are commonly determined using the ACF or its primary derivative termed Partial Autocorrelation Function (PACF) and/or other derivatives [48,49], [50]. For a suitably long data set, an useful estimate of the sample ACF or the PACF is obtained using $\frac{1}{4}$ of the total time-series [27]. This avoids issues of overfitting or coloured residuals [51]. The lags at which the correlation is lost in this sample PACF or the ACF is identified as the model order for the Autoregressive (AR) process or the Moving Average (MA) process respectively [27] [52]. In case both the ACF and PACF trail off without hard zero crossing lags, these are considered indicative of stationary ARMA or non-stationary Autoregressive Integrated Moving average (ARIMA) processes and their model orders are determined accordingly [27]. Some studies use Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) etc. to develop the AR, MA, ARMA models of time-series signals [53]. The AIC and BIC based approaches also use ACF and/or PACF to obtain a preliminary model order, then perform multiple trials over a range of orders to determine a suitable model order [15] [54].

In the present work, the sample ACF has been revisited from a fresh per-

spective to determine the model order of real-valued stochastic signals.

(B) Deterministic or tonic component

Once the system order has been determined, various techniques such as state-space methods, transfer function based approaches and polynomial models are used for modelling the deterministic trend or periodicity of the time-series. The most popular technique among them, that is also simple and easy to implement, is that of the polynomial regression models [55]. Some of the widely used polynomial models are AR, ARX (Autoregressive with external input), MA, ARMA, ARMAX (ARMA with exogenous input), ARIMA models and their combinations or derivatives [56, 57].

Time-series polynomial models have wide applications in the field of biomedical research and human condition monitoring also. In [58] and [59], the estimated parameters of ARMA models have been used respectively for human gait analysis and for obtaining a physiologically understandable relation between blood pressure and flow signals. Christini et al. [60] showed the importance of non-linear models by applying linear AR and ARMA models and non-linear polynomial AR (PAR) and bilinear (BL) models for the analysis of heart rate dynamics.

Several other researchers have also applied time-series analysis on ECG, EEG or even EMG signals for predicting various human conditions [61–63]. A high accuracy ECG signal prediction method has been established based on Autoregressive Integrated Moving Average (ARIMA) model and discrete wavelet transform (DWT) [53]. An efficient multi-level stress detection method has been developed based on EMG and ECG signals, where the effectiveness of EMG signal is investigated in response to stressful conditions and their efficacy is compared for binary and multi-level stress detection [64]. In [65], an AI-based stroke disease prediction system using both machine learning (Random Forest) and deep learning (Long Short-Term Memory) algorithms is developed for real-time EMG signals that can predict stroke, heart disease accurately. An adaptive online procedure is presented for AR modelling of non-stationary multivariate time series signals to analysis the respiratory movement, heart rate fluctuation, blood pressure and multichannel EEG signals in [66]. A personal recognition method based on EMG Signal is also developed in [67].

Another modelling approach is that of Kalman filter (KF) models, which have the ability to estimate and predict the true state of a dynamic system in the presence of system disturbances as well as sensor noise. It is also used in diverse applications like financial time series data modelling [68], economic forecasting [69], climate and weather modelling, environmental monitoring,

signal processing [70]. When applied to audio and speech processing, the KF denoises the signals and thus enhances the quality of audio recordings or speech communication. Kalman filter is applied in several other applications also for speech enhancement [71–74] and noise reduction [75]. In [76], a Bayesian estimator is used as a Bayesian Kalman filter update step for single-channel speech enhancement in the modulation domain.

Kalman filters have also been used in the analysis of biomedical time-series data, such as ECG or EEG, for filtering noise and extracting relevant information. Bayesian filters such as the extended Kalman filter (EKF), extended Kalman smoother (EKS) and unscented Kalman filter (UKF) have been proposed for ECG denoising in [77, 78]. EEG artifacts have been detected using the EKF in [79], while parameter estimation using the UKF is used for dynamic EEG system modeling [80]. A KF based method is also used for EEG spike enhancement in [81].

In a similar way, there is a scope to identify the deterministic tonic components of the DDP signals using both these approaches and then use the model parameters for human condition monitoring.

(C) Stochastic or phasic component

In time-series models, residuals represent the differences between the actual values of the time series data and the predicted values generated by the model. It is widely assumed that a model with residuals that are random and normally distributed provides accurate predictions of the underlying system dynamics. Generally in time series model, autocorrelation, independence tests and cross covariance analysis are done for checking the characteristics of the residuals [82]. As mentioned earlier, the ACF, PACF methods typically attempt to yield residuals with minimum noise variance.

However, it is being accepted in recent times that the nature of stochasticity of a system also contains valuable system information. Specifically, there are two major types of stochasticity in systems, namely external perturbations due to random variations in the environmental conditions and internal system fluctuations [44], [45]. This is evident from the residual diagnostics studied for cross-section time-series regression models [83]. In a particular application in astronomical light curves, an irregular AR model is applied to the residuals of a transit of an extrasolar planet, identified based on the signal autocorrelation, to illustrate errors that remain with temporal structure [84]. In another application, a robust method is developed for background noise identification and subsequent actual signal detection [85]. It is also well accepted that bio-potentials, and particularly EDA signals, have a well-defined

stochastic *phasic* component which are often a composite effect of internal and/or external processes [50], [15], [86]. A time integration of the continuous measure of phasic activity is proposed as a straightforward indicator of event-related sympathetic activity in [87]. It has further been established that the phasic component of EDA is more correlated to brain activity than the tonic component [88].

Hence, an effective model of a real signal should have proper representation of the stochastic phasic component that is low in magnitude, but not necessarily minimal, and can aid in interpreting system behavior.

1.1.2 Electrodermal activity (EDA)

Electrodermal activity (EDA) is a physiological measure that reflects the dynamically changing electrical properties of the skin. The changes of the skin conductance are primarily influenced by the activity of sweat glands, which is controlled by the sympathetic nervous system [86]. It is established that when an individual experiences emotional arousal, stress, or other physiological changes, the sympathetic nervous system is activated, leading to an increase in sweat gland activity.

The electrodermal signals have a *tonic* component that is the slow changing or quasi-static electrodermal level (EDL) of the acquired signal. The low basal skin potential level (SPL) is one EDL based score proposed by Christie and Venables [89]. It is a measure of the basal sympathetic activity and is influenced by factors like temperature, humidity, and emotional state. Lykken et al. [90] observed inter-individual differences for the minimum SPL even in fully relaxed individuals.

The other component of the EDA is its *phasic* component, which is also called the electrodermal response (EDR). However, it is not always necessary to have a distinct relationship between a stimulus and an EDR. Various EDR studies have been done by analyzing the latency time, window, amplitude and shape of response [15]. SCR is a specific type of phasic component of EDA that is often associated with emotional responses. It captures rapid changes in skin conductance in response to specific internal or external stimuli or events, such as emotional triggers, cognitive tasks, or external stimuli.

The EDA is also classified into two categories based on the absence of presence of external electrical stimuli. When EDA is recorded without any application of an external current, it is termed as *endosomatic* EDA. In this case, only the potential differences originating in the skin itself are recorded. In the alternative *exosomatic* EDA recording, either direct current (DC) or alternating current (AC) is applied to the skin [15]. While the EDA measure-

ments are primarily influenced by endosomatic factors, yet the exosomatic factors affect the physiological processes or contribute to the external conditions and these in turn influence the EDA. Thus, understanding the interplay between internal physiological processes and external environmental factors and their effect on the bio-potentials or skin conductance is important for accurate interpretation of EDA measurements.

The time-series modelling approaches like ARIMA models, hidden Markov models (HMM) and dynamic Bayesian networks have been used to capture the dependencies and transitions within EDA data over time [91]. Machine learning (ML) algorithms and techniques like support vector machines (SVM), random forests (RF), or neural networks (NN), predictive models based on EDA features have also been developed to classify emotional states, predict stress levels and detect anomalies [92]. A new methodological pipeline is stated in [93] to recognize acute stress conditions using EDA modelling. A biophysical model using an extended Kalman filter is also implemented for EDA [94]. Edelberg (1983) established a simplified electrical equivalent resistive circuits with capacitive elements for EDA modelling by [15].

These EDA signals are also used in various feature extraction applications to determine the signal amplitude, frequency, slope, and variability. Time-domain analysis, frequency-domain analysis or wavelet analysis are also involved in this extraction [95]. As mentioned earlier, their regression models are used to establish the relationship between EDA and other independent variables such as emotional stimuli [16] , cognitive load [96], or environmental factors. Exosomatic EDA is used in psycho-sociological states research where non-electrical stimuli like audio, visual even different odorant stimuli are also used along with the electrical stimuli [97–100]. Arousal assessment by measuring tonic EDA signals [101], multiple arousal studies based on daily life EDA asymmetry [102], habituation [103] and dishabituation [104] are also studied. In [105], psychophysiological stress involving heart rate variability (HRV) and EDA reflecting sweat gland activity were reviewed and the utility in health care simulation research is also demonstrated. EDA studies for assessing emotional states and stress (by applying different stressful stimuli) [106–108], anxiety disorder [109–111] and depression [112] confirmed that both specific and non-specific responses are useful for the assessment of stress. SCL (skin conductance level), SCR and NS.SCR (non-specific skin conductance response) frequency are considered to be very useful for depression identification [113,114]. Subjects with some phobia, like spider phobia, have been identified by recording SC (skin conductance) and fMRI (Functional magnetic resonance imaging) together [115]. In the field of physiological research also, tests have been conducted to understand behavior [116], gender and hemispheric dominance using mirror image stimulus

morphometry [117, 118]. EDA is also used for assessment of physiological signals during happiness or sadness [119], to characterize fear conditioning and fear extinction [120] or even to study the gender-wise visual attention performance under mental stress [121].

EDA signifies the autonomic brain system activity [122] which has wide applications in the psychological research domain and EDA is applied to analyze marketing data where, artificial intelligence (AI), virtual reality (VR), and neuropsychological tools are used [123]. In particular, the sympathetic nervous system activity is also measured continuously and non-invasively using EDA [124]. In particular, the quantitative analysis of wrist EDA during sleep, EDA patterns in sleep stages and their utility for sleep versus wake classification have also been studied in [125], [126].

Hence, DDP signals have various applicability in emotions study, human physiology analysis, biometric authentication, clinical and psychiatric research, neuromarketing research, gaming and virtual reality and sleep research.

Another aspect of the utility of EDA signals is in assessing the effects of diabetes and other physiological effects and/or diseases. A relation between EDA and diabetic foot is discussed in [127]. EDA is analyzed for quantification of sweat gland activity in diabetic kidney disease [128]. EDA is also used for detection of peripheral abnormalities in type-2 diabetes. The feasibility and usefulness of EDA and HRV measurements for the detection of diabetic foot disease in patients with type-2 diabetes mellitus (DM) are also reviewed in [129]. Nonlinear analysis of these signals is also done for healthy subjects and patients with chronic obstructive pulmonary disease [130].

However, due to the unstable and irregular response characteristics of traditional endosomatic EDA [15], there are very few applications using this type of EDA [131]. Therefore, once these issues are addressed, there is significant scope for more applications using that particular type of EDA signal in future.

1.1.3 Differential Dermal Potentials (DDP)

Differential dermal potentials (DDP) are a special type of endosomatic EDA signals which differ from most such signals since these are acquired in the *differential* mode between two active sites [132, 133]. Thus the sites and mode of acquisition of this signal are similar to that of exosomatic EDA signals but without any applied stimuli [15, 16, 134]. These differential signals are recorded by connecting a pair of sensor electrodes across the intermediate phalanges of the middle finger and the forefinger of one hand or feet [25], [23]. In order to acquire these biopotentials reliably, a signal conditioning

circuit [22] and a dedicated 4-channel data acquisition system (DAS) [133] have been developed.

Detailed study, validation and characterization of these biopotentials have been done by Bhattacharya [23] and Sarkar [1]. In [23], the determination of the qualitative and quantitative features and parameters of DDP signals are measured to characterize healthy human subjects at rest and a process is developed using a minimal set of the identified parameters to synthesize realistic pairs of DDP. The conditions and methodology for acquiring DDP signals are also stated here briefly. In a further study of DDP signals [1], the calibration of the proposed DAS, different experiments for acquiring these signals and applications of the features of DDP signals in human condition monitoring have been discussed. Since these skin potentials are generated by all cells and tissues, including the nerves, blood vessels and muscles in or under the skin, hence they are controlled by both the sympathetic and the parasympathetic activities unlike most other EDA signals. It has been successfully demonstrated in an application that the DDP signals acquired simultaneously from both hands using 2 channels of the 4-channel DAS can effectively discriminate cognitive load [26]. In this study, the variation of DDP for differing levels of cognitive load are studied and the results are used to classify the low and high mental workload. The performance of the DAS is compared with that of conventional exosomatic device (GSR). In another study [21], a non-invasive technique has been used to measure these signals and statistical parameters that are associated with the health of the human system have been identified along with their ranges. In yet another study, a comparative study of the DDP signals acquired from left and right hands of human subjects has been performed [25]. Here, the qualitative as well as statistical characteristics of these signals are compared to determine the bilateral relationship and based on these findings, a new parameter has also been defined.

These real-time signals are typically slowly varying with low magnitudes, typically within $\pm$ 300mV. Based on this characteristic, a 1^{st} order model has been developed which fits a straight line of suitable slope through the zero crossing instant (ZCI) of the mean subtracted deviation signals [23], [24]. However, it is evident from the studies that the residuals have higher order dynamics.

Hence, there is a need to determine the model order properly and to obtain suitable dynamic models of these DDP signals. This is expected to provide a scope for wide applicability of these signals.

1.1.4 Restfulness, hypertensive classification and posture change studies

Restfulness assessment is an important topic of research since proper rest helps to reduce stress, overcome fatigue and restore normal physical and mental state [135, 136]. Apart from these aspects, the study of the restful posture without any externally applied stimuli is also useful as a baseline for several other studies, whether static or dynamic or involving application of electrical and/or non-electrical stimuli [26].

Rest and sleep are essentially different from each other but no-nap rest is very similar to stage 1 sleep that lasts not more than 10 minutes [137, 138]. In particular, EMG signals exhibit lower amplitudes [63] and EEG signals show steadier characteristics [139] during stage 1 sleep than while a subject is awake. This is indicative of the relaxation of the muscles and nerves during this transition period between wakefulness and sleep. Sleep detection in drivers during highly automated driving are done using potential physiological measures [140] and driver drowsiness detection [141]. In fact, to reduce the higher number of car accidents, driver fatigue and stress detection and evaluation are done using ECG, EMG, EDA [142]. Recently, it has been shown that DDP signals acquired from the left hand (LH) of subjects have statistical and spectral features that are affected by long rest [24].

Classification is a popular technique nowadays in the field of biomedical condition monitoring and detection purposes. In few recent studies, the classification technique is used for brain disease [143, 144] and tumour detection from medical image data [145, 146], analysis of chest disease [147], EMG signal classification [148], pneumonia disease classification [149], detection of epileptic seizure from EEG signals [150], arrhythmia detection [151], even detection of Covid-19 from X-ray [152].

In the backdrop of the increased incidences of heart attack and stroke diseases, the study of hypertension-normotension plays an important role in recent research. There are different hypertensive and normotensive classification methods using ECG and photoplethysmograph (PPG) signals [153]. Several indexes have also been proposed for hypertension detection using ECG signals and electrodermal signals [154]. In a recent study, an automated technique is proposed for detection of severity of hypertension in human subjects [155]. In another study, difference in races in cardiovascular and non-cardiovascular sympathetic nervous system (SNS) activity was studied for hypertensive and normotensive subjects. The mentioned studies used classification technique in machine learning for prediction of hypertension [153–156] from different physiological featured data [157, 158] or medical data [159].

Another topic meriting interest in recent research is posture detection. Study of posture has been done for behaviour assessment from long term posture and activity records [160], study of efficiency [161] and engagement [162] of students or scholars, remote health care monitoring using sensors as in WBAN (Wireless Body Area Network) [163], fall detection of lone elderly subjects using key frame extraction from video data and then SVM (support vector machine) classification [164] and mobile robotic applications [165] to name a few. It is well established from several studies that posture and cardiac activity are dependent on each other [166] in terms of cardiac output [167], heart rate or R-R interval [168] or even systolic blood pressure [169]. For this reason, posture estimation has reliably been performed from cardiac monitoring also [170]. Machine learning approach is typically used for this approach in which the data for different postures are considered as different classes for the classification [170,171].

There is a scope to study the efficacy of the features of the models of the DDP signals to be proposed in this work in these specific applications.

1.1.5 Different conditions for data acquisitions [1]

The DDP signals are acquired from human subjects hence, this study was approved by the Institute Ethical Committee (IEC) of Jadavpur University, Kolkata, West Bengal, India and the approval number is JU/IEC/2020/2312/01. In accordance with 1964 Helsinki Declaration and its later amendments or comparable ethical standards, all the participants were informed about the study briefly and consent forms were signed by each of the participants before participating in the study. A common overall questionnaire was filled up by each of the participants to record their general health condition. The subjects also filled a daily questionnaire for recording their self assessment of their daily health. These consent form and the two sets of questionnaires are affixed in Appendix.

(A) Subject recruitment

Total 32 dextral subjects were selected for DDP signals acquisition and they were volunteers, either students, scholars or teaching and non-teaching staff members from the Salt Lake campus of Jadavpur University. The ages of all male and female subjects were 24 to 58 years and the weight, height were 55 to 105 kg, 148 to 185 cm respectively. They were from Eastern India, both urban and rural localities in India with qualification of under graduate, post graduate, pursing Ph.D. or awarded. Subjects with Hypertension, diabetes mellitus, high cholesterol and/or thyroid problems are allowed to participate

in this study. They were addicted in smoking citrate/biri, jarda paan, tea and coffee and the informations were noted down in the consent forms.

(B) Materials

It was mentioned in [1] that all DDP signals were acquired using the RISH Multi 18S with the RISH Multi SI232 assembly and the 4-channel (labelled as Ch0, Ch1, Ch2 and Ch3) DAS. 8mm round Ag-AgCl velcro mounted snap type electrodes were used for connection along with single core coaxial cables as connectors and aqueous ultrasonic gel as the electrode-skin interface medium. Data were displayed and stored in a PC or a laptop. These signals were acquired from LH (left hand), RH (right hand), LL (left leg) and RL (right leg) locations respectively. The oxygen saturation level (SpO2) and the systolic and diastolic BP were measured using a pulse oximeter (Make: CONTAC, Model: CMS50D) and a sphygmomanometer (Make: Rossmax , Model: GB101) respectively while the PR was measured manually using a stop watch in all experiments.

(C) Experimental protocol and method

There are 4 experiments (DS1, DS2, DS3, DS4) are followed for acquiring the DDP signals. For each of experiments, some protocols are maintained. Those are discussed in step by step thereafter.

Step 1: Before starting an experiment, every subject was asked whether it is first time or not and if the answer is yes then they are asked to give their consent in a predefined format and if no, the next step is proceed. The detailed of consent form is stated in [1].

Step 2: As the DDP signals are acquired in different days so, a set of daily questionnaires regarding the physical and mental status of a subject on that day were noted before starting the experiment.

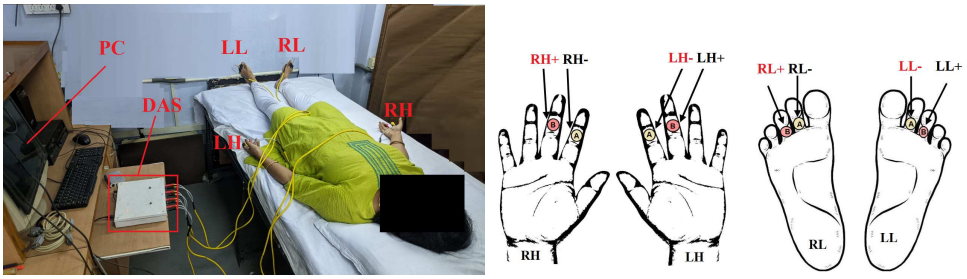
Step 3: Then the subject was asked to lie down on the experimental bed in supine posture or sit in the chair in sitting posture for relaxation upto 10 minutes. The form of the posture for data acquisition depends upon the experiment type.

Step 4: After relaxation, the BP, temperature, PR and/or SpO2 of the subject was measured and recorded.

Step 5: Thereafter, the electrodes were connected in specific places of fingers in differential mode and data are recorded. The connection types of the electrodes are different for different experiments. After the recording, BP, PR of the subject were also checked again.

1.1.6 DDP signal acquisition experiments [1]

In order to study the three aforementioned applications, DDP signals were acquired by Sarkar [1] from 4 types of experiments labelled as DS1, DS2, DS3 and DS4 using the 4-channel DAS or the Rish Multi 18S [172] multimeter and Rish Multi SI232 [173]. These signals were acquired from one or more of the four locations, namely LH (left hand), RH (right hand), LL (left leg) and RL (right leg). Total 32 dextral volunteers within the ages of 24 to 58 years participated in these experiments for acquiring the DDP signals. There were 21 male and 11 female subjects in this set. The experimental set up and electrode connections for a 4-channel data acquisition are shown in Figure1.1a and 1.1b respectively. In standard recommendations for electrodermal measurements [134], it is recommended that exosomatic EDA should be recorded in differential mode across two active sites while endosomatic EDA should be acquired in referenced mode across an active and an inactive site (neutral point). However, as stated in [1, 24], the (endosomatic) DDP signals are acquired in differential mode, so there is no neutral point in these signal acquisitions.



(a) (b)

Figure 1.1: (a) Complete experimental setup (b) Placement of electrodes on both hands and both feet [1]

The protocol for acquiring DDP signals are common for all experiments. Since the typical bandwidth of the EDA signals is known to be $0 - 3$ Hz, hence the sampling frequency has been set to 20 samples/s for DDP signal acquisition [25]. Furthermore, it has been shown in [1, 24] that the settling time of the DDP signals is less than 2 minutes, specifically the mean and SD are 81.17s and 59.63s respectively. Hence, the acquired signals, which are of either 20 minutes or 10 minutes duration in the 4 experiments DS1 to DS4, are subdivided into 10 or 5 subsequent 2-minute duration sets respectively. However, the posture of the subjects, duration of acquiring DDP signals and the connection of electrodes are different for different experiments. The details of the four experiments, DS1 to DS4, are stated hereafter.

DS1

In the DS1 experiment, DDP signals were acquired only from the LH of individual subjects for 20 minutes in a predefined sequence. The DDP was acquired between the intermediate phalanges of the middle and the index fingers. The data set was acquired for 2 minutes in sitting posture. The data recording continued in the next 2 minutes while subjects changed their posture from sitting to supine. In the subsequent 16 minutes, the acquisition continued while the subjects remained in supine posture. A total of 57 datasets of 20 minutes duration was acquired from 24 subjects within a period of 2 months.

DS2

In the DS2 experiment, DDP signals were acquired between the intermediate phalanges of the middle and the index fingers of both LH and RH of individual subjects for 10 continuous minutes in supine posture. A total number of 98 sessions of data were acquired from 16 subjects within 4 months duration for this experiment.

DS3

The experimental set up and the electrode connections in DS3 are the same as in DS2. 3 subjects participated in this experiment over a period of 3 months and a total of 66 datasets were recorded. The electrode connection was same as DS2 in supine posture and data acquisition with SpO2 recording was continued for 2 minutes. After that the pulse oximeters were removed and BP and PR were recorded from both hands. In next 2 minutes, the pulse oximeters were further connected to both the thumbs and no data were recorded. Again the data as well as BP, PR were recorded for 2 minutes in supine posture. After this 2 minutes, the subject was asked to get up from the supine position to a sitting posture on the chair beside the bed and relax without any body movement. In this posture, data and SpO2 were recorded for 2 minutes and the same recordings were noted in standing posture for next 2 minutes again. The subjects were then asked to go out, climb down a fixed flight of stairs to the lower floor, climb back up the same way and return to the cubicle, all within 1 minute. After that, the subject was asked to sit in the chair while the electrodes and pulse oximeters are reconnected. Then the recording was restarted with the subject in sitting posture for 2 minutes. The BP was measured thereafter from left hand and compared with the previous record when subject was in sitting posture. Once these match, then the acquisition is stopped. Else the data recording in sitting posture is resumed again for 2 minutes followed by the BP check for the left hand

until the condition is met or till 10 minutes, whichever is earlier. After the acquisition is finally stopped, the electrodes are disconnected, all the records are saved and subjects are asked to resume their regular work.

DS4

In the DS4 experiment, DDP signals are recorded continuously for 10 minutes from all 4 locations, namely LH, RH, LL and RL, of a subject in supine posture. The acquisition technique is same as DS2 except for the electrode connections, 4 in case of DS4 and 2 in case of DS2. A total of 66 datasets of 10 minutes duration were acquired from 12 subjects in a duration of 2 months.

The independence of the signals acquired simultaneously from the 4 limbs in DS4 have been established in [1] from the results of their cross-correlation. Hence, considering the signal from each channel as independent and then subdividing them into 2 minutes long non-overlapping sets, a total of 5706 numbers of DDP signals have been obtained from all the four (4) experiments. Each of these 2 minutes long signals have a sampling time of 50 ms, and hence 2400 discrete samples. All these signals have been considered for each of the three model validation studies presented in this work.

In case of the restfulness assessment studies and the two classification studies, only the signals from the suitable experiment(s) have been considered. Further details of the signal pre-processing and processing in these cases have been discussed in the relevant Sections.

1.2 Aim and Organization of the thesis

1.2.1 Aim and objectives

The aim of this work is to develop various models of the differential dermal potentials (DDP) and to determine the use and efficacy of the parameters derived from these models in studying and/or classifying human conditions. This has been addressed in this thesis in various stages.

The first step for developing the model of any real time-series signal is to determine its order. For this, the traditional approach based on the sample autocorrelation function (ACF) of the signal is revisited in this work. A new perspective based on the derivation of closed form solutions of the ACF of incrementally complex signal models from first principles is explored.

The next step is to develop alternative models of the signal based on the model order determined using the sample ACF. In the first stage, a procedure to obtain a model with a deterministic polynomial time-series component

and a stochastic component, which can represent the dynamics of the signal adequately, is investigated. In order to clean the system and sensor noises in the signal, the development of two other models based on the Kalman filter (KF) approach are studied. Two distinct approaches are investigated for obtaining the state space model and the subsequent KF model: a) using the linear prediction coefficients (LPC) obtained from the popular autoregressive (AR) model approach and b) deriving it ab-initio from the deterministic polynomial time-series component developed in the first model.

Finally, the validity of the models developed are ascertained by studying the efficacy of their parameters in analyzing and/or classifying the three applications related to the DDP signals that were studied in [1], namely a) study the effect of long duration of rest, b) classify the presence or absence of hypertension and c) differentiate between the sitting and supine postures of a subject.

1.2.2 Organization of the thesis

Chapter 1 contains the review of existing literature and the aim and organization of the thesis. The review of literature in Section 1.1 contains a detailed study of the existing state of art in electrodermal activity (EDA). A background of the recent development of the special type of EDA, namely the differential dermal potentials (DDP), is stated in the same Section. This includes a brief overview of 4 sets of experiments conducted and reported in [1] for acquiring the DDP signals since these are used for the studies presented in this thesis. Section 1.2 contains the aim and objectives as well as the organization of this thesis and finally, the list of contributions is stated in Section 1.3.

Chapter 2 deals with the development and use of the stochastic ACF-ZCL model, or ACF model in short. The concept of zero crossing lag (ZCL) of sample ACF as well as the methodology for its use in determining the model order and the ACF model are proposed in Section 2.2. The expressions for obtaining the ZCL of power series signals, polynomial time-series signals, stochastic signals and a general signal from first principles are stated in this Section along with some key observations. Finally, an algorithm is proposed for the model order determination and obtaining the stochastic ACF-ZCL model, or ACF model in short, of a signal in this Section. Section 2.3 contains the modelling protocol, method and analysis of previously acquired [1] DDP signals. The effect of noise and coefficient interactions on the model order are illustrated in this Section. The analysis of the tonic and phasic components of the ACF models of the DDP signals and a comparison of the

ACF model features with the standard PACF model features are also presented in this Section. The assessment of various aspects of restfulness in terms of the effect of rest on the characteristics of the tonic and the phasic parameters of the ACF model is stated in Section 2.4. The chapter discussions are stated in Section 2.5.

The formulation of two distinct KF models are proposed in Chapter 3. The standard Kalman Filter model is denoted here as the KF_{LPC} model since its underlying state space model is obtained using the standard AR based Linear Prediction Coefficients (LPC). The formulation of the state space model from the available signal, its LPC and covariances is proposed in Section 3.2. In Section 3.3, a time-series (TS) state space model is derived ab-initio from the ACF model of the time-series signal y_k. The derived TS state space model forms the basis of the proposed KF_{TS} model. The standard KF algorithm that is applied to the LPC and TS state space models to obtain the standard KF_{LPC} and proposed KF_{TS} models of the DDP signals is described in Section 3.4. The robustness (J_1) and sensitivity (J_2) metrics of KF models have been proposed in [174]. These metrics, their associated parameters and their implications in terms of the two KF models are also discussed in this Section. The KF models of the DDP signals are discussed in detail in Section 3.5 including their modelling protocol and the method used for obtaining the filter parameters. This Section also contains a comparison of the various parameters and the metrics of the two KF models and the ACF model. Thereafter, the restfulness assessment performed in Section 2.4 is extended in Section 3.6 to compare the ACF and the two KF models systematically in terms of the spectral entropy and the various tonic and phasic components of the 3 models as well as the noise residuals and the key metrics parameters of the two KF models. The chapter discussions are presented in Section 3.7.

The effectiveness of these models in classifying two human conditions is discussed in Chapter 4. The experimental methodology used for the classification studies is described in Section 4.1. At the outset, the type of data sets used and their pre-processing, as discussed in [1], have been stated. The details of the signal quantization performed in this study for the two classifications, the model parameters used, the feature selection and ranker methods tested, the classification metrics evaluated as well as the cross-validation protocols used for the studies are also discussed in this Section. The details of the hypertensive and normotensive classification are stated in Section 4.2 and the posture classification in Section 4.3. In each case, the selection of ranker, the attributes selection and the cross validation results for the 3 models and

their analysis are discussed in detail in the respective subsections. Section 4.4 contains the overall discussions.

Thesis conclusions and future scope of work are presented in Chapter 5.

1.3 List of contributions

The list of contributions of this study presented chapter-wise are as follows:

1: Chapter 2: The concept of zero crossing lag (ZCL) of sample ACF, its closed-form solution for various types of signals as well as the methodology for its use in determining the model order of these signals is proposed.

2: Chapter 2: A stochastic ACF-ZCL model of general time-series signals is proposed along with an algorithm to determine its tonic and phasic components. This is applied to analyze several DDP signals.

3: Chapter 3: The KF_{LPC} model of DDP signals is developed using the standard LPCs obtained from the AR modelling approach. A methodology is proposed for obtaining the state space model for the KF_{LPC} model.

4: Chapter 3: The KF_{TS} model along with its underlying state space model is derived ab-initio from the proposed ACF model.

5: Chapter 2 and 3: The effect of rest on various features of the tonic and phasic parameters of the ACF, KF_{LPC} and KF_{TS} models as well as on the metric parameters of the 2 KF models are analyzed.

6: Chapter 4: The effectiveness of ACF, KF_{LPC} and KF_{TS} model parameters in two classification studies have been established. These are
(i) hypertensive and normotensive classification or detecting the presence or absence of hypertension in subjects
and (ii) sitting and supine posture classification or detecting the change in posture of human subjects.

Chapter 2

Stochastic ACF-ZCL model of DDP signals

2.1 Introduction

In this chapter, a suitable time-series model of the DDP signal is established using its ACF characteristics. As mentioned in Section 1.1.1, Chapter 1, it is known that several real time signals like EDA have a slowly varying *tonic* component along with a fast changing *phasic* component [15, 44–46]. Similarly, it is to be expected that the model for the DDP signals, will have the deterministic component and the residual error time-series will be representative of the tonic and phasic components of the signal dynamics respectively. Thus, the objective of modelling these signals is ascertain the deterministic trend using time series regression methods like a p^{th}-order polynomial of time. This results in a residual error time series, which is typically serially correlated and represents the stochastic component [175].

Hence, the primary task for modelling any system is its order identification [15] [54]. For this purpose, the most popular method is to study its ACF characteristics. As mentioned in Chapter 1 in [1], it is reported that the ACF characteristics of DDP signals are of two types: one having a deterministic linear trend and others having higher order dynamics. However, contrary to prevalent ACF based approaches, a direct method has been proposed in this work to fit the model to the acquired DDP signals. This method is based on a closed form solution of the 1^{st} zero crossing lag (ZCL) of the sample autocorrelation function (ACF) of an underlying ideal signal. This and some other notations are listed in Table 2.1. Since the trend of the DDP signals is linear, so a 1^{st} order ideal affine model was primarily assumed for these signals. However, higher order dynamics have been observed in several of

the residuals of the stochastic affine models of the signals. So, this same approach has been extended to obtain general order models, henceforth termed as stochastic ACF-ZCL models (Table 2.1).

The work presented in this Chapter aims to

a) Develop a method based on the ZCL of the ACF to determine the model order for any real time-series signal, alternatively referred as time-series or series.

b) Determine a stochastic polynomial regression model of the actual signal and validate the model properties in comparison with those of the standard PACF model.

c) Establish the correlation of the deterministic trend and the stochastic residual time-series with the tonic and phasic components of the real signal and interpret the physiological relevance of the model features. The specific application of no-nap restfulness has been used for this purpose.

Table 2.1: List of notations

Symbol	Description
T	Sampling time=50 ms.
k	Sample (or time instant).
N	Number of samples=2400.
i	Order of the particular term.
n	Overall polynomial model order.
a_i	Coefficient of the i-th order term.
w_k	Zero mean, finite variance Gaussian noise.
l_{zc}, ZCL	1^{st} zero crossing lag of sample ACF.
r_k	Residual, alternatively phasic component of signal.
RMS	Root of mean of square of a variable.
CDF	Cumulative distribution function
SSE	Sum of squares of error
IQR	Inter-quartile range
Stochastic ACF-ZCL model	ZCL based ACF model with additive stochastic residuals

In this work, closed form solutions are determined for the ZCL of a general power law series, both deterministic and stochastic, as well as of a polynomial time-series. These time-series are denoted as y_{dk}, y_{sk} and y_{pk} respectively as listed in Table 2.2. The power law series are considered to be of order i while the polynomials are considered to be of order n, unless otherwise mentioned. It must be mentioned here that in the seminal text by Box and Jenkins et al. [27], the dying out of the sample ACF of stochastic processes is discussed

in brief but no solution is provided for the ZCL of power law or polynomial time-series. The ZCL is used to determine the order of the deterministic polynomial regression fit, hereafter also referred simply as polynomial or polyfit, component of a stochastic model. Thereafter, an algorithm is proposed to fit the proposed stochastic ACF-ZCL model, also referred to in short as ACF model, to actual time-series signals y_k. Finally, the ACF model components are utilized for restfulness analysis.

The Chapter is organized as follows. The concept of ZCL of sample ACF and the structure of the stochastic ACF-ZCL model are proposed in Section 2.2. Section 2.2.1 contains the expressions for obtaining the ZCL of the stochastic power series y_{sk}, the deterministic power series y_{dk}, the stochastic 1^{st} order polynomial fit (polyfit) y_{k1}, the polynomial time series y_{pk} and a general signal y_k from first principles in theorem and corollary format. The key observations from the theorems, presented in Section 2.2.2, are used in Section 2.2.3 to propose an algorithm for the model order determination and obtaining the stochastic ACF-ZCL model of a signal. The modelling protocol, method and analysis of all acquired DDP signals of 1^{st} order and higher order are presented in Section 2.3. The effect of noise and coefficient interactions as discussed in Section 2.2.2 are illustrated in Section 2.3.3 while the analysis of the tonic and phasic components of DDP signals are presented in Sections 2.3.4 and 2.3.5 respectively. A comparison of the proposed stochastic ACF-ZCL model features with the standard PACF model features is discussed in Section 2.3.6. The assessment of various aspects of restfulness using the stochastic ACF-ZCL model features is stated in Section 2.4. The details of the assessment methodology as well as the effect of rest on the characteristics of the tonic and the phasic model parameters are discussed in Sections 2.4.1, 2.4.2 and 2.4.3 respectively. Discussions are stated in Section 2.5.

2.2 Stochastic ACF-ZCL model

It is well known that any bio signal contains higher order dynamics. As studied in [1], it is observed that the DDP signals also reflect the same nature. This is evident in the ACF characteristics of these signals also. The value of the sample ACF of the DDP signal starts from 1 and then decreases via 0 to a negative minimum. As evident in Figure 2.1, the ACF value thereafter increases again and may undergo further undulations with or without additional zero crossings till it finally ends at 0 [176]. In this work, the lag at which the sample ACF crosses the zero axis *for the first time* has been used to determine the model order of the signal. As marked in Figure 2.1, this *first zero crossing* of the signal ACF is denoted as l_{zc} and termed

Table 2.2: List of terms and their mathematical forms

S.No.	Terms	Description	Mathematical forms
1.	$(k+l)^i$		$= ({}^iC_0k^i + {}^iC_1k^{i-1}l + ... + {}^iC_il^i)$
2.	S_{kl}		$\doteq ({}^iC_1k^{i-1}l + ... + l^i) = (k+l)^i - (k)^i$
3.	S_i (S_n)		$\doteq \sum_{k=1}^{N} k^i$ $(= \sum_{k=1}^{N} k^n)$
4.	S_j		$\doteq [\sum_{k=1}^{N}(\sum_{i=0}^{n-1} a_i(kT)^i)]/N$
5.	$X_{k,l}$	Cross covariance of noise w_k	$\doteq \frac{1}{N}\sum_{k=1}^{N-l} w_k.w_{k+l} = E[w_k.w_{k+l}]$
6.	X_1		$= \frac{24X_{k,l}}{(N-l_{zc})(a_1T)^2}$
7.	y_{dk}	Deterministic (i^{th} order) power law series	$= a_i(kT)^i, \ \forall k \in [0,N]$
8.	y_{sk}	Stochastic (i^{th} order) power law series	$= a_i(kT)^i + w_k = y_{dk} + w_k, \ \forall k \in [0,N]$
9.	y_{pk}	Deterministic (n^{th} order) polynomial or trend or tonic component	$= \sum_{i=0}^{n} a_i(kT)^i, \ \forall k \in [0,N]$ $= y_{dk}\vert_{i=n} + \sum_{i=0}^{n-1} a_i(kT)^i, \ \forall k \in [0,N]$
10.	y_{pk1}	Deterministic 1^{st} order polynomial	$= a_1(kT) + a_0, \ \forall k \in [0,N]$
11.	$\overline{y_{dk}}$,$\overline{y_{sk}}$, $\overline{y}$	Mean of the power law series	$= \frac{a_iT^i}{N}\sum_{k=1}^{N} k^i + E[w_k] = \frac{a_iT^i}{N}S_i$
12.	$\overline{y_{pk}}$,$\overline{y_k}$	Mean of the polynomial y_{pk} or signal (or its model) y_k	$= \frac{\sum_{k=1}^{N}(\sum_{i=0}^{n} a_i(kT)^i)}{N}$ $= \frac{\sum_{k=1}^{N}(a_n(kT)^n)}{N} + \frac{\sum_{k=1}^{N}(\sum_{i=0}^{n-1} a_i(kT)^i)}{N}$ $= \frac{a_nT^n}{N}S_n + \frac{S_j}{N}$
13.	y_k	Stochastic ACF-ZCL model	$= y_{pk} + w_k \ \ \forall k \in [0,N]$ $= y_{sk}\vert_{i=n} + \sum_{i=0}^{n-1} a_i(kT)^i, \ \forall k \in [0,N]$
14.	y_{k1}	Stochastic 1^{st} order ACF-ZCL model	$= y_{pk1} + w_k \ \ \forall k \in [0,N]$
15.	$\overline{y_{k1}}$	Mean of y_{k1}	$= \frac{\sum_{k=1}^{N}(a_1kT + a_0 + w_k)}{N}$ $= \frac{a_1T(N+1)}{2} + a_0 + E[w_k] = \frac{a_1T(N+1)}{2} + a_0$
16.	r_k, w_k	Residuals of signals	$= y_k - y_{pk}, \ \forall k \in [0,N]$
17.	r_{k1}	Residuals of 1^{st} order signals	$= y_{k1} - y_{pk1}, \ \forall k \in [0,N]$

the zero crossing lag (ZCL)(Table 2.1).

As mentioned earlier, a stochastic polynomial regression fit of any general time-series signal is proposed as the stochastic ACF-ZCL model in this work, while this study has been validated on the DDP signals. Thus, the actual signal y_k is modelled as $y_{pk} + w_k$, where w_k is an additive time-correlated noise which adds on to a deterministic high-order polynomial time-series y_{pk}.

In order to assess the ACF characteristics of this total signal, we take a block-by-block build up approach. Each term in the polynomial component is of the form of a deterministic power law process $y_{dk} = a_i(kT)^i$(Table 2.2). So we systematically observe the characteristics of this y_{dk}, then the corresponding stochastic series $y_{sk} = a_i(kT)^i + w_k$. Then we consider the ACF of the deterministic polynomial y_{pk} and infer the effects of these individual components on the ACF of the actual acquired signal y_k.

The nature of l_{zc}, of $y_{dk} = a_i(kT)^i \ \forall k \in [0,N]$, $N = 2400$ for $i = 1,\ 4,\ 7$ is shown in Figure 2.1. It is to be noted that the ACF of y_{dk} is independent of the coefficient a_i, as will also be established in 2.2.1 and 2.2.2. Figure 2.1, it is observed that the l_{zc} values of y_{dk}, marked as red circles, decrease as the order i of the signal increases. Specifically, the l_{zc}s for $i = 1, 4, 7$ are 878, 780 and 675 respectively. It is further observed that in this case, the ACF crosses 0 once and then ends at 0 at the final time instant N. There are no

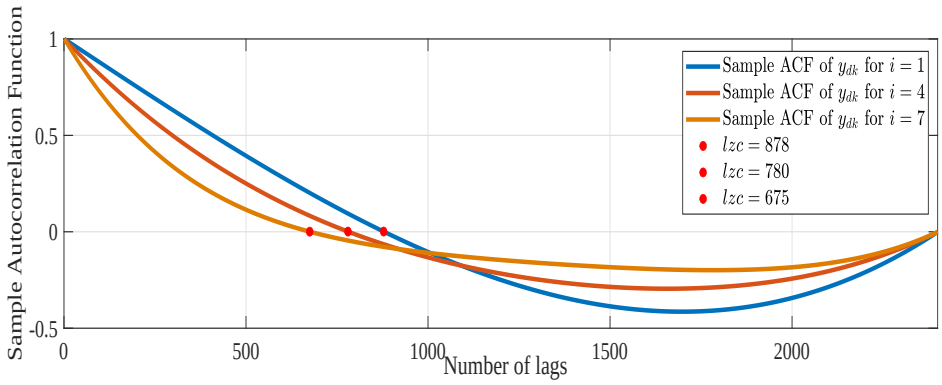


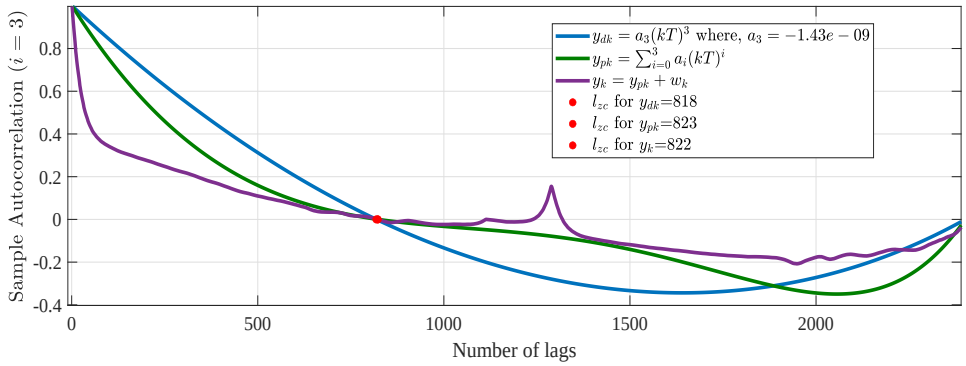
Figure 2.1: Sample ACF of $y_{dk} = a_i(kT)^i$ for any value of a_i where, $i = 1, 4, 7$ $\forall k \in [0, N],\ N = 2400$

other zero crossings. Hence, it may be said that the order of the process y_{dk} can be determined from the l_{zc} of the corresponding sample ACF.

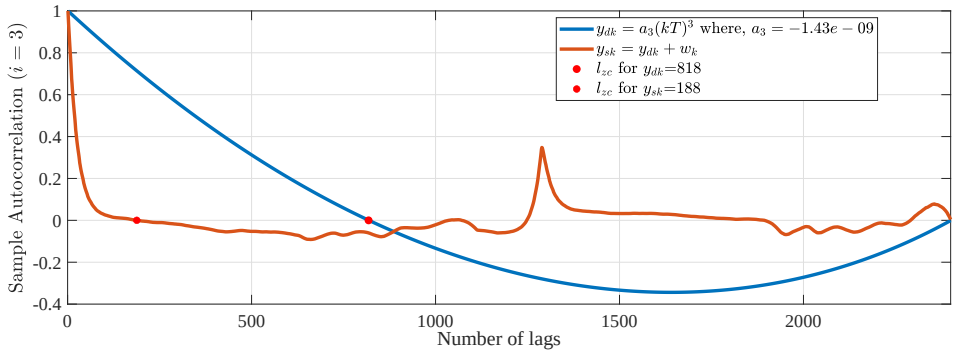
An acquired DDP signal, y_k, is chosen as a test case and its sample ACF (violet colour) is plotted in Figure 2.2a. It has been determined using the modelling algorithm developed in Section 2.2.3 that the best fit polynomial for this signal is of order 3. The stochastic component w_k, practically taken to be the residual r_k $(= y_k - y_{pk})$, has power -27.16dB. It is to be noted that w_k is used interchangeably with r_k henceforth. The ACFs for the corresponding fitted power series $y_{dk} = a_i(kT)^i$ (blue colour) and the fitted polynomial $y_{pk} = \sum_{i=0}^{3} a_i(kT)^i$ (green colour) are shown in Figure 2.2a. To get a preliminary idea of the effect of noise on the ACF characteristics of y_{dk}, the corresponding fitted stochastic power series $y_{sk} = y_{dk} + w_k$ $\forall k \in [0, N]$ (as given in Table 2.2) obtained by adding the residual r_k to y_{dk}, is shown in Figure 2.2b in red colour. In order to appreciate these ACF characteristics, the corresponding original signals for y_k, y_{dk}, y_{sk} and y_{pk} are shown in Figure 2.3a while the residual series r_k is shown in Figure 2.3b.

As shown in Figure 2.2a, when additional lower order terms pertaining to a polynomial regression fit are added to the power law series y_{dk}, the characteristics of the sample ACF may remain the same but the ZCL value changes as expected from a comparison of y_{dk} and the corresponding y_{pk}. However, in this case, merely adding w_k to y_{dk} causes a large change in its ACF characteristics and the l_{zc} as evident from the sample ACF of the stochastic power series $y_{sk}(= y_{dk} + w_k)$ in Figure 2.2b.

Based on these observations, a method to model a signal is proposed henceforth in this Chapter. For this, the first task is to ascertain the order



(a)



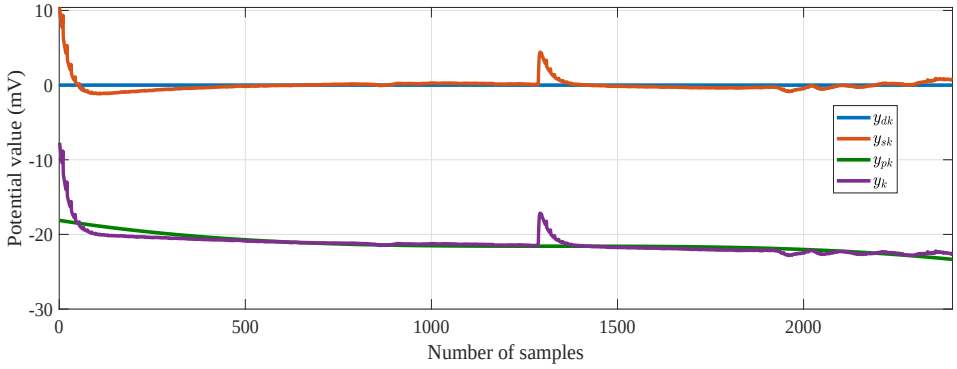
(b)

Figure 2.2: Sample ACFs of (a) $y_{dk} = a_i(kT)^i, i = 3, a_3 = -1.43e-09$, $y_{pk} = \sum_{i=0}^{3} a_i(kT)^i$ and $y_k = y_{pk} + w_k$, power of $w_k = -27.16dB$ (b) y_{dk} and $y_{sk} = y_{dk} + w_k \quad \forall k \in [0, N], \; N = 2400$

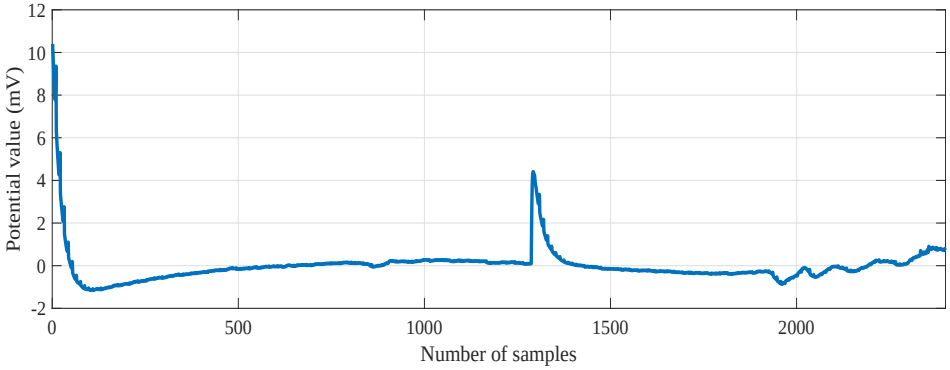
of the deterministic component of the time-series from the value of its ZCL.

2.2.1 ZCL of general time-series

In this subsection, the expressions to obtain the l_{zc} of 3 types of time-series:
a) the stochastic power law series y_{sk}, which is considered mathematically as y_{dk} corrupted by an additive zero mean, finite variance Gaussian noise w_k,
b) the stochastic 1^{st} order polyfit y_{k1} and
c) the deterministic polynomial series y_{pk}
have been derived systematically from first principles in theorem format. The expressions for the l_{zc} of the deterministic power law series y_{dk} and the deterministic 1^{st} order polyfit y_{pk1} are stated as corollaries.
The noise w_k is not assumed to be totally uncorrelated or white, but is merely



(a)



(b)

Figure 2.3: Time-series plots of (a) y_{dk}, y_{sk}, y_{pk}, y_k and (b) residual r_k for the sample ACFs shown in Figure 2.2

assumed to be zero mean and uncorrelated with the tonic component y_{dk} (or y_{pk}) in order to mimic the phasic component of the real time-correlated DDP signal.

As mentioned earlier, the notations that are commonly used in this work have been listed at the outset in Table 2.1 while the mathematical forms of y_{dk}, y_{sk}, y_{pk}, y_{pk1}, y_k, y_{k1} as well as some additional expressions used for the derivations are stated in Table 2.2.

The expressions for l_{zc} of the sample ACF for y_{sk}, y_{k1} and y_{pk} have been stated here as Theorems I, II and III respectively. It is to be noted that y_{dk} and y_{pk1} can be treated as special cases of y_{sk} and y_{k1} respectively. So the closed form expressions for l_{zc} of y_{dk} and y_{pk1} are provided as Corollaries 1 and 2 of Theorems I and II respectively.

It is to be noted that since the noise w_k is assumed to be zero-mean, hence $E[w_k] = 0$. So the mean $\overline{y}$ of y_{dk} and y_{sk} are identical. The expression

for $\overline{y} = \overline{y_{sk}} = \overline{y_{dk}}$ is stated in Table 2.2.

Theorem I:
Considering an additive zero mean Gaussian noise w_k that is uncorrelated with the deterministic component and with finite variance $X_{k,l}$ as defined in Table 2.2, the l_{zc} of the sample ACF of the stochastic power law series y_{sk} is obtained as the solution of

$$\left[\left(\frac{S_i}{N}\right)^2 - \frac{S_i \sum_{k=1}^{N-l_{zc}} k^i}{N(N-l_{zc})} + \frac{\sum_{k=1}^{N-l_{zc}} k^i (k+l_{zc})^i}{(N-l_{zc})}\right] - \left[\frac{S_i \sum_{k=1}^{N-l_{zc}} (k+l_{zc})^i}{N(N-l_{zc})} + \frac{N}{N-l_{zc}} \frac{X_{k,l}}{(a_i T^i)^2}\right] = 0 \tag{2.1}$$

Proof: Let $l \leq N$ be the lag between k and $(k+l)$ samples of stochastic power law series y_{sk} with total number of samples N. Then, the sample ACF of a general biased stochastic series $\hat{R}_x(l\Delta k)$ is [27]

$$\hat{R}_x(l\Delta k) = \frac{\frac{1}{N}\sum_{k=1}^{N-l}(y_{sk} - \overline{y})(y_{sk+l} - \overline{y})}{\frac{1}{N}\sum_{k=1}^{N}(y_{sk} - \overline{y})^2} \tag{2.2}$$

Since the denominator of (2.2) is a square term, hence it is always positive. So, only the numerator of (2.2) is set to 0 in order to solve for the l_{zc} value. Using definitions of y_{sk} and $\overline{y} = \overline{y_{sk}}$ in Table 2.2 and thereafter substituting the relation $(k+l)^i = k^i + S_{kl}$ obtained from the expression for S_{kl} in Table 2.2 in the numerator of (2.2), we obtain

$$\begin{aligned}
&\frac{1}{N}\sum_{k=1}^{N-l}(y_{sk} - \overline{y})(y_{sk+l} - \overline{y}) \\
&= \frac{1}{N}\sum_{k=1}^{N-l}\left[(a_i(kT)^i + w_k - \overline{y_{sk}})(a_i(k+l)^i T^i + w_{k+l} - \overline{y_{sk}})\right] \\
&= \frac{1}{N}\sum_{k=1}^{N-l}\left[(a_i(kT)^i - \overline{y_{sk}} + w_k)(a_i T^i(k^i + S_{kl}) + w_{k+l} - \overline{y_{sk}})\right] \\
&= \frac{1}{N}\sum_{k=1}^{N-l}\left[(a_i(kT)^i - \overline{y_{sk}} + w_k)(a_i(kT)^i - \overline{y_{sk}} + a_i T^i S_{kl} + w_{k+l})\right]
\end{aligned} \tag{2.3}$$

Multiplying the terms in (2.3) and using the relation $X_{k,l}$ as defined in Table

2.2 yields

$$\begin{aligned}
&\frac{1}{N}\sum_{k=1}^{N-l}(y_{sk}-\overline{y})(y_{sk+l}-\overline{y}) \\
&=\frac{1}{N}\sum_{k=1}^{N-l}[(a_i(kT)^i-\overline{y_{sk}})^2+(a_i(kT)^i-\overline{y_{sk}})a_iT^iS_{kl}] \\
&+\frac{1}{N}\sum_{k=1}^{N-l}[(a_i(kT)^i-\overline{y_{sk}})(w_k+w_{k+l})+w_ka_iT^iS_{kl}]+X_{k,l} \qquad (2.4)
\end{aligned}$$

Let us define a term A in which the relation $\overline{y_{sk}}=\frac{a_i}{N}T^iS_i$ from Table 2.2 is substituted to yield

$$A \doteq (a_i(kT)^i-\overline{y_{sk}})=(a_i(kT)^i-\frac{a_i}{N}T^iS_i)=(a_iT^i)(k^i-\frac{S_i}{N}) \qquad (2.5)$$

Substituting for A in (2.7) and using the fact $\frac{1}{N}\sum_{k=1}^{N-l}[(a_i(kT)^i-\overline{y_{sk}})(w_k+w_{k+l})+w_ka_iT^iS_{kl}]=0$ since the noise w_k is assumed to be uncorrelated with y_{dk} yields

$$\begin{aligned}
&\frac{1}{N}\sum_{k=1}^{N-l}(y_{sk}-\overline{y})(y_{sk+l}-\overline{y}) \\
&=\frac{1}{N}\sum_{k=1}^{N-l}[(a_i(kT)^i-\overline{y_{sk}})^2+(a_i(kT)^i-\overline{y_{sk}})a_iT^iS_{kl}]+X_{k,l} \qquad (2.6)\\
&=\frac{1}{N}\sum_{k=1}^{N-l}[A^2+(a_iT^i)S_{kl}A]+X_{k,l} \qquad (2.7)\\
&=\frac{1}{N}\sum_{k=1}^{N-l}[(a_iT^i)^2(k^i-\frac{S_i}{N})^2+(a_iT^i)S_{kl}(a_iT^i)(k^i-\frac{S_i}{N})]+X_{k,l} \\
&=\frac{1}{N}\sum_{k=1}^{N-l}(a_iT^i)^2[(k^i-\frac{S_i}{N})^2+S_{kl}(k^i-\frac{S_i}{N})]+X_{k,l}=0 \qquad (2.8)
\end{aligned}$$

The LHS of (2.8) is next divided by $(a_iT^i)^2/N$. Then using the expression $S_{kl}=(k+l)^i-k^i$ as in Table 2.2, it is expanded term by term to obtain

$$\sum_{k=1}^{N-l}\left[\left(k^i-\frac{S_i}{N}\right)^2+((k+l)^i-k^i)\left(k^i-\frac{S_i}{N}\right)\right]+\frac{N}{(a_iT^i)^2}X_{k,l} \qquad (2.9)$$

$$
\begin{aligned}
&= \sum_{k=1}^{N-l}(k^i)^2 + \sum_{k=1}^{N-l}\left(\frac{S_i}{N}\right)^2 - \sum_{k=1}^{N-l}2k^i\frac{S_i}{N} + \sum_{k=1}^{N-l}(k)^i(k+l)^i - \sum_{k=1}^{N-l}(k^i)^2 \\
&- \sum_{k=1}^{N-l}\frac{S_i}{N}(k+l)^i + \sum_{k=1}^{N-l}\frac{S_i}{N}k^i + \frac{N}{(a_iT^i)^2}X_{k,l} \\
&= \sum_{k=1}^{N-l}\left(\frac{S_i}{N}\right)^2 - \sum_{k=1}^{N-l}k^i\frac{S_i}{N} + \sum_{k=1}^{N-l}(k)^i(k+l)^i \\
&- \sum_{k=1}^{N-l}\frac{S_i}{N}(k+l)^i + \frac{N}{(a_iT^i)^2}X_{k,l} = 0
\end{aligned} \tag{2.10}
$$

Further simplifying (2.10) and taking the term $(N-l)$ common yields

$$
\begin{aligned}
(N-l)\left[\left(\frac{S_i}{N}\right)^2 - \frac{S_i\sum_{k=1}^{N-l}k^i}{N(N-l)} + \frac{\sum_{k=1}^{N-l}k^i(k+l)^i}{(N-l)}\right] - \\
(N-l)\left[\frac{S_i\sum_{k=1}^{N-l}(k+l)^i}{N(N-l)} - \frac{N}{N-l}\frac{X_{k,l}}{(a_iT^i)^2}\right] = 0
\end{aligned} \tag{2.11}
$$

$l = N$ is the end point and hence the l_{zc}, which is the 1^{st} zero crossing lag, is obtained as the closed form solution of setting the term within the square bracket to zero. This is identical to (2.1).

Corollary 1:
The l_{zc} of the deterministic power law series y_{dk} is obtained as the solution of

$$
\left[\left(\frac{S_i}{N}\right)^2 - \frac{S_i\sum_{k=1}^{N-l_{zc}}k^i}{N(N-l_{zc})} + \frac{\sum_{k=1}^{N-l_{zc}}k^i(k+l_{zc})^i}{(N-l_{zc})} - \frac{S_i\sum_{k=1}^{N-l_{zc}}(k+l_{zc})^i}{N(N-l_{zc})}\right] = 0 \tag{2.12}
$$

by substituting $X_{k,l} = 0$ in equation (2.1).

Theorem II:
Considering an additive zero mean Gaussian noise w_k that is uncorrelated with the deterministic component and with finite covariance $X_{k,l}$ as defined in Table 2.2, $X_1 = \frac{24X_{k,l}}{(N-l_{zc})(a_1T)^2}$ and number of samples N, the l_{zc} of the 1^{st} order stochastic ACF-ZCL model $y_{k1} = a_1kT + a_0 + w_k$ is obtained as

$$
l_{zc} = \frac{-N + \sqrt{3N^2 - 2 + X_1}}{2} \tag{2.13}
$$

Proof: Let $l \leq N$ be the lag between k and $(k+l)$ samples of the 1^{st} order stochastic ACF-ZCL model y_{k1} with total number of samples N. Then, equation (2.2) can be rewritten for the sample ACF of the first order stochastic model y_{k1} as

$$\hat{R}_x(l\Delta k) = \frac{\frac{1}{N}\sum_{k=1}^{N-l}(y_{k1}-\overline{y_{k1}})(y_{k1+l}-\overline{y_{k1}})}{\frac{1}{N}\sum_{k=1}^{N}(y_{k1}-\overline{y_{k1}})^2} \tag{2.14}$$

To identify the l_{zc} of y_{k1}, the numerator of 2.14 is set to 0 as in Theorem I yielding

$$\begin{aligned}&\frac{1}{N}\sum_{k=1}^{N-l}(y_{k1}-\overline{y_{k1}})(y_{k1+l}-\overline{y_{k1}}) = 0\\ &=\sum_{k=1}^{N-l}\left((a_1kT+a_0+w_k)-\overline{y_{k1}}\right)\left(a_1(k+l)T+a_0+w_{k+l}-\overline{y_{k1}}\right)\end{aligned} \tag{2.15}$$

The noise is assumed to be zero-mean, hence $E[w_k]=0$. So, $\overline{y_{k1}}$ is obtained (as defined in Table 2.2) as

$$\overline{y_{k1}} = \frac{\sum_{k=1}^{N}(a_1kT+a_0+w_k)}{N} = \frac{a_1T(N+1)}{2}+a_0+E[w_k] = \frac{a_1T(N+1)}{2}+a_0 \tag{2.16}$$

Substituting the value of $\overline{y_{k1}}$ in (2.15) and simplifying yields

$$\begin{aligned}&\sum_{k=1}^{N-l}\left(a_1kT+w_k-\frac{a_1T(N+1)}{2}\right)\left(a_1(k+l)T+w_{k+l}-\frac{a_1T(N+1)}{2}\right)\\ &=[\sum_{k=1}^{N-l}(\frac{a_1T(2k-N-1)}{2})^2+a_1lT(\frac{a_1T(2k-N-1)}{2})]+X_{k,l}\\ &=[\sum_{k=1}^{N-l}\frac{a_1^2}{4}T^2((2k-N-1)^2+2l(2k-N-1))]+X_{k,l}=0\end{aligned} \tag{2.17}$$

It is to be noted that $l=l_{zc}$ for equation (2.17) to hold. Thus, using l_{zc} and simplifying equation (2.17) leads to

$$[\sum_{k=1}^{N-l_{zc}}\frac{a_1^2}{4}\{(4k^2+(N+1)^2-4k(N+1))+2l_{zc}(2k-(N+1))\}]+\frac{X_{k,l}}{T^2}=0 \tag{2.18}$$

On expanding the relevant terms within the square bracket for $k = 1$ to $(N - l_{zc})$, equation (2.18) can be written as

$$\frac{a_1^2}{4}[\frac{2(N-l_{zc})(N-l_{zc}+1)(2N-2l_{zc}+1)}{3} + (N-l_{zc})(N+1)^2 - 2(N-l_{zc})(N-l_{zc}+1)(N+1) + 2l_{zc}(N-l_{zc})(N-l_{zc}+1) - 2l_{zc}(N+1)(N-l_{zc})] + \frac{X_{k,l}}{T^2} = 0 \quad (2.19)$$

Dividing equation (2.19) all over by $(N - l_{zc})$ and then simplifying yields

$$(\frac{a_1^2}{4})(\frac{N^2 - 2l_{zc}N - 2l_{zc}^2 - 1}{3}) + \frac{X_{k,l}}{(N-l_{zc})T^2} = 0$$

$$\text{So, } (a_1^2/12)(2l_{zc}^2 + 2l_{zc}N - N^2 + 1) - \frac{X_{k,l}}{(N-l_{zc})T^2} = 0 \quad (2.20)$$

Dividing equation (2.20) by $(a_1^2/6)$ and using $X_1 = \frac{24X_{k,l}}{(N-l_{zc})(a_1T)^2}$, a quadratic equation for l_{zc} is obtained as follows.

$$l_{zc}^2 + l_{zc}N - (\frac{N^2-1}{2} + \frac{X_1}{4}) = 0 \quad (2.21)$$

Solving this yields the closed form solution of l_{zc} as

$$l_{zc} = \frac{-N \pm \sqrt{3N^2 - 2 + X_1}}{2} \quad (2.22)$$

Since l_{zc} cannot be negative, so the solution which yields a practically feasible value of l_{zc} is obtained as equation (2.13) provided the term $\sqrt{3N^2 - 2 + X_1} > N$.

Corollary 2:
The closed form solution of l_{zc} of the 1^{st} order polyfit $y_{pk1} = a_1kT + a_0$ is

$$l_{zc} = \frac{-N + \sqrt{3N^2 - 2}}{2} \quad (2.23)$$

Theorem III:
Considering S_n and S_j as defined in Table 2.2 and defining $A1 = (a_n(kT)^n - \frac{a_n}{N}T^nS_n)$, $C1 = \sum_{i=0}^{n-1} a_i(kT)^i - \frac{S_j}{N}$, the l_{zc} of the sample ACF of the deterministic polynomial series y_{pk} is obtained as the solution of

$$\left[\frac{1}{N}\sum_{k=1}^{N-l_{zc}}[A1^2 + a_nT^nS_{kl}A1] + \frac{1}{N}\sum_{k=1}^{N-l_{zc}}[2A1C1 + A1\sum_{i=0}^{n-1}a_iT^iS_{kl}]\right] + \left[\frac{1}{N}\sum_{k=1}^{N-l_{zc}}[C1^2 + C1\sum_{i=0}^{n}a_iT^iS_{kl}]\right] = 0 \quad (2.24)$$

Proof:
The sample ACF $\hat{R}_x(l\Delta k)$ of y_{pk} can be written as [27]

$$\hat{R}_x(l\Delta k) = \frac{\frac{1}{N}\sum_{k=1}^{N-l}(y_{pk}-\overline{y_{pk}})(y_{pk+l}-\overline{y_{pk}})}{\frac{1}{N}\sum_{k=1}^{N}(y_{pk}-\overline{y_{pk}})^2} \tag{2.25}$$

As in Theorem I, in this case also the numerator of $\hat{R}_x(l\Delta k)$ is set to 0 for obtaining the l_{zc}. Accordingly, using the expressions for y_{pk} and S_{kl} from Table 2.2, we have

$$\begin{aligned}
&\frac{1}{N}\sum_{k=1}^{N-l}(y_{pk}-\overline{y_{pk}})(y_{pk+l}-\overline{y_{pk}}) \\
&= \frac{1}{N}\sum_{k=1}^{N-l}\left[\left(\sum_{i=0}^{n}a_i(kT)^i-\overline{y_{pk}}\right)\left(\sum_{i=0}^{n}a_i(k+l)^iT^i-\overline{y_{pk}}\right)\right] \\
&= \frac{1}{N}\sum_{k=1}^{N-l}\left[\left(a_n(kT)^n+\sum_{i=0}^{n-1}a_i(kT)^i-\overline{y_{pk}}\right)\left(\sum_{i=0}^{n}(a_iT^ik^i+a_iT^iS_{kl})-\overline{y_{pk}}\right)\right] \\
&= 0
\end{aligned} \tag{2.26}$$

Using the terms S_n, S_j and the expression $\overline{y_{pk}} = \frac{a_nT^n}{N}S_n + \frac{S_j}{N}$ from Table 2.2 in (2.26) yields

$$\begin{aligned}
&= \left[\frac{1}{N}\sum_{k=1}^{N-l}\left(a_n(kT)^n+\sum_{i=0}^{n-1}a_i(kT)^i-\frac{a_nT^n}{N}S_n-\frac{S_j}{N}\right)\right] \\
&\quad \frac{1}{N}\sum_{k=1}^{N-l}\left[\left(\sum_{i=0}^{n}(a_iT^ik^i+a_iT^iS_{kl})-\frac{a_nT^n}{N}S_n-\frac{S_j}{N}\right)\right]
\end{aligned} \tag{2.27}$$

Let the terms dependent on model order n and the remaining terms respectively be denoted as

$$A1 \doteq (a_n(kT)^n - \frac{a_n}{N}T^nS_n),\ C1 \doteq \sum_{i=0}^{n-1}a_i(kT)^i - \frac{S_j}{N} \tag{2.28}$$

Then (2.27), and hence (2.26), can be expressed as

$$\frac{1}{N}\sum_{k=1}^{N-l}\left[(A1+C1)(A1+a_nT^nS_{kl}+C1+\sum_{i=0}^{n-1}a_iT^iS_{kl})\right] = 0 \tag{2.29}$$

The condition for l_{zc} as in (2.24) is obtained by expanding (2.29).

2.2.2 Key Observations from Theorems

In several cases, the ACF of real time-series signals cross the zero axis multiple times due to interactions between the signal and noise coefficients before finally ending at zero. So, the expressions obtained in the aforementioned theoretical derivations for the l_{zc} of y_{dk}, y_{sk}, y_{pk} can be used for determining the model order n.

The key observations from the three theorems and the two corollaries stated in subsection 2.2.1 are described as follows:

1. The n^{th} order stochastic ACF-ZCL model of the actual time-series signal, as well as the signal itself, is represented as y_k. It is composed of the deterministic and noise components of y_{sk} (at $i = n$) along with the 0 to $(n-1)$ order components of y_{pk}. Hence it is to be expected that the l_{zc} of y_k will be affected by the noise interactions as in y_{sk} as well as the coefficient interactions as in y_{pk}.

2. The sample ACF is zero at the final instant for all cases, as is to be expected for a finite time series data.

3. The l_{zc} of the deterministic 1^{st} order polynomial model y_{pk1} is uniquely obtained and does not depend upon the coefficients a_0 or a_1 or the sampling time T but is dependent only on the number of observations N as seen in Corollary 2.

4. The l_{zc} of the general i^{th} order power law series y_{dk} of order i is uniquely obtained and is decided by i and N, but is independent of a_i or T as seen in Corollary 1.

5. The l_{zc} of the stochastic 1st order model y_{k1} is uniquely obtained provided the particular noise sequence w_k and hence X_1 is known. Furthermore, it is independent of a_0 but varies with a_1, T, and N as seen in Theorem II.

6. The l_{zc} of the stochastic series y_{sk} or the deterministic polynomial y_{pk} or the stochastic ACF-ZCL model y_k cannot be determined uniquely. Furthermore, the expressions for the l_{zc} in these cases depend on all the terms i, N, a_i and T as seen in Theorems I and III and Observation 1.

7. The first two terms in the expression in (2.24) for y_{pk} and in (2.7) for y_{sk} are identical. This is so since the underlying terms A in (2.5) and $A1$ in (2.28) are identical, being terms in the ACF of the common

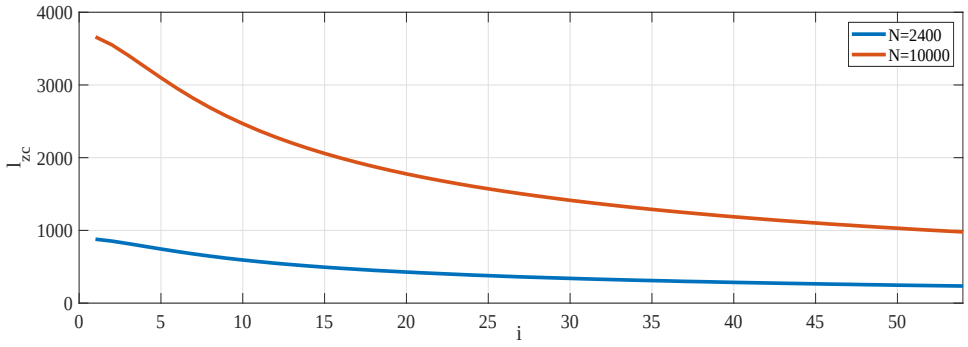


Figure 2.4: Zero crossing lag l_{zc} vs. order i of deterministic power law series y_{dk} for N=2400 and N=10000

Table 2.3: Order and ZCL (l_{zc}) of the series y_{dk} and y_{sk}, y_{pk} or y_k for N=2400

Order	l_{zc} of y_{dk}	l_{zc} of y_{sk}, y_{pk}, y_k	Order	1^{st} l_{zc} of y_{dk}	1^{st} l_{zc} of y_{sk}, y_{pk}, y_k
1	878	$[865, \infty)$	6	708	$[691, 725)$
2	852	$[834, 865)$	7	675	$[660, 691)$
3	817	$[798, 834)$	8	645	$[631, 660)$
4	780	$[761, 798)$	9	617	≤ 630
5	743	$[725, 761)$	36	305	

deterministic component y_{dk}. It is to be noted that i in y_{sk} is the same as n in y_{pk}. So the nature of l_{zc} of y_{sk} and y_{pk} can be expected to be similar since these two terms are augmented by interaction terms in both cases.

8. The l_{zc} of y_{sk} will be higher or lower than that of the corresponding y_{dk} depending on the contributions of the net noise correlation $X_{k,l}$, as shown in (2.7).
   In case of the 1^{st} order models, this is evident from the effect of the $X_{k,l}$ dependent term X_1 on the value of the l_{zc} in (2.13). The l_{zc} shifts towards higher lag values than that for the corresponding y_{pk1} if X_1 is positive. Else, if X_1 is negative but is such that $(3N^2 - 2 + X_1) \geq N$, then l_{zc} shifts towards lower lag values. The other solutions are unrealistic yielding negative or complex solutions for the l_{zc}.
   A similar interpretation exists for the l_{zc} of y_{pk} and the effect of net coefficient interactions. In this scenario, the net coefficient interactions obtained from the 3rd to 6th terms in (2.24) play an analogous role as

that of $X_{k,l}$ in (2.7).

9. Taking both types of interactions into account, it is necessary to judiciously ascribe a range of the l_{zc} values to indicate a particular system order n since this crucially affects the efficacy of the system model.
In case of y_{sk}, y_{pk} or y_k, the range of l_{zc} values are considered to vary from the mid-value of the l_{zc} of y_{dk} for orders $(i-1)$ and i to that of orders i and $(i+1)$.
Table 2.3 lists the deterministic l_{zc} of y_{dk} for $i = [1, 9], 36$ and the corresponding range of l_{zc} considered for y_{sk}, y_{pk} and/or y_k for $i = [1, 9]$.

10. Figure 2.4 shows a plot of l_{zc} vs. i for the series y_{dk} for $i = [1, 54]$ for $N = 2400$ and $N = 10000$. It is observed that the plots asymptotically reach a fixed value in both cases. Furthermore, in both cases, as i increases beyond 9, the change in the values of l_{zc} drops significantly and beyond $i = 36$, the l_{zc} values are almost identical.

11. It is observed from Figure 2.4 and Table 2.3 that the values of l_{zc} drop to less than $l_{zc}/2$ as the series order increases from $n = 9$ to $n = 36$. The values of l_{zc} for $n = 9$ are 617 and 2572 while those for $n = 36$ are 305 and 1267 for $N = 2400$ and $N = 10000$ respectively. Hence, any time-series with an l_{zc} less than half of that at $n = 9$ will have an order $n = 36$ or more. It is known that for real time-series signals, the ACF of residuals reveal further structure, or additional dynamics, in the system and lower ACF lags indicate lower noise variance [177, 178]. Combining these, it can be said that a large separation in the orders (or l_{zc}) of y_{pk} and w_k (or r_k) is related to a large SNR.

2.2.3 Algorithm

Based on the observations, an algorithm has been developed that uses the l_{zc} of the actual time-series signal y_k for identifying the primary system order n. n is then updated using the ZCL of the residuals r_k, till a sufficiently large SNR is assured. The updated n is finally used to determine the standard best fit polynomial y_{pk} and the residual series r_k of the stochastic ACF-ZCL model y_k. The flowchart for developing the mathematical model is provided in Figure 2.5. The steps of the algorithm are stated hereafter.

Step 1: The sample ACF of the acquired time-series signal y_k $\forall k \in [0, N]$ is calculated using any available standard technique. In this study, $N = 2400$ in most cases.

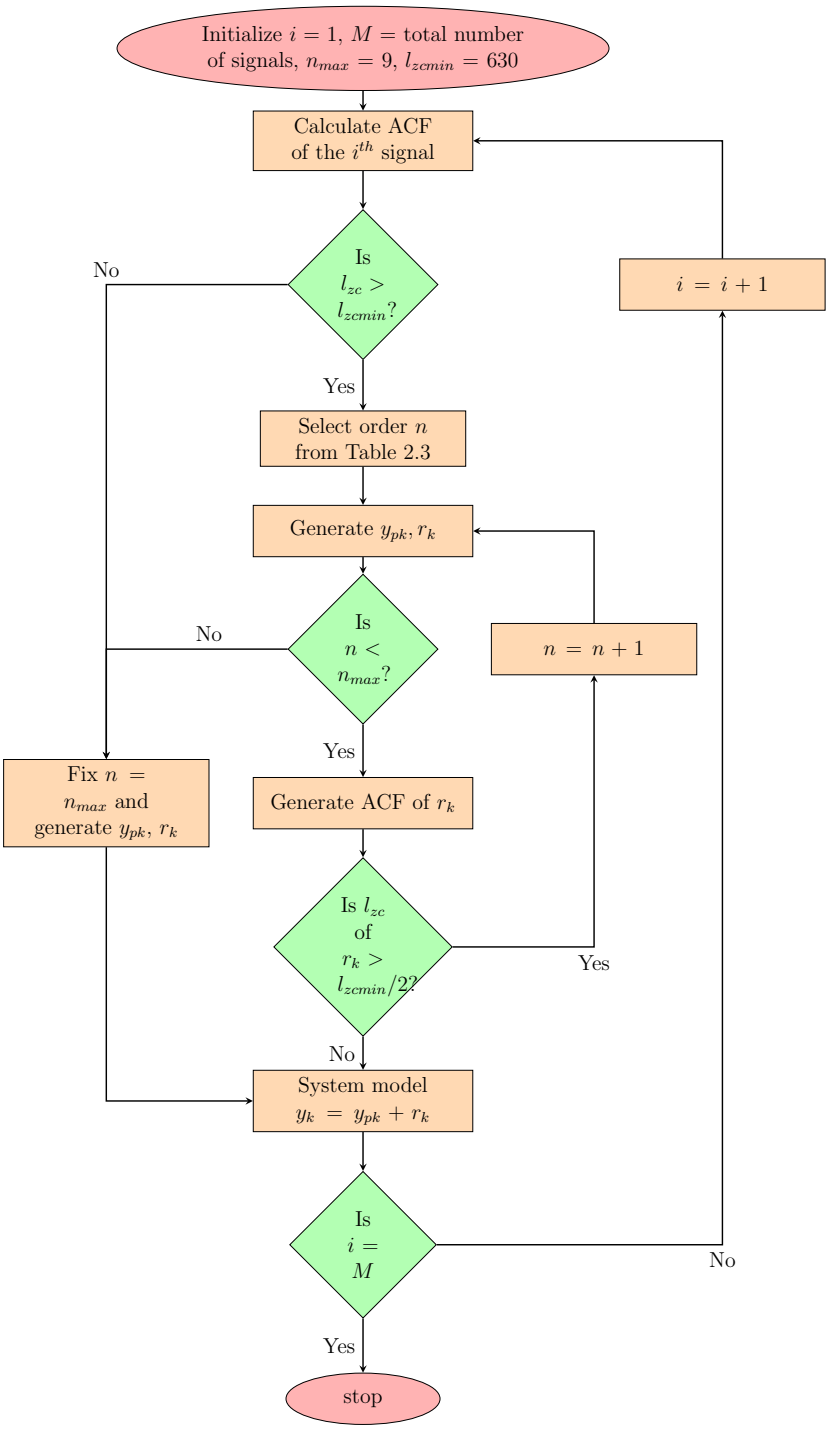


Figure 2.5: Flowchart for determining stochastic ACF-ZCL model

Step 2: The l_{zc} value is identified from the sample ACF and the signal order n is determined using Table 2.3 for cases where $l_{zc} \geq l_{zcmin}$. For signals with $l_{zc} < l_{zcmin}$, n is set as n_{max}. In this study, $n_{max} = 9$ and since $N = 2400$, so $l_{zcmin} = 630$.

Step 3: A best fit polynomial y_{pk} of order n is generated using standard least-squares technique [179] and the corresponding residual series is generated as $r_k = (y_k - y_{pk})\ \forall k \in [0, N]$.

Step 4: If $n < n_{max}$, the ACF of r_k is calculated and its l_{zc} value is determined. The system order and model is adjusted till l_{zc} of r_k is less than $l_{zcmin}/2$, indicating a sufficiently large SNR as discussed in Observation 11 listed above.

Step 5: If the l_{zc} of r_k is larger than $l_{zcmin}/2$, increase model order to $n+1$ and repeat Steps 3 and 4. Then check the condition for the new r_k until its $l_{zc} < l_{zcmin}/2$ or $n = n_{max}$.

Step 6: If the l_{zc} value of r_k is within $l_{zcmin}/2 = 315$ or the model order $n = n_{max}$, retain the system polynomial as y_{pk} with order n.

Thus, $y_k = y_{pk} + r_k\ \forall k \in [0, N]$ is the stochastic ACF-ZCL model of the time series signal y_k, where y_{pk} and r_k represent its deterministic trend or tonic and stochastic residual or phasic components.

2.3 Modelling and analysis of DDP signals

Table 2.4: Mean correlation matrix of all 4 acquired signals

Signals	LH	RH	LL	RL
LH	1	0.25	-0.01	0.02
RH	0.25	1	-0.08	-0.08
LL	-0.01	-0.08	1	-0.06
RL	0.02	-0.08	-0.06	1

In the DS4 experiment stated in Section 1.1.6, total 68 sets of 4-channel DDP signals are available. Each set of signals comprises of simultaneously recorded 10-minute long LH, RH, LL and RL signals at a sampling time of $T = 50$ ms. The cross-correlation coefficients of the 4 simultaneously acquired LH, RH, LL and RL signals in each data set has been determined and then their mean over all 68 sets of data has been tabulated in Table 2.4. As is to be expected for short duration biosignals, the cross-correlation

values are typically very low, within ±0.1. Only the LH and RH signals show a larger correlation coefficient of 0.25. This value is significant only for long-range processes [175, 180] but since a duration of 10 minutes does not qualify as such, hence this too can be regarded as low.

Thus the results indicate that although these signals have been acquired simultaneously from 4 locations on the body of the same subject, yet they can be treated as independent signals. Accordingly, all acquired signals from all the different experiments have been treated as independent signals in the present study.

2.3.1 Modelling protocol and method

The flow chart of the overall protocol for the data modelling and analysis is shown in Figure 2.6. A detailed description of the steps are given thereafter.

Step 1: Each signal to be modelled is considered to be of 2-minutes duration. Hence, any acquired signal of longer duration is divided into suitable number of non-overlapping subsets of 2-minutes duration each. The modelling and analysis is then done on each of the 2-minute potentials.

Step 2: The sample ACF of the 2-minute DDP signals is calculated and the 1^{st} zero crossing lag (l_{zc}) value is identified from it.

Step 3: Using the algorithm stated in Section 2.2.3, the stochastic ACF-ZCL model y_k of each 2 minute data is determined. In this model, the signal order n is determined from the ZCL of the signal ACF, the corresponding deterministic polynomial component is y_{pk} and the residual r_k is expectedly the corresponding stochastic component.

Step 4: The next stage involves the analysis of the various model components of the 2-minute long signals.

Step 5: Finally, the models of the signals acquired in DS4 experiment for the restfulness assessment application have been studied in detail. In this, the tonic and phasic characteristics of the signal have been inferred from the model parameters for the consecutive 5 non-overlapping 2 minute segments of each 10 minute long signal.

In view of the finding regarding the independence of these signals, the stochastic ACF-ZCL models of the DDP signals that have been acquired by Sarkar [1] from all 4 types of experiments, namely DS1, DS2, DS3 and DS4 as stated in Section 1.1.6, have been determined. For this, each long signal has been subdivided into 2 minutes segments as described above. Hence,

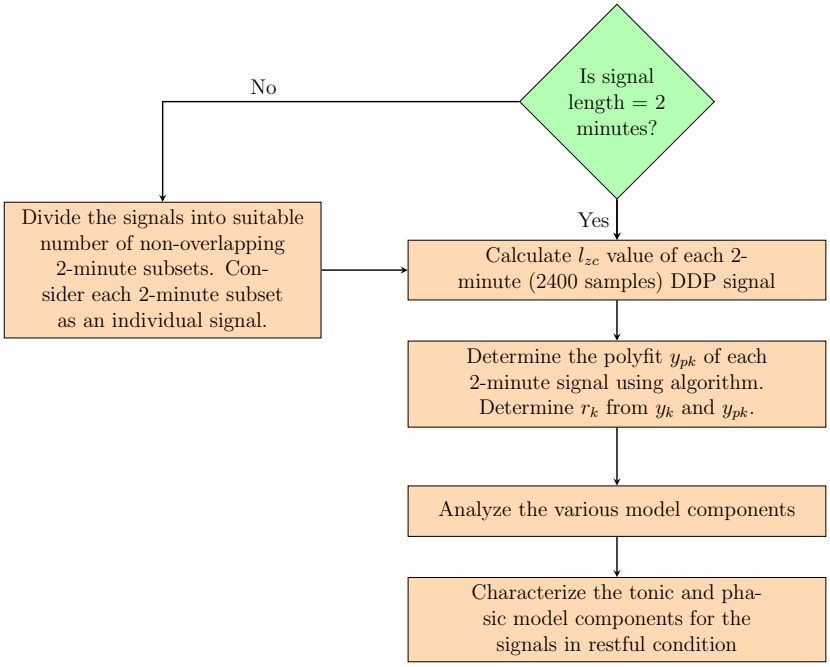


Figure 2.6: Steps of modelling protocol of DDP signals

there are 5706 numbers of 2 minutes long DDP signals with a sampling time of 50 ms, that is 2400 discrete samples.

2.3.2 Fitting order 1 and higher order models

The first stage involves the order determination and modelling of all the 5706 acquired DDP signals. From the ACF of the acquired signals, it is observed that in 902 signals, the l_{zc} of the noisy signal lies within 865 and 1183. Thus these are fitted with the 1^{st} order models $y_{k1} = y_{pk1} + r_{k1}$. It is interesting to note that of these, 76 signals have their l_{zc} equal to 878, which is the same as the l_{zc} for y_{dk} for $n = 1$ as stated in Table 2.3.

The l_{zc} of the remaining 4804 signals are either out of this range or their residuals have higher order dynamics and hence, these signals are fitted with respective n-th order models $y_k = y_{pk} + r_k \ \forall \ n > 1$. However, since all signals seem to exhibit a predominantly linear trend, so these 4804 signals have also been fitted with the 1^{st} order models $y_{k1} = y_{pk1} + r_{k1}$ considering $y_k = y_{k1}$. This is done in order to verify the efficacy of the proposed algorithm in determining the suitable signal order and corresponding model.

In accordance with the stochastic ACF-ZCL model formulation, the residual signals r_k (or r_{k1}) have been generated from the stochastic model y_k (or

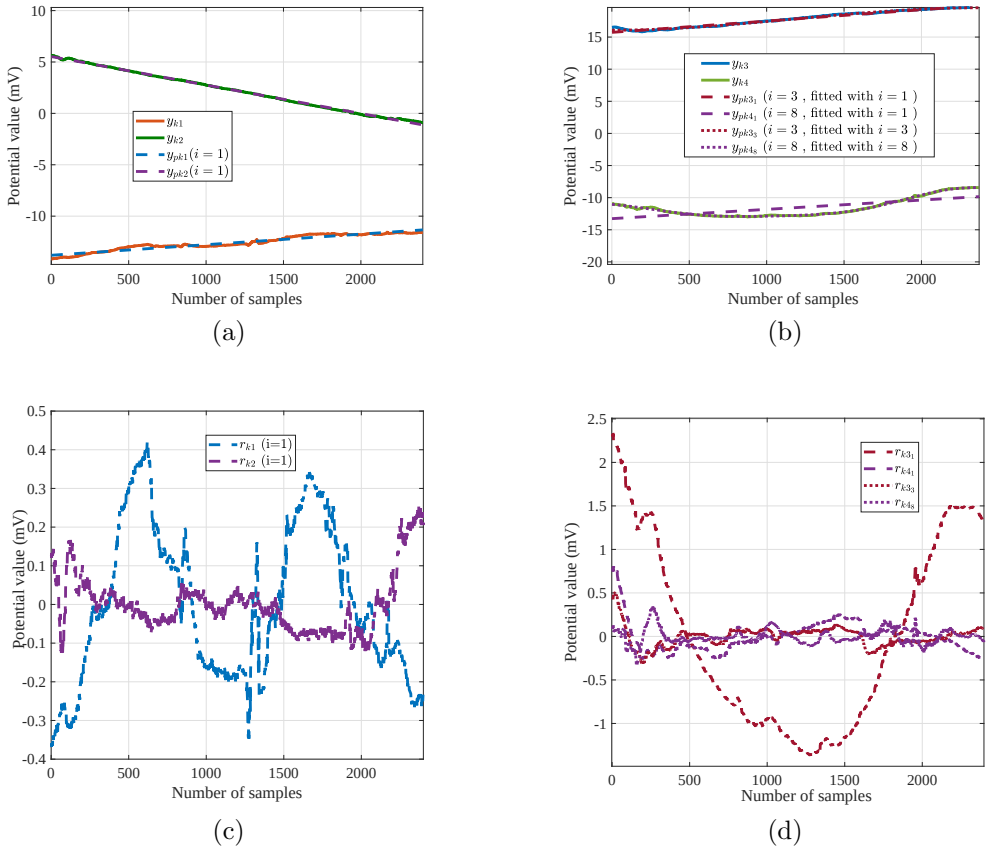
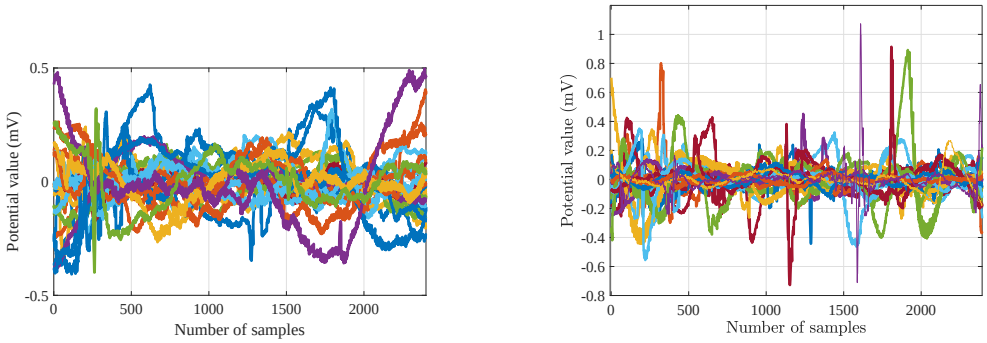


Figure 2.7: Time-series plots of (a) 1st order signals y_{k1}, y_{k2} and their polyfits y_{pk1}, y_{pk2}, (b) higher order signals y_{k3}, y_{k4} and their polyfits y_{pk3_1}, y_{pk4_1}, y_{pk3_3} and y_{pk4_8} (c) residuals r_{k1}, r_{k2} and (d) residualsr_{k3_1}, r_{k4_1}, r_{k3_3} and r_{k4_8}

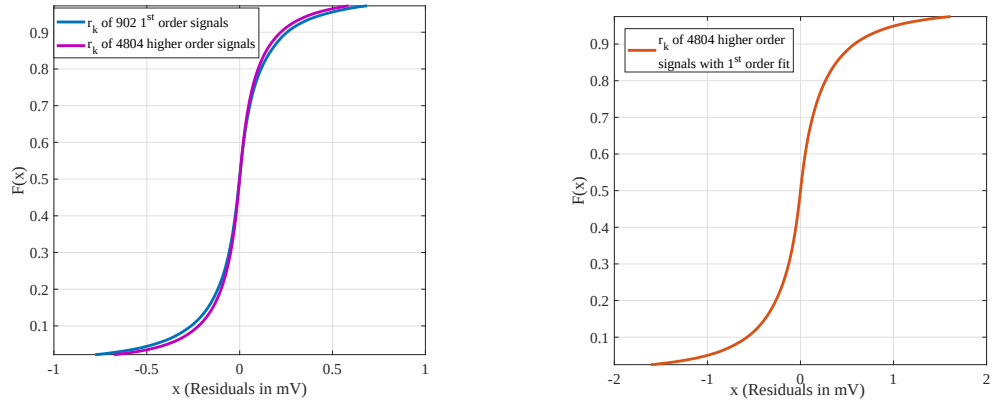
y_{k1}) by subtracting the corresponding polyfit y_{pk} (or the 1^{st} order fit y_{pk1}) from it. The characteristics of these residuals help in evaluating the goodness of fit of the respective model. Figure 2.7a shows the time-series plots of 1^{st} order signals y_{k1} and their respective y_{pk1} while Figure 2.7b shows the time-series plots of the higher order signals y_k and their respective 1^{st} order polyfits y_{pk1} as well as their higher order polyfits y_{pk}. The corresponding residuals are shown in Figures 2.7c and 2.7d respectively.

As expected, it is observed that the residuals of the proper order fits are within ± 0.5 mV range while the residuals r_{k1} that pertain to 1^{st} order models of the higher order signals are much larger (within ± 2.5 mV). Higher order dynamics are also clearly observed in the residuals in these cases indicating the need for fitting higher order y_{pk} models.



(a) (b)

Figure 2.8: Time plot of residuals of 25 sample signals: (a) 1st order signals (b) higher order signals



(a) (b)

Figure 2.9: Cdf plot of 95% of the residuals (about median) of (a) 902 1st order and 4804 higher order signals (b) 4804 higher order signals fitted with 1st order model

The next step is to study the characteristics of the residuals in more detail. Figure 2.8a and Figure 2.8b show the residuals of 25 representative 1^{st} order and higher order signals respectively. It is observed that in all cases the residuals are random in nature with amplitudes typically within ± 0.5 mV with some sudden spikes.

Thereafter, a statistical analysis is performed of the instantaneous residuals of all the signals. The zoomed cdf plots pertaining to 95% (about the median) of the residuals r_{k1} and r_k of the 1^{st} order signals y_{k1} and the higher

Table 2.5: Statistical analysis of 1^{st} order residuals

Parameters	902 1^{st} order signals		4804 higher order signals with 1^{st} order fit		4804 higher order signals	
	Overall	95% about median	Overall	95% about median	Overall	95% about median
Range (mV)	$(-20, 10)$	$(-0.8, 0.8)$	$(-250, 250)$	$(-2, 2)$	$(-223.37, 71.08)$	$(-0.8, 0.8)$
Mean, median (mV)	0.00	0.00	0.00	0.00	0.00	0.00
SD (mV)	0.43	0.20	1.25	0.48	0.58	0.18
Skewness	-1.45	-0.02	10.85	0.01	-22.08	0.02
Kurtosis	98.37	5.81	2.80×10^3	6.23	7.89×10^3	6.67

order signals y_k are shown in Figure 2.9a while that of the residuals r_{k1} of the 1^{st} order models y_{k1} of the higher order signals y_k are shown in Figure 2.9b. The corresponding statistical parameters are tabulated in Table 2.5.

It is observed in Figure 2.9a that 95% of the r_k and r_{k1} lie within ± 0.8 mV, of which more than 91% lie within ± 0.5 mV. It is to be mentioned that the overall spread of the r_k and r_{k1} is however much larger owing to some instantaneous spikes in some signals, as evident from Table 2.5. It is also evident from the Figure and the Table that the residuals are zero mean and 95% of the r_k and r_{k1} are unskewed with standard deviation within 0.2 mV. Further, most of these residuals are clustered very close to 0 with a high kurtosis of more than 5.5, even considering the 95% population.

As evident from Figure 2.9b and the corresponding statistics in Table 2.5, the characteristics of the residuals are much worse in all cases for 1^{st} order fits of higher order signals further validating the efficacy of the proposed algorithm for determining the model order and in fitting a stochastic ACF-ZCL model.

2.3.3 Effect of interactions on signal models

It is mentioned in Observations 7 and 8 in Section 2.2.2 that the l_{zc} of y_{sk} (alternatively, y_{pk}) is affected by the $X_{k,l}$ in (2.7) (alternatively, the polynomial coefficient interactions). The effects of these interactions on the shift in l_{zc} from that in the deterministic power law cases are studied hereafter in terms of the signals and their models.

(A) Effect of noise interactions

Since $y_k = y_{pk} + r_k$, hence y_k differs from its deterministic component y_{pk} only in the noise component r_k. The effect of noise interactions in y_k is thus observed by comparing the ACFs of y_k and its corresponding y_{pk}. The ACF

plots of two representative DDP signals, labeled as signal 1 and signal 2, are shown in Figure 2.10a along with their respective polyfit series y_{pk}. In the zoomed plot shown in Figure 2.10b, the l_{zc} of signal 1 and its polyfit are marked as black dots while those of signal 2 and its polyfit are marked as red dots. Both the signals chosen have 3rd order polynomial fits, but in case of signal 1, the l_{zc} of y_k shifts towards higher lag than that of the corresponding y_{pk}, whereas the reverse happens in case of signal 2. Thus Observation 8 in Section 2.2.2 that the contribution of X_{kl} being positive or negative will cause the l_{zc} to shift to the left or right is validated.

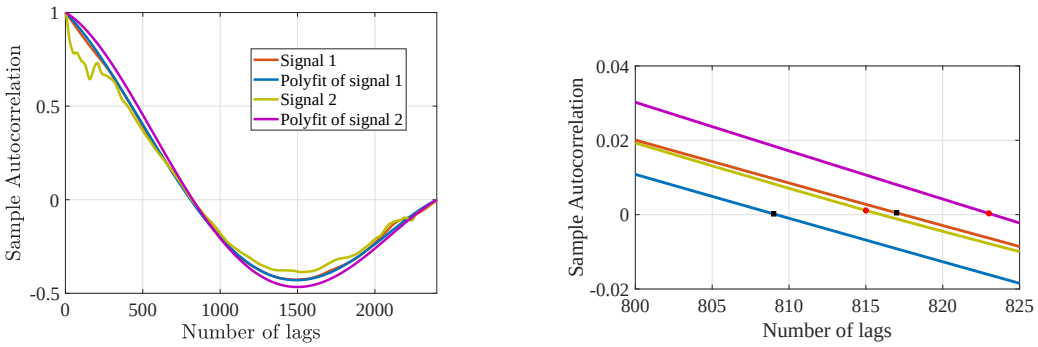


Figure 2.10: (a) ACF with l_{zc} of two representative signals and their 3rd order polyfits (b) zoomed plot near the l_{zc} values

(B) Effect of coefficient interactions

In the next stage, the combined effect of the noise and coefficient interactions on the l_{zc} of y_k is studied. It is observed in Table 2.3 that the l_{zc} of the 3^{rd} order deterministic power law series y_{dk} is 817, while the corresponding range of l_{zc} for y_k, y_{sk} and/or y_{pk} is considered to be $[798, 834)$. Considering all signals y_k identified to have 3rd order polyfit components, it is observed that the l_{zc} of the polyfits y_{pk} lie in the range $[648, 940]$ while the l_{zc} of all the actual signals y_k lie in the range $[798, 1028]$. In order to visualize this, polyfit models y_{pk} of 50 representative 3^{rd} order signals are shown in Figure 2.11.

It is thus evident that the range of l_{zc} for the y_{pk} is smaller than the range of l_{zc} for the actual signals. This variation is necessarily due to the additive effect of noise interactions since $y_k = y_{pk} + r_k$. However, as expected, the lower limit of l_{zc} for all the actual signals y_k is maintained at 798. The increase in the upper limit of l_{zc} from a theoretical value of 834 to a practical value of 1004 is because some signals with l_{zc} in the range of 1^{st} or 2^{nd} order

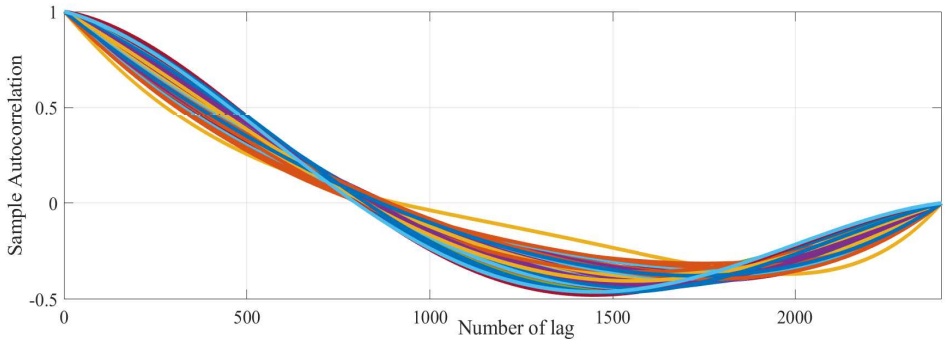


Figure 2.11: ACF plot of 3^{rd} order polyfit signal without noise

Table 2.6: Orderwise coefficients and residual ranges

Order n of signal y_k	Number of signals	Range of coefficient (a_n)	Range of residuals r_k(mV)
1	902	$(-0.01, 0.01)$	$(-0.8, 0.8)$
2	1587	$(-4.16, 4.03) \times 10^{-6}$	$(-0.6, 0.6)$
3	1192	$(-2.61, 1.12) \times 10^{-8}$	$(-0.6, 0.8)$
4	329	$(-0.57, 4.78) \times 10^{-11}$	$(-0.8, 0.8)$
5	198	$(-1.42, 1.82) \times 10^{-14}$	$(-0.8, 0.8)$
6	152	$(-1.80, 2.66) \times 10^{-17}$	$(-0.6, 0.8)$
7	122	$(-0.24, 2.97) \times 10^{-19}$	$(-0.6, 0.8)$
8	104	$(-1.80, 3.75) \times 10^{-23}$	$(-0.6, 0.6)$
9	1120	$(-1.03, 0.94) \times 10^{-25}$	$(-0.8, 0.8)$

signals also get included in this set due to the condition on the residuals to account for higher order dynamic content in them. Thus Observations 6, 8 and 9 in Section 2.2.2 are validated.

2.3.4 Analysis of tonic components of the model

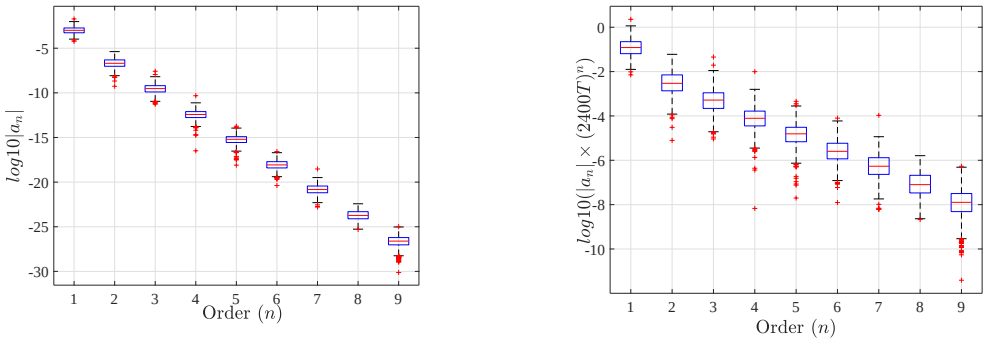
It has already been shown that the mV order magnitude DDP signals can be represented suitably using their deterministic polyfits y_{pk}. These y_{pk} can be considered as the tonic components of the models as mentioned in literature [15], [46] and are studied hereafter.

Analysis of coefficients

The order n_{ACF} of the ACF model has been identified using the l_{zc} and thereafter the polyfit y_{pk} is determined using the algorithm proposed in Section 2.2.3. The range of the highest order coefficients a_n of all the y_{pk} signals for different orders n and the corresponding range of r_k for the ACF model have been calculated for all 5706 signals and tabulated in Table 2.6.

In order to appreciate the contribution of the tonic components of y_{pk}, box plots of the logarithm of the magnitudes of the highest order coefficients of all polyfit signals y_{pk}, specifically $log10(|a_n|)$, are shown in Figure 2.12a. It is observed from both Table 2.6 and Figure 2.12a that the magnitude of a_n drops significantly from 10^{-2} to 10^{-25} as the order increases from $n = 1$ to $n = 9$.

In order to verify the maximum contribution of these terms in y_{pk}, the box plots of $log(|a_n| \times (kT)^n)$ at the terminal instant, that is for $t = 2400 \times 50ms = 120s$, are shown in Figure 2.12b. Since the plots of $log(|a_n|(kT)^n) = log|a_n| + nlog(kT)$, hence these differ from the corresponding $log|a_n|$ plots by a fixed gain in magnitude of $nlog(120) = 2.08n$. It is observed that there is a significant difference in magnitudes of the components at $n = 9$ from those of the lower order components in general as well as at the terminal instant. This justifies the truncation of the ACF modelling algorithm at $n = 9$ for this class and length of signals.



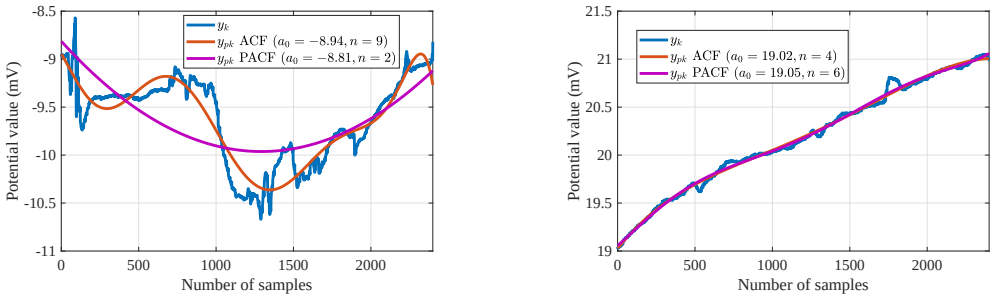
(a) (b)

Figure 2.12: Box plots of highest order coefficients of all polyfit signals y_{pk}: (a) $log10(|a_n|)$ vs. n (b) $log10(|a_n|(2400T)^n)$ vs. n

Analysis of y_{pk}

In order to validate the results, the standard order identification technique using PACF is also applied and the results of modelling are compared with those of the proposed ACF method. In the PACF model, the signal order n_{PACF} has been identified using the function *parcorr* in MATLAB version R2021a. Thereafter, for both the models, the MATLAB function *fit* which uses standard least-squares polynomial regression [179] has been used to fit the polynomial y_{pk} and obtain r_k.

Two representative 2-minute signals y_k, their ACF based and PACF based polyfits y_{pk} are shown in Figures 2.13a and 2.13b. The respective model order n and its primary coefficient a_0 is also listed in Figures 2.13a and 2.13b. In signal 1, the ACF model order is much higher than that of the PACF model while in signal 2, the ACF model order is slightly lower than that of the PACF model but the a_0 values of the two models are almost identical in both cases. In both cases, the ACF model captures the trend component reliably while the PACF model fails to do so in signal 1.



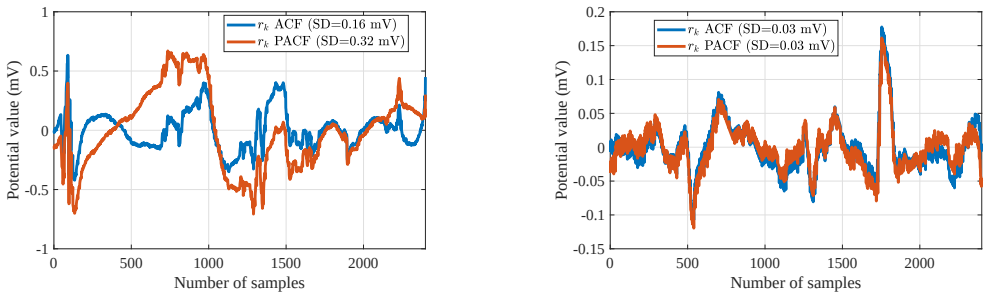
(a) (b)

Figure 2.13: Time-series plots of (a) Signal 1 and its ACF and PACF polyfits (b) Signal 2 and its ACF and PACF polyfits

2.3.5 Analysis of the phasic component of the model

The time-correlated stochastic r_k represents the phasic component of the model. The corresponding r_k of the 2 signals y_k and their polyfits y_{pk} shown in Figures 2.13a and 2.13b are shown in Figures 2.14a and 2.14b. The time-correlated character of r_k is evident in all cases and the range of the residuals are observed to be similar for both the ACF and the PACF models. All 4 residuals are zero mean. However, the standard deviation (SD) of the ACF and PACF model residuals of signal 1 are 0.16 mV and 0.32 mV respectively

while it is 0.03 mV for both residuals of signal 2 as expected from their close match in the plots.



(a) (b)

Figure 2.14: Time-series plots of (a) both residuals for Signal 1 (b) both residuals for Signal 2

From the analysis of tonic and phasic component, it is clear that the proposed stochastic ACF-ZCL model of DDP signals is representative of the acquired DDP signals.

2.3.6 Comparison between stochastic ACF-ZCL model and standard PACF model

In order to further study the suitability of the proposed model, the ACF models of all the 5706 signals are compared with the corresponding standard PACF models on the basis of the order of the trend components as well as the statistical characteristics, Sum of squared errors (SSE) and Root mean square (RMS) of the residuals.

As mentioned earlier, all 5706 signals have been fitted with stochastic ACF-ZCL and PACF models. Thereafter, the cumulative distribution function (CDF) of the ratios of the orders of the two polyfits n_{ACF}/n_{PACF} and that of the corresponding SSE residuals SSE_{ACF}/SSE_{PACF} are calculated. These as well as their zoomed plots are shown in Figure 2.15a -2.15d. It is observed from Figure 2.15b that $n_{PACF}/n_{ACF} \geq 1$ that is, the detected order using PACF is equal or larger than that obtained using the proposed ACF technique for 70% cases while SSE_{ACF}/SSE_{PACF} is less than 1 in 30% cases. The subscripts ACF and $PACF$ refer to the ACF and PACF models respectively.

The RMS of the r_k in both cases, RMS_{PACF} and RMS_{ACF} have been compared and the cdf of their ratio plotted in Figure 2.16. It is observed

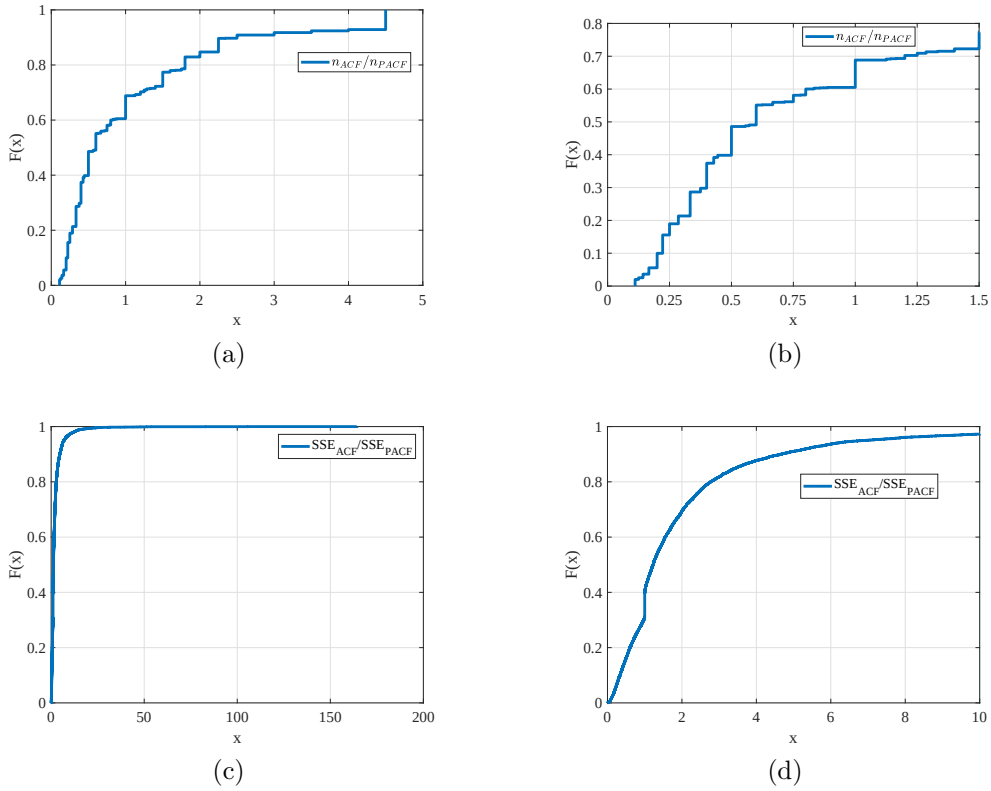


Figure 2.15: (a) CDF plot (b) Zoomed CDF plot of n_{ACF}/n_{PACF} (c) CDF plot (d) Zoomed CDF plot SSE_{ACF}/SSE_{PACF}

that in 90% cases, the RMS_{ACF} (RMS_{PACF}) is within twice (1.78 times) of RMS_{PACF} (RMS_{ACF}) and is at most 12.82 (15.21) times that of the RMS_{PACF} (RMS_{ACF}) considering all the signals. Thus, the ACF model typically provides a lower to equal order of fit as compared to the PACF model but this is associated with a marginally higher RMS of the residuals.

In view of this finding, a statistical analysis of the model residuals is done to further study the efficacy of the proposed approach. The statistical analysis of 95% (about median) of the residuals for the proposed ACF models and the standard PACF models is shown in Table 2.7. All residuals of both models are zero mean and the basic statistics of 95% of all the residuals about the median of both models are observed to be similar. More specifically, the residuals r_k lie within ± 0.8 mV for both models, while the corresponding SD is 0.19 mV (0.18 mV) for the ACF (PACF) model. Furthermore, the ACF (PACF) model residuals are non-skewed with kurtosis of 6.52 (8.02). This indicates that in both types of models, less than 5% of the potentials have

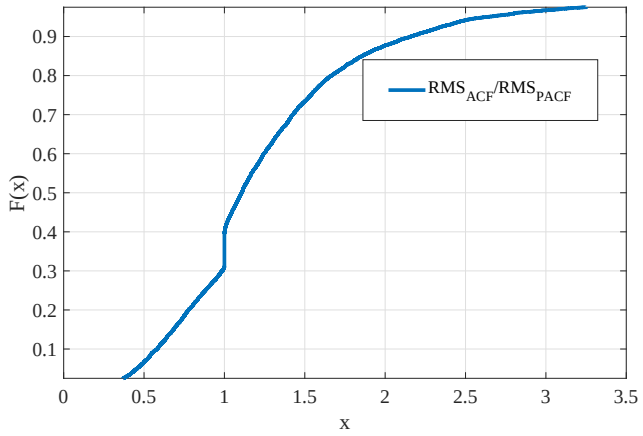


Figure 2.16: Cdf plot of ratio of residual errors:: RMS_{PACF} and RMS_{ACF}

Table 2.7: Statistical analysis of 95% (about median) of the residuals

	ACF method	PACF method
Range	$(-0.8, 0.8)$ mV	$(-0.8, 0.8)$ mV
Mean, median	0.00 mV	0.00 mV
SD	0.19 mV	0.18 mV
Skewness	-0.01	0.04
Kurtosis	6.52	8.02

local abrupt large spikes. This is to be expected since the data were acquired mostly from restful supine or sitting subjects who were not subject to any external stimuli.

In the next stage, the correlations of the residuals are studied. The Durbin-Watson test results lie within 0 to 2 for residuals of both ACF and PACF models indicating positive series correlation. This validates the consideration of the finite cross-covariance $X_{k,l}$ of the noise w_k in y_{sk} or y_k.

The Pearson coefficient test of the residuals show that 60.31% of the residuals of the ACF models follow beta distribution (both 4-parameter and F-location types or alternatively, of 1st and 2nd kind [181]) while 39.69% of the residuals are not related to any distribution. In case of the residuals of the PACF models, these proportions are 53.62% and 46.37% respectively.

This comparative study thus establishes that the conventional PACF model and the proposed ACF model are comparable and the residuals of both models exhibit positive series correlations. However, the ACF model typically captures the deterministic tonic component with a lower order poly-fit component than the corresponding PACF model. Furthermore, the ACF

model residuals follow the beta distribution more consistently and so can be considered representative of the phasic component of the signal.

2.4 Restfulness Assessment

The effect of rest on the parameters of the tonic and phasic components of the 4-channel DDP signal is observed in this Section. As mentioned in Chapter 1, DDP signals were acquired simultaneously from 4 limbs of the volunteers while they were in restful supine posture in the DS4 experiment. For purposes of statistical variety, the data was collected from the subjects in multiple (at least 3 or more) sessions that were conducted on different days and at different times within working hours. The acquired 10 minute long data from each limb of a subject in a particular session was then subdivided into 5 subsequent 2-minute duration sets. These denote 5 successive stages of restfulness of the subject. The proposed algorithm mentioned in Section 2.2.3 was then applied for modelling each 2-minute duration set of these signals as stated in the modelling protocol flowchart in Section 2.3.1.

2.4.1 Assessment Methodology

In the DS4 experiment [1], total 68 numbers of 10 minute 4-channel data sets were recorded from 16 volunteers over a total duration of 6 months. Thus, there are total $68 \times 4 = 272$ sets of independent signals from subjects in restful supine condition for 10 minutes. Thereafter, each of these signals have been segmented into 5 non-overlapping sets of 2-minutes signals. These are denoted as (rest) state 1, 2 and so on till state 5. So, $272 \times 5 = 1360$ numbers of 2-minute duration signals with $N = 2400$ discrete samples each are available for the analysis since the sampling time is $T = 50$ ms. The model features of these 5 stages of each DDP signal are analyzed henceforth to assess the systematic stage-wise changes in the deterministic and stochastic components of the ACF model due to restfulness.

One of the parameters considered in this study is the spectral entropy [182]. It is known that for a signal $x(n)$ with discrete Fourier transform $X(k)$, its power spectrum is $S(k) = |X(k)|^2$. Then, the probability distribution $P(k)$ and the spectral entropy E are defined as

$$P(k) = \frac{S(k)}{\sum_i S(i)}, \quad E = -\sum_{k=1}^{N} P(k) log_2 P(k). \tag{2.30}$$

For the calculation of the spectral entropy E, the 10 minute DDP signal is

first z-normalized as $(y_k - \overline{y_k})$/SD of (y_k) [24] and then the signal is subdivided into the 5 non-overlapping 2-minute subsets for further analysis.

The comprehensive list of parameters considered for the restfulness assessment are l_{zc}, n, mean of y_k and y_{pk}, spectral entropy of y_k, y_{pk} and r_k as well as the SD of y_k and r_k.

2.4.2 Effect on tonic parameters

For the assessment of restfulness, the box plots of the l_{zc} of the signal ACF and the order n of the ACF models for each of the 5 rest states are shown in Figures 2.17a and 2.17b. In case of the l_{zc}, it is observed that the upper

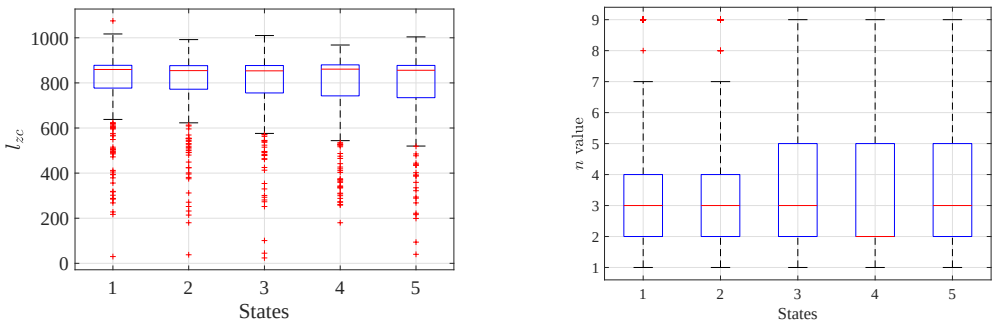


Figure 2.17: Box plots of (a) l_{zc}, (b) order n of ACF models for 5 rest states

quartile and median values remain almost same but the lower quartile decreases with rest. In case of the order of the ACF models, the median value is 3 in all states except state 4 where it drops to 2. This is also the lower quartile for all states. However, the upper quartile of n increases with rest indicating an emerging determinism of the system.

A comparison of the box plots of the mean of y_k and y_{pk} in the 5 rest states, as shown in Figures 2.18a and 2.18b, as well as the spectral entropy of y_k and y_{pk} in the 5 rest states, as shown in Figures 2.18c and 2.18d, show that the plots of these y_k and y_{pk} parameters are almost identical in both cases. This further validates that y_{pk} represents the tonic component of the DDP signal.

It is observed that the median as well as the inter-quartile range (IQR) of the mean of the tonic component show minimal change with rest. But the box plots of the spectral entropy of y_k and y_{pk} show that state 3 exhibits the maximum entropy in terms of the median as well as the IQR and then these decrease systematically till state 5. It is known that in the restful condition,

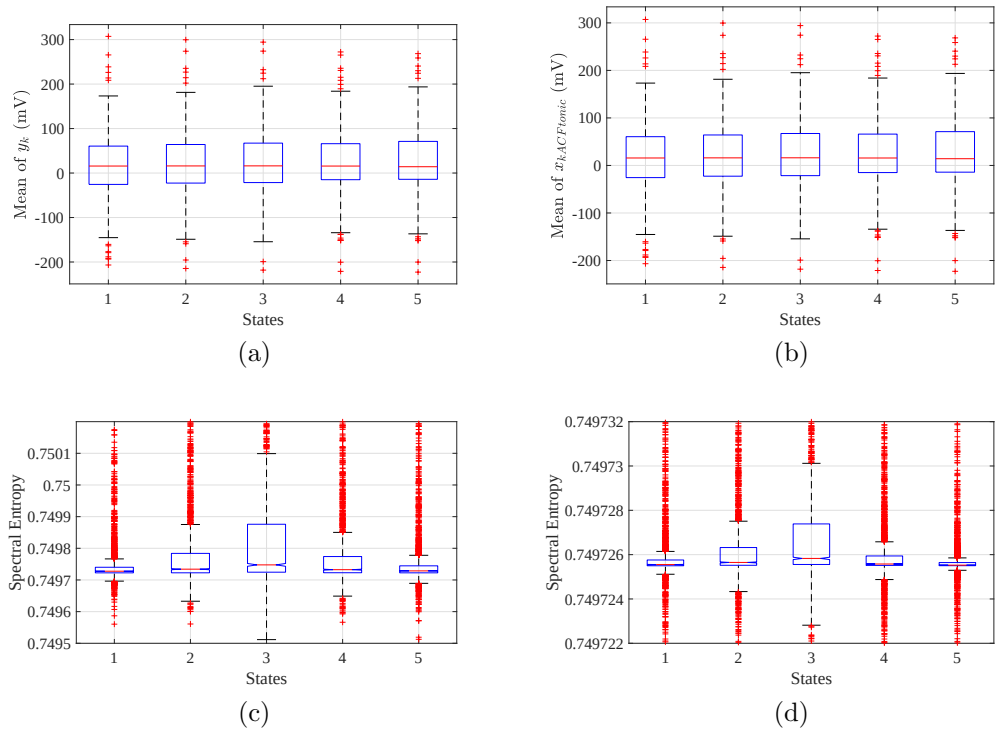


Figure 2.18: Box plots of mean of (a) y_k (b) $y_{pk} = x_{kACFtonic}$, spectral entropy of (c) y_k (d) y_{pk} of ACF models in 5 rest states

the body distributes the power over a larger frequency range [183] and so the spectral entropy also increases [184]. The maximum entropy in state 3 thus indicates that a supine subject achieves fully rested condition typically within 4-6 minutes of no-nap rest and this phenomena markedly affects the tonic component of the DDP signal.

2.4.3 Effect on phasic parameters

The box plots of the SD of the signals y_k and the residual r_k, as shown in Figure 2.19a and 2.19b, are observed to vary similarly from state 1 to state 5 and are almost identical for states 3, 4 and 5. Since the phasic component is the fast changing component, it is more likely to be represented by the SD of the signal. Hence, the box plots validate that the stochastic residual series r_k of the ACF model represents the phasic component of the signal.

It is further observed that the IQR as well as the median of the residuals decrease significantly towards zero with rest. This supports the finding in [24]

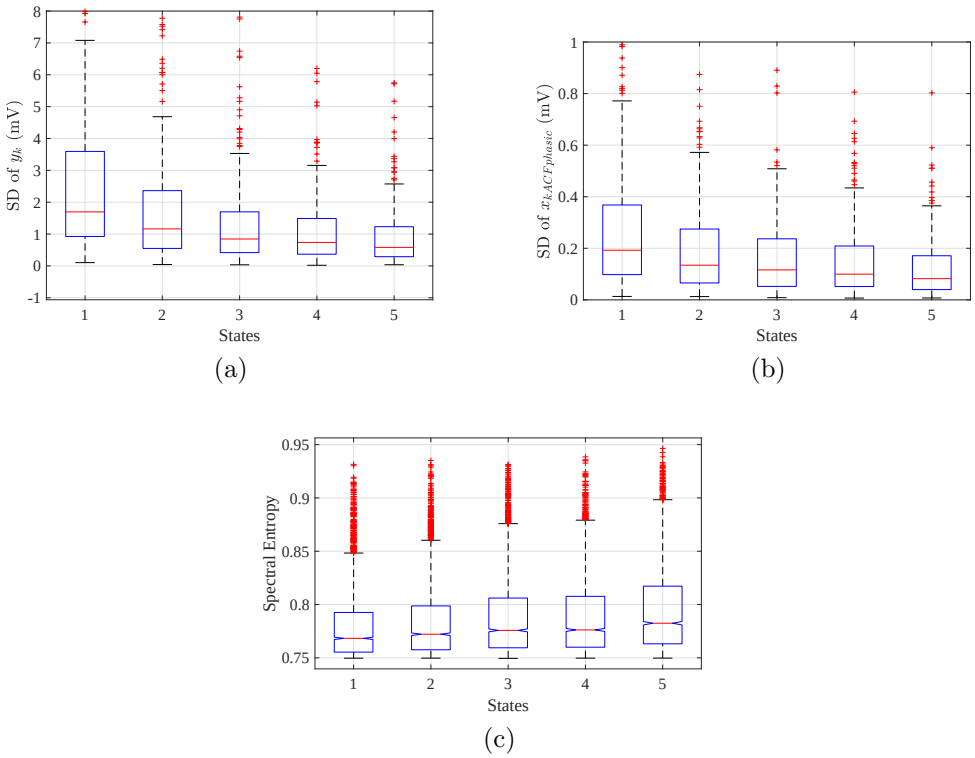


Figure 2.19: (a) SD of y_k and (b) $r_k = x_{kACFphasic}$ (c) Spectral entropy of r_k

that the average SD of the signals decrease with rest. Additionally, the median and the lower quartile of the box plots of the spectral entropy shown in Figure 2.19c also decrease with rest. Both these findings further support the emerging determinism of the signal with rest.

2.5 Discussions

Three objectives that are stated at the outset of this Chapter have been established. The first objective is to use the ZCL value of sample ACF of general time series to determine the order of general time-series signals. Based on this order, the polynomial regression models of the signals are developed. The ZCL of stochastic as well as deterministic (i) straight line (ii) power law series ($n = 1, 2, 3...$) and (iii) a deterministic polynomial series are stated as Theorems and Corollaries and derived. Based on the results and observations from these derivations, an algorithm is developed to obtain the stochastic ACF-ZCL model, or ACF model in short, of a finite time-series.

The theoretical observations from the Theorems and Corollaries are vali-

dated using the results of modelling of DDP signals. Here, total 5706 signals are taken for modelling. It is observed that of all the signals, 902 signals can be fitted with linear models. It is then established that the remaining 4804 signals do not satisfy the l_{zc} criterion and/or contain higher dynamics in their residuals thus requiring higher order model fits. Accordingly, higher order ACF models of these DDP signals are obtained using the proposed algorithm. From the theoretical derivations, it is found that the l_{zc} value can be shifted towards left and right lag for noise interactions and also coefficient interactions. These findings are also validated from the results of ACF modelling. Hence, the stochastic ACF-ZCL model of DDP signals are well established.

In the next objective, the proposed ACF models of the signals are compared with the models whose orders are determined using the standard PACF method. It is observed that the order of PACF model is greater or equal than ACF model in 70% cases. To observe the comparison results in more detail, SSE of all signals are calculated for both cases and it is found that in 30% cases, SSE of ACF model is less. This may be due to the local large spikes observed in several of the signals. RMS of the residuals of all signals are also calculated. It is observed that the RMS_{ACF} (RMS_{PACF}) is within twice (1.78 times) of RMS_{PACF} (RMS_{ACF}) in 90% cases and is at most 12.82 (15.21) times that of the RMS_{PACF} (RMS_{ACF}) considering all the signals. From the statistical analysis, it is found that the value of parameters are almost similar in both cases and the Durbin-Watson test results show that both sets of residuals are positively correlated. However, it is observed that the deterministic tonic component of the ACF model typically has a lower order polyfit component than the corresponding PACF model and the ACF model residuals follow the beta distribution more consistently. So, both the tonic and the phasic components of the signal can be said to be represented better using the ACF model features.

The final objective involves the assessment of the effect of restfulness on the ACF model features. This study is performed on the 5 numbers of consecutive 2-minute subsets of the original 10-minutes long signals acquired using DS4 experiment. The median of l_{zc} is almost equal in all the consecutive 5 states but the order n is low in state 4. It is also observed that the statistical characteristics of the mean and SD of the acquired signals are almost identical with those of the mean of the trend polynomial and the SD of the stochastic residuals respectively. This establishes that these represent the tonic and phasic components of the DDP signals. It is observed that the maximal entropy of the tonic component occurs in state 3 and then decreases indicating that the subject achieves a fully rested condition typically within 4-6 minutes of no-nap supine rest. Furthermore, the variations of the ZCL,

system order, mean of the tonic component, SD and spectral entropy of the residuals support the emerging determinism of the system with rest. The proposed stochastic ACF-ZCL model thus provides physiologically interpretable tonic and phasic components of a time-series signal. Additionally, the restful condition provides a useful baseline for further condition monitoring studies of DDP or similar biosignals involving applied stimuli.

The observations and interpretations from this study confirm that the proposed stochastic ACF-ZCL model of a time-series is well established.

Chapter 3

Kalman Filter model of DDP signals

3.1 Introduction

In Chapter 2, a stochastic ACF-ZCL model has been proposed for a general time-series signal. Thereafter, it has been shown to be effective in obtaining the tonic and phasic components of the DDP signals. This enables a detailed characterization of the signals and how they are affected by external or internal phenomena.

As discussed in Chapter 1, Kalman filter (KF) models of signals are also used for similar purposes. Traditionally, these utilize the Autoregressive (AR) modelling approach to obtain the required linear prediction coefficients (LPC) of the state space model for the KF algorithm [75], [71].

In this Chapter, the KF model of the DDP signals are developed using the stochastic polynomial form ($y_k = y_{pk} + w_k$) obtained in the ACF model. The order of the state space model of the DDP signals is determined using the l_{zc} and the polyfit y_{pk} is used to develop the state space model systematically. The KF algorithm is then implemented using this state space model of the DDP signals.

It is to be noted that the KF model assumes a process noise and a measurement noise. Accordingly, methods for determining the two sets of noises and their covariances have been devised for the state-space descriptions of the conventional and proposed KF models. Thereafter, using the assumed noise covariances, the filtered estimates of the signals are obtained using the KF algorithm. The filtered signals are then partitioned into the tonic and phasic components.

In order to ascertain the validity and efficacy of the KF models, a com-

parative study is done of the stochastic ACF-ZCL model, the proposed KF model and the conventional KF model of the DDP signals. They are compared in terms of the respective fit signals, their parameters and the statistical characteristics of the phasic components. The robustness and sensitivity metrics [174] of the KF models are also determined for these signals. All these parameters of the KF models are evaluated for the various states identified in the restfulness study discussed in Section 2.4 for the ACF models in Chapter 2.

The objectives of this Chapter are to

a) Derive the state space model of DDP signals using the time-series (TS) polyfit obtained in the ACF model approach along with the standard Linear Prediction Coefficients (LPC) based state space model of the signals.

b) Develop the corresponding KF_{LPC} and KF_{TS} models of DDP signals using standard KF equations. Determine the deterministic and stochastic components of the proposed and conventional KF models of DDP signals and study their properties. This includes the robustness and sensitivity metrics for the KF models also.

c) Compare the properties of the tonic and phasic components of the proposed and conventional KF models and the earlier proposed stochastic ACF-ZCL model. Observe and compare the effect of long duration of (no-nap) rest on the parameters of the 3 models.

The standard AR based Linear Prediction Coefficient (LPC) state space model formulation that forms the basis of the standard Kalman Filter model, denoted here as the KF_{LPC} model, is discussed in Section 3.2. In Section 3.3, a time-series (TS) state space model is derived ab-initio from the ACF model of the time-series signal y_k. The derived TS state space model forms the basis of the proposed KF_{TS} model. The standard KF algorithm that is applied to the LPC and TS state space models to obtain the standard KF_{LPC} and proposed KF_{TS} models of the DDP signals is described in Section 3.4. The robustness (J_1) and sensitivity (J_2) metrics, as obtained from these KF models, associated metrics parameters and their implications in terms of the KF models are discussed in Section 3.4.1. The KF models of the DDP signals and their analysis is discussed in detail in Section 3.5. This includes the modelling protocol and the method used for obtaining the filter parameters in Section 3.5.1, followed by fitting tonic and phasic components of the KF models to two signals of different orders in Section 3.5.2. In all

common cases, the two KF model parameters are compared and validated in terms of those of the ACF model, while the KF residuals and the KF metrics are compared among themselves. The detailed statistical analysis of these 3 models is performed in Section 3.5.3 while the metrics related parameters of the two KF models are discussed and compared in Section 3.5.4. The restfulness assessment performed in Section 2.4 is extended in Section 3.6 to compare the ACF as well as the two KF models. This is done systematically in terms of the spectral entropy of the 3 models across all states of rest in Section 3.6.1, followed by a comparison of the characteristics of the noise residuals of the two KF models in Section 3.6.2 and finally a detailed comparison of various tonic and phasic components and the key metrics parameters in Section 3.6.3. The discussions are presented in Section 3.7.

3.2 Developing the KF_{LPC} model

The standard KF model, denoted henceforth as KF_{LPC} model, relies on AR models in which it is assumed that the current value of a signal depends linearly on its past values. Accordingly, the k^{th} sample of a n^{th} order signal x_{kLPC} is represented in terms of its n number of past values x_{k-iLPC}, the linear prediction coefficients (LPCs) of the past values b_i and the AR process noise w_{kLPC} as [71]. The AR coefficients b_i have been determined using the command 'ar' in MATLAB, version 2021a.

$$x_{kLPC} = -\sum_{i=1}^{n} b_i x_{k-iLPC} + w_{kLPC} \tag{3.1}$$

The method for obtaining the LPC based state space description of the KF_{LPC} model is described hereafter.

3.2.1 State space description of LPC model

The n^{th} order state space model of the system described in equation 3.1 can be represented as

$$\begin{bmatrix} x_{kLPC} \\ x_{k-1LPC} \\ x_{k-2LPC} \\ . \\ . \\ . \\ x_{k-n+2LPC} \\ x_{k-n+1LPC} \end{bmatrix} = \begin{bmatrix} -b_1 & -b_2 & -b_3 & . & . & . & -b_{n-1} & -b_n \\ 1 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & 0 & . & . & . & 0 & 0 \\ . & . & . & . & & & & . \\ . & . & . & & . & & & . \\ . & . & . & & & . & & . \\ 0 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{k-1LPC} \\ x_{k-2LPC} \\ x_{k-3LPC} \\ . \\ . \\ . \\ x_{k-n+1LPC} \\ x_{k-nLPC} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ 0 \end{bmatrix} w_{kLPC}$$

or

$$X_{kLPC} = \phi X_{k-1LPC} + G w_{kLPC} \tag{3.2}$$

In this state space model, X_{kLPC} is the $(n \times 1)$ state vector matrix, ϕ is the $(n \times n)$ state transition matrix, G is the $(n \times 1)$ input matrix and w_{kLPC} is the noise corrupting the input signal at the k^{th} instant. The standard AR algorithm provides the LP-Coefficients and the process noise covariance σ^2_{wLPC} of the state space model. w_{kLPC} is then generated as a random, uncorrelated Gaussian sequence of required length and with standard deviation σ_{wLPC}.

Considering the $(1 \times n)$ observation matrix H and the zero mean, Gaussian measurement noise v_{kLPC}, the actual measured signal y_k is obtained as

$$\begin{aligned} Y_k &= HX_{kLPC} + v_{kLPC} = [1 \;\; 0 \;\; ... \;\; 0]X_{kLPC} + v_{kLPC} \\ &= y_k = x_{kLPC} + v_{kLPC} \end{aligned} \tag{3.3}$$

In order to satisfy this relation, the measurement noise is generated using the actual signal y_k and the process x_{kLPC} described in equation 3.1 as $v_{kLPC} = y_k - x_{kLPC}$. σ^2_{vLPC} is the covariance of this measurement noise sequence.

The medians of the two covariances, σ^2_{wLPC} and σ^2_{vLPC}, for all the acquired signals are considered to be the effective process and measurement noise covariances, q_{LPC} and r_{LPC}, of the DDP signals. For the LPC model, the process noise covariance matrix Q_{LPC} is generated as $q_{LPC}GG^T$, while the measurement noise being scalar, the measurement noise covariance matrix $R_{LPC} = r_{LPC}$. The standard KF_{LPC} model is obtained by using this n^{th} order state space description along with the noise covariances Q_{LPC} and R_{LPC} in the standard KF algorithm described in Section 3.4.

3.3 Proposed Time-Series state space model

As discussed earlier, the DDP time-series signal y_k can be modelled as a sum of a time-series polynomial regression y_{pk} and a time-correlated stochastic noise w_k. This representation of the signal is used to derive ab-initio a Time-Series (TS) based state-space model as an alternative to the conventional LPC based state-space model discussed in Section 3.2.1.

Let ${}^pC_j$ and ! denote the combination and factorial operators respectively. In order to develop the state-space model, let us consider the k-th instant of the signal y_k. The corresponding state, denoted as x_{kTS}, has a deterministic component x^d_k and the instantaneous process noise component w_{kTS}. This model is formulated such that y_{pk} of the ACF model is identical to x^d_k while w_k of the ACF model is a combination of w_{kTS} and the measurement noise v_{kTS} associated with this proposed model. All the noises are assumed to be zero mean with Gaussian distribution. Thus,

$$\begin{aligned} y_k &= y_{pk} + w_k = \sum_{i=0}^{n} a_i (kT)^i + w_k \quad \forall k \in [0, N], \quad i \in [0, n] \\ &= x_{kTS} + v_{kTS} = x_k^d + w_{kTS} + v_{kTS} \qquad (3.4) \\ x_{kTS} &\doteq x_k^d + w_{kTS}; \quad x_k^d \doteq y_{pk} = \sum_{i=0}^{n} a_i (kT)^i \qquad (3.5) \\ w_k &\doteq w_{kTS} + v_{kTS} \qquad (3.6) \end{aligned}$$

As in case of the ACF model, in this case also T is the (constant) sampling time, k denotes the sample (or time instant), N denotes the number of samples, n denotes the order of the signal y_k (and its corresponding y_{pk}), i is the variable representing the local order of the term $t_k = kT$ and a_i is the coefficient of the i^{th} order term.

The state space model is developed from the successive discrete time-derivatives, or finite differences, of the deterministic state component x_k^d. The general p^{th} difference $\Delta_d^p x_k \; \forall p \in [1, n]$ can be derived from first principles as follows.

The first difference $\Delta_d^1 x_k$ is obtained using equation (3.5) as

$$\begin{aligned} \Delta_d^1 x_k &= (x_{k+1}^d - x_k^d)/T \qquad (3.7) \\ &= [\sum_{i=0}^{n} a_i [(k+1)T]^i - (\sum_{i=1}^{n} a_i (kT)^i)]/T \\ &= \sum_{i=0}^{n} a_i T^{i-1} [(k+1)^i - k^i] \qquad (3.8) \\ &= \sum_{i=0}^{n} a_i T^{i-1} [\sum_{j=0}^{1} (-1)^{j+1} \{{}^1C_j (k+j)^i\}] \qquad (3.9) \end{aligned}$$

Using equation (3.8), the 2^{nd} difference $\Delta_d^2 x_k$ is obtained similarly as

$$\begin{aligned} \Delta_d^2 x_k &= (\Delta_d^1 x_{k+1} - \Delta_d^1 x_k)/T \qquad (3.10) \\ &= \sum_{i=0}^{n} a_i T^{i-2} [(k+2)^i - (k+1)^i] - \sum_{i=0}^{n} a_i T^{i-2} [(k+1)^i - k^i] \\ &= \sum_{i=0}^{n} a_i T^{i-2} [(k+2)^i - 2(k+1)^i + k^i] \qquad (3.11) \\ &= \sum_{i=0}^{n} a_i T^{i-2} [\sum_{j=0}^{2} (-1)^{j+2} \{{}^2C_j (k+j)^i\}] \qquad (3.12) \end{aligned}$$

By mathematical induction, the p^{th} difference $\Delta_d^p x_k \quad \forall p \in [1, n]$ can be represented as

$$\begin{aligned} \Delta_d^p x_k &= (\Delta_d^{p-1} x_{k+1} - \Delta_d^{p-1} x_k)/T & (3.13) \\ &= \sum_{i=0}^{n} a_i T^{i-p} [\sum_{j=0}^{p} (-1)^{j+p} \{{}^pC_j (k+j)^i\}] & (3.14) \end{aligned}$$

Depending on the order n of the polynomial, the term $\Delta_d^n x_{k-1}$ can be further simplified. For a polynomial of order $n = 1$, the maximum value of $p = n = 1$ and so we have

$$\begin{aligned} x_k^d &= \sum_{i=0}^{n} a_i (kT)^i = a_0 + a_1 kT & (3.15) \\ x_{k-1}^d &= \sum_{i=0}^{n} a_i (k-1)^i T^i = a_0 + a_1 (k-1)T \\ \Delta_d^1 x_{k-1} &= [a_0 - a_0 + a_1 \{k - (k-1)\}T]/T = a_1 = 1! a_1 & (3.16) \end{aligned}$$

Similarly, for a polynomial of order $n = 2$,

$$\begin{aligned} x_k^d &= \sum_{i=0}^{2} a_i (kT)^i = a_0 + a_1 kT + a_2 (kT)^2 & (3.17) \\ x_{k-1}^d &= a_0 + a_1 (k-1)T + a_2 (k-1)^2 T^2 \\ \Delta_d^1 x_{k-1} &= (x_k^d - x_{k-1}^d)/T \\ &= [a_0 - a_0 + a_1 \{k - (k-1)\}T + a_2 \{k^2 - (k^2 - 2k + 1)\}T^2]/T \\ &= a_1 + a_2 (2k-1)T & (3.18) \\ \Delta_d^1 x_k &= a_1 + a_2 [2(k+1) - 1]T & (3.19) \\ \Delta_d^2 x_{k-1} &= [a_1 - a_1 + a_2 \{2(k+1) - 1 - (2k-1)\}T]/T \\ &= 2a_2 = 2! a_2 & (3.20) \end{aligned}$$

By mathematical induction, it can be shown that for a signal of order n,

$$\Delta_d^n x_{k-1} = n! a_n \quad (3.21)$$

Let the system states for the state-space model be $x_{kTS}, \Delta_d^1 x_k, \Delta_d^2 x_k, ..., \Delta_d^{n-1} x_k$. Then, using equation (3.5) and the definitions of the finite differences in equations (3.7), (3.10) and (3.13), we have

$$
\begin{aligned}
x_{kTS} &= x_k^d + w_{kTS} = x_{k-1}^d + \Delta_d^1 x_{k-1} T + w_{kTS} && (3.22)\\
\Delta_d^1 x_k &= \Delta_d^1 x_{k-1} + \Delta_d^2 x_{k-1} T && (3.23)\\
\Delta_d^2 x_k &= \Delta_d^2 x_{k-1} + \Delta_d^3 x_{k-1} T && (3.24)\\
&\vdots\\
\Delta_d^{n-1} x_k &= \Delta_d^{n-1} x_{k-1} + \Delta_d^n x_{k-1} T && (3.25)
\end{aligned}
$$

Further, equation (3.25) can be written using equation (3.21) as

$$\Delta_d^{n-1} x_k = \Delta_d^{n-1} x_{k-1} + n! a_n T \tag{3.26}$$

Hence, using equations (3.22)-(3.26) and denoting $n!a_n$ as the known input u_k to the system, the proposed state space model of order n is obtained as

$$
\begin{bmatrix} x_{kTS} \\ \Delta_d^1 x_k \\ \Delta_d^2 x_k \\ \vdots \\ \Delta_d^{n-1} x_k \end{bmatrix} = \begin{bmatrix} 1 & T & 0 & \cdots & 0 \\ 0 & 1 & T & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1}^d \\ \Delta_d^1 x_{k-1} \\ \Delta_d^2 x_{k-1} \\ \vdots \\ \Delta_d^{n-1} x_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ T \end{bmatrix} u_k + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} w_{kTS}
$$

or

$$
\begin{aligned}
X_{kTS} &\doteq F X_{k-1TS} + G_1 u_k + G_2 w_{kTS} && (3.27)\\
Y_k &= H X_{kTS} + v_{kTS} = y_k = x_{kTS} + v_{kTS} && (3.28)
\end{aligned}
$$

As mentioned earlier, for the n-th order signal, the $(n \times 1)$ state vector $X_{kTS} \doteq [x_{kTS} \;\; \Delta_d^1 x_k \;\; \Delta_d^2 x_k \;\; \ldots \;\; \Delta_d^{n-1} x_k]^T$, F is the $(n \times n)$ state transition matrix, the known (constant) input $u_k \doteq n! a_n$ where a_n is the n-th order coefficient of the polynomial x_k^d (or y_{pk}), w_{kTS} is the process noise and G_1, G_2 are the $(n \times 1)$ input vectors for u_k and w_{kTS} respectively. Y_k is the actual measured signal, v_{kTS} is the zero mean Gaussian measurement noise that is assumed to be uncorrelated with w_{kTS} and H is the $(1 \times n)$ observation vector.

In order to determine the noises w_{kTS} and v_{kTS}, we first consider the calibration study of the data acquisition system (DAS) used for acquiring the signals of this magnitude [132]. The sensor noise v_{kTS} is generated as

a random zero-mean Gaussian noise with the variance as obtained from the calibration chart. Thereafter, using equation (3.6), we determine w_{kTS} as $(w_k - v_{kTS})$, where w_k is the noise in the ACF model.

Let the variances of the noise sequences w_{kTS} and v_{kTS} for a particular acquired signal y_k be denoted as σ^2_{wTS} and σ^2_{vTS} respectively. As in the LPC state space model, in this case also, the medians of σ^2_{wTS} and σ^2_{vTS} for all the acquired signals are considered to be the effective process and measurement noise variances, q_{TS} and r_{TS} respectively, of the DDP signal. For the proposed model, the process noise variance matrix Q_{TS} is then generated as $q_{TS}G_2G_2^T$, while the measurement noise being scalar, the measurement noise variance matrix $R_{TS} = r_{TS}$.

This proposed n^{th} order TS state space description and the noise covariances Q_{TS} and R_{TS} are used to obtain the proposed KF_{TS} model as discussed hereafter.

3.4 KF models of DDP signals

In this section, the KF standard equations [174], [75] are used to develop two different KF models of DDP signals. The 1^{st} KF model is the conventional KF model, denoted henceforth as KF_{LPC}, which uses the LPC state space model stated in Section 3.2.1 while the 2^{nd} KF model, denoted henceforth as KF_{TS}, uses the proposed Time-Series (TS) state space model described in Section 3.3.

The Kalman filter systematically calculates the apriori and aposteriori state estimates, $\hat{X_k^-}$ and $\hat{X}_k$, as well as the corresponding state error covariances, P_k^- and P_k, while using the Kalman gain K_k. Two other intermediate terms determined for calculating K_k and $\hat{X}_k$ are the innovation and its covariance, denoted as s_k and S_k respectively. These equations of the KF algorithm are stated sequentially hereafter.

In case of the two KF models KF_{LPC} and KF_{TS}, the difference is only in the form of the apriori state estimate $\hat{X_k^-}$. Hence, this estimate is suitably denoted as $\hat{X}^-_{kLPC}$ and $\hat{X}^-_{kTS}$ for the two models respectively. The other relations are identical in form for both the KF models and hence are stated in general form. The two noise covariances, represented in general as Q and R, will refer to KF_{LPC} or KF_{TS} depending on the KF model being determined.

$$\begin{aligned}
\hat{X}^-_{kLPC} &= F\hat{X}_{k-1} && (3.29)\\
\text{or } \hat{X}^-_{kTS} &= F\hat{X}_{k-1} + G_1 u_k &&\\
P^-_k &= FP_{k-1}F^T + Q && (3.30)\\
s_k = Y_k - H\hat{X}^-_k &= y_k - H\hat{X}^-_k && (3.31)\\
S_k &= HP^-_k H^T + R && (3.32)\\
K_k &= P^-_k H^T S^{-1}_k && (3.33)\\
\hat{X}_k &= \hat{X}^-_k + K_k s_k && (3.34)\\
P_k &= P^-_k - P^-_k H^T S^{-1}_k H P^-_k && (3.35)
\end{aligned}$$

It is to be noted that the input $u_k = n!a_n$ remains constant in all iterations of the KF_{TS}.

It is reiterated that $Y_k = y_k$ is considered to be the k-th instant of the DDP signal being modelled. In both the KF algorithms, the Kalman gain K_k, the aposteriori state vector estimate $\hat{X}_k$ and its corresponding error covariance P_k are continually updated in every iteration. The initialization of $\hat{X_{k-1}}$ and P_{k-1} in both the KF models have been done using the same approach. The first n number of discrete samples of the DDP signal being modelled have been used for initialization of $\hat{X_{k-1}}$, where n is the order of the signal as determined using the l_{zc} approach. In order to initialize P_{k-1}, the two point differencing method [185] has been used. This requires an additional set of data and so the first $(n+1)$ number of discrete samples are used for this.

3.4.1 Robustness and Sensitivity Metrics

It is observed that the noisy, measured output is y_k and the KF provides its cleaned estimate as the first scalar component $\hat{x}_k$ of the state vector $\hat{X}_k$, as already shown in case of the conventional Kalman filtering of the speech signal [72], [71]. In order to interpret the relation between y_k and $\hat{x}_k$, the Kalman recursion equations are rewritten in scalar form by pre-multiplying the vector equations by H and pre- and post- multiplying the covariance equations by H and H^T respectively. The only exceptions are the relations for the innovation and its covariance, which remain as they are. Further, since $HG_1 = 0$, so the apriori estimate $\hat{x}^-_k = H\hat{X}^-_k$ assumes a common form for both the KF models. Further, for the respective KF model, the scalar process noise variance $q = HQH^T$ is assumed suitably while the measurement noise variance being essentially scalar, $R = r$.

The recast KF algorithm with the scalar state estimates and corresponding covariances can thus be written as follows.

$$
\begin{aligned}
\hat{x}_k^- &= H\hat{X}_k^- = HF\hat{X}_{k-1} && (3.36)\\
HP_k^-H^T &= HFP_{k-1}F^TH^T + HQH^T \doteq \alpha_k + q && (3.37)\\
\alpha_k &\doteq HFP_{k-1}F^TH^T;\ \ q = HQH^T && (3.38)\\
s_k &= y_k - \hat{x}_k^- && (3.39)\\
S_k &= HP_k^-H^T + R = \alpha_k + q + r && (3.40)\\
HK_k &= HP_k^-H^TS_k^{-1} = HP_k^-H^T(HP_k^-H^T + R)^{-1} && (3.41)\\
&= \frac{\alpha_k + q}{\alpha_k + q + r} && (3.42)\\
\hat{x}_k &= H\hat{X}_k = \hat{x}_k^- + HK_ks_k = \hat{x}_k^- + HK_k(y_k - \hat{x}_k^-)\\
&= (1 - HK_k)\hat{x}_k^- + HK_ky_k && (3.43)\\
HP_kH^T &= HP_k^-H^T - HP_k^-H^TS_k^{-1}HP_k^-H^T\\
&= (1 - HP_k^-H^TS_k^{-1})HP_k^-H^T = (1 - HK_k)HP_k^-H^T && (3.44)\\
&= \frac{r(\alpha_k + q)}{\alpha_k + q + r} && (3.45)
\end{aligned}
$$

α_k represents the contribution of the aposteriori mean squared error from the previous instant to the total apriori mean prediction squared error. It can be further simplified for an n^{th} order system considering P_k is real symmetric.

$$
\alpha_k = H(FP_{k-1}F^T)H^T
$$

$$
= \begin{bmatrix}1\\0\\0\\\cdot\\\cdot\\\cdot\\0\end{bmatrix}^T \begin{bmatrix}1 & T & 0 & \cdot & \cdot & \cdot & 0\\ 0 & 1 & T & \cdot & \cdot & \cdot & 0\\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & & & \cdot\\ \cdot & \cdot & \cdot & & \cdot & & \cdot\\ \cdot & \cdot & \cdot & & & \cdot & \cdot\\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 1\end{bmatrix} P_{k-1} \begin{bmatrix}1 & T & 0 & \cdot & \cdot & \cdot & 0\\ 0 & 1 & T & \cdot & \cdot & \cdot & 0\\ 0 & 0 & 1 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & & & \cdot\\ \cdot & \cdot & \cdot & & \cdot & & \cdot\\ \cdot & \cdot & \cdot & & & \cdot & \cdot\\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 1\end{bmatrix}^T \begin{bmatrix}1\\0\\0\\\cdot\\\cdot\\\cdot\\0\end{bmatrix}
$$

$$
\begin{aligned}
&= \begin{bmatrix}1 & T & 0 & \cdot & \cdot & \cdot & 0\end{bmatrix} \begin{bmatrix}p_{11} & p_{12} & p_{13} & \cdot & \cdot & \cdot & p_{1n}\\ p_{12} & p_{22} & p_{23} & \cdot & \cdot & \cdot & p_{2n}\\ p_{13} & p_{23} & p_{33} & \cdot & \cdot & \cdot & p_{3n}\\ \cdot & \cdot & \cdot & \cdot & & & \cdot\\ \cdot & \cdot & \cdot & & \cdot & & \cdot\\ \cdot & \cdot & \cdot & & & \cdot & \cdot\\ p_{1n} & p_{2n} & p_{3n} & \cdot & \cdot & \cdot & p_{nn}\end{bmatrix} \begin{bmatrix}1\\T\\0\\\cdot\\\cdot\\\cdot\\0\end{bmatrix}\\
&= p_{11} + 2Tp_{12} + T^2p_{22} && (3.46)
\end{aligned}
$$

It is thus apparent that only the effects of the state estimation errors for the present instant k and its immediate previous instant $k-1$ contribute to α_k. The noise variances q and r can be typically considered to be time-invariant for a particular acquired signal since biosignals do not normally vary rapidly or instantaneously.

These parameters allow for the evaluation of the sensitivity metric J_1 and the robustness metric J_2, as proposed in Saha et al. [174] and used in the speech applications [72], [71], [75], for the KF model. If the KF model is operating in the sensitive zone, this indicates that its ability and/or adaptability to any dynamic changes is high whereas a robust KF is useful for discarding the effect of dynamic model errors or disturbances.

The instantaneous robustness metric J_{2k} and the corresponding instantaneous sensitivity metric J_{1k} are defined in terms of the trace operator denoted as $tr(\cdot)$. In this one-dimensional case with particular scalar noises q and r, the trace reduces to a scalar form [75] as follows.

$$J_{2k} = tr[(HP_k^- H^T)^{-1} HQH^T] = \frac{q}{\alpha_k + q} \tag{3.47}$$

$$J_{1k} = tr[(S_k)^{-1} R] = \frac{r}{\alpha_k + q + r} \tag{3.48}$$

For the evaluation of the overall filter performance, the instantaneous performance metrics are time-averaged in terms of the total time horizon N to obtain the KF sensitivity and robustness metrics J_1 and J_2 as

$$J_1 = \frac{1}{N}\sum_{k=1}^{N} J_{1k}; \quad J_2 = \frac{1}{N}\sum_{k=1}^{N} J_{2k} \tag{3.49}$$

In practice, the parameters of the two metrics and the corresponding q values are defined for the time-averaged values of J_2 and J_1 obtained for a particular signal. The values of J_2 and J_1 and the KF estimates vary if the underlying process noise covariance q is varied. As evident from their definitions in equations (3.47) and (3.49), the value of J_{2k}, and hence J_2, increases from a minimum value, typically close to 0, to a high value as q increases. In a similar manner, but in a reverse trend, it can be said from equations (3.48) and (3.49) that J_{1k}, and hence J_1, decreases from its maximum value to a minimum value as q increases. The various metric related parameters are listed and defined in Table 3.1.

Further representing the scalar form of the KF algorithm listed in equa-

Table 3.1: Parameters of the Robustness and Sensitivity Metrics

Parameters	Definition
J_1	The sensitivity metric
J_2	The robustness metric
J_{1max}, J_{2max}	The respective maximum values of J_1 and J_2
J_{1min}, J_{2min}	The respective minimum values of J_1 and J_2
J_{com}	The value of J_1 and J_2 for which $J_1 = J_2$
q_{min}	The maximum value of q for which $J_1 = J_{1max}$
q_{max}	The minimum value of q for which $J_2 = J_{2max}$
q_{com}	The value of q for which $J_1 = J_2 = J_{com}$

tions (3.36)-(3.45) in terms of the instantaneous metrics, we have

$$HP_k^- H^T = \alpha_k + q = q/J_{2k} \tag{3.50}$$

$$\alpha_k = \frac{(1 - J_{2k})q}{J_{2k}} \tag{3.51}$$

$$1 - HK_k = \frac{r}{\alpha_k + q + r} = J_{1k} \tag{3.52}$$

$$\hat{x}_k = (1 - HK_k)\hat{x}_k^- + HK_k y_k = J_{1k}\hat{x}_k^- + (1 - J_{1k})y_k \tag{3.53}$$

$$HP_k H^T = (1 - HK_k)HP_k^- H^T = \frac{r(\alpha_k + q)}{\alpha_k + q + r} = \frac{J_{1k}q}{J_{2k}} \tag{3.54}$$

The following observations can be made from the above relations.

- By definition, as q increases, J_{2k} of the system increases till it saturates to a maximum value J_{2kmax}. The minimum value of q for which the system would operate at J_{2kmax} is denoted as q_{max}.

- Also by definition, as q decreases, J_{1k} of the system increases till it saturates to a maximum value J_{1kmax}. The maximum value of q for which this saturation is obtained is denoted as q_{min}.

- As per equation (3.53), the value of J_{1k} decides the proportion of the measured output y_k and its apriori KF estimate $\hat{x}_k^-$ in the aposteriori KF estimate $\hat{x}_k$.

- The aposteriori state error covariance HP_kH^T is affected by both the metrics, but in opposing ways. As per equation (3.54), it is directly related to J_{1k} and inversely to J_{2k}.

Hence an interim choice of the metrics, denoted as J_{com}, where $J_{1k} = J_{2k}$ provides a compromise choice between the apriori estimate and the measurement thus allowing the system to balance between responding to valid external changes and rejecting disturbances. The corresponding value of q for which the system would operate in this manner is denoted as q_{com}.

- It is observed from equation (3.50) that any further increase of q beyond q_{max} causes the apriori state error covariance $HP_k^- H^T$ to increase proportionately.

  As evident from equations (3.54) and (3.53), the aposteriori state error covariance HP_kH^T of the system will be minimized in this scenario to yield a noise-optimized estimate $\hat{x}_k$ of the signal y_k if the system operates at a small value of J_{1k} and a corresponding value of q.

  If the system operates in such a scenario, it is expected to respond suitably, or sensitively, to dynamic changes affecting the system.

- Alternatively, as q decreases below q_{min}, J_{2k} decreases till it saturates to a minimum value J_{2kmin}. From equation (3.50), it is also observed that for even smaller q, the corresponding $HP_k^- H^T$ decreases further.

  In this scenario, the value of J_{1k} for the system will necessarily be large to trust the apriori estimate and thus obtain an accurate aposteriori estimate $\hat{x}_k$ and to keep HP_kH^T low.

  A system operating in this regime will tend to treat any dynamic changes affecting it as disturbances and thus respond robustly by rejecting it.

3.5 Modelling and analysis of DDP signals

The same 5706 signals that have been considered for the analysis of the ACF model in Chapter 2 are modelled in this Chapter using the KF_{LPC} and KF_{TS} algorithms. It is to be recalled that each 2 minutes long signal is acquired with a sampling time of 50 ms yielding 2400 discrete samples.

In the proposed ACF model, the measured signal y_k is directly partitioned into a tonic and a phasic component as $y_k = y_{pk} + w_k$. Using the notations $x_{kACF_{tonic}}$ and $x_{kACF_{phasic}}$ for these two components, the ACF model can thus be expressed as

$$\begin{aligned} y_k &= x_{kACF} = x_{kACF_{tonic}} + x_{kACF_{phasic}} \\ y_{pk} &\doteq x_{kACF_{tonic}}, \quad w_k \doteq x_{kACF_{phasic}} \end{aligned} \tag{3.55}$$

As discussed in Section 3.4, the Kalman filter is used to remove, or reduce, the process and measurement noise (residuals) corrupting the signal y_k to obtain the cleaned signal estimate $\hat{x}_k$. This $\hat{x}_k$ is then partitioned into the tonic and phasic components. Just as $x_{kACF_{tonic}} = y_{pk}$ is obtained in the ACF model, the tonic components in the KF_{LPC} and KF_{TS} models, $x_{kLPC_{tonic}}$ and $x_{kTS_{tonic}}$ respectively, are also obtained by fitting suitable polynomial regression models to the respective $\hat{x}_k$. Thereafter, the phasic components $x_{kLPC_{phasic}}$ and $x_{kTS_{phasic}}$ are obtained as the difference of the tonic component from its respective cleaned estimate. Thus,

$$\hat{x}_{kLPC} = x_{kLPC_{tonic}} + x_{kLPC_{phasic}} \tag{3.56}$$

$$\hat{x}_{kTS} = x_{kTS_{tonic}} + x_{kTS_{phasic}} \tag{3.57}$$

$$y_k = \hat{x}_{kLPC} + r_{kLPC} = \hat{x}_{kTS} + r_{kTS} \tag{3.58}$$

3.5.1 Modelling protocol and method

The detailed description of the steps followed to develop the KF models, obtain its parameters and metrics are given hereafter.

Step 1: The 2 minutes duration signals considered for the ACF models are considered for obtaining the KF_{LPC} and KF_{TS} models.

Step 2: The order n of each 2 minute signal y_k are identified using the l_{zc} method discussed in Section 2.2.3 for the ACF models. These are used to obtain the system matrices ϕ and G of the LPC model and F, G_1, G_2 of the TS model and the output matrix H for both LPC and TS state space models as described in Sections 3.2 and 3.3 respectively.

Step 3: The process and measurement noise sequences generated using the particular state space model are used to obtain the process and measurement noise covariances σ_w^2 and σ_v^2 respectively. The effective time-invariant fixed process and measurement noise covariances, q and r respectively, of the DDP signal are obtained as the medians of σ_w^2 and σ_v^2 respectively for all the acquired signals.

The process noise covariance matrix Q of the respective model is generated using the noise input matrix G as qGG^T, while the measurement noise being scalar, the measurement noise covariance matrix $R = r$.

Step 4: Using the nominal process noise variance q, a variable process noise variance is defined as $q_v = q \times 10^l$ where l varies from $(-10, 10)$.

Step 5: The initial state vector and state error covariance matrix are determined from the initial $(n+1)$ samples of the signal y_k using the two-point differencing method [174]. Using these along with the system matrices and the noise covariances Q and R discussed in Steps 2 and 3, the respective KF_{LPC} and KF_{TS} model parameters are obtained for each q_v. These include the cleaned state vector $\hat{X}_k$ and its error covariance matrix P_k. The cleaned estimate $\hat{x}_k$ of the signal y_k is obtained as the first value of $\hat{X}_k$.

Step 6: The instantaneous metrics J_{1k}, J_{2k} are determined for the nominal q as well as for each q_v. Thereafter the overall sensitivity and robustness metrics, J_1 and J_2, of a particular signal are calculated for the varying q_v. From the plots of J_1 and J_2 for varying q, the parameters J_{1max}, J_{com}, J_{2max}, q_{min}, q_{com}, q_{max}, J_{1min} and J_{2min} are identified for the particular signal.

Step 7: The tonic ($x_{kLPC_{tonic}}$ and $x_{kTS_{tonic}}$) and phasic ($x_{kLPC_{phasic}}$ and $x_{kTS_{phasic}}$) components for both KF models are determined from the respective cleaned signal estimate $\hat{x_k}$ using the same approach as in the ACF model.

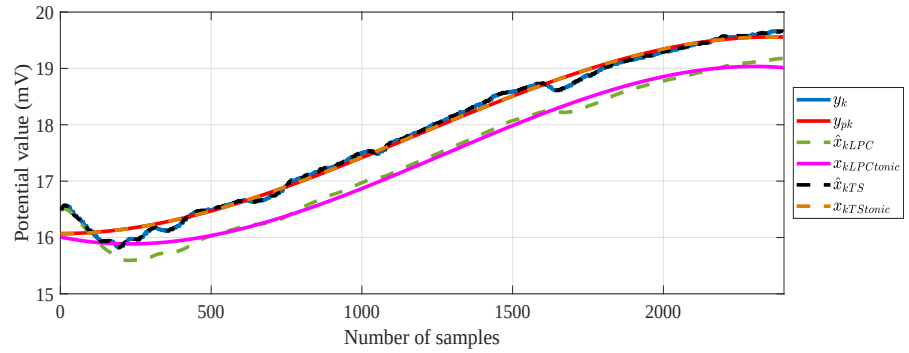
Step 8: A comparative analysis is done of the various parameters of the tonic and phasic components of the ACF, the KF_{LPC} and the KF_{TS} models including a comparison of the noise residuals of the two KF models. The performance metrics J_1 and J_2 of the KF_{LPC} and the KF_{TS} models and their various parameters are also compared.

Step 9: Finally, the effect of restfulness on the tonic and phasic components and the metrics of the KF_{LPC} and the KF_{TS} models are studied and interpreted using the same restfulness data discussed in Section 2.4.

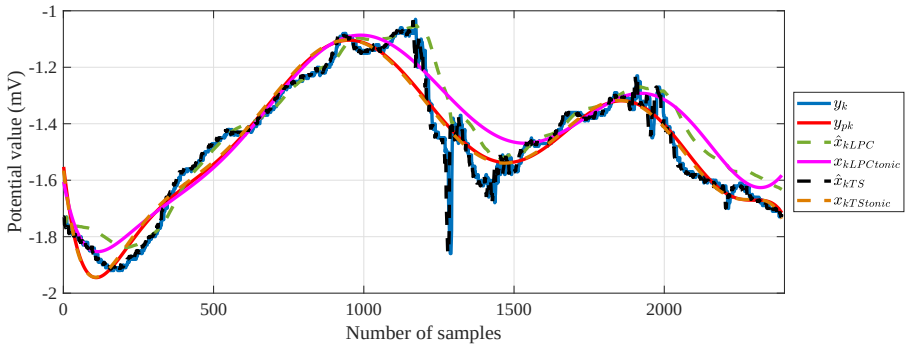
3.5.2 Fitting KF_{LPC} and KF_{TS} models

The time plots of 2 representative signals, y_{k3} of order $n = 3$ and y_{k9} of order $n = 9$, along with their corresponding cleaned KF estimates $\hat{x}_{kLPC}$ and $\hat{x}_{kTS}$ using the KF_{LPC} and KF_{TS} algorithms are shown in Figures 3.1a and 3.1b respectively. These figures also contain the corresponding tonic components y_{pk}, $x_{kLPC_{tonic}}$ and $x_{kTS_{tonic}}$ of the ACF, KF_{LPC} and KF_{TS} models respectively.

It is observed that the signal y_{k3} as well as both its cleaned estimates and all 3 tonic components vary within a span of 5 mV in this duration while the



(a)



(b)

Figure 3.1: Time-series plots of y_k, y_{pk}, $\hat{x}_{kLPC}$, $x_{kLPC_{tonic}}$, $\hat{x}_{kTS}$ and $x_{kTS_{tonic}}$ for (a) an order 3 signal y_{k3} and (b) an order 9 signal y_{k9}

spans of the signal y_{k9}, its cleaned estimates and their tonic components lie within 1 mV in the 2-minutes duration. The cleaned KF_{TS} model estimates of the original signals y_{k3} and y_{k9} are almost identical to the respective signals but the KF_{LPC} estimates show a bias shift as time increases in both cases. Furthermore, $\hat{x}_{kLPC}$ of KF_{LPC} model in both cases is much smoother than the corresponding y_k or $\hat{x}_{kTS}$. As expected, the tonic components y_{pk}, $x_{kLPC_{tonic}}$ and $x_{kTS_{tonic}}$ are smoothened versions of the respective signals or their estimates.

The time plots of the corresponding phasic components w_k, $x_{kLPC_{phasic}}$ and $x_{kTS_{phasic}}$ of the ACF, KF_{LPC} and KF_{TS} models for the y_{k3} and y_{k9} signals are shown in Figures 3.2a and 3.2b respectively. The respective noise residuals r_{kLPC} and r_{kTS} that have been eliminated using the KF algorithms from both the signals y_{k3} and y_{k9} are shown in Figure 3.2c.

All the phasic components lie within ± 0.6 mV. As in the case of the orig-

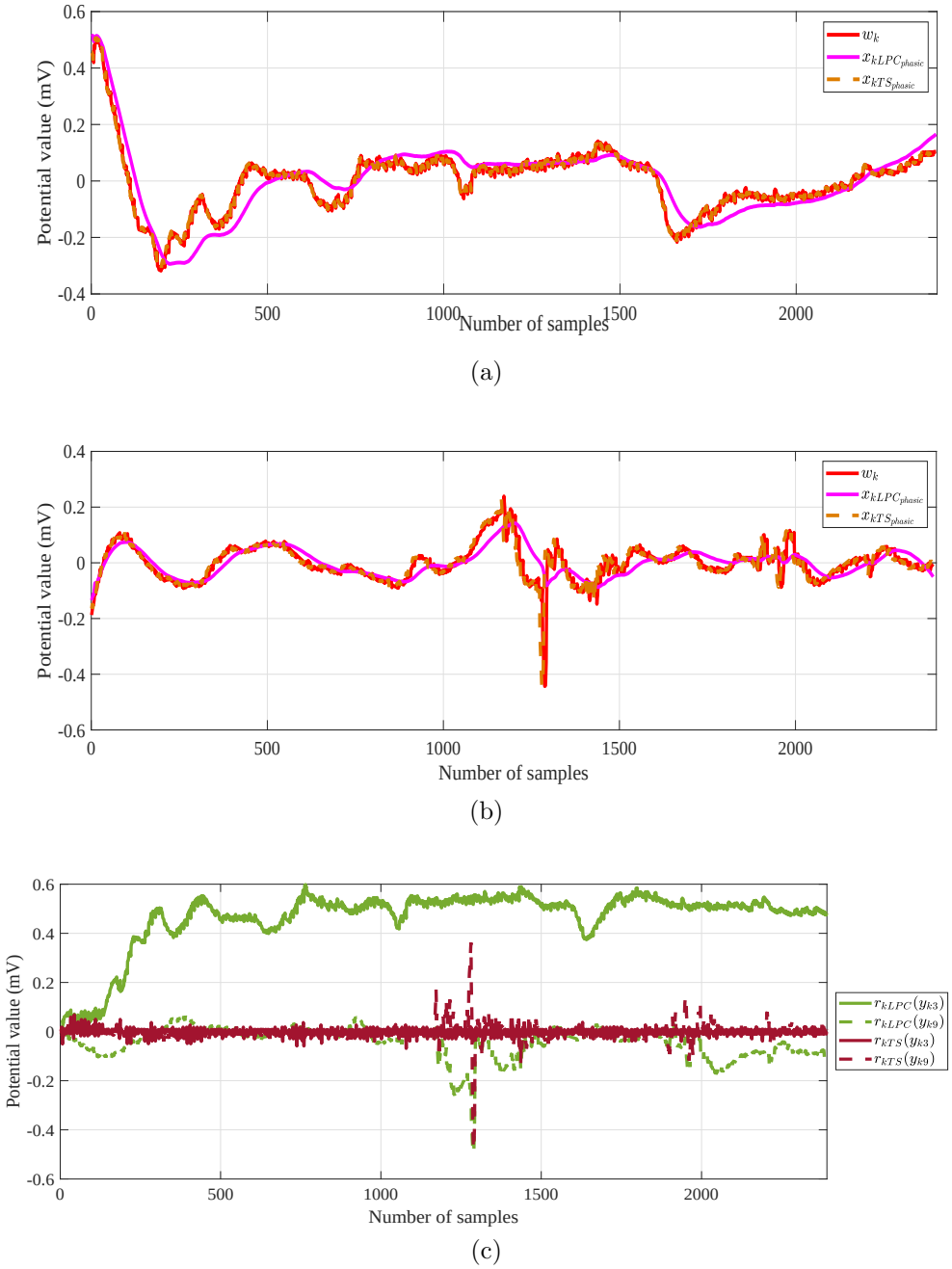


Figure 3.2: Time-series plots of w_k, $x_{kLPC_{phasic}}$ and $x_{kTS_{phasic}}$ of signals (a) y_{k3} (b) y_{k9}; (c) noise residuals r_{kLPC} and r_{kTS} of y_{k3} and y_{k9}

inal signal and its estimates, in these cases also, the ACF model component w_k and the KF_{TS} model component $x_{kTS_{phasic}}$ are almost identical while the

KF_{LPC} model component $x_{kLPC_{phasic}}$ is much smoother than the other two phasic components and distinguishably separate.

Thus, as expected, the residuals r_{kLPC} are much larger as compared to r_{kTS}. In fact, r_{kLPC} of y_{k3} exhibits a large bias of 0.6 mV. However, r_{kLPC} of y_{k9} varies within ±0.2 mV. Both the KF_{TS} model residuals r_{kTS} are primarily limited to ±0.05 mV except for few instantaneous spikes.

3.5.3 Comparison of ACF, KF_{LPC} and KF_{TS} models

In order to validate these preliminary findings, the phasic components of the ACF, KF_{LPC} and KF_{TS} models of 25 other representative signals are shown in Figure 3.3a -3.3c. In all these cases, it is observed that the phasic components vary within ±1 mV. Further, the $x_{kLPC_{phasic}}$ characteristics are much smoother than those of the corresponding w_k and $x_{kTS_{phasic}}$.

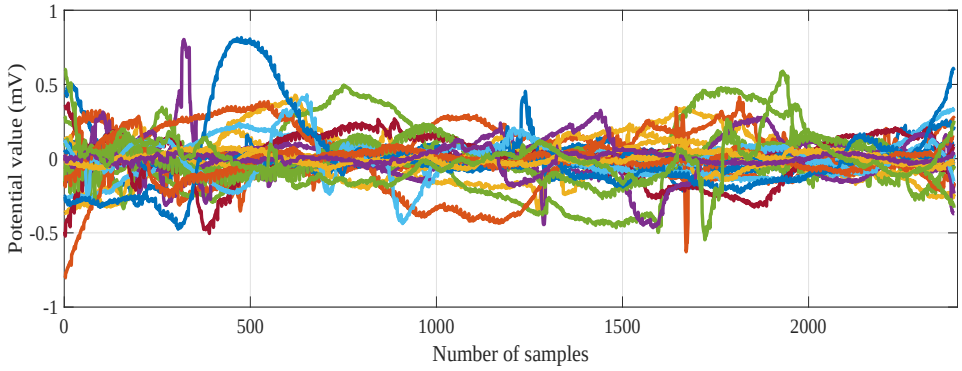
The corresponding KF residuals r_{kLPC} and r_{kTS} of the 25 signals are shown in Figures 3.4a and 3.4b respectively. As expected, these characteristics are significantly different for the two models. The KF_{TS} model residuals are very close to zero, typically within ±0.2 mV and exhibit random, stochastic characteristics. However, the KF_{LPC} model residuals exhibit relatively gradual changing characteristics about an overall bias value. These residuals and their biases typically vary within (−0.5) to (+1) mV.

The statistical characteristics of the tonic and phasic components as well as the noise residuals of the ACF, KF_{LPC} and KF_{TS} models for all 5706 signals are stated in Table 3.2. These parameters are also tabulated in the same Table for 95% of the data about the median. In addition to statistical analysis, the zoomed cdf plots (within 95%) of the tonic components of the three models are shown in Figures 3.5a, 3.5c and 3.5e while those of the corresponding phasic components are shown in Figures 3.5b, 3.5d and 3.5f. The zoomed cdf plots (within 95%) of the noise residuals of the KF_{LPC} and the KF_{TS} models are shown in Figures 3.6a and 3.6b.

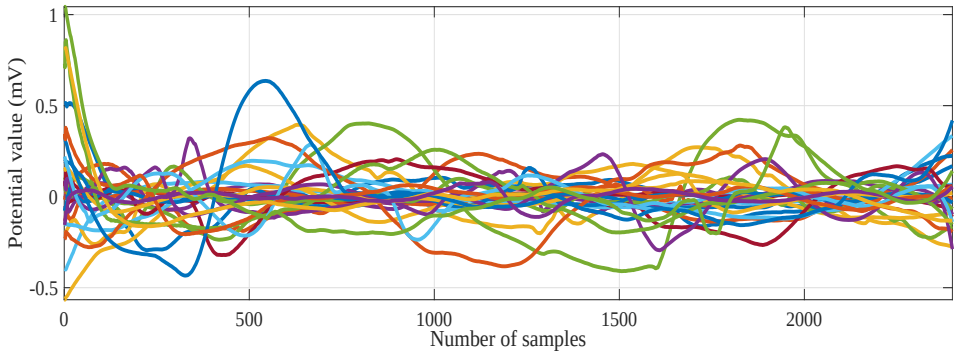
The observations from the statistical analysis and the cdf plots are listed hereafter.

1. All the parameters of the tonic and the phasic components of the ACF and the KF_{TS} models are very close while those of the KF_{LPC} models are distinguishably separate. This is true for the overall as well as the 95% populations.

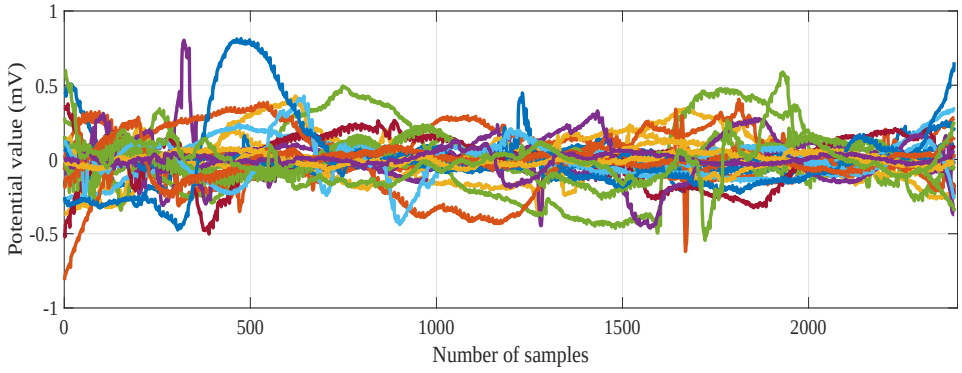
   It is however observed that the 3 cdf plots of the tonic components are similar in nature while the 3 cdf plots of the phasic components are also similar to each other. Furthermore, the phasic component cdf



(a)



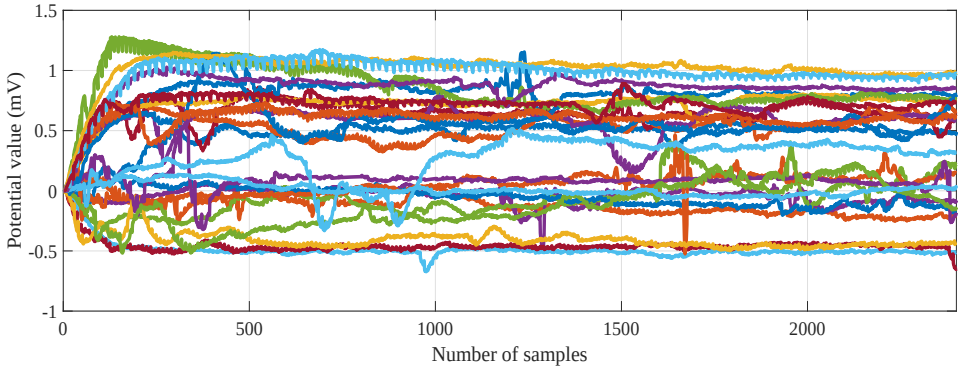
(b)



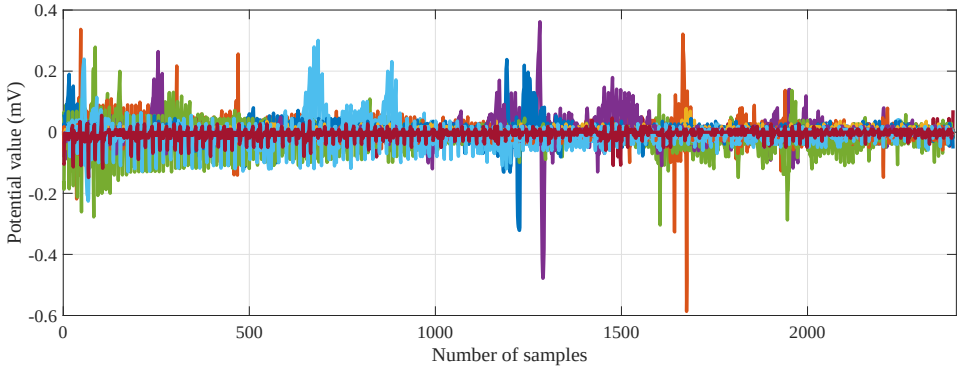
(c)

Figure 3.3: Time-series plots of of (a) $x_{kACF_{phasic}}$ (b) $x_{kLPC_{phasic}}$ and (c) $x_{kTS_{phasic}}$ of 25 representative signals

plots are symmetric and equally skewed about the median with 80% values within ±0.2 mV.



(a)



(b)

Figure 3.4: Time-series plots of (a) r_{kLPC} and (b) r_{kTS} of 25 representative signals

In case of the KF_{LPC} and KF_{TS} model residuals, it is observed that while the cdf plot of r_{kLPC} seems to follow a continuous probability distribution characteristics, the data distribution of r_{kTS} is not smooth as might be expected for noise residuals obtained from cleaning the recorded signals.

2. The minimum and maximum values of the tonic components are almost twice that of the 95% population in all 3 models indicating the presence of significant outliers in 5% of the data.

   This is even more notable in case of the phasic components where 95% of the data in all models are limited within ± 0.8 mV while the overall minimum and maximum values range within $[-250, 72]$ mV for the ACF and the KF_{TS} models and within ± 40 mV for the KF_{LPC} model.

Table 3.2: Statistical analysis of tonic and phasic components and noise residuals

Parameters	Models	Tonic Component		Phasic Component		Noise Residuals	
		Overall	95%	Overall	95%	Overall	95%
Min (mV)	ACF	-224.03	-127.26	-223.37	-0.8	-	-
	KF_{LPC}	-219.73	-122.41	-34.74	-0.6	-231.93	-0.6
	KF_{TS}	-224.03	-126.81	-221.54	-0.8	-89.81	-0.07
Max (mV)	ACF	451.31	198.27	71.08	0.8	-	-
	KF_{LPC}	438.46	195.87	37.93	0.8	58.77	0.8
	KF_{TS}	451.32	197.80	71.69	0.8	221.13	0.07
Mean (mV)	ACF	21.46	20.28	0.00	0.00	-	-
	KF_{LPC}	20.97	20.21	0.00	0.00	0.48	0.45
	KF_{TS}	21.43	20.25	0.00	0.00	0.00	0.00
SD (mV)	ACF	75.98	62.28	0.58	0.19	-	-
	KF_{LPC}	74.34	61.00	0.46	0.17	1.71	1.34
	KF_{TS}	75.93	62.17	0.58	0.19	0.16	0.01
Skewness (mV^3)	ACF	0.49	0.41	-22.08	-0.01	-	-
	KF_{LPC}	0.48	0.44	-0.71	0.43	0.14	0.46
	KF_{TS}	0.49	0.42	-21.25	0.01	549.28	-0.05
Kurtosis (mV^4)	ACF	4.34	3.15	7.89×10^3	6.52	-	-
	KF_{LPC}	4.33	3.16	1.32×10^3	6.03	98.96	3.23
	KF_{TS}	4.34	3.15	7.50×10^3	6.54	6.43×10^5	6.75

Since the origin of the models are different, time series for the KF_{LPC} model and finite differences for the KF_{TS} model, hence the residuals obtained are also different. The corresponding noise residuals of the KF_{LPC} model for 95% population lie within $[-0.6, 0.8]$ mV. This is almost an order larger than that of the KF_{TS} model, which is within ± 0.07 mV. This is indicative of the difference in the noise removal processes using the two KF modelling approaches.

3. CDF plots of the powers of the acquired signals y_k as well as the phasic components of their ACF models ($x_{kACF_{phasic}}$) and the noise residuals (r_{kLPC} and r_{kTS}) of both KF models are shown in Figure 3.7. It is observed that the power of y_k primarily lies within -32dB to +22dB. The power of $x_{kACF_{phasic}}$ lies within the range -64dB to -20dB. The power in the residuals of the KF_{LPC} model r_{kLPC} is typically greater than even that of the ACF phasic component $x_{kACF_{phasic}}$ and lies within -56dB to -10dB. In comparison, the power of r_{kTS} varies within -90dB to -34dB, which is distinctly lesser than that of y_k. Hence, it can be said that of the 2 proposed KF models, maximal signal dynamics are captured in the tonic and phasic components of the KF_{TS} model.

4. The mean of the overall and 95% of the tonic components in all 3 models are quite close to each other and lie typically within 20 to 21.5 mV.

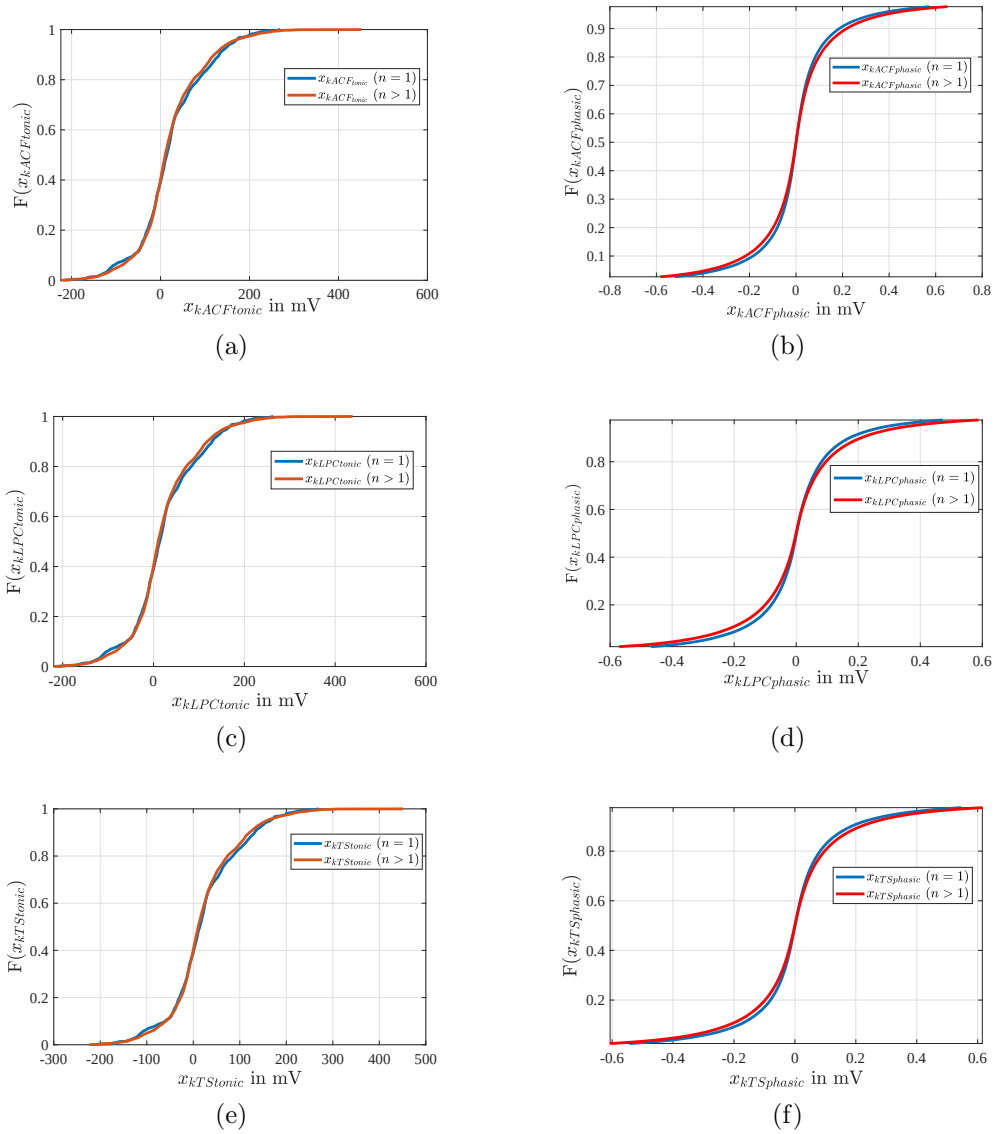
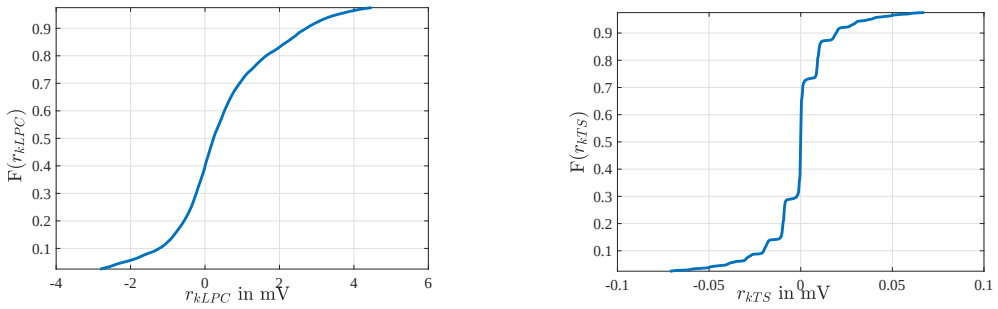


Figure 3.5: CDF plot (95% about median) of 1^{st} and higher order signals: (a) $x_{kACF_{tonic}}$ (c) $x_{kLPC_{tonic}}$ and (e) $x_{kTS_{tonic}}$; (b) $x_{kACF_{phasic}}$(d) $x_{kLPC_{phasic}}$ and (f) $x_{kTS_{phasic}}$

Furthermore, all the phasic components are zero mean, as expected from the approach adopted for determining the phasic components. The noise residuals of the KF_{TS} model are also zero mean but the mean of the noise residuals of the KF_{LPC} model show a bias of 0.45 to 0.48 mV. This indicates the removal, or alternatively an addition, of a



(a) (b)

Figure 3.6: CDF plot (95% about median) of r_{kLPC} and r_{kTS}

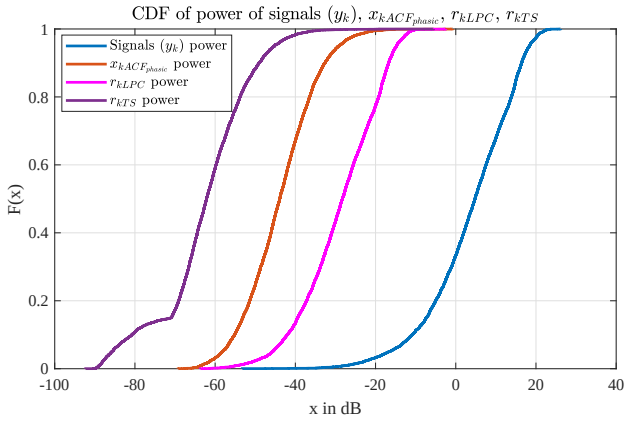


Figure 3.7: CDF plots of the powers of the signal y_k, the ACF model phasic component $x_{kACF_{phasic}}$ and the KF model residuals r_{kLPC} and r_{kTS}

bias in the cleaned KF estimate due to the KF_{LPC} approach.

5. The SD of the overall and 95% of the tonic components in all 3 models are observed to be close to each other. The SD values are almost thrice that of the corresponding mean and lie within 60 to 76 mV.

   The overall SD of the phasic components are 0.58 mV for the ACF and the KF_{TS} models and 0.46 mV for the KF_{LPC} model. However, this reduces to almost 1/3rd when considering 95% of the population with the corresponding values being 0.19 mV for the ACF and the KF_{TS} models and 0.17 mV for the KF_{LPC} model.

   The SD of the noise residuals vary significantly by almost 3 orders of magnitude for the two KF models. The SD values specifically are 1.34 mV and 0.01 mV for 95% of the KF_{LPC} and the KF_{TS} model data

respectively.

A point to note is that the noise residual in case of the KF_{TS} model is of the order of 0.01 mV or 10 μV, which is close to the resolution of the DAS used. These observations are substantiated from the cdf plots of the noise residuals in Figures 3.6a and 3.6b.

6. It is observed that the overall as well as 95% population of the tonic components in all three models are significantly positively skewed with values in the range of 0.41 to 0.49. Considering the 95% population, it is observed that the skewness of the phasic component as well as noise residuals of the KF_{LPC} model are also 0.43 and 0.46 respectively. But the corresponding phasic components of the ACF and the KF_{TS} models as well as the noise residuals of the KF_{TS} model are almost unskewed with values of -0.01, 0.01 and -0.05 respectively.

7. The kurtosis of the tonic components of all 3 models (and the noise residuals of the KF_{LPC} model) are close to that of a normal distribution at 3.15 (3.23) for 95% population. The corresponding 95% phasic components of the ACF and the KF_{TS} models as well as 95% of the noise residuals of the KF_{TS} model exhibit more peakedness with kurtosis values within 6 to 7.

3.5.4 Analysis of Metrics of the KF models

The process noise covariance matrix Q in case of the KF_{LPC} and the KF_{TS} models have the form $Q_{LPC} = q_{LPC}GG^T$ and $Q_{TS} = q_{TS}G_2G_2^T$ respectively. Here G and G_2 are both $(n \times 1)$ constant input vectors of the form

$$G \;=\; G_2 = [1\ 0\ 0\ \cdots\ 0\ 0]^T$$

Thus in order to observe the characteristics of J_1 and J_2 for changing Q, it suffices to vary only the respective scalar process noise covariance q. Let the nominal value of q, whether q_{LPC} or q_{TS}, be denoted as q_{nom}. Then the variable scalar process noise covariance is obtained as $q_v = q_{nom} \times 10^l$ and J_1 and J_2 values are determined for changing l.

The representative plots of J_1 and J_2 of the 2 signals y_{k3} and y_{k9} for varying l are shown in Figure 3.8 for the KF_{LPC} and KF_{TS} models. It is observed that $J_{1min} = J_{2min} = 0$ in all cases. Further, $J_{1max} = J_{2max} = 1$ for the KF_{LPC} model. However, in case of the KF_{TS} model, while $J_{2max} = 1$ for both signals, the values of J_{1max} for y_{k3} and y_{k9} are 0.97 and 0.88 respectively.

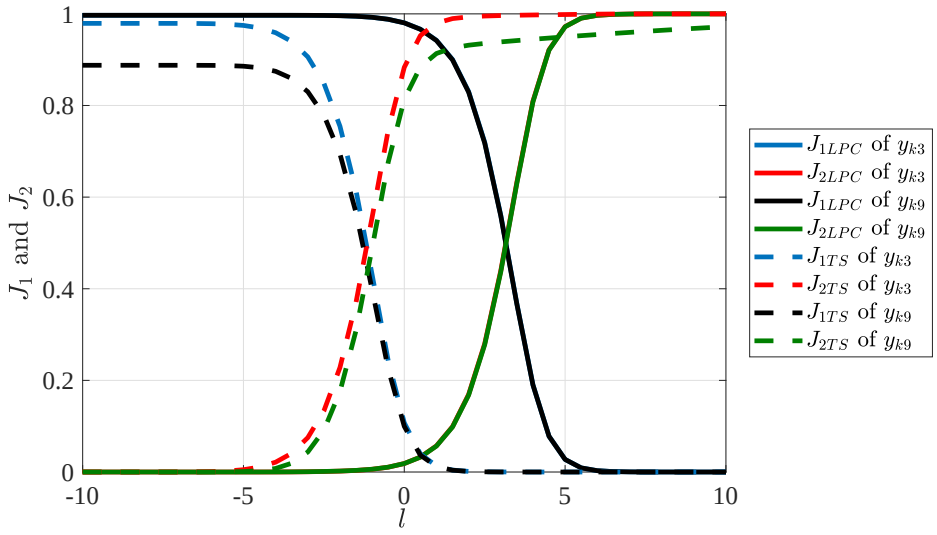


Figure 3.8: Plots of J_1, J_2 for varying l of KF_{LPC} and KF_{TS} models for signals y_{k3} and y_{k9}

Table 3.3: Statistical parameters of J_{1max}, J_{com} and l values of q_{min}, q_{com}, q_{max} of both KF models

Parameters	Models	J_{1max}	J_{com}	l of q_{min}	l of q_{com}	l of q_{max}
Mean (mV)	KF_{LPC}	0.99	0.50	-4.97	3.15	9.00
	KF_{TS}	0.96	0.48	-8.30	-1.16	6.42
SD (mV)	KF_{LPC}	$1.81 \times 10^{-}4$	0.0042	0.40	0.0040	0.01
	KF_{TS}	0.04	0.02	0.37	0.01	5.22
$J_{1min} = J_{2min} = 0$, $J_{2max} = 1$ in all cases						

In all cases, the J_{com} values are almost half of the respective J_{1max} values and vary similarly. Accordingly, q_{com} of the KF_{LPC} model is a fixed value while that of the KF_{TS} model varies slightly. It is also observed that the q_{com} of the KF_{LPC} model is associated with and $l = 3.15$. However, the q_{com} values for the KF_{TS} models are slightly negative at -1.17 and -1.13 for y_{k3} and y_{k9}. This further highlights the difference in the two KF models.

In view of these findings, the mean and SD of J_{1max}, J_{com} as well as the l values of q_{min}, q_{com} and q_{max} for the KF_{LPC} and KF_{TS} models of all 5706 signals have been evaluated and are stated in Table 3.3. It is to be noted that $J_{2max} = 1$ and $J_{1min} = J_{2min} = 0$ in all cases.

As expected, $J_{1max} = 0.99$ in all cases of the KF_{LPC} models as evident from its very low SD value (1.81×10^{-4}). In case of the KF_{TS} models, the average value of J_{1max} is lower (0.96) while the SD of 0.04 ensures that in

some cases, the value reaches 1. Accordingly, while $J_{com} \simeq 0.5$ for all KF_{LPC} models, J_{com} in case of KF_{TS} models has an upper bound of 0.5 but is also lower in several cases, as evident from the mean and SD of 0.48 and 0.02 respectively.

As observed for the 2 representative signals, the l values of q_{min}, q_{com} and q_{max} for the KF_{LPC} models are shifted almost by +3 to +4 as compared to the corresponding values of the KF_{TS} models. The SD of the respective l values are similar for the two models in case of q_{min} (0.40 and 0.37). The respective values for q_{com} (0.004 and 0.01) of both models and that of q_{max} (0.01) of the KF_{LPC} model are almost of the same order. However, it is observed that SD of l for q_{max} (5.22) of the KF_{TS} model is much larger. q_{max} is indicative of J_2 reaching its maximum value and the value of J_2 is governed by the process description of the underlying KF model. Hence, the large variation in the value of q_{max} is reflective of a wide variation in the underlying process dynamics of the KF_{TS} models of the 5706 DDP signals.

The findings from the q_{com} (or its corresponding l value) of the KF models of the DDP signals regarding the underlying system behaviour are opposing in nature. In case of the KF_{LPC} models, mean of l at q_{com} is 3.15. Hence, the nominal system with $l = 0$ lies to the left of this compromise point. This indicates that the nominal system is robust and hence resists, or responds minimally to, any internal or external inputs. On the other hand, since mean of l at q_{com} for the KF_{TS} model is -1.16, hence the nominal system is slightly shifted towards the right of this point. This indicates a sensitive responsive behaviour of the system to incoming internal and/or external inputs.

Thus it may be expected that the metric parameters of the KF_{TS} model will be more responsive to changes in the human condition.

3.6 Restfulness Assessment

In Section 2.4 of Chapter 2, the effect of long term rest on the parameters of the tonic and phasic components of the ACF model have been analyzed on the data acquired in the DS4 data sets. This study is extended in this Section in order to perform a comparative analysis of the effect of restfulness on the ACF model parameters with the parameters of the tonic and phasic model components and the metrics of the KF models.

As mentioned earlier, the DS4 data sets contain total 68 numbers of 10 minute long DDP signals acquired from LH, RH, LL and RL of 16 volunteers over a total duration of 6 months. Thus a total of $68 \times 4 = 272$ sets of independent signals acquired from subjects in restful supine condition for 10 minutes are considered. For restfulness assessment, each 10 minutes signal is

Table 3.4: Statewise median of spectral entropy of tonic and phasic components and noise residuals

Parameters	Models	Median				
		State 1	State 2	State 3	State 4	State 5
Spectral Entropy of phasic component	ACF	0.767	0.772	0.776	0.776	0.782
	KF_{LPC}	0.765	0.771	0.767	0.767	0.766
	KF_{TS}	0.846	0.856	0.859	0.860	0.863
Spectral Entropy of r_k	KF_{LPC}	0.750	0.753	0.753	0.751	0.751
	KF_{TS}	0.942	0.943	0.943	0.943	0.943
Medians of $SE_{ACF_{tonic}}$= 0.750, $SE_{KF_{LPCtonic}} = SE_{KF_{TStonic}}$ =0.756						

Table 3.5: Statewise IQR of spectral entropy of tonic and phasic components and noise residuals

Parameters	Models	IQR				
		State 1	State 2	State 3	State 4	State 5
Spectral Entropy of phasic component	ACF	0.037	0.041	0.047	0.048	0.054
	KF_{LPC}	0.012	0.024	0.027	0.025	0.025
	KF_{TS}	0.058	0.058	0.057	0.054	0.054
Spectral Entropy of r_k	KF_{LPC}	0.002	0.012	0.010	0.005	0.003
	KF_{TS}	0.039	0.045	0.054	0.055	0.058
IQR of $SE_{ACF_{tonic}} \doteq SE_{KF_{LPCtonic}} \doteq SE_{KF_{TStonic}} \doteq 0$						

divided into 5 subsets (or states) of 2 minutes window, yielding $272 \times 5 = 1360$ numbers of 2-minute duration signals (with $N = 2400$ discrete samples each) for the analysis. KF_{LPC} and KF_{TS} models are generated for each of these 2 minutes signals and their tonic and phasic components as well as their metrics are evaluated as discussed in Section 3.5.1.

3.6.1 Spectral entropy of models

Spectral entropy (SE) is known to be able to differentiate regular and irregular system dynamics [186]. It is further established that SE of regular or ordered systems are lower in value and as disorder or chaos increases in a system, its SE value increases and so the upper threshold of SE is useful in distinguishing states with varying stochasticity [187], [188].

The box plots of SE of the tonic and phasic components of the ACF, KF_{LPC} and KF_{TS} models for all 1360 signals in the 5 rest states are shown in Figures 3.9a -3.9f while the box plots of the SE of the noise residuals of the KF_{LPC} and KF_{TS} models are shown in Figures 3.10a and 3.10b respectively. The medians and Inter Quartile Range (IQR) of all these parameters are listed in Table 3.4 and Table 3.5 respectively.

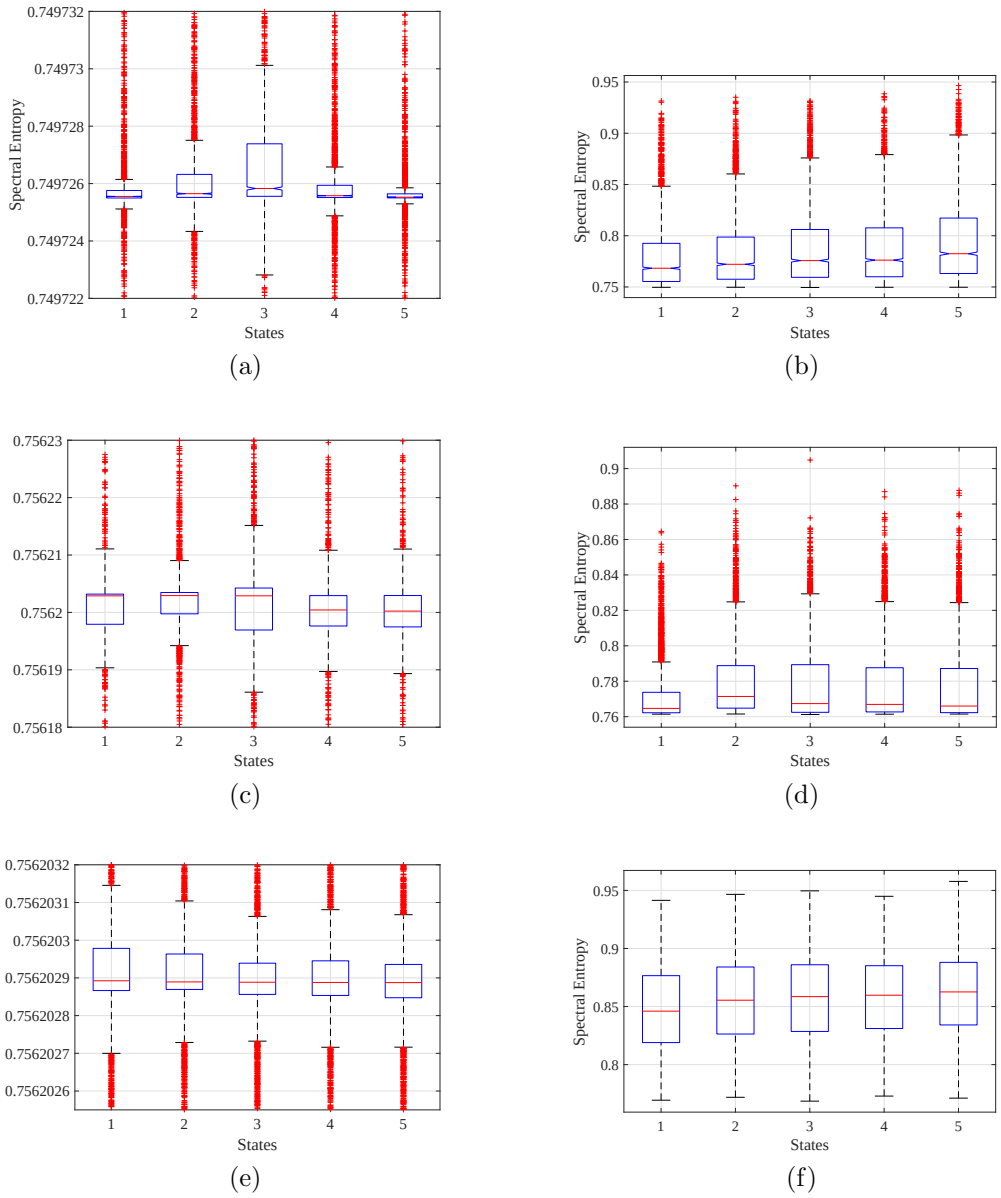


Figure 3.9: Box plots of spectral entropy (SE) in 5 rest states of (a) $x_{kACF_{tonic}}$, (b) $x_{kACF_{phasic}}$, (c) $x_{kLPC_{tonic}}$, (d) $x_{kLPC_{phasic}}$, (e) $x_{kTS_{tonic}}$ and (f) $x_{kTS_{phasic}}$

The observations from the box plots of the SE of the various model components and from the Tables listing the corresponding median and IQR values are stated hereafter.

1. The SE of the tonic components of the ACF, KF_{LPC} and KF_{TS} models are henceforth denoted as $SE_{ACF_{tonic}}$, $SE_{KF_{LPCtonic}}$ and $SE_{KF_{TStonic}}$

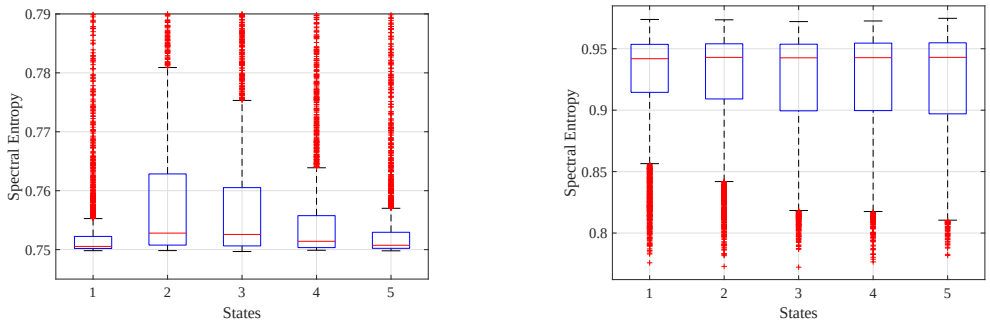


Figure 3.10: Box plots of SE in 5 rest states of (a) r_{kLPC} (b) r_{kTS}

respectively while that of the corresponding phasic components will be denoted as $SE_{ACF_{phasic}}$, $SE_{KF_{LPCphasic}}$ and $SE_{KF_{TSphasic}}$ respectively. The SE of the noise residuals of the KF_{LPC} and KF_{TS} models will be denoted as $SE_{KF_{LPCresiduals}}$ and $SE_{KF_{TSresiduals}}$ respectively.

2. In all 3 models, it is observed that the SE of the tonic component is almost a fixed value close to 0.75 with the IQR = 0 in all cases. Specifically, $SE_{ACF_{tonic}}$=0.750 and $SE_{KF_{LPCtonic}}$=$SE_{KF_{TStonic}}$=0.756.

   Since the SE of the corresponding phasic components are typically more than 0.75, thus in this case also it is inferred that the tonic components are representative of the regular dynamics of the system while the phasic components are indicative of the system stochasticity in all 3 models for all acquired signals.

3. As mentioned at the outset, the SE of all tonic parameters are practically fixed values across all states.

   However, it is evident from the zoomed box plots in Figures 3.9c and 3.9a that the IQR of $SE_{KF_{LPCtonic}}$ and $SE_{ACF_{tonic}}$ are both maximum in state 3. The corresponding box plots of $SE_{KF_{TStonic}}$ in Figure 3.9e are almost invariant across all states.

4. In case of $SE_{ACF_{phasic}}$, its lower threshold is fixed at 0.75. However, as discussed in Section 2.4.2, its median, IQR and upper threshold increase monotonically. The median increases from 0.767 in state 1 till 0.782 in state 5. The corresponding IQR also increases from 0.037 in state 1 till 0.054 in state 5.

5. The median of $SE_{KF_{LPCphasic}}$ increases from 0.765 in state 1 to 0.771 in state 2 but thereafter drops to 0.767 in state 3 and stabilizes.

   The corresponding IQR increases from 0.012 in state 1 to 0.027 in state 3 but then drops and stabilizes at 0.025. This pattern of change is reflected in the upper threshold too since the lower threshold is fixed at 0.76 across all states.

6. The lower threshold of $SE_{KF_{LPCresiduals}}$ is fixed at 0.75 across the 5 states while the upper threshold, 3rd quartile and median increase sharply from state 1 to state 2 and then decrease till state 5. The median changes from 0.750 in state 1 to 0.753 in state 2 and 3 and then drops and stabilizes at 0.751.

   Since these values overlap those of the tonic and phasic components, this indicates that certain significant regular dynamics are eliminated as residuals in the KF_{LPC} approach.

   The associated IQR increases from 0.002 in state 1 to 0.012 in state 2 and 0.010 in state 3 but then decreases sharply till 0.003 in state 5. It is to be noted that these IQR values are almost one order lower than those of the corresponding $SE_{KF_{TSresiduals}}$ indicating lesser variability.

7. In case of the KF_{TS} model, it is observed that the box plots as well as the median and the IQR range are fairly constant across all the states, irrespective of whether it is that of the tonic component or the phasic component or the noise residual.

   Thus in case of the KF_{TS} model, the SE value is not a differentiating parameter for the 5 states of rest. It is instead an identifier of a particular component irrespective of the stage of rest. This is most evident in case of the phasic component where the box plot has no outliers in any of the 5 states and yet the overall ranges are fairly constant.

8. Furthermore, as mentioned earlier, the $SE_{KF_{TStonic}}$ is typically 0.756 for all states with IQR practically 0.

   The $SE_{KF_{TSphasic}}$ also has an almost fixed lower threshold of 0.76 and upper threshold close to 0.95. The median increases from 0.846 in state 1 to 0.863 in state 5.

   In case of $SE_{KF_{TSresiduals}}$, its upper threshold is fixed at 0.975. The median is fixed at 0.943. The associated IQR varies from 0.039 in state 1 to 0.058 in state 5.

   The medians and associated IQR ranges (as evident from the box plots) of the SE of the tonic, phasic and residuals of the KF_{TS} model are

distinct, non-overlapping and in increasing order. This validates the increasing randomness of the three components.

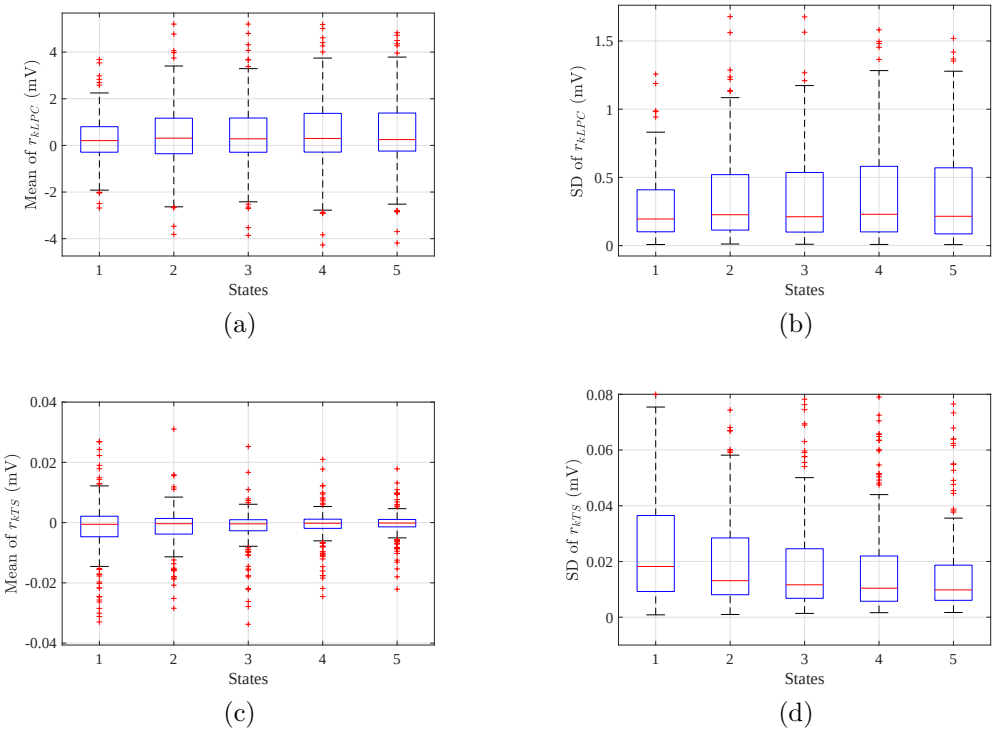


Figure 3.11: Box plots of mean of (a) r_{kLPC} and (c) r_{kTS} and SD of (b) r_{kLPC} and (d) r_{kTS} in 5 rest states

3.6.2 Effect on noise residuals of KF models

It is expected that the noise residuals of the two KF models will effectively eliminate the process and measurement noises in affecting the signal acquisition while retaining all pertinent information. Hence, it is expected that the bulk of the noise residuals will have significantly lower thresholds than the corresponding signal and will be random in nature and will be uncorrelated with the increasing duration of rest from state 1 to state 5.

Thus, prior to the detailed study of the model parameters, the suitability of the KF models for this application are studied in terms of the characteristics of their noise residuals. The box plots of the mean and SD of r_{kLPC} and r_{kTS} in the 5 rest states are shown in Figures 3.11a - 3.11d and their

Table 3.6: Statewise median of noise residuals of KF models

Parameters	Models	Median				
		State 1	State 2	State 3	State 4	State 5
Mean of r_k (mV)	KF_{LPC}	0.21	0.31	0.28	0.29	0.24
	KF_{TS}	~ 0	~ 0	~ 0	~ 0	~ 0
SD of r_k (mV)	KF_{LPC}	0.19	0.22	0.21	0.22	0.22
	KF_{TS}	0.018	0.013	0.011	0.010	0.009

Table 3.7: Statewise IQR of noise residuals of KF models

Parameters	Models	IQR				
		State 1	State 2	State 3	State 4	State 5
Mean of r_k (mV)	KF_{LPC}	1.09	1.51	1.46	1.95	1.62
	KF_{TS}	0.006	0.004	0.001	0.003	0.002
SD of r_k (mV)	KF_{LPC}	0.3	0.41	0.44	0.48	0.49
	KF_{TS}	0.027	0.02	0.015	0.016	0.012

corresponding median values and IQR are stated in Tables 3.6 and 3.7 respectively.

The observations from the box plots and the statistics of the median and IQR of the noise residuals are stated hereafter.

1. The mean of r_{kLPC} exceeds those of r_{kTS} by 3 orders of magnitude. While mean values of all r_{kTS} lie within $\pm$0.4 mV, majority of the mean values of r_{kLPC} lie within $\pm$ 4 mV. This is evident from the values of the respective median and IQR values also across all states.

2. The medians of mean of r_{kTS} across all states are practically 0 and the IQR values lie within 1 to 6 μV in all the states. It is reiterated that this matches the noise threshold of the DAS used for the signal acquisition.

3. The medians of the mean of r_{kLPC} lie within 0.20 mV to 0.31 mV and the IQR in all states exceed 1 mV and is 1.95 mV for state 4. These values are commensurate with actual signal levels in several cases.

4. In case of the SD values, the SD of r_{kLPC} exceeds those of r_{kTS} by at least 1 order of magnitude. While most SD values of r_{kTS} lie within 0 to 0.08 mV, those of r_{kLPC} lie within 0 to 1.2 mV. This is evident from the values of the respective median and IQR values also across all states.

5. The medians of the SD of r_{kLPC} are typically 0.22 mV across all states. The IQR increases from 0.3 mV in state 1 to 0.49 mV in state 5. The overall noise residual range increases state 1 till state 5 for this model.

6. The medians of the SD of r_{kTS} decrease from 0.018 mV to 0.009 mV from state 1 till state 5 and the corresponding IQR also decreases from 0.027 mV in state 1 to 0.012 mV in state 5. The overall noise residual range, barring the outliers, also decreases similarly from state 1 till state 5. SD of r_{kTS} indicates dynamics in the signal.

   It is to be recalled that the residuals of the two models are obtained as the difference of the true signal y_k and the KF estimates of the two models $\hat{x_k}$ as stated in equations (3.56)-(3.58). So this difference in the characteristics of the SD of the two noise residuals may be due to the fact that while the KF_{LPC} model utilizes the dynamic change of an n-th order time-series, the KF_{TS} model is based on the n-th order finite differences, where n denotes the model order in both cases. Furthermore, it can be hypothesized that it is the characteristic change of the neglected higher order dynamics ($n_{max} = 9$ in this study) due to rest that is evident in the SD of r_{kTS}.

3.6.3 Effect on tonic, phasic and metric parameters

Finally, the effect of increasing duration of rest on tonic and phasic components of the ACF, KF_{LPC} and KF_{TS} models as well as the metric parameters of the two KF models are compared and analyzed.

The box plots of the mean and SD of the tonic components y_{pk}, $x_{kLPC_{tonic}}$, $x_{kTS_{tonic}}$ of the 3 models in the 5 rest states are shown in Figure 3.12a to 3.12f. The mean of all the phasic components are identically zero by design. So in case of the phasic components, only the box plots of the SD of $w_k = x_{kACF_{phasic}}$, $x_{kLPC_{phasic}}$ and $x_{kTS_{phasic}}$ of the 3 models in the 5 rest states are shown in Figure 3.13a to 3.13c. From the preliminary findings in Section 3.5.4, it is observed that J_{com} and l of q_{com} are the 2 key parameters to be observed from the metrics. The box plots of these parameters of the two KF models are shown in Figures 3.14a to 3.14d. The median and IQR of these parameters for all 5 states of rest are tabulated in Tables 3.8 and 3.9.

The observation from these box plots and associated tables of the medians and IQR are stated hereafter.

1. The ranges of y_{pk}, $x_{kLPC_{tonic}}$ and $x_{kTS_{tonic}}$ are similar as evident from the box plots. The medians for all 3 models and across all 5 states lie within 13.88 to 16.10 mV.

Table 3.8: Statewise median of tonic, phasic and metric components

Parameters	Models	Median				
		State 1	State 2	State 3	State 4	State 5
Mean of tonic component (mV)	ACF	15.72	15.96	16.10	15.69	14.17
	KF_{LPC}	15.50	15.70	15.82	15.40	13.88
	KF_{TS}	15.69	15.93	15.08	15.68	14.14
SD of tonic component (mV)	ACF	1.68	1.15	0.82	0.71	0.57
	KF_{LPC}	1.66	1.15	0.82	0.72	0.57
	KF_{TS}	2.83	2.22	1.79	1.39	1.51
SD of phasic component (mV)	ACF	0.19	0.13	0.11	0.09	0.08
	KF_{LPC}	0.19	0.14	0.14	0.12	0.11
	KF_{TS}	0.18	0.13	0.11	0.09	0.08
J_{com}	KF_{LPC}	0.49	0.49	0.49	0.49	0.49
	KF_{TS}	0.493	0.493	0.493	0.498	0.493
l of q_{com}	KF_{LPC}	2.68	2.97	2.98	3.05	3.03
	KF_{TS}	-1.01	-0.93	-0.75	-0.60	-0.57

Table 3.9: Statewise IQR of tonic, phasic and metric components

Parameters	Models	IQR				
		State 1	State 2	State 3	State 4	State 5
Mean of tonic component (mV)	ACF	86.09	86.73	88.67	80.92	85.22
	KF_{LPC}	84.80	85.17	87.07	79.23	83.68
	KF_{TS}	86.06	86.66	88.61	80.75	85.10
SD of tonic component (mV)	ACF	2.62	1.8	1.27	1.11	0.95
	KF_{LPC}	2.58	1.77	1.27	1.06	0.92
	KF_{TS}	3.51	3.03	2.67	2.33	2.21
SD of phasic component (mV)	ACF	0.27	0.21	0.18	0.15	0.14
	KF_{LPC}	0.25	0.21	0.17	0.23	0.17
	KF_{TS}	0.27	0.21	0.18	0.15	0.14
J_{com}	KF_{LPC}	~ 0	~ 0	~ 0	~ 0	~ 0
	KF_{TS}	0.01	0.01	0.02	0.02	0.02
l of q_{com}	KF_{LPC}	0.001	0.001	0.001	0.00	0.00
	KF_{TS}	0.01	0.01	0.01	0.01	0.01

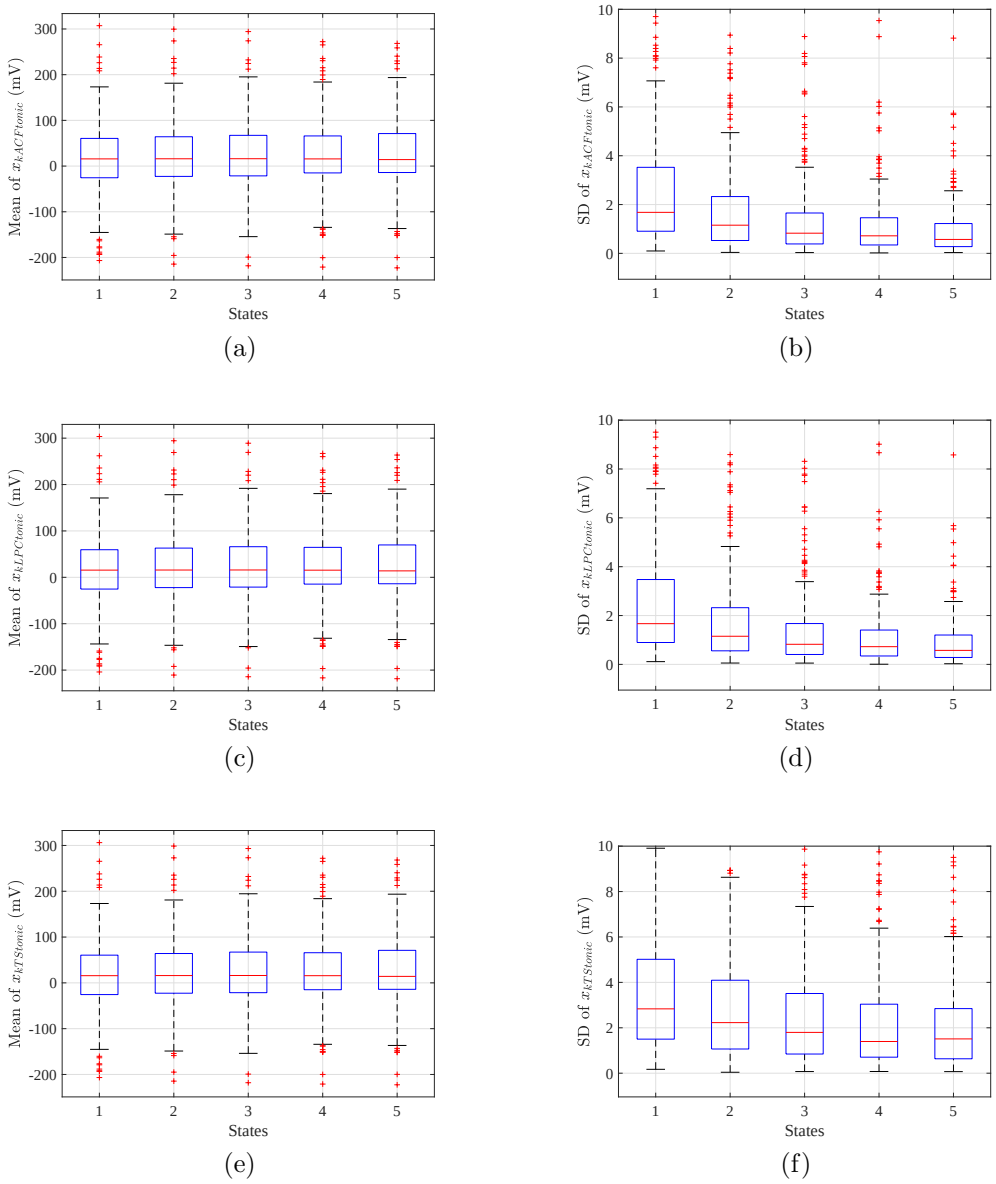


Figure 3.12: Box plots of mean of (a) $x_{kACF_{tonic}}$ (c) $x_{kLPC_{tonic}}$ (e) $x_{kTS_{tonic}}$ and SD of (b) $x_{kACF_{tonic}}$ (d) $x_{kLPC_{tonic}}$ (f) $x_{kTS_{tonic}}$

It was inferred in Section 2.4 that y_k and y_{pk} are identical in ACF model and y_{pk} represents the tonic component of DDP signals. The similarity of the box plots of all tonic components leads to the inference that $x_{kLPC_{tonic}}$ and $x_{kTS_{tonic}}$ are representative of the tonic components of DDP signals in the KF_{LPC} and KF_{TS} models.

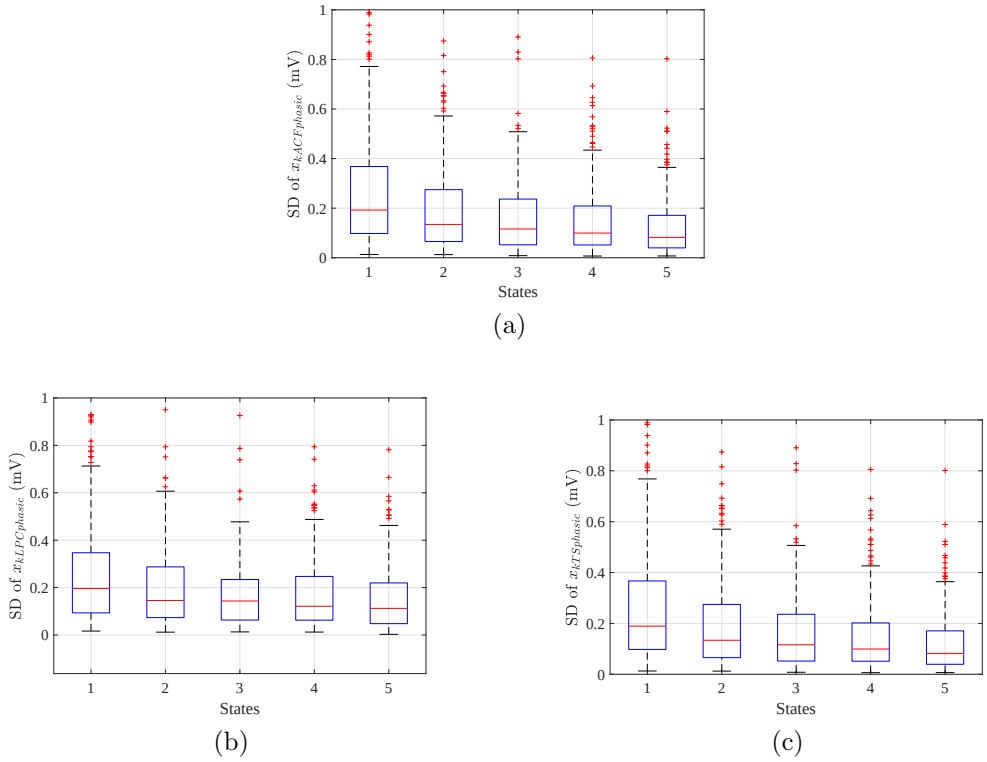


Figure 3.13: Box plots of SD of (a) $w_k = x_{kACF_{phasic}}$, (b) $x_{kLPC_{phasic}}$, and (c) $x_{kTS_{phasic}}$ in 5 rest states

2. The nature of change of y_{pk} and $x_{kLPC_{tonic}}$ are similar across all states. Their medians increase till state 3 and then decrease quite sharply till state 5.

   However, $x_{kTS_{tonic}}$ varies in a slightly different manner. Its median increases in state 2 from state 1, then decreases in state 3, then increases in state 4 and then decreases again in state 5.

   The IQR values in all 3 models vary similarly. These values increase till state 3, then drop in state 4 but then increase again in state 5.

3. The SD of y_{pk} and $x_{kLPC_{tonic}}$ vary similarly within 0 to 7.5 mV. The 1st and 3rd quartiles, median and upper threshold decrease almost identically from state 1 to state 5 in an exponential manner. The lower threshold is very close to 0 in all cases.

   The corresponding SD values for $x_{kTS_{tonic}}$ also decrease monotonically quasi-linearly from state 1 to state 5 with the exception of the median in state 5 which is slightly larger than that in state 4. In general, all

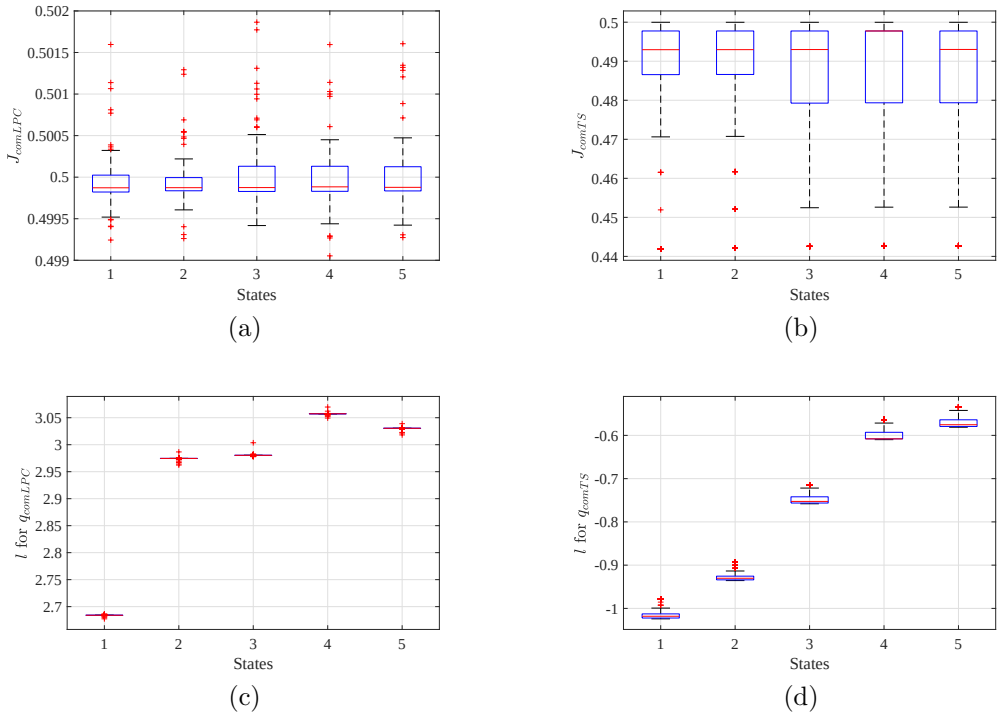


Figure 3.14: Box plots of (a) J_{comLPC} (b) J_{comTS}, l of (c) q_{comLPC} and (d) q_{comTS}

these values are typically larger by 1 mV and their ranges of variation are also larger, ranging from 0 to 10 mV in state 1 till 0 to 6 mV in state 5. In this case also, the lower threshold is practically 0 in all states.

4. It is reiterated that owing to the method of determining the tonic and phasic components, the mean of the phasic component is zero for all 3 signal models.

5. It has been discussed in Section 2.4.3 that the SD of actual DDP signals [24] and that of $w_k = x_{kACF_{phasic}}$ change in the same manner due to the effect of long duration of rest. It is further observed from the box plots as well as the median and IQR that the SD of $x_{kACF_{phasic}}$ and $x_{kTS_{phasic}}$ vary almost identically across all states.

   Thus the information of the SD of the DDP signals is retained as the SD of the phasic component in the ACF and KF_{TS} models.

   The nature of change of the SD of $x_{kLPC_{phasic}}$ is, however, quite different.

This is evident from the overall range as well as the IQR of this SD. These decrease till state 3 and then increase slightly in state 4 and then decrease to the level of that in state 3 yet again.

6. Comparing the box plots of J_{com} of the KF_{LPC} and KF_{TS} models, it is observed that the median of J_{comLPC} across all states is 0.49 and this is practically a fixed value with an IQR and overall range less than 0.001.

   In case of J_{comTS}, its upper threshold across all states is 0.5 while its median is 0.493 in all states except in state 4 (0.498). The IQR is 0.01 in states 1 and 2 and thereafter increases to 0.02 for the remaining states. The lower threshold also changes accordingly.

   Thus it is observed that this parameter is quite close to $n/2 = 0.5$ in both models and almost invariant across all states.

7. In case of the l of q_{com} in the KF_{LPC} model, it is observed that l of q_{comLPC} is close to +3, varying from +2.68 in state 1 to 2.97 in state 2 till it finally reaches 3.03 in state 5.

   The l of q_{comTS} increases systematically from -1.01 till -0.57. These values are practically constant as evident from the box plots as well as the very low IQR values.

   Thus the inference of the system behaviour from the two models remains the same as discussed in Section 3.5.4.

8. In case of l of q_{comTS}, it is further observed that the medians as well as the IQR around the median are distinct and non-overlapping and increasing from state 1 till state 5. Except in the final stable rest state, the overall ranges are also distinctly non-overlapping.

   This feature is thus a strong demarcator of the stages of rest.

3.7 Discussions

It is known that recorded signals are corrupted by sensor, or measurement, noise and the sensor noise cleaned signal may further be corrupted by system noises or inherent process disturbances. Two Kalman Filter models are developed in this Chapter to obtain cleaned estimates of the DDP signals that can be used for various applications. The first step for this is to formulate the corresponding state space models of the system outlining the process and measurement noises affecting the system. The first KF model uses the standard Linear Prediction Coefficients (LPC) based state space model and

the corresponding KF model is labeled the KF_{LPC} model. The second state space model is derived ab-initio from the time-series (TS) polyfit obtained in the ACF model. The corresponding KF model is termed the KF_{TS} model.

The LPC and TS state space models differ not only in the states and the system matrices, but also in the way the noises are defined. In case of the LPC model, the process noise covariance is obtained from the LPC determination process and the measurement noise is derived as the difference between the signal and the corresponding process output. In the TS model, the measurement noise is obtained from the sensor noise calibration data and the process noise is obtained as a difference between the ACF model residual and the measurement noise. It is to be mentioned that the tonic component of the ACF model and the TS state space model are assumed to be identical and forms the basis for deriving the TS state space description.

Using the respective forms of the state space models, the KF_{LPC} and KF_{TS} models of the DDP signals are obtained using standard KF equations. Since the DDP signal is a one-dimensional signal, hence the simplified scalar forms of the robustness and sensitivity metrics are derived and interpreted for these signals. It is observed that while the robustness metric decides the goodness of the process model description, the sensitivity metric is a direct measure of the Kalman gain. The error covariance of the aposteriori, or post-measurement, cleaned estimate is related directly to the sensitivity metric and inversely to the robustness metric. Thus the value at which these two metrics are balanced, denoted as J_{com}, is a critical KF parameter. The error covariance of the aposteriori, or post-measurement, cleaned estimate is also proportional to the scalar process noise covariance q and hence the value of q for J_{com}, denoted as q_{com} is another critical KF parameter.

Based on these model and metric descriptions, the two KF models are fitted to the DDP signals using a well-defined systematic approach. The order of the signals are determined based on the l_{zc} as detailed in the ACF model. Thereafter, the KF models and their robustness and sensitivity metrics are determined. Thereafter, the tonic and phasic components of the respective cleaned estimates are determined using the same approach as used for the ACF model.

These KF models are fitted to the same 5706 signals as discussed in the ACF model. It is observed that the tonic and phasic components of the KF_{TS} models are almost identical to those of the ACF model, while those of the KF_{LPC} models are distinguishably separate. In all cases, 80% of the phasic components lie within ± 0.2 mV. The KF_{LPC} residuals show a bias of 0.45 to 0.48 mV with SD of 1.34 mV for 95% population, while the KF_{TS} residuals are zero mean with 0.01 mV SD for 95% population as expected and SD of r_{kTS} indicates dynamics in the signal. The SD of the phasic components in

all 3 models lie within 0.19 mV for 95% of the population.

The sensitivity metric J_1 and its maximum value J_{1max} show a variability in case of the KF_{TS} models while those of the KF_{LPC} models are fixed in nature. Accordingly, the two key parameters J_{com} and q_{com} are variable for the KF_{TS} models and invariant for the KF_{LPC} models. Furthermore, the values of q_{com} for the two KF models are opposing in nature. The positive values of q_{com} for the KF_{LPC} models indicating robust nominal system behaviour while the negative values of q_{com} for the KF_{TS} models indicate sensitive model behaviour.

Finally, the effect of rest on the various parameters of the two KF models have been studied using the same 272 number of 10 minute long signals and methodology described in Section 2.4. These include the spectral entropies of the tonic and phasic components and the noise residuals of the KF models, the mean and SD of the tonic components, SD of the phasic components and the two metric parameters J_{com} and q_{com}.

The spectral entropy (SE) of the tonic and phasic components of the KF_{LPC} models are observed to increase till the second or third state of rest and then decrease. However, the spectral entropies of the tonic and phasic components and the noise residuals of the KF_{TS} models are observed to be fairly constant across all stages of rest. Further, their medians and ranges are observed to increase from tonic components to phasic components and are maximum for the noise residuals indicating sequential departure from regularity to irregularity as expected. Thus the SE of the KF_{TS} model components are useful identifiers.

The mean of the tonic components of both the KF models vary similarly as that of the ACF model. The SD of the tonic components of the ACF and KF_{LPC} models vary identically, decreasing exponentially with increasing rest. While the SD of the tonic components of the KF_{TS} models also decrease monotonically with increasing rest, yet their nature of decrease is quasilinear. The SD of the phasic components of the ACF and KF_{TS} models vary identically, decreasing exponentially with increasing rest while that of the KF_{LPC} models decrease initially and then increase and decrease slightly to become almost stationary with increasing rest.

As expected, the nature of change of the two metric parameters J_{com} and q_{com} with rest are distinctly different for the two KF models. Both these parameters remain fairly constant over changing duration of rest for the KF_{LPC} models. J_{com} of the KF_{TS} models have a fixed upper bound of 0.5 for all stages of rest but the lower bound drops sharply in the third state and then becomes constant. The corresponding q_{com} increases sharply and distinctly with increasing rest providing a distinguishing parameter for the stages of rest.

Thus, the KF_{TS} model developed ab-initio using the system order and polyfit obtained in the earlier proposed ACF model provides consistent tonic and phasic components. The key features of these components and the KF metrics can be used to reliably differentiate the stages of rest. The standard KF_{LPC} model also performs creditably in several aspects though not so markedly.

Chapter 4

Efficacy of Models in classifying Human Conditions

In Chapters 2 and 3, the ACF, KF_{LPC} and KF_{TS} models have been developed ab initio and the effect of restfulness on their tonic, phasic and other parameters have been studied. As mentioned in Chapter 1, this is one of the human conditions studied in [1], [24], [189]. The classification of two human conditions, namely (i) hypertensive and normotensive subjects and (ii) sitting and supine postures, have also been tested and discussed in [1], [189].

In this Chapter, the objective is to study the efficacy of the parameters of the ACF, KF_{LPC} and KF_{TS} models of the DDP signals in the two aforementioned classification applications. The open source data mining software WEKA, version 3.9.4 [1, 190, 191] is used for these binary classification studies. The same DDP signal datasets used for the hypertensive-normotensive and the posture change classifications, namely DS2 and DS1 respectively [1], [189], are used in the present study also.

This Chapter is organized as follows. The experimental methodology used for the classification studies is described in Section 4.1. The type of data sets used and their pre-processing, as discussed in [1] are stated at the outset. Thereafter, the details of the signal quantization performed for the two studies, the model parameters used for the classifications, the understanding of the feature selection and ranker methods tested, the background of the classification metrics evaluated as well as the cross-validation protocols used for the studies are discussed in Sections 4.1.1 to 4.1.5. This is followed by the details of the hypertensive and normotensive classification in Section 4.2 and the posture classification in Section 4.3. In each case, the selection of ranker, the attributes selection and the cross validation results for the 3 models and their analysis are discussed in detail in the subsequent subsections. Section 4.4 contains the overall discussions.

4.1 Experimental Methodology

The objective of this study is to ascertain whether the presence or absence of a human condition like hypertension or the change in human dynamics like posture change affects one or more parameters of each of the 3 models. Is the effect significant enough to enable viable classification and if so, to what extent? Thus the classification is done by considering the parameters of any one, out of the three, models at a time.

Classification 1: Hypertensive and Normotensive Subjects
The DS2 dataset as discussed in Chapter 1 is used for this binary classification. In this, 10 minutes long signals have been acquired simultaneously from the left hand (LH) and right hand (RH) of a subject. This data has been collected from 5 hypertensive and 17 normotensive subjects, that is a total of 22 subjects, in supine posture.

Classification 2: Sitting and Supine Postures
In this case, the dataset used is DS1 in which 20 minutes long signals have been acquired only from the left hand (LH) of a subject. This data has been collected from 24 subjects. As described earlier in Chapter 1, the data is acquired for the initial 2 minutes in sitting posture, then 2 minutes during change in posture from sitting to supine and last 16 minutes in supine posture.

For the purpose of the binary classification done in this study, only the sitting posture component during the first 2 minutes and the initial supine posture component during 4 to 6 minutes of the DS1 datasets have been considered. These two sets are labelled as the Sitting and Sup 1 states respectively.

4.1.1 Data Preprocessing [1]

As mentioned as the outset, the data sets used for the classification studies are those that were acquired and used in the study by Sarkar [1]. The acquisition methodology of these DDP signals is stated in Chapter 1 for ready reference. For the classification study, the acquired datasets have been preprocessed in a definite manner as stated hereafter. The data cleaning is first done to remove anomalies or to address missing data and then this cleaned data is z-normalized. For z-normalization, the bias of the signals (temporal mean of the signals) are subtracted from the signals and then the values are divided by the temporal SD of the respective signals. The sampling frequency has

been set to 20 samples/second.

In the next stage, the datasets have been suitably quantized in order to obtain multiple subsets in a particular class. Each 20 (or 10) minute long DS1 (or DS2) DDP signal is first divided into 10 (or 5) numbers of 2 minute signals. Then each 2 minute signal is further sub-divided into 12 small subsets of 10s each as discussed hereafter.

DS1: As mentioned earlier, DS1 signals are acquired from LH in sitting posture in the first 2-minutes. These are denoted as the Sitting state data. Since the subjects change posture from sitting to supine in next 2-minutes, so this 2-minutes signals are not considered for the binary classification. After that, the signals are acquired for the remaining 16-minutes in supine posture. The first 2-minutes subset of these 16-minutes is denoted as Sup 1 and this is the supine state considered for the binary classification.

Hence for posture classification, the Sitting state is compared with the Sup 1 state. Here, total 57 numbers of 20 minute long data sets were considered after data cleaning. Of these, only the 2-minutes Sitting and the 2-minutes Sup 1 signal sets are considered for the classification. Each 2-minutes subset is further divided into 12 small subsets. Hence each state contains 12 *small sets* $\times$ 57 *sets* $=$ 684 *small sets*. The data for the two states are arranged in a 684×2 data matrix $A1$ as follows.

$$A1 = \begin{bmatrix} Sitting_1 & Sup\ 1_1 & \\ Sitting_2 & Sup\ 1_2 & \\ \vdots & \vdots & \vdots \\ Sitting_{684} & Sup\ 1_{684} & \end{bmatrix} \tag{4.1}$$

As mentioned earlier, each 2-minutes data set is divided into 12 small subsets, so the time duration of each subset is 10s. Thus, $Sitting_1$ to $Sitting_{12}$ are the 12 small subsets of 10s duration of the original 2-minutes long *Sitting* subset of the first data set. The other terms are also to be considered in the same way.

DS2: In DS2 experiment, the signals are acquired from both LH and RH in supine posture and the length of the signal is 10-minutes long. In the first stage, the signals are divided into 5 subsets of 2-minutes each. These subsets form the 5 states that are termed as Sup 1 to Sup 5.

In this case, a total of 156 bilateral signals of 10 minutes duration are considered. Here also the 2 minute subsets of the DS2 data set are further

quantized into 12 small sets of 10s length. Hence, the total number of data sets for each of the 5 states are 12 *small sets* $\times$ 156 *sets* = 1872. The sets and small sets of the LH and RH signals are labelled accordingly. The data for the five states of any one hand are arranged in a 1872 $\times$ 5 data matrix $A2$ as follows.

$$A2 = \begin{bmatrix} Sup\ 1_1 & Sup\ 2_1 & Sup\ 3_1 & Sup\ 4_1 & Sup\ 5_1 \\ Sup\ 1_2 & Sup\ 2_2 & Sup\ 3_2 & Sup\ 4_2 & Sup\ 5_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Sup\ 1_{1872} & Sup\ 2_{1872} & Sup\ 3_{1872} & Sup\ 4_{1872} & Sup\ 5_{1872} \end{bmatrix}$$

As discussed for the DS1 dataset, in this case also $Sup\ 1_1$ to $Sup\ 1_{12}$ are the 12 small subsets of 10s duration of the original 2-minutes long $Sup\ 1$ subset of the first data set. The other terms are also to be considered in the same way.

4.1.2 Model Parameters used in Classification

For each preprocessed 10s signal, the l_{zc} of the ACF is calculated and this is used to determine the signal order n following the method mentioned in Section 2.2.3. Thereafter, n-th order ACF, KF_{LPC} and KF_{TS} models are developed for each of the 10s signals following the steps outlined in Chapters 2 and 3. The statistical parameters and spectral entropy of the tonic (or deterministic) and the phasic (or stochastic) components of the 10s signals have been calculated for all three models using MATLAB. Each 10s signal has been z-normalized again to further calculate their normalized variance and kurtosis. In case of the two KF models, the J_{com} and q_{com} parameters have also been determined from their robustness and sensitivity metrics, J_1 and J_2 respectively. It is to be recalled that J_{com} is the compromise metric value and q_{com} is the index of the multiplier for the corresponding process noise covariance Q_{com}. The list of all the features calculated for the three models are given in Table 4.1.

4.1.3 Feature Selection (Ranker) Algorithm

The next step in classification is to reduce the attribute space from total feature (or parameter) set and identify a subset of features by removing irrelevant or redundant features. This is a critical step that helps in reducing the classification model dimension and hence its training time, improves the

Table 4.1: List of Model Parameters used for the Classifications

Parameters	ACF model	KF_{LPC}	KF_{TS}
General	l_{zc}, a_0, n	-	-
Statistical parameters	Mean of $x_{kACF_{tonic}}$	Mean of $x_{kLPC_{tonic}}$	Mean of $x_{kTS_{tonic}}$
	SD, Variance, Skewness, Kurtosis of $x_{kACF_{tonic}}$ and $x_{kACF_{phasic}}$	SD, Variance, Skewness, Kurtosis of $x_{kLPC_{tonic}}$ and $x_{kLPC_{phasic}}$	SD, Variance, Skewness, Kurtosis of $x_{kTS_{tonic}}$ and $x_{kTS_{phasic}}$
	Normalized variance, Normalized kurtosis of $x_{kACF_{tonic}}$ and $x_{kACF_{phasic}}$	Normalized variance, Normalized kurtosis of $x_{kLPC_{tonic}}$ and $x_{kLPC_{phasic}}$	Normalized variance, Normalized kurtosis of $x_{kTS_{tonic}}$ and $x_{kTS_{phasic}}$
Spectral parameters	$SE_{ACF_{tonic}}$, $SE_{ACF_{phasic}}$	$SE_{LPC_{tonic}}$, $SE_{LPC_{phasic}}$	$SE_{TS_{tonic}}$, $SE_{TS_{phasic}}$
Metrics parameters	-	J_{com}, q_{com}	J_{com}, q_{com}
Total number of parameters for one hand	$18 \times 5 = 90$	$17 \times 5 = 85$	$17 \times 5 = 85$

efficiency and accuracy of the classification algorithm and also avoids the pitfalls of over-fitting of the data. Although various attribute evaluators are available in the WEKA software, only the *AttributeSelection* filter was used in this experiment. It is a very flexible supervised attribute filter and allows the use of various combinations of search and evaluation methods.

Out of the various techniques available in the WEKA platform for selecting machine learning data sets using various Attribute Evaluators, three of them are mostly used in classification analysis [192]. These are the CorrelationAttributeEval, InfoGainAttributeEval and ReliefFAttributeEval technique. All of these are used with a Ranker Search Method that evaluates each attribute and lists the results in decreasing rank order [193]. The details of these Attribute Evaluators [194] are as follows.

1. The CorrelationAttributeEval technique requires the use of a Ranker Search method. This is one of the popular techniques for most relevant attribute selection and depends on the Pearson's correlation coefficient. In this technique, the correlation between each attribute and the output variable is calculated. The attributes with the highest positive or negative correlation (close to -1 or 1) are ranked first and selected. The ranking score decreases as the correlation amplitude reduces. Hence attributes with low absolute correlation are omitted.

2. As the name suggests, the InfoGainAttributeEval attribute evaluator supports feature selection via information gain, or entropy, calculations. As in the correlation technique, the Ranker Search Method must be

used in this case also. In this technique, entry values vary from 0 (no information) to 1 (maximum information). The attributes that contribute more information have a higher entropy value and are likely to be selected while the others which contribute low information or lower score are removed.

3. The ReliefFAttributeEval attribute evaluator is another type of feature selection algorithm that evaluates the worth of an attribute by repeatedly sampling an instance. The Monte Carlo Approach is used in this evaluator for randomized selection of instances [195]. This is usable on both discrete and continuous class data. In this case, the value of the given attribute for the nearest instance of the same as well as different class is evaluated and the total value obtained for a particular attribute is used as a ranker.

4.1.4 Classification Metrics

Both the classifications studied in this Chapter use the random forest (RF) classifier, although the underlying random tree (RT) classifier has been used occasionally to validate the choice of the attribute(s) identified by the ranker. RT is a classifier that selects a decision tree randomly to evaluate various subsets of the data set, while RF is a classifier that evaluates a number of such decision trees on the various subsets of the given data set. In both cases, the average of the obtained values is calculated to improve the predictive accuracy of that data set and control over-fitting. The advantages of RF classifier are it takes less training time as compared to other algorithms, predicts output with high accuracy, runs efficiently even for large data sets and can maintain accuracy despite lack of a large proportion of data.

Some traditional classification metrics that are used to evaluate the classification models are described below.

1. **Confusion matrix**: The confusion matrix of a binary classification comprises of two dimensions, one of actual values and another of predicted values, as shown in Table 4.2. The actual values are the true values of the given observations while the predicted values are the values that are predicted by the classification model. This matrix is used to evaluate the performance of the classification models, make predictions on test data and estimate the goodness of the classification model.
   The two classes of the binary classification are labelled as Positives and Negatives. The matrix elements expressed in terms of these classes are the following.

Table 4.2: Structure of confusion matrix

		Predicted output		
		Positive	Negative	Total
True/Actual instances	Positive	TP	FN	$TP + FN$
	Negative	FP	TN	$FP + TN$
	Total	$TP + FP$	$FN + TN$	N

True Positive (TP): The total number of cases where the model has predicted Positive and the actual value is also Positive.
False Negative (FN): The total number of cases where the model has predicted Negative, but the actual value is Positive.
False Positive (FP): The total number of cases where the model has predicted Positive, but the actual value is Negative.
True Negative (TN): The total number of cases where the model has predicted Negative and the actual value is also Negative.
Thus, the total number of cases $N = TP + FP + TN + FN$.

2. **Accuracy**: It is one of the important parameters that defines how often the model predicts the correct output and is evaluated as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

3. **Precision or Positive Prediction Value (PPV)**: It is defined as the number of correct positive outputs provided by the model out of all positive classes that have been predicted by the model. It is calculated using the formula:

$$Precision = \frac{TP}{TP + FP}$$

4. **Recall or True Positive Rate (TPR) or Sensitivity**: This is defined as the number of cases in which the model predicts Positives correctly out of the total actual instances of Positives. The recall must be as high as possible.

$$Recall = \frac{TP}{TP + FN}$$

5. **Specificity and False Positive Rate (FPR)**: Specificity is defined as the number of cases in which the model predicts Negatives correctly

out of the total actual instances of Negatives. As for the recall, the specificity must also be as high as possible.
On the other hand, False Positive Rate (FPR) is the number of cases in which the model predicts Negatives incorrectly out of the total actual instances of Negatives. Hence,

$$\begin{aligned} Specificity &= \frac{TN}{TN+FP} \\ FPR &= 1 - Specificity = \frac{FP}{FP+TN} \end{aligned}$$

6. **F1 score**: In case of unbalanced data sets where the actual instances of Positives or Negatives vary significantly from actual instances for the other class, the accuracy or some of the other metrics get skewed as opposed to calculations based on balanced data sets. For example, if two models have low precision and high recall or vice versa, it is difficult to compare these models. In this scenario, the F1-score is used as the metric that is desensitized to data imbalances. This is evaluated using the recall and the precision as follows.

$$F1score = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

7. **Receiver Operating Characteristics (ROC) Curves and Precision-Recall Curves**: ROC Curves summarize the trade-off between the TPR (or Sensitivity) and FPR (or reverse of Specificity) for a predictive model using different probability thresholds. On the other hand, Precision-Recall Curves summarize the trade-off between the Precision (or PPV) and the Recall (or TPR or Sensitivity) for a predictive model using different probability thresholds. Higher the Area under the Curve (AUC) values for these two curves, more the reliability of the particular classification model being tested.

4.1.5 Cross Validation Protocols

Both the classifications are validated using three different cross validation techniques, namely 10 fold cross validation (10FCV), leave one out cross validation (LOOCV) and leave one subject out cross validation (LOSOCV).

In 10FCV, the total dataset is grouped into 10 number of subsets with equal number of populations. Of these 10 subsets, 9 subsets are used to train the model and the remaining 1 subset is taken for testing in the first cycle.

In the next cycle, a different subset is used for testing and the remaining 9 are applied for training. In this way, the process is repeated for 10 iterations. The overall performance and the classification metrics are then calculated by averaging all 10 prediction values.

LOOCV is very similar to 10FCV but in this technique, the total N number of data sets is considered in each iteration in place of partitioning it into 10 groups. Of all the N data sets, only 1 set of data is kept aside for testing while all the other N-1 sets are used for training. Hence, the number of the iterations in LOOCV is N which is equal to the number of sets of data. In this case, the overall performance and the classification metrics are calculated by averaging all N prediction values.

In LOSOCV, the objective is to eliminate the subject bias and so, subjects are eliminated one after other. In this technique, all the datasets of one subject are taken as testing set in a particular iteration, while the rest are considered as training sets. Hence, the number of iterations is same as the number of subjects in this case. Furthermore, the overall confusion matrix and the classification metrics are calculated by averaging all S prediction values, where S is the total number of subjects.

The confusion matrix, accuracy and all other metrics as well as the ROC and PRC AUCs are determined automatically in WEKA software once the CV method is chosen as 10FCV or LOOCV.

However this is not so for the LOSOCV since there are likely to be data imbalances for the various subjects. Furthermore, in case of LOSOCV, the class definitions of Positive and Negative in WEKA software need to be stated explicitly in order to obtain the proper confusion matrices for the data set of each subject. If all data sets of a subject pertains to one class while all data sets of another subject pertains to another class, then the class definitions of Positive and Negative in the two cases need to be reversed suitably. In case some of the data sets of a subject belong to one class while the other data sets belong to the second class, then this reversal has to be done for the respective data sets. This is necessary in order to obtain the proper weightages in the calculations of the metrics for the particular class. Thereafter, the confusion matrices calculated for all the conditions are added to obtain the overall confusion matrix. The metrics are then determined individually for both the classes and from these, the weighted average values are obtained.

In case of LOSOCV, the ROC or PRC AUCs have not been calculated in this study.

4.2 Hypertensive and normotensive classification

Using the methodology detailed in Section 4.1, the binary classification of hypertensive and normotensive subjects have been done using the parameters of each model, as listed in Table 4.1. The class definitions of Positive and Negative in this case refer to Hypertensive or Normotensive subjects. The identification of the state of a subject is based on the medical history of the person. Hence for a particular subject, all data sets belong to one class, either Hypertensive or Normotensive. The DS2 data sets have been acquired from 5 hypertensive subjects and 17 normotensive subjects.

As already detailed in Section 4.1.1, this study is based on the total 156 DS2 data sets that have been acquired simultaneously from the LH and the RH of human subjects in supine posture for 10 minutes. Each 10 minutes signal is subdivided to obtain 5 sets of non-overlapping 2 minutes signals. These form the 5 states, defined as Sup 1 to Sup 5, of the particular LH or RH data set. After the quantization of each of these 2 minutes signals belonging to a state, total 1872 numbers of 10s signals are generated in each state of one data set. As mentioned in Section 4.1.2, the ACF model, the KF_{LPC} model and the KF_{TS} model are fitted to each of these 10s signals and all the model parameters listed in Table 4.1 are calculated for each quantized signal. As stated in the Table, the number of parameters calculated from the ACF model, the KF_{LPC} model and the KF_{TS} model in one state of one hand are 18, 17 and 17 respectively and so the total number of parameters considering all 5 states of one hand are 90, 85 and 85 respectively. In DS2 data sets, LH and RH both hands are considered. So the total number of parameters from ACF model are $90 \times 2 = 180$ and from KF_{LPC} and KF_{TS} models are $85 \times 2 = 170$.

4.2.1 Selection of Attribute Evaluator

A preliminary classification considering all parameters is done using *AttributeSelection* filter. This provides the ranking of the parameters. Based on the ranking, trial and error method is used to select the top few relevant attributes beyond which the overfitting leads to either reduced accuracy, precision or a saturation of the results. These selected attributes are then used for the classification study.

The accuracy of classification using the 3 different ranker methods discussed in Section 4.1.3 for each model is stated in Table 4.3. It is observed

that using CorrelationAttributeEval, number of attributes for ACF, KF_{LPC} and KF_{TS} models are 7, 6, 1 respectively which are reasonable in number and the classification accuracies are 91.18%, 99.94% , 99.51% respectively. The number of attributes for the ReliefFAttributeEval are the same for ACF and KF_{TS} models but less for KF_{LPC}, specifically 3. In this case, the accuracy at 97.38% is significantly higher for the ACF model, but decreases marginally for KF_{TS} model to 99.19% while it remains the same for the KF_{LPC} model.

Table 4.3: Rank wise selected attributes and accuracy for hypertensive-normotensive classification

Ranker	Number of attributes selected					
	ACF		KF_{LPC}		KF_{TS}	
	Number of attributes	Accuracy (%)	Number of attributes	Accuracy (%)	Number of attributes	Accuracy (%)
Correlation AttributeEval	7	91.18	6	99.94	1	99.51
InfoGain AttributeEval	**3**	**98.07**	**3**	**99.89**	**1**	**99.51**
Relief AttributeEval	7	97.38	2	99.94	1	99.19

Using InfoGainAttributeEval, the accuracies of the classification results are uniformly high at 98.07%, 99.89% and 99.51% for the ACF, KF_{LPC} and KF_{TS} models respectively while the numbers of attributes, specifically 3, 3 and 1 respectively, are uniformly low for all models. Hence, this evaluator is used in this study to classify the two classes of hypertensive and normotensive signals.

4.2.2 Selected attributes and their analysis

Applying the Supervised Attribute Selection filter in WEKA, 29 (out of total 180 parameters) attributes from ACF model, 32 (out of total 170 parameters) attributes from KF_{LPC} model and 22 (out of total 170 parameters) attributes from KF_{TS} model are selected. On using the InfoGainAttributeEval attribute evaluator on this subset of features, finally 3, 3 and 1 attributes respectively are selected for the 3 models as stated in Table 4.4.

In order to appreciate the characteristics of all the 7 selected attributes, their box plots are shown in Figures 4.1a, 4.1b and 4.1c for the ACF model, Figures 4.2a, 4.2b and 4.2c for the KF_{LPC} model and Figure 4.3 for the KF_{TS} model. The corresponding box plot parameters for the 3 models are listed in Tables 4.5, 4.6 and 4.7. The observations from these selected features and their box plots are listed hereafter.

Table 4.4: List of the selected attributes using ACF, KF_{LPC}, KF_{TS} model parameters for hypertensive-normotensive classification

Model	Sl No.	Selected attributes as per rank	Cross validation accuracy (%)		
			10FCV	LOOCV	LOSOCV
ACF	1	Norvar $x_{kACF_{phasic}}$ of LH3	98.07	98.13	83.01
	2	Norvar $x_{kACF_{phasic}}$ of LH4			
	3	Norvar $x_{kACF_{phasic}}$ of LH2			
KF_{LPC}	1	q_{com} of LH5	99.89	100	69.87
	2	q_{com} of LH1			
	3	q_{com} of RH1			
KF_{TS}	1	q_{com} of LH4	99.51	99.51	72.70

Table 4.5: Box plot parameters for hypertensive normotensive classification using ACF model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
Norvar $x_{kACF_{phasic}}$ of LH3	Hypertensive	-0.21	-0.16	-0.03	-0.03	0.12
	Normotensive	-0.23	-0.12	-0.06	-0.04	0.07
Norvar $x_{kACF_{phasic}}$ of LH4	Hypertensive	-0.21	-0.16	-0.03	-0.03	0.15
	Normotensive	-0.21	-0.11	-0.06	-0.04	0.06
Norvar $x_{kACF_{phasic}}$ of LH2	Hypertensive	-0.21	-0.17	-0.03	-0.03	0.13
	Normotensive	-0.23	-0.12	-0.06	-0.05	0.06

Table 4.6: Box plot parameters for hypertensive normotensive classification using KF_{LPC} model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
q_{com} of LH5	Hypertensive	2.34	2.34	2.35	2.73	2.88
	Normotensive	1.48	2.21	2.75	2.86	3.41
q_{com} of LH1	Hypertensive	1.92	2.35	2.35	2.73	2.88
	Normotensive	1.27	2.21	2.75	2.85	3.45
q_{com} of RH1	Hypertensive	2.25	2.35	2.35	2.73	2.88
	Normotensive	1.5	2.21	2.74	2.86	3.41

Table 4.7: Box plot parameters for hypertensive normotensive classification using KF_{TS} model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
q_{com} of LH4	Hypertensive	-0.64	-0.61	-0.49	-0.24	-0.20
	Normotensive	-0.46	-0.34	-0.29	-0.23	-0.15

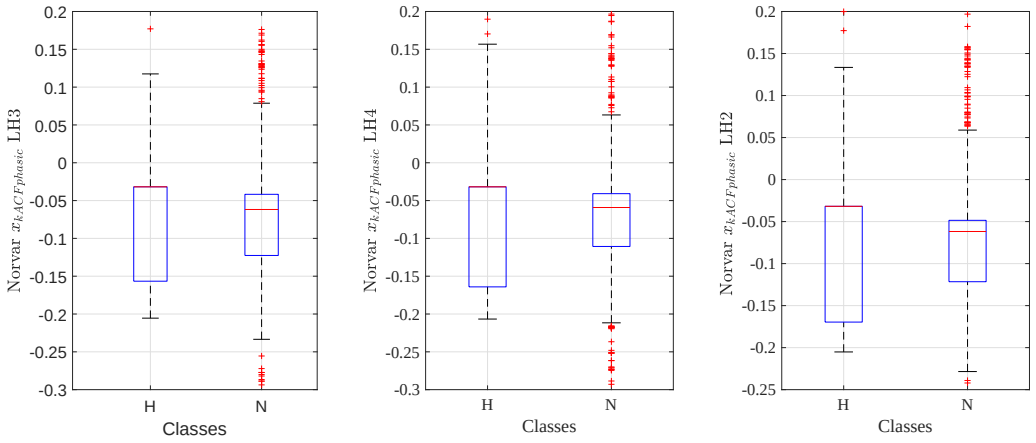


Figure 4.1: H: Hypertensive, N: Normotensive. Box plots of (a) Norvar $x_{kACF_{phasic}}$ of LH3 (b) Norvar $x_{kACF_{phasic}}$ of LH4 (c) Norvar $x_{kACF_{phasic}}$ of LH2.

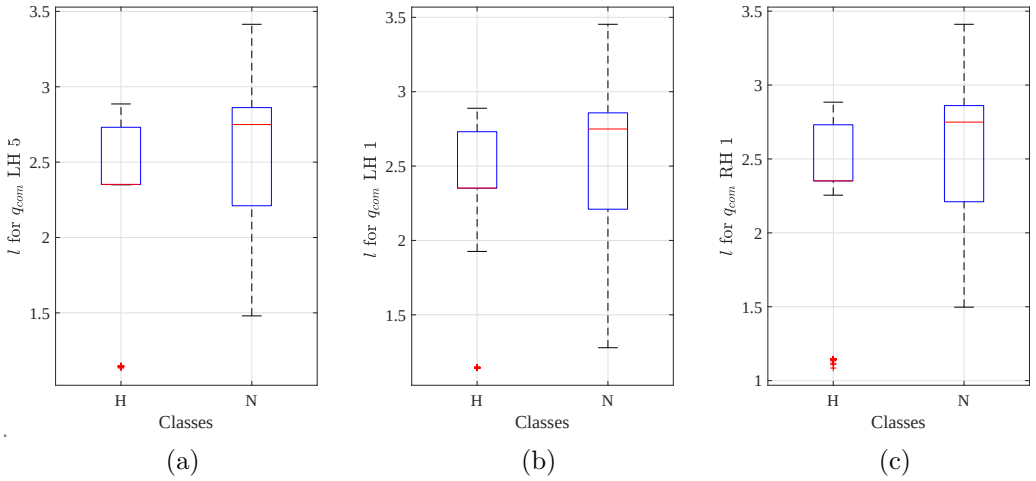


Figure 4.2: H: Hypertensive, N: Normotensive. Box plots of (a) l of q_{com} of LH5 (b) l of q_{com} of LH1 (c) l of q_{com} of RH1 for KF_{LPC} model

1. Of the total 7 attributes selected for the 3 models, it is observed that 6 attributes are from LH and only 1 is from RH. These are the normalized variance (Norvar) of $x_{kACF_{phasic}}$ of 3 states (state 3, 4, 2) from LH for ACF model, q_{com} of 2 states from LH (state 5, 1) and 1 state from RH

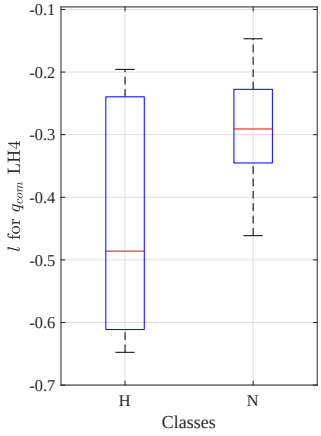


Figure 4.3: H: Hypertensive, N: Normotensive. Box plots of l of q_{com} of LH4 for KF_{TS} model

(state 1) for KF_{LPC} model and q_{com} of 1 state from LH (state 4) for KF_{TS} model.

Thus, as already observed in [1], the presence or absence of hypertension is most reflected in the LH parameters.

2. In case of the ACF model, all 3 attributes are selected from only the phasic component of LH and RH signals. In case of both KF based models, the parameters selected are the indices q_{com} of the process noise covariances at compromise Q_{com} [174] for various states. These are all obtained from the performance metrics specific to the KF model. Also, in terms of the nominal value q_{nom}, $q_{com} = 10^l q_{nom}$ where, l is termed as the multiplier, as stated earlier.

3. The box plots of all 3 normalized variance features selected in the ACF model are similar. In general, the medians are quite high for hypertensive signals and lower for normotensive signals. The lower adjacent values are similar but the upper adjacent values are quite high for the hypertensive signals.

4. The box plots of all 3 q_{com} features selected in the KF_{LPC} model are also similar. The overall range of the individual classes are almost similar for all 3 box plots. In all cases, the median value is low and is the same as the lower quartile in hypertensive signals. The overall range for the normotensive class is much larger and the median is also significantly higher in all 3 cases.

5. In case of the KF_{TS} model, there is only 1 selected feature. Although this is also a q_{com} value, yet it is to be recalled that the natures of the q_{com} for the KF_{LPC} and KF_{TS} models are very different, differing also in the sign. Hence the characteristics are expectedly significantly different.

   In particular, the range of q_{com} of LH4 in case of hypertensive subjects is much larger than that for normotensive subjects and the median is low for hypertensive signals as compared to that for normotensive subjects.

4.2.3 Confusion Matrices and Metrics

The confusion matrices for 10FCV, LOOCV and LOSOCV for hypertensive-normotensive classification using the selected attributes listed in Table 4.4 for the 3 models are stated in Table 4.8. As stated earlier, the total instances for this binary classification is 1872.

The corresponding classification metrics of precision, recall (or TPR or sensitivity), specificity, F1-score as well as the ROC and PRC AUC for 10FCV, LOOCV and LOSOCV of the ACF, KF_{LPC}, KF_{TS} models are tabulated in Tables 4.9, 4.10 and 4.11 respectively.

Table 4.8: Confusion matrix (in %) of RF classifier using ACF, KF_{LPC}, KF_{TS} model parameters for hypertensive-normotensive classification

CV Method	Actual	ACF		KF_{LPC}		KF_{TS}	
		HT	NT	HT	NT	HT	NT
10FCV	HT	10.79	1.39	12.18	0	11.81	0.37
	NT	0.53	87.29	0.11	87.71	0.11	87.71
LOOCV	HT	10.84	1.34	12.18	0	11.81	0.37
	NT	0.53	87.29	0	87.82	0.11	87.71
LOSOCV	HT	2.40	9.78	1.23	10.95	6.09	6.09
	NT	7.21	80.61	19.18	68.64	21.21	66.61
DS2 data set: N=1872, HT: Hypertensive, NT: Normotensive							

The observations from the confusion matrices and the classification metrics for the 3 models using the 3 different cross-validation techniques are stated hereafter.

1. There are 5 hypertensive (HT) subjects with 228 sets of data and 17 normotensive (NT) subjects with 1644 sets of data. So this is a heavily unbalanced binary classification with 12.18% HT and 87.82% NT data. Hence, more credence is to be given to F1-score than to accuracy.

Table 4.9: Results for Hypertensive and Normotensive classification using ACF model parameters

CV Method	Class	Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV	HT	0.953	0.886	0.994	0.981	0.991	0.965
	NT	0.984	0.994	0.886	0.989	0.991	0.998
	WA	0.980	0.981	0.899	0.980	0.991	0.994
LOOCV	HT	0.953	0.890	0.994	0.921	0.991	0.967
	NT	0.985	0.994	0.89	0.989	0.991	0.998
	WA	0.981	0.981	0.903	0.981	0.991	0.995
LOSOCV	HT	0.25	0.19	0.91	0.22	-	-
	NT	0.89	0.91	0.19	0.90	-	-
	WA	0.81	0.82	0.28	0.81	-	-

Table 4.10: Results for Hypertensive and Normotensive classification using KF_{LPC} model parameters

CV Method	Class	Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV	HT	0.991	1.00	0.99	0.996	1.00	1.00
	NT	1.00	0.99	0.99	0.99	1.00	1.00
	WA	0.99	0.99	1.00	1.00	1.00	1.00
LOOCV	HT	1.00	1.00	1.00	1.00	1.00	1.00
	NT	1.00	1.00	1.00	1.00	1.00	1.00
	WA	1.00	1.00	1.00	1.00	1.00	1.00
LOSOCV	HT	0.60	0.10	0.78	0.17	-	-
	NT	0.86	0.78	0.10	0.82	-	-
	WA	0.88	0.69	0.18	0.74	-	-

2. The classification metrics of the precision (or PPV), Recall (or Sensitivity or TPR) and Specificity are also evaluated for all 3 cross-validation methods. The ROC and PRC AUC are also evaluated for the 10FCV and LOOCV methods.

   In all these cases, the values are calculated twice, once by considering the hypertensive (HT) case to be the Positive class and normotensive (NT) as the Negative class and then by reversing this class definition. Thereafter, their weighted average (WA) is calculated which is expectedly more biased towards the NT class.

3. As is expected, in each of the 3 models, the confusion matrices for the 10FCV and LOOCV are identical or almost so, as evident in Table 4.8.

Table 4.11: Results for Hypertensive and Normotensive classification using KF_{TS} model parameters

CV Method	Class	Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV, LOOCV	HT	0.991	0.969	0.99	0.980	0.986	0.975
	NT	0.996	0.99	0.97	0.99	0.986	0.996
	WA	0.995	0.995	0.973	0.995	0.996	0.994
LOSOCV	HT	0.22	0.50	0.76	0.31	-	-
	NT	0.92	0.76	0.50	0.83	-	-
	WA	0.83	0.73	0.53	0.77	-	-

Accordingly, the metrics are also very similar or identical for these two CV methods for a particular model.

Hence, of these two methods, only the 10FCV results will be discussed henceforth.

4. From Table 4.8, it is observed that for 10FCV method, 202 HT and 1634 NT are predicted correctly out of a total of 1872 instances, that is 98.07% instances are predicted correctly using the ACF model features. The corresponding prediction accuracies using the KF_{LPC} and KF_{TS} model features are significantly high at 99.89% and 99.51% respectively as also stated in Table 4.4.

5. A comparison of the other metrics for the 10FCV across the 3 models shows that except specificity (97.3%), all other metrics including the AUCs are almost 99.5% for the KF_{TS} model. These high values are even more significant since the KF_{TS} model metrics are obtained on the basis of just one feature.

   The KF_{LPC} model exhibits very high values of 99% for the precision and recall and 100% for the others. Yet, as also evident from the saturation of the corresponding accuracy and all others metrics at 100% for the LOOCV, this might be due to overfitting or subject data bias.

   In comparison, the WA values of the metrics are lower in case of the ACF model. Yet they are significantly high with the precision, recall and F1-score values at 98% and the ROC and PRC AUCs larger than 99.1%. Only the specificity is slightly lower at 89.9%.

6. In order to study the effect of subject bias, the LOSOCV results for the metrics have also been tabulated for all 3 models. As might be

expected, significant drops are observed in the metrics in all 3 models for all the metrics.

The drop in F1-score is maximum for the KF_{LPC} model from 100% to 74% indicating steep subject bias in the data.

In case of the KF_{TS} model also, the F1-score drops from 99.5% to 77%. This indicates the need to evaluate the other model parameters as selectable features instead of a dependence on just one single feature.

In case of the ACF model, the F1-score drops by 17% to 81%. Thus, the features selected for this model provide the most balanced F1-score result.

7. A comparison of the specificity values for the LOSOCV in all 3 models show that the KF_{TS} model results (WA: 0.53) are the most balanced and highest of all 3 models being almost double that of the other two models with WA of 0.28 and 0.18 for the ACF and KF_{LPC} models respectively.

8. A comparison of the accuracy or F1-scores of 10FCV of the 3 models with those stated in [1] **and reproduced as Table 5.2** shows that all 3 models provide competitive scores for this binary classification. A comparison table is given in Chapter 5, Section 5.1.5.

4.3 Posture change classification

The second binary classification is that of two postures, the Sitting state and the supine state Sup 1. For this, the model parameters of the LH, as listed in Table 4.1, are used to detect the two different postures (states) assumed by subjects during the DS1 data set acquisition as discussed in Section 4.1.1. In this case, the class definitions of Positive and Negative refer to Sitting or Sup 1 postures.

In this case, a total of 57 DS1 data sets acquired for 20 minutes from the LH of human subjects in first sitting, then supine posture are considered. Of these long data sets, the initial 2 minutes data in sitting posture and the data in the 4 to 6 minutes duration when the subject is in initial supine posture are considered. These 2 minutes signals form the 2 states, defined as Sitting and Sup 1. After the quantization of each of these 2 minutes signals belonging to a state, total 684 numbers of 10s signals are generated in each state of one data set. As in Section 4.2, in this case also the ACF model, the KF_{LPC} model and the KF_{TS} model are fitted to each of these 10s signals and all the model parameters listed in Table 4.1 are calculated for each quantized

Table 4.12: Rank wise selected attributes and accuracy for posture change classification

Ranker, AttributeEval	Number of attributes selected					
	ACF		KF_{LPC}		KF_{TS}	
	Number of attributes	Accuracy (%)	Number of attributes	Accuracy (%)	Number of attributes	Accuracy (%)
Correlation, InfoGain, Relief	5	96.34	1	100	1	100

signal. Thus in this case, the number of parameters calculated from the ACF model, the KF_{LPC} model and the KF_{TS} model in each state are 18, 17 and 17 respectively.

4.3.1 Selection of Attribute Evaluator

As mentioned in Section 4.2.1, in this case also a preliminary classification is done using *AttributeSelection* filter considering all 17 or 18 parameters in order to obtain their ranking. Based on the ranking, the top few relevant attributes are selected for use in the classification study.

The results for the 3 different ranker methods discussed in Section 4.1.3 for each model is stated in Table 4.12. It is observed that using all 3 attribute evaluators, the number of attributes for ACF, KF_{LPC} and KF_{TS} models are 5, 1 and 1 respectively and their classification accuracies are 96.34%, 100% and 100% respectively. In case of the KF_{LPC} model, increasing the number of parameters does not affect the 100% accuracy but in case of the KF_{TS} model, the accuracy decreases as the number of features are increased. In fact, a choice of 5 attributes in the KF_{TS} model causes the accuracy to drop to 99.85%.

In case of the ACF model, the best results are obtained for 5 attributes irrespective of the ranker algorithm chosen. These attributes are the same across all the rankers but their ranking orders vary in the 3 ranker algorithms. Thus the choice of 5 attributes yields the same and maximum accuracy for the ACF model irrespective of the ranker algorithm chosen. In order to differentiate the 3 ranker algorithms, only the top ranked attribute in each ranker is chosen. On performing the posture classification using this single attribute, the InfoGainAttributeEval ranker provides the highest accuracy as compared to the other two. So this ranker is henceforth chosen for the posture classification study.

4.3.2 Selected attributes and their analysis

As already stated earlier, the original number of features is limited to 17 or 18 in case of the posture change classification. In all the 3 models, the application of the Supervised Attribute Selection filter in WEKA identifies 5 numbers of features. On using the InfoGainAttributeEval attribute evaluator on this subset of features, finally 5, 1 and 1 attributes respectively are selected for the 3 models as detailed in Table 4.13.

Table 4.13: List of the selected attributes using ACF, KF_{LPC}, KF_{TS} model parameters for posture change classification

Model	Sl No.	Selected attributes as per rank	Cross validation accuracy (%)		
			10FCV	LOOCV	LOSOCV
ACF	1	Norvar of $x_{kACF_{phasic}}$	96.34	96.49	84.80
	2	Norvar of $x_{kACF_{tonic}}$			
	3	Norkur of $x_{kACF_{tonic}}$			
	4	a_0			
	5	$SE_{ACF_{phasic}}$			
KF_{LPC}	1	q_{com}	100	100	90.49
KF_{TS}	1	q_{com}	100	100	65.39

In order to appreciate the characteristics of all the 7 selected attributes, their box plots are shown in Figures 4.4a, 4.4b, 4.4c, 4.4d and 4.4e for the ACF model and in Figures 4.5a and 4.5b for the KF_{LPC} and KF_{TS} models. The corresponding box plot parameters for the 3 models are listed in Tables 4.14 and 4.15. The observations from these selected features and their box plots are listed hereafter.

1. Of the total 5 attributes selected for the ACF, it is observed that there are 2 phasic parameters: normalized variance of $x_{kACF_{phasic}}$ and spectral entropy of ACF_{phasic}, 2 tonic parameters: normalized variance and kurtosis of $x_{kACF_{tonic}}$ and 1 signal coefficient a_0.

2. In case of the KF models, only the performance metric parameter q_{com} is selected. This exemplifies the effect of posture change on the process noise covariance at compromise Q_{com}.

3. The box plots of the 5 features selected in the ACF model are quite distinct from each other. The lower adjacent till upper quartile of the box plots of the normalized variance of $x_{kACF_{phasic}}$ for the two classes

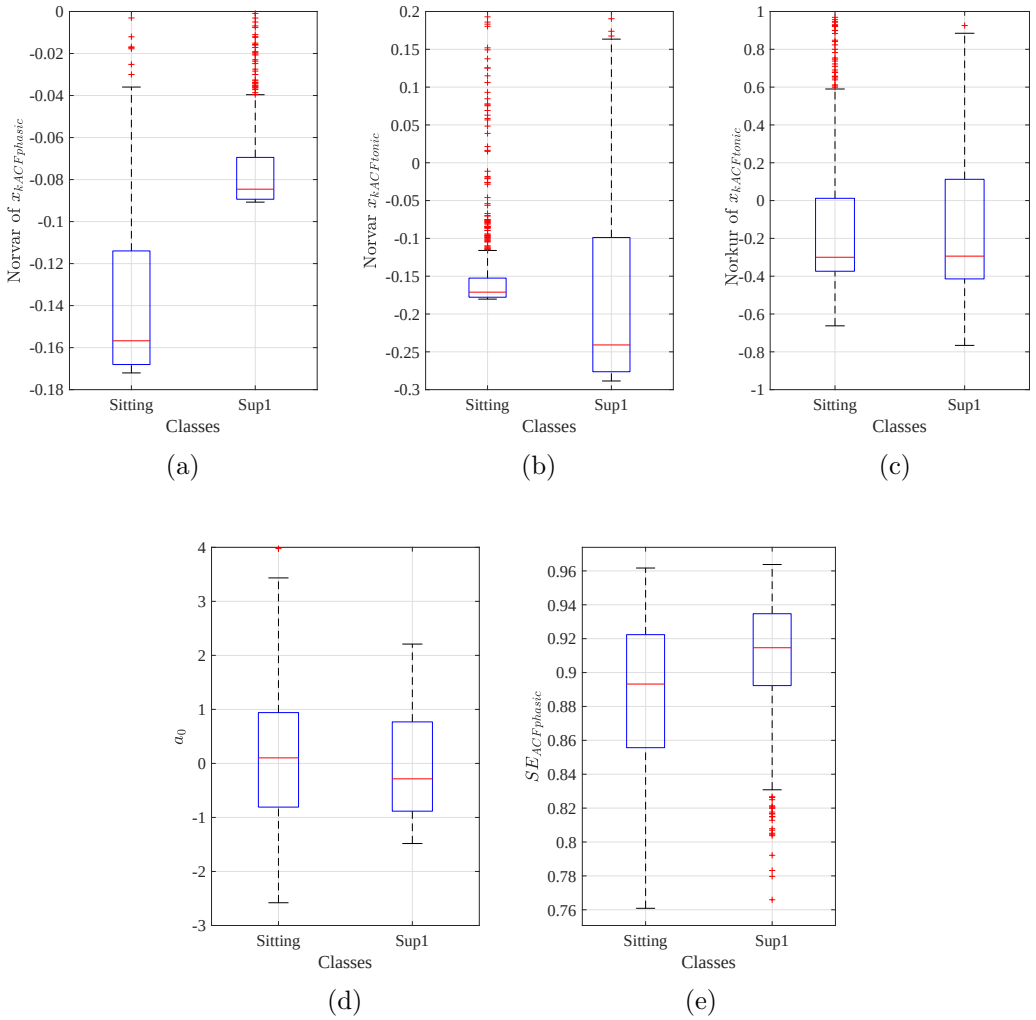


Figure 4.4: Box plot of (a) Norvar of $x_{kACF_{phasic}}$ (b) Norvar of $x_{kACF_{tonic}}$ (c) Norkur of $x_{kACF_{tonic}}$ (d) a_0 (e) $SE_{ACF_{phasic}}$

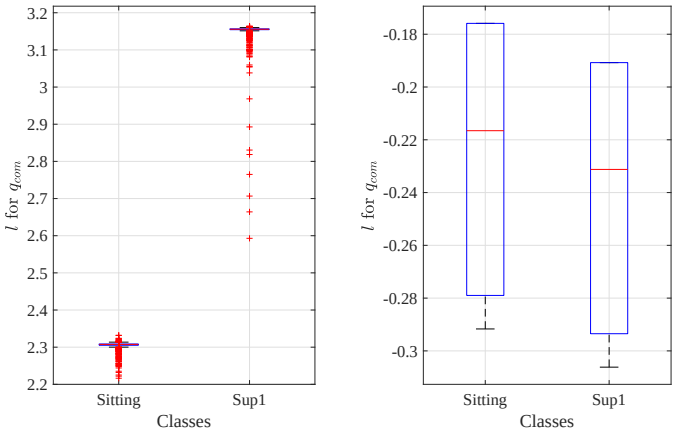
are distinct with the Sitting state values lower but spread over a much larger range than those of the Sup 1 state.

This is almost the reverse in case of the normalized variance of $x_{kACF_{tonic}}$. The range for the Sitting class is much smaller and its median is much larger than those of the Sup 1 class but the range for the Sitting state falls within the range for the Sup 1 state.

Hence, the same variable for the tonic and phasic components exhibit almost reverse natures. For the other 3 features, namely the normal-

Table 4.14: Box plot parameters for posture classification using ACF model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
Norvar of $x_{kACF_{phasic}}$	Sitting	-0.17	-0.16	-0.15	-0.11	-0.03
	Sup 1	-0.09	-0.09	-0.08	-0.07	-0.04
Norvar of $x_{kACF_{tonic}}$	Sitting	-0.18	-0.18	-0.17	0.15	-0.11
	Sup 1	-0.28	-0.27	-0.24	-0.09	0.16
Norkur of $x_{kACF_{tonic}}$	Sitting	-0.66	-0.37	-0.30	0.01	0.59
	Sup 1	-0.77	-0.41	-0.29	0.11	0.88
a_0	Sitting	-2.57	-0.80	0.10	0.94	3.43
	Sup 1	-1.48	-0.88	-0.29	0.77	2.21
$SE_{ACF_{phasic}}$	Sitting	0.76	0.85	0.89	0.92	0.96
	Sup 1	0.83	0.89	0.91	0.93	0.96



(a) (b)

Figure 4.5: Box plot of l of q_{com} for (a) KF_{LPC} model (b) KF_{TS} model

ized kurtosis denoted as Norkur in the figure, a_0 and SE of the phasic component, the overlap in the box plots of the two states are more and their median values are also closer.

4. The box plots of the q_{com} for the two classes (or states, in this case) in the KF_{LPC} model are absolutely segregated. The values are clustered around fixed values in both states. The values in case of the Sitting state are distinctly less than that for the Sup 1 state. This justifies the high accuracy of 100% for 10FCV using this model. The q_{com} for the

Table 4.15: Box plot parameters for posture classification using KF_{LPC} and KF_{TS} model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
$q_{com_{LPC}}$	Sitting	2.29	2.300	2.305	2.307	2.313
	Sup 1	3.151	3.154	3.155	3.156	3.16
$q_{com_{TS}}$	Sitting	-0.29	-0.27	-0.21	-0.17	-0.17
	Sup 1	-0.30	-0.29	-0.23	-0.19	-0.19

KF_{LPC} model are clearly independent for the two postures.

This can be ascribed to the fact that any change in posture leads to distinct change in the pattern of the time-series signal which forms the basis for the KF_{LPC} model. This pattern change is reflected in the changed q_{com} value.

5. For the KF_{TS} model, the median of q_{com} for the Sitting state is higher than that for the Sup 1 state. However, the range of the two states are significantly overlapping. Although the box plots are not distinct for the two states, yet this model also provides 100% accuracy for 10FCV. This might be due to overfitting or subjective bias.

4.3.3 Confusion Matrices and Metrics

The confusion matrices for 10FCV, LOOCV and LOSOCV for Sitting and Sup 1 posture classification using the selected attributes listed in Table 4.13 for the 3 models are stated in Table 4.16.

The corresponding classification metrics of precision, recall (or TPR or sensitivity), specificity, F1-score as well as the ROC and PRC AUC for 10FCV, LOOCV and LOSOCV of the ACF, KF_{LPC}, KF_{TS} models are tabulated in Tables 4.17, 4.18 and 4.19 respectively.

1. The number of instances for the Sitting state and the Sup 1 state are 684 each, with a total of 1368 instances. Thus this is a balanced binary classification unlike the hypertensive- normotensive classification discussed in Section 4.2. So in this case accuracy as well as F1-score have same relevance.

   However, in the LOSOCV classification of all 3 models, the data for the true Sup 1 state in case of 3 subjects could not be classified, whether correctly or incorrectly. Consequently, the total instances in case of

Table 4.16: Confusion matrix (in %) of RF classifier using ACF, KF_{LPC}, KF_{TS} model parameters for posture change classification

CV Method	Actual	ACF		KF_{LPC}		KF_{TS}	
		ST	SP 1	ST	SP 1	ST	SP 1
10FCV	ST	47.88	2.12	50	0	50	0
	SP 1	1.54	48.46	0	50	0	50
LOOCV	ST	47.95	2.05	50	0	50	0
	SP 1	1.46	48.54	0	50	0	50
LOSOCV	ST	41.37	8.63	49.93	0.07	26.75	23.25
	SP 1	2.70	21.86	7.02	17.54	2.56	22.00
DS1 data set: N=1368, ST: Sitting, SP 1: Sup 1							

Table 4.17: Results for posture classification using ACF model parameters

Application		Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV	ST	0.969	0.958	0.96	0.963	0.995	0.995
	SP 1	0.958	0.969	0.96	0.964	0.927	0.995
	WA	0.964	0.963	0.97	0.963	0.995	0.995
LOOCV	ST	0.970	0.959	0.97	0.965	0.995	0.995
	SP 1	0.960	0.971	0.96	0.965	0.965	0.994
	WA	0.965	0.965	0.97	0.965	0.995	0.995
LOSOCV	ST	0.94	0.83	0.89	0.88	-	-
	SP 1	0.72	0.89	0.83	0.80	-	-
	WA	0.83	0.86	0.86	0.84	-	-

LOSOCV is 1020 and the corresponding accuracy and other metrics are calculated accordingly on this basis.

2. As for the hypertensive-normotensive classification, in this case also the accuracy, Precision (or PPV), Recall (or Sensitivity or TPR) and Specificity are evaluated for all 3 cross-validation methods while the ROC and PRC AUC are evaluated for the 10FCV and LOOCV methods. The method for determining the metrics for the 2 classes and the weighted average (WA) is the same as that used for the hypertensive-normotensive classification.

3. The confusion matrices for the 10FCV and LOOCV are almost identical for the ACF model as evident in Table 4.16. Accordingly, the metrics are also very similar or identical for these two CV methods for a particular model.

Table 4.18: Results for posture classification using KF_{LPC} model parameters

Application		Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV, LOOCV	ST	1.00	1.00	1.00	1.00	1.00	1.00
	SP 1	1.00	1.00	1.00	1.00	1.00	1.00
	WA	1.00	1.00	1.00	1.00	1.00	1.00
LOSOCV	ST	0.88	0.99	0.71	0.93	-	-
	SP 1	0.99	0.71	0.99	0.83	-	-
	WA	0.94	0.85	0.85	0.88	-	-

Table 4.19: Results for posture classification using KF_{TS} model parameters

Application		Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV, LOOCV	ST	1.00	1.00	1.00	1.00	1.00	1.00
	SP 1	1.00	1.00	1.00	1.00	1.00	1.00
	WA	1.00	1.00	1.00	1.00	1.00	1.00
LOSOCV	ST	0.91	0.54	0.90	0.68	-	-
	SP 1	0.49	0.90	0.54	0.63	-	-
	WA	0.70	0.72	0.72	0.66	-	-

Hence, of the 2 methods, it suffices to study the 10FCV results only.

4. From Table 4.13, it is observed that the accuracy of the ACF model for the 10FCV and LOOCV are 96.34% and 96.49% respectively. It is further observed from the confusion matrices that the numbers of TP and TN instances are balanced in both methods while the numbers of FP and FN instances are also similar to each other.

   Since this is a balanced binary classification, hence as expected, the WA precision, recall, specificity as well as F1-score are typically around 96% while the ROC and PRC areas are typically 99.5%. These high values coupled with similar statewise values are indicative of high reliability of this classification.

5. The confusion matrices for the 10FCV and LOOCV are exactly identical for the two KF models with 100% cases correctly identified. Hence, all the metrics are also 100%.

   As evident from the box plot characteristics, this result for the KF_{LPC} model may be due to the clear differentiation evident in the box plots of the q_{com} for the two states. However, this differentiation is not evident in the q_{com} box plots for the KF_{TS} model. So it needs to be investigated whether there is any overfitting or subjective bias in this result.

6. The LOSOCV results for the metrics have been tabulated for all 3 models in order to study the effect of subject bias. As might be expected, drops are observed in all the metrics in all 3 models. The typical characteristic of the reversal of the values of recall and specificity for the two classes is valid for all 3 models.

   The drop in F1-score is minimum for the KF_{LPC} model from 100% to 88% indicating marginal subject bias in the data. In case of the ACF model, the F1-score drop is similar, from 96.3% to 84%. However, for this posture change classification, the KF_{TS} model is subject to steep subject bias and hence the F1-score drops sharply from 100% to 66%. So in this case also, it is necessary to evaluate the other model parameters too instead of a dependence on just one single feature.

4.4 Discussions

The ACF, KF_{LPC} and KF_{TS} models that have been developed ab initio in Chapters 2 and 3, and more specifically the respective model parameters that been identified, have been used in this Chapter to study whether the presence or the absence of hypertension affects the some of the model parameters, which ones most significantly and how effective are these parameters in discriminating the two conditions.

The second application studied using the same sets of model parameters is that of posture change. The questions looked into are - whether a change of posture from sitting to supine leads to significant effect on some of the model parameters of the three models and how reliably can this effect be used to differentiate the two postures?

Both these studies have been performed on the three models separately using the open source data mining software WEKA, version 3.9.4 [1,190,191]. The same DS2 and DS1 DDP signal datasets and the methods for signal quantization that have been used for the hypertensive-normotensive and the posture change classifications respectively in [1], [189] are used in the present study also.

Both of these are binary classification studies involving two states, namely Hypertensive and Normotensive in the first study and Sitting and Sup 1 (supine) in the second study. The first study classifies subjects in one of the two classes according to their underlying health condition. This is an unbalanced classification since it is performed with a total of 156 sets of data acquired from 5 hypertensive and 17 normotensive subjects. In the second study, all 24 subjects undergo the physical posture change leading to a balanced classification on the basis of a total of 57 acquired sets of data.

The choice of a suitable classification method and associated attribute evaluator and/or ranker algorithm is critical to the performance of the classifications. Since the random forest (RF) classifier was established as the most viable one in [1], hence it has been used in these studies too. Thereafter, it has established in both these applications that the InfoGainAttributeEval evaluator cum ranker which utilizes the spectral entropy or information gain method provides the most balanced and among the highest accuracies for all the three models. So both the studies have been performed using this ranker algorithm.

In the hypertensive study, the total number of parameters used are 180 for the ACF model and 170 each for the two KF models. However, using the ranker, it is observed that only 3, 3 and 1 parameters are identified to be most significant for the classifications based on the ACF, KF_{LPC} and KF_{TS} models respectively. The 3 ACF parameters are all normalized variances of the phasic components of 3 different states, while all the parameters of both KF models are the indices related to the process noise covariance matrix at compromise Q_{com} that has been identified from the performance metrics derived from the KF algorithm [174]. Furthermore, it is observed that of these 7 parameters selected across the 3 models, 6 are derived from the left hand (LH) data. This is in accordance with the observation in [1] that the LH signals being closer to the heart seem to respond more significantly to hypertension or its absence.

In the posture change study, 5 parameters are selected for the ACF model. 2 each of the 5 ACF model parameters are related to the phasic and the tonic components while 1 additional parameter is the nominal ACF model coefficient a_0. The box plots of the normalized variance of the phasic and tonic components in the 2 classes are well differentiated, while those of the other 3 parameters are somewhat similar.

As in case of the hypertensive classification, for the posture change classification also, only the parameter related to the process noise covariance at compromise Q_{com} [174] is selected for both the KF models. This exemplifies the effect of posture change on this performance metric parameter.

The box plots of the q_{com} for the two classes in the KF_{LPC} model are absolutely segregated while those in case of the KF_{TS} model have a significant overlap. This is expected to affect the classification results for these two models.

Three standard cross-validation methods have been applied on the selected model parameters of the 3 models in both the classification studies. These are 10 fold cross-validation (10FCV), leave one out cross-validation (LOOCV) and leave one subject out cross-validation (LOSOCV). In all cases, it is observed that the results of 10FCV and LOOCV are either identical or

very close. However, since the subject bias effect is expectedly removed in LOSOCV, the results for the method differ significantly from those of the other two methods.

In case of the hypertensive classification, the accuracies and the weighted average (WA) F1-scores of the ACF, KF_{LPC} and KF_{TS} models for 10FCV method are (98.07%, 98%), (99.89%, 100%) and (99.51%, 99.5%) respectively. Thus, despite the unbalanced data sets, the 10FCV results for the accuracy and WA F1-score are almost identical for the ACF and KF_{TS} models, with those for the KF_{TS} model being higher. A comparison of the 10FCV results for the accuracy and F1-scores of the 3 models with those stated in [1] establish that all 3 models provide competitive scores for this binary classification. However, the F1-score at 100% in case of the KF_{LPC} model needs to be investigated for overfitting and/or subjective bias. The WA Precision and Recall values match with the accuracy for all 3 models while the Specificity is slightly lower in case of the ACF and KF_{TS} models at 89.9% and 97.3% respectively.

The effect of subject bias on the classification is studied using the LOSOCV approach. The values of all the metrics in all 3 models reduce significantly. Since this is an unbalanced classification, hence the F1-score is most relevant. It is observed that the drop in F1-score is maximum for the KF_{LPC} model from 100% to 74% indicating steep subject bias in the data. On the other hand, the F1-score drops minimally for the ACF model by 17% to 81%. The final LOSOCV WA F1-score value is also maximum for this model. Thus the features selected for the ACF model provide the most reliable and best F1-score result.

In case of the KF_{TS} model, the F1-score drops quite steeply from 99.5% to 77% but the scores for the individual states are most balanced and the maximum score of 33% for the hypertensive state is obtained using this model. The specificity values for the KF_{TS} model (WA: 0.53) are also the most balanced and highest of all 3 models and is almost twice that of the other two models, specifically WA of 0.28 and 0.18 for the ACF and KF_{LPC} models respectively. These factors establish the utility of the KF_{TS} model in this classification despite the overall drop in F1-scores.

In case of the sitting and supine posture classification, the accuracies and the weighted average (WA) F1-scores of the ACF model for 10FCV method are almost identical at 96.34% and 96.3%. This is to be expected since this is a balanced classification. The respective values for the two classes (or states, in this case) are also similar. All these factors are indicative of high reliability of this classification. This is justified by the relatively medium drop in the F1-score for LOSOCV by 12.3% to 84%.

In case of the KF_{LPC} and KF_{TS} models, these values are identically 100%.

However, as mentioned earlier, the characteristics of the selected feature in case of the KF_{LPC} model justifies this result but that is not so evident in case of the KF_{TS} model. The LOSOCV result for the F1-score of the KF_{LPC} model shows a drop of 12% to 88% indicating marginal subject bias in the data. In case of the KF_{TS} model, this drop is much sharper from 100% to 66% indicating a significant subjective bias and/or lack of sufficient number of discriminating features.

Thus, it is evident that the KF_{LPC} and ACF models show high 10FCV accuracies and relatively low or medium effects of subjective bias for the classification of posture change. The KF_{TS} model with its single selected feature is not suitable for this differentiation.

Chapter 5

Conclusions and Future Scope of Work

5.1 Conclusions

The aim of this work is to develop various models of the differential dermal potential (DDP) signals and to determine the use and efficacy of the parameters derived from these models in studying and/or classifying human conditions.

As mentioned at the outset of this work, this DDP signal is a specific type of endosomatic electrodermal activity (EDA) that is acquired in differential mode unlike the standard referenced mode [1], [24]. Since these signals have been proposed and are in use only recently, hence their characteristics as well as their modelling are yet to be investigated in detail. The study presented in this work utilizes the signals acquired for the work in [1] and hence suitable details of the acquisition of these DDP signals are provided in Section 1.1.6.

One of the first steps for developing a model is to determine the signal order. This aspect has been addressed in this work from a new perspective based on the derivation of the sample autocorrelation function (ACF) of three types of signals. The signal order thus obtained has been used to develop three different models from basic principles.

DDP signals have been acquired by Sarkar [1] from 4 types of experiments labelled as DS1, DS2, DS3 and DS4 using a specific 4-channel data acquisition system (DAS) and RISH multimeter, as stated in Section 1.1.6. The results of the cross-correlation of simultaneously acquired signals from the 4 limbs in DS4 establish the independence of these signals. Hence, a total of 5706 numbers of DDP signals acquired from all the 4 experiments have been considered for all the three modelling studies. Each signal is 2 minutes long

with a sampling time of 50 ms and thus has 2400 discrete samples.

The parameters obtained from each of the three models have been used individually in the three applications studied in [1] related to the DDP signals, namely a) study the effect of long duration of rest, b) classify the presence or absence of hypertension and c) differentiate between the sitting and supine postures of a subject.

5.1.1 Model order identification

In this work, the lag value at which the 1^{st} zero crossing of the sample ACF of the signal occurs is termed as the zero crossing lag (ZCL), or the l_{zc}. The model order of a signal is popularly determined in terms of the dying out of the sample ACF of stochastic processes, as discussed in Box and Jenkins et al. [27]. However, no closed form solution is provided for the ZCL of any time-series signal. The same sample ACF is used in this work for the same purpose but here a method to determine the model order for a real time-series signal is developed based on the closed form solution of the ZCL.

Closed form solutions for the l_{zc} are determined for various types of signals. The first of these is the deterministic i^{th} order power law series, denoted as y_{dk}. The l_{zc} of this general series is uniquely obtained and is decided by i and the total time instant N, but is independent of the coefficients a_i or sampling time T.

This has been determined as a special case of the stochastic i^{th} order power law series. This series, denoted as y_{sk}, is considered as the sum of the deterministic series y_{dk} and an additive zero mean Gaussian noise series w_k that is uncorrelated with the deterministic component. Another signal type considered is the deterministic polynomial series y_{pk}. It has been shown that the l_{zc} in both these cases are not unique and furthermore, their expressions depend on all the terms i, N, a_i and T.

The closed form solutions of the l_{zc} of the 1st order stochastic polynomial model y_{k1} as well as its deterministic counterpart y_{pk1} have also been determined. It is established that the l_{zc} of y_{pk1} is unique. In case of y_{k1}, the l_{zc} is unique provided the particular noise sequence w_k and its correlation is known. It is further observed that the l_{zc} of y_{pk1} does not depend upon the coefficients a_0 or a_1 or the sampling time T but is dependent only on the number of observations N while that of y_{k1} is independent of a_0 but varies with a_1, T, and N.

The n^{th} order stochastic ACF-ZCL model of the actual time-series signal, as well as the signal itself, is represented as y_k. This model is composed of the deterministic and noise components of y_{sk} (at $i = n$) along with the 0 to $(n-1)$ order components of y_{pk}. It is found from the theoretical derivations that

due to the noise interactions as in y_{sk} as well as the coefficient interactions as in y_{pk}, the l_{zc} of y_k can be shifted within limits towards left or right lag as compared to that of the underlying y_{dk}. These findings have been validated from the results of the stochastic ACF-ZCL models of acquired DDP signals.

It is established that there exists a one-to-one map between the ZCL and the order n of the deterministic power law series y_{dk}. This in conjunction with the shift in the ZCL within limits due to noise or polynomial coefficient interactions forms the basis for determining the order n of the time-series signal y_k for all three types of models proposed in this work.

5.1.2 Stochastic ACF-ZCL model

It is known that real time EDA signals can be decomposed into a slowly varying *tonic* component and a fast changing *phasic* component [15]. In the model for the DDP signal also, the deterministic trend or tonic competent of these signals is ascertained as the polynomial regression time series with the order n as determined using the ZCL. The resultant error time series, obtained as the residual of the signal and the polynomial fit component, is typically serially correlated and represents the stochastic phasic component.

In order to obtain this model for the DDP signals, an algorithm is proposed to fit the proposed stochastic ACF-ZCL model, also referred to in short as ACF model, to actual time-series signals y_k. The algorithm utilizes the ZCL of the actual signal and the findings relating the order n of the polynomial fit to the ZCL. In this, the preliminary model order n determined from the ZCL of the signal is updated using the ZCL of the residuals r_k, till a sufficiently large SNR is assured. The updated n is finally used to determine the standard best fit polynomial y_{pk} and the residual series r_k of the proposed additive ACF model $y_k = y_{pk} + r_k$.

As mentioned at the outset, a total of 5706 DDP signals have been considered for determining their ACF models. It is observed that using the proposed algorithm, 902 signals can be fitted with linear models while the remaining 4804 signals are fitted with higher order models. The efficacy of the ACF model is tested by comparing these models with corresponding models for these signals whose orders are determined using the standard partial ACF (PACF) method.

It is observed that the order of PACF model is greater or equal than ACF model in 70% cases means the model complexity is high. To observe this comparison results more preciously, Sum of squared errors (SSE) of all signals are calculated for both cases and it is found that in 30% cases, SSE of ACF model is less. This may be due to the local large spikes observed in several of the signals. The root of mean of squared (RMS) of the residuals of

all signals are also calculated. It is observed that the RMS_{ACF} (RMS_{PACF}) is within twice (1.78 times) of RMS_{PACF} (RMS_{ACF}) in 90% cases and is at most 12.82 (15.21) times that of the RMS_{PACF} (RMS_{ACF}) considering all the signals. The results of the statistical analysis of the residuals are almost similar in both cases and the Durbin-Watson test results show that both sets of residuals are positively correlated. From the Pearson coefficient tests, it is observed that 60.31% (53.62%) of the residuals of the ACF (PACF) models follow beta distribution.

It is thus observed that the deterministic tonic component of the ACF model typically has a lower order polyfit component than the corresponding PACF model and the ACF model residuals follow the beta distribution more consistently. The observations and interpretations from this study confirm that the proposed stochastic ACF-ZCL model of a time-series is well established with well defined tonic and phasic components.

5.1.3 KF_{LPC} and KF_{TS} models

It is well known that acquired signals are corrupted with sensor noise and furthermore, some process disturbances or noises may also distort the true nature of the signal being acquired. The Kalman filter (KF) is a popular technique used to clean these noises and obtain dynamic system models that respond to internal and external valid inputs. This KF algorithm requires a suitable state space model description with well-defined states, system matrices and output relations as well as process and measurement noises and their statistical descriptions. The Autoregressive (AR) modelling approach that is typically adopted for this purpose provides the required linear prediction coefficients (LPC) of the state space model. The KF model developed in this work using this standard approach is termed the KF_{LPC} model.

It has already been established in this work that the ACF model represents the tonic and phasic components of the signal adequately. Hence it is used to derive ab-initio another KF model. y_{pk}, which is the polynomial time-series (TS) developed in the ACF model, is used to obtain the description of the state space model of this model and so it is termed the KF_{TS} model.

Once the respective state space model formulations are obtained, these are used in the standard KF algorithm to obtain the KF_{LPC} and KF_{TS} models. Hence these models differ only in the state space model descriptions. However, the differences in the LPC and TS state space models are manifold - in the states, the system matrices and also in the way the noises are defined.

In case of the LPC state space model, the LP coefficients obtained from the AR model of the signal are used to obtain the states and system matrix.

The process noise covariances are also obtained from the AR process. These make up the process model of the state space description. Thereafter, the measurement noise is derived as the difference between the signal y_k and the corresponding process model output.

In case of the TS model, the tonic component y_{pk} of the ACF model is assumed to be identical to the deterministic component of the TS process model. Accordingly, the states and system matrix are derived ab-initio from this polynomial time-series y_{pk}. As for the noises, the measurement noise is obtained from the sensor noise calibration data and using this, the process noise is obtained as a difference between the ACF model residual and the measurement noise.

Once the respective state space models are obtained, the KF_{LPC} and KF_{TS} models of the DDP signals and their output in terms of the cleaned signal estimates are obtained using the standard KF algorithm. As mentioned earlier, the order of the signals are determined based on the l_{zc} obtained using the signal ACF. These cleaned KF estimates are then partitioned into the tonic and phasic components following the same approach used in the ACF model. The tonic components in the KF_{LPC} and KF_{TS} models are obtained by fitting suitable polynomial regression models to the respective cleaned estimates and their difference is ascribed as the phasic components.

Thereafter, the robustness and sensitivity metrics [174] of the KF models are determined. Since the DDP signal is a one-dimensional signal, the robustness and sensitivity metrics for these two KF models are scalar in form. These simplified scalar forms of the metrics for the DDP signals are derived and their relation to the model performance is interpreted in this work. It is observed that while the robustness metric decides the goodness of the process model description, the sensitivity metric is a direct measure of the Kalman gain. Thus J_{com}, which is the value at which these two metrics are balanced, is a characteristic parameter of the KF model. Furthermore, the error covariance of the aposteriori, or post-measurement, cleaned estimate is related directly to the sensitivity metric and inversely to the robustness metric. This error covariance is also proportional to the scalar process noise covariance q. Hence the value of q_{com}, which is the q for J_{com}, is another characteristic KF parameter.

The stochastic ACF-ZCL model, the KF_{LPC} and the KF_{TS} models of the DDP signals are compared to ascertain the validity and efficacy of these three proposed models. For this, both types of KF models are fitted to the same 5706 signals as are used to obtain the ACF models. They are compared in terms of the respective fit signals, their parameters and the statistical characteristics of the phasic components. The various parameters of the robustness and sensitivity metrics of the two KF models are also compared.

It is observed that the tonic and phasic components of the KF_{TS} models are almost identical to those of the ACF model, while those of the KF_{LPC} models are distinguishably separate. In all cases, 80% of the phasic components lie within ± 0.2 mV. The KF_{LPC} residuals show a bias of 0.45 to 0.48 mV with SD of 1.34 mV for 95% population, while the KF_{TS} residuals are zero mean with 0.01 mV SD for 95% population as expected. The SD of the phasic components in all 3 models lie within 0.19 mV for 95% of the population.

The sensitivity metric J_1 and its maximum value J_{1max} show a variability in case of the KF_{TS} models while those of the KF_{LPC} models are fixed in nature. Accordingly, the two key parameters J_{com} and q_{com} are variable for the KF_{TS} models and invariant for the KF_{LPC} models. Furthermore, the values of q_{com} for the two KF models are opposing in nature. The positive values of q_{com} for the KF_{LPC} models indicate robust nominal system behaviour while the negative values of q_{com} for the KF_{TS} models indicate sensitive model behaviour.

Thus, the KF_{TS} model developed ab-initio using the system order and polyfit obtained in the earlier proposed ACF model as well as the KF_{LPC} model obtained using the standard AR approach provide consistent tonic and phasic components, although the KF_{LPC} model estimates show a distinct bias. Since the nature of the metrics in the two KF models are distinctly separate, hence it is expected to lead to different interpretations in the applications of these models.

Here, three specific applications of the DDP signals have been studied using the datasets recorded in all four experiments, namely DS1, DS2, DS3 and DS4 as stated in Section 1.1.6.

Application 1: Determination of the effective duration of rest in supine posture.

Application 2: Classification of hypertensive and normotensive subjects.

Application 3: Classification of sitting and supine postures.

Three standard cross-validation methods have been applied on the selected model parameters of the 3 models in the classification studies performed in Applications 2 and 3. These are 10 fold cross-validation (10FCV), leave one out cross-validation (LOOCV) and leave one subject out cross-validation (LOSOCV). In all cases, it is observed that the results of 10FCV and LOOCV are either identical or very close. However, since the subject bias effect is expectedly removed in LOSOCV, the results for the method differ significantly from those of the other two methods.

Of the 3 proposed models, the stochastic ACF-ZCL model provides the basis for order determination and also some key features and understanding of the tonic and phasic components of the recorded DDP signals, while any one of the two KF based models, KF_{LPC} or KF_{TS}, is very effective in distinguishing the various classes in the aforementioned 3 applications. These two KF models are distinctly separate in terms of their characteristics (as evident from Table 3.2). Hence, the efficiency of the KF models depend on the particular application. It is observed that while the KF_{LPC} model is very effective in distinguishing posture change (as evident in Figure 4.5 of Section 4.3, Chapter 4), the KF_{TS} model is more efficient in differentiating stages of rest (as seen in Figure 3.10 of Section 3.6.1, Chapter 3) as well as for the Hypertensive-Normotensive classification (as evident in Figure 4.3 of Section 4.2.2, Chapter 4).

A detailed description of the 3 applications vis-a-vis the 3 models are presented in the 3 subsequent subsections. Furthermore, each subsection contains a segment in which the results obtained using the 3 proposed models are compared with existing results for the particular application.

5.1.4 Determination of the effective duration of rest

In the DS4 experiment designed by Sarkar [1], total 68 numbers of 10 minute 4-channel data sets were recorded from 16 volunteers over a total duration of 6 months. Since it has been established in the same work that the signals acquired simultaneously from the four limbs can be treated to be uncorrelated to each other, hence a total of $68 \times 4 = 272$ sets of independent signals have been obtained from subjects in restful supine condition for 10 minutes. Thereafter, each of these signals have been segmented into 5 non-overlapping sets of 2-minutes signals. These are denoted as (rest) state 1, 2 and so on till state 5. So, $272 \times 5 = 1360$ numbers of 2-minute duration signals with $N = 2400$ discrete samples each are available for the analysis since the sampling time is $T = 50$ ms.

The effect of rest on the various parameters of the ACF model have been studied. In order to compare the acquired signal and the ACF model, a comparison is done of the mean of the acquired signal y_k and its polynomial fit y_{pk} as well as the standard deviations (SD) of y_k and r_k. Further studies on the effect of rest are done in terms of the ZCL or l_{zc}, the model order n and the spectral entropies of y_k, y_{pk} and the residual r_k.

It is observed that the statistical characteristics of the mean and SD of the acquired signals are almost identical with those of the mean of the tonic polynomial and the SD of the stochastic residuals respectively for all the states. This establishes that these represent the tonic and phasic components

of the DDP signals.

In [1], the effects of restfulness have been observed in 3 specific spectral characteristics of the DDP signals. These are central frequency, power at central frequency and spectral entropy. Furthermore, from the maximum entropy analysis, the effective duration of rest is ascertained to be 4 to 6 minutes. This is supported by the nature of changes of the box plots of the 3 spectral parameters also.

As for the effect of rest on the ACF model parameters, it is observed that the median of l_{zc} is almost equal in all the consecutive 5 states but the order n is low in state 4. However, as observed in [1] for the entropy of the overall signal, it is observed that the mean as well as the IQR of the entropy of the *tonic component* of the ACF model is maximum in state 3 and then decreases indicating that the subject achieves a fully rested condition typically within 4-6 minutes of no-nap supine rest. Furthermore, the variations of the ZCL, system order, mean of the tonic component and the SD of the residuals support the emerging determinism of the system with rest.

The proposed stochastic ACF-ZCL model thus provides tonic and phasic components that match those of the original time-series signal. Specific parameters of the model can also be identified that are affected consistently due to physiological change as in the case of increased duration of rest.

The effect of rest on the various parameters of the two KF models have also been studied. The model parameters considered are the spectral entropies of the tonic components, the phasic components and the noise residuals of the KF models; the mean and SD of the tonic components; SD of the phasic components and the two metric parameters J_{com} and q_{com}.

In case of the KF_{TS} model, the spectral entropy (SE) is an useful identifier of the tonic components, the phasic components and the noise residuals. The SE values of each of these components are observed to be fairly constant across all stages of rest. However, their medians and ranges are observed to increase from the tonic components to the phasic components and are maximum for the noise residuals indicating sequential departure from regularity to irregularity as expected.

The spectral entropy (SE) of the tonic and phasic components of the KF_{LPC} models are quite different. The SE values are observed to increase till the second or third state of rest and then decrease.

The mean of the tonic components of both the KF models vary similarly as that of the ACF model.

The SD of the tonic components of all 3 models show a decreasing nature. However, while this parameter varies identically for the ACF and KF_{LPC} models by decreasing exponentially with increasing rest, that for the KF_{TS} models decreases in a quasilinear manner.

In case of the SD of the phasic components of the models, those of the ACF and KF_{TS} models vary identically, decreasing exponentially with increasing rest while that of the KF_{LPC} models decrease initially and then increase and decrease slightly to become almost stationary with increasing rest.

The metric parameter J_{com} for the two KF models are both quite close to $n/2 = 0.5$ and almost invariant across all states. However, one distinguishing feature in case of the KF_{TS} models is that the lower bound of J_{com} drops sharply in the third state and then becomes constant.

As observed earlier in the general case, the q_{com} for the two KF models are intrinsically different. This parameter is represented in terms of its nominal value q_{nom} as $q_{com} = 10^l q_{nom}$ where l is termed as the multiplier. The typical multiplier values are fairly constant and around $+3$ in case of the KF_{LPC} models, while in case of the KF_{TS} models, it varies within -1 and -0.5. More specifically, the q_{com} of the KF_{TS} model increases sharply and distinctly with increasing rest providing a distinguishing parameter for the stages of rest.

Thus, the key features of the model components of the ACF model and the metrics of the KF_{TS} model can be used to reliably differentiate the stages of rest. The standard KF_{LPC} model also performs creditably in several aspects though not so markedly.

Comparison with existing methods

Rest in terms of *short nap* has been a topic of research in several studies as tabulated in Table 5.1. In most of these studies, sophisticated standard techniques like ECG and EEG have been used for sleepiness or alertness detection, while in some cases EOG has also been used to track slow eye movement that occurs due to fatigue or drowsiness along with EMG. Self reporting or performing definite tasks have been used in almost all cases to check change in efficiency and the output is measured using any standard scale.

In the present case, the objective is to determine a physiologically determinable effective duration of rest while the subject was in daytime no-nap supine condition in a controlled environment. Hence, the subjects were not disturbed during the experiment and neither was any self reporting done by them. It has already been established in Section 2.4 and Section 3.6 that the effect of restfulness is obtained using the spectral entropy of tonic component, SD of phasic component of ACF-ZCL model and SD of phasic component, q_{com} of KF_{TS} model in 4 to 6 minutes. For KF_{LPC} model, restfulness is predicted in 2 to 4 minutes using the same parameters as in KF_{TS} model.

Table 5.1: Estimation of effective duration of rest

Serial	Authors & year	Experiment	Measurement	Outcome
1	Mats Gillberg *et al.* (1996) [196]	Effect of short nap at day time	Self reported sleepiness/alertness, ECG and EEG	19.8 minutes
2	J. A. Horne *et al.* (1996) [197]	Effect of nap, coffee and placebo on drivers sleepiness/alertness	Self reported sleepiness, ECG indicates alertness	10.8 minutes
3	Masaya Takahashi *et al.* (2000) [198]	Effect of 15 minute nap after a short night sleep	Polygraph monitoring for sleep, Visual analog scale and ECG for subjective sleepiness and logical reasoning test for alertness	10.2 minutes
4	Amber J. Tietzel *et al.* (2001) [199]	Estimation of effective duration of short term rest after noctural sleep restriction	ECG and EOG measurement	10 minutes
5	Amber J. Tietzel *et al.* (2002) [200]	Estimation of effective duration of short term and ultra short rest after noctural sleep restriction	EEG and EOG measurement, different tests	10 minutes
6	Mitsuo Hayashi *et al.* (2005) [201]	Corrective power of short day time nap	Subjective mood, visual detection and symbol-digit substitution tasks, and slow eye movements	9.1 minutes

Table 5.1 – continued from previous page

Serial	Authors & year	Experiment	Measurement	Outcome
7	Mitsuo Hayashi *et al.* (1999) [202]	Effect of 20 minute nap and no nap at noon	EEG, mood, performance, and self-ratings of performance level	20 minutes
8	Mitsuo Hayashi *et al.* (1999) [203]	Effect of 20 minute afternoon nap on mood, performance and cardiac activity	EEG, mood, performance, and self-ratings of performance level	20 minutes
9	Mitsuo Hayashi *et al.* (2003) [204]	Effect of short daytime nap with caffeine, bright light and face washing on day time sleepiness	event related potentials (ERP) and EEG power spectra	15 minutes
10	Masaya Takahashi *et al.* (1998) [205]	Effect of post lunch short nap on alertness, performance, and autonomic balance	event-related potential, subjective sleepiness and ECG	7.3 minutes
11	Amber Brooks *et al.* (2006) [206]	Comparison of different length of afternoon nap	ECG, EMG, EOG	10 minutes
12	Nasser Al-Busaidi (2018) [207]	Subjective and objective effects of post lunch short nap	EEG,EOG and work place questionnaire	10 minutes
13	Cassie J. Hilditch (2016) [208]	10 or 30 minutes nap, which one associated with sleep inertia	The Samn-Perelli Fatigue Scale for fatigue measurement, polysomnography (PSG), EEG, EMG, EOG	10 minutes

Table 5.1 – continued from previous page

Serial	Authors & year	Experiment	Measurement	Outcome
14	Mohamed Romdhani (2020) [209]	Effects of nap after sleep loss on reaction time, mood, and biochemical response in athletes	Anaerobic sprint test, reaction time, Hooper index, Epworth Sleepiness Scale, biomarkers and antioxidant status.	20 minutes
15	Sarkar (2023) [1]	Spectral entropy (SE) and other relevant features of DDP signals	Maximum entropy	4 - 6 minutes
16	Proposed ACF method	1. SE of tonic component, 2. SD of phasic component	1. Maximum of Median of SE and 2. Median of SD drops below 0.10 mV	4 - 6 minutes
17	Proposed KF_{LPC} model	1. SD of phasic component, 2. Level of process noise q_{com} at compromise metric (in terms of nominal q, q_{nom}) [174]	1. Median of SD $< 0.15mV$, 2. Median of $q_{com} > 800q_{nom}$	2 - 4 minutes,
18	Proposed KF_{TS} model	1. SD of phasic component, 2. q_{com}	Median of SD $< 0.10mV$, Median of $q_{com} > 0.15q_{nom}$	4 - 6 minutes,

The outcomes of these experiments in terms of the duration that provides subsequent enhancement in subject efficiency vary between 7.3 minutes to 20 minutes, but most of these studies converge to an effective duration of around 10 minutes.

5.1.5 Classification of hypertensive and normotensive subjects

One of the human conditions studied in [1] is the hypertensive and normotensive classification, or detecting the presence or absence of hypertension in subjects. The open source data mining software WEKA, version 3.9.4 [190] has been used for this binary classification study. In this, the random forest (RF) classifier has been used since it is established as the most viable one.

This is a binary classification involving two classes, the Hypertensive and the Normotensive classes, and the subjects are accordingly labeled in one of the two classes according to their underlying health condition. Since the study is performed on the DS2 dataset which comprises of a total of 156 sets of data acquired from 5 hypertensive and 17 normotensive subjects, hence it is an unbalanced classification.

In the present work, the efficacy of the parameters of the ACF, KF_{LPC} and KF_{TS} models of the DDP signals has been studied in this classification. The same DDP signal datasets and methodology used in [1] are used in the present study also. This study ascertains whether the presence or the absence of hypertension affects some of the model parameters, which ones most significantly and how effective are these parameters in discriminating the two conditions.

The total number of parameters used are 180 for the ACF model and 170 each for the two KF models from each hand. The RF classifier is used in this case also but the associated attribute evaluator and/or ranker algorithm has been chosen after testing three viable alternatives since this is critical to the performance of the classifications. The InfoGainAttributeEval evaluator cum ranker which utilizes the spectral entropy or information gain method has been chosen in this application. This uniformly provides the most balanced and among the highest accuracies for all the three models.

Using the ranker, it is observed that only 3, 3 and 1 parameters of the ACF, KF_{LPC} and KF_{TS} models respectively are identified to be most significant. The 3 ACF parameters are all normalized variances of the phasic components of 3 different states, while all the parameters of both KF models are the indices related to q_{com}. Thus it can be said that q_{com} plays a characteristic role in capturing information related to changes in the signal

dynamics in this classification.

Furthermore, it is observed that of these 7 parameters selected across the 3 models, 6 are derived from the left hand (LH) data. This is in accordance with the observation in [1] that the LH signals being closer to the heart seem to respond more significantly to hypertension or its absence.

In the hypertensive classification, the accuracies and the weighted average (WA) F1-scores of the ACF, KF_{LPC} and KF_{TS} models for 10FCV and LOOCV method are (98.07%, 98%), (99.89%, 100%) and (99.51%, 99.5%) respectively. Thus, despite the unbalanced data sets, the 10FCV results for the accuracy and WA F1-score are almost identical for the ACF and KF_{TS} models, with those for the KF_{LPC} model being higher. A comparison of the 10FCV results for the accuracy and F1-scores of the 3 models with those stated in [1] establish that all 3 models provide competitive scores for this binary classification. However, the F1-score at 100% in case of the KF_{LPC} model indicates over fitting and/or subjective bias. The WA Precision and Recall values match with the accuracy for all 3 models while the Specificity is slightly lower in case of the ACF and KF_{TS} models at 89.9% and 97.3% respectively.

The effect of subject bias on the classification is studied using the LOSOCV approach. The values of all the metrics in all 3 models reduce significantly. Since this is an unbalanced classification, hence the F1-score is most relevant. It is observed that the drop in F1-score is maximum for the KF_{LPC} model from 100% to 74% indicating steep subject bias in the data. On the other hand, the F1-score drops minimally for the ACF model by 17% to 81%. The final LOSOCV WA F1-score value is also maximum for this model.

It is further noted that in case of the KF_{TS} model, the F1-score drops quite steeply from 99.5% to 77% but the scores for the individual states are most balanced and the maximum score of 33% for the hypertensive state is obtained using this model. The specificity values for the KF_{TS} model (WA: 0.53) are also the most balanced and highest of all 3 models and is almost twice that of the other two models, specifically WA of 0.28 and 0.18 for the ACF and KF_{LPC} models respectively.

These results can be compared with the classification accuracy of 0.765 and F1-scores using 10FCV, LOOCV and LOSOCV of 0.752, 0.756 and 0.738 respectively achieved in [1] using the selected features of the acquired DDP signals. This shows that the selected features of all the 3 models, and in particular the ACF and KF_{TS} models, are more effective for this hypertensive classification. In particular, the features selected for the ACF model provide the most reliable and best F1-score result, while the balanced and highest specificity establishes the utility of the KF_{TS} model in this classification despite the overall drop in F1-scores.

Thus the features selected for the ACF model of the DDP signals provide the most reliable and best F1-score results for the hypertensive-normotensive classification, followed by the feature for the KF_{TS} model.

Comparison with existing methods

Hypertension causes many severe health complications like heart problem, stroke or even death. According to the WHO report [210], 17.9 million people have died in 2019 due to cardiovascular diseases. Among these, 85% were due to heart attack and stroke [211]. 63% of the total deaths in India are due to noncommunicable diseases, of which 27% are due to cardiovascular diseases [212]. The different stages of hypertension are detailed in [213]. Since this fatality occurs due to lack of awareness, lack of primary care and follow up, so it is essential to have a simple, reliable automated screening system along with the usual BP monitoring system.

A comparison of the existing methods for classifying hypertensive subjects is presented in Table5.2. It is observed that there exist methods based on ECG with accuracies well above 90%. The classification accuracy of the DDP signals based method is at most 80.60%. In comparison, the present method shows 83.01%, 69.87%, 72.70% accuracy using LOSOCV for ACF-ZCL, KF_{LPC}, KF_{TS} models respectively.

Yet, the simplicity of its acquisition even by nominally skilled health workers and the minimal subject discomfort during its acquisition using the simple 10 minutes rest protocol can be useful for primary monitoring and screening purposes.

Table 5.2: Comparison of the existing methods for classifying hypertension [1]

Serial	Authors & year	Experiment	Classification technique	Accuracy
1	Yongbo Liang *et al.* (2018) [153]	Hypertension assessment using ECG and PPG	Logistic Regression, AdaBoost Tree, Bagged Tree, and k-Nearest Neighbur classification	F1 score: 94.84%
2	Gabor Kovacs *et al.* (2016) [156]	Pulmonary hypertension assessment using ECG and other non-invasive tools	Multivariate Logistic Regression	PPV:92%, NPV:97%
3	J. S. Rajput *et al.* (2020) [155]	Automatic detection of severity of hypertension ECG signals	Optimal bi-orthogonal Wavelet Filter bank	Average accuracy: 99.95%
4	Hongbo Ni *et al.* (2015) [213]	Severity of hypertension	Multiscale fine-grained HRV analysis	Precision: 95.1%
5	M. Simjanoska *et al.* (2018) [214]	Blood pressure classification from ECG signals	Complexity analysis-based Machine Learning	Precision: 96.68%
6	M.G.Poddar *et al.* (2019) [215]	Automated Hypertension classification technique from HRV analysis	Probabilistic Neural Network, k-Nearest Neighbour, and Support Vector Machines classifiers	96.67% in SVM

Table 5.2 – continued from previous page

Serial	Authors & year	Experiment	Classification technique	Accuracy
7	Tim Seidler *et al.* (2019) [216]	Pulmonary hypertension detection from ECG signals	Random Forest of classification trees and Regression Trees, Lasso penalized Logistic Regression, Boosted classification trees, Support Vector Machines	Accuracy: 95%, AUC: 0.87
8	Sarkar (2023) [1]	1. SD of left hand (LH) of DDP signals 2. Gap and Pairsum (PS) of LH and right hand (RH) signals and 3. lateralization coefficients of mean of LH and RH signals	Random Forest of classification trees and Regression Trees	Accuracy: 80.60%
9	Proposed ACF model	Normalized variance of phasic component	RF classification using Leave one subject out (LOSOCV)	Accuracy: 83.01%
10	Proposed KF_{LPC} model	q_{com}	RF classification using LOSOCV	Accuracy: 69.87%
11	Proposed KF_{TS} model	q_{com}	RF classification using LOSOCV	Accuracy: 72.70%

5.1.6 Classification of sitting and supine postures

Another human condition studied in [1] involves a sitting and supine posture classification or detecting the change in posture of human subjects. This is also a binary classification of the Sitting and Sup 1 (supine) states. This study is performed on the DS1 dataset in which all 24 subjects undergo the physical posture change leading to a balanced classification on the basis of a total of 57 acquired sets of data. Here, the issues studied are whether a change of posture from sitting to supine leads to significant effect on some of the model parameters of the three models and how reliably can this effect be used to differentiate the two postures.

Unlike the hypertensive-normotensive classification, the total number of parameters used are 180 for the ACF model and 170 each for the two KF models from each hand. The RF classifier is also used here and InfoGainAttributeEval evaluator is used as the ranker.

Using the ranker, it is observed that 5 parameters are selected for the ACF model while only 1 parameter is selected in case of both the KF models. 2 each of the 5 ACF model parameters are related to the phasic and the tonic components while 1 additional parameter is the nominal ACF model coefficient a_0. The box plots of the normalized variance of the phasic and tonic components in the 2 classes are well differentiated, while those of the other 3 parameters are somewhat similar. The characteristic posture change parameter selected in case of both the KF models is the metric parameter q_{com}. However, the box plots of q_{com} for the two postures in the KF_{LPC} model are absolutely segregated while those in case of the KF_{TS} model have a significant overlap.

Here also three cross-validation methods 10FCV, LOOCV and LOSOCV have been used on the selected model parameters of the 3 models. In this classification, the results of 10FCV and LOOCV are identical or similar where as, the results of LOSOCV are quite different.

In case of the sitting and supine posture classification, the accuracies and the weighted average (WA) F1-scores of the ACF model for 10FCV method are almost identical at 96.34% and 96.3%. This is to be expected since this is a balanced classification. The respective values for the two classes (or states, in this case) are also similar. All these factors are indicative of high reliability of this classification. This is justified by the relatively medium drop in the F1-score for LOSOCV by 12.3% to 84%.

In case of the KF_{LPC} and KF_{TS} models, these values are higher and actually saturated at identically 100%. The characteristics of the box plot of the selected feature in case of the KF_{LPC} model justifies this result but that is not so evident in case of the KF_{TS} model. The LOSOCV result for

the F1-score of the KF_{LPC} model shows a drop of 12% to 88% indicating marginal subject bias in the data. In case of the KF_{TS} model, this drop is much sharper from 100% to 66% indicating a significant subjective bias and/or lack of sufficient number of discriminating features.

In comparison, it is observed that the overall accuracy obtained in [1] was 0.99, 0.98 and 0.93 for the 10FCV, LOOCV and LOSOCV experiments respectively. Hence, in this particular scenario, the basic statistical features of the acquired signal perform better in the classification with comparable reliable results obtained from the ACF model and moderate results for the KF_{LPC} model. It is evident that the KF_{LPC} and ACF models show high 10FCV accuracies and relatively low or medium effects of subjective bias for the classification of posture change. The KF_{TS} model with its single selected feature is not suitable for this differentiation.

Comparison with existing methods

A comparison of recent methods for posture classification is presented in Table 5.3. In a study, Foubert *et al.* proposed a sitting and supine classification technique using bed-based pressure sensor array and videography [217]. Results show very low miss rate with significantly high accuracy. In another study, a smart chair system has been developed combining six infra-red reflective distance with pressure sensor array for posture classification [218]. Use of IMU (inertial measurement unit) sensor in posture classification shows mixed performances in terms of accuracy [166, 219]. Overall accuracy in the experiment by Lockhart *et al.* is greater than 0.95 [219], while Vo states the classification accuracy of sitting posture to be 0.60 only, which is very low [166]. A common method of posture classification is by using data from multiple accelerometers such as light weight thigh worn activePal accelerometers [220–222]. The classification accuracy of this accelerometer data are very high in almost all cases, except the standing/light activity in upright posture in the experiment performed by Bassett *et al.* [221]. Furthermore, electrocardiogram (ECG) is also used for posture classification in many experiments. In these experiments, signals were acquired using normal ECG monitor [223], or Holter monitor [170] or even downloaded data available in online repository [224]. In [1], the posture classification was established using features of DDP signal with an 10FCV accuracy of 99.9% in 2-level. Comparing the results of these existing methods for posture classification with the present 3 types of model parameters, it is found that the accuracy values are 84.80%, 90.49% and 65.39% using LOSOCV for ACF-ZCL, KF_{LPC}, KF_{TS} models respectively. Therefore, it can be said that the classification results of the present approach (using LOSOCV) for KF_{LPC} model is close to

(or perhaps better than) that of the existing (10FCV) posture classification results using DDP signals [1].

Thus the validity of all three proposed models are established and the characteristic features of each of the models have been identified exhaustively. In case of the two KF models, their metric parameters have also been determined. It is observed that several of these identified features show relatable changes as the duration of rest of a supine subject is increased. All the three models show identifiable changes, yet of the three models, the changes in the ACF and KF_{TS} models are more significant. This observation holds in case of the hypertensive classification too. But in case of the posture change classification, the KF_{LPC} model classifies the postures most reliably and that too in terms of its single metric parameter. This also exemplifies the utility of the robustness and sensitivity metrics in such applications.

Table 5.3: Comparison of the existing methods for classifying sitting and supine posture

Serial	Authors & year	Experiment	Classification technique	Accuracy
1	Nicholas Foubert *et al.* (2012) [217]	Pressure sensor array	SVM, k-NN	0.94
2	Thurmon E. Lockhart *et al.* (2013) [219]	IMU signal	Wavelet based approach	0.968
3	David R. Bassett *et al.* (2014) [221]	2 accelerometers (activePal) data used	Coken kappa statistics	1.00
4	Nhat Nguyen Vo (2014) [166]	Accelerometer based sensor	ANN	Supine: 0.972 sitting: 0.608
5	Saeid Wahabi *et al.* (2015) [225]	ECG signal	SVM	Sitting: 0.98 and Supine: 0.94
6	Kate Lyden *et al.* (2016) [220]	thigh-worn, tri-axial activPAL3 accelerometer	their novel method	supine: 0.967, sitting: 0.929
7	Dai Sasakawa *et al.* (2018) [171]	Height and a Doppler radar cross section (RCS)	kNN	Sitting 0.965, supine 1.00
8	Nindynar Rikatsih (2018) [226]	Movement data of UCI Machine Learning Repository	kNN	0.995
9	Angel D. Ruiz1 *et al.* (2019) [224]	Electrocardiography and respiratory flow	linear SVM	0.995
10	Yutaka Yoshida *et al.* (2019) [170]	Holter electrocardiographic (ECG) recorder with built-in accelerometer	RF	Supine: 0.98, Sitting: 0.717
11	Sarkar (2023) [1]	RF classification using mean, SD and normalized variance	binary RF	Supine: 1.00, Sitting: 0.997

Table 5.3 – continued from previous page

Serial	Authors & year	Experiment	Classification technique	Accuracy
12	Proposed ACF model	Normalized variance of tonic, phasic and Normalized kurtosis of tonic components, a_0 and Spectral entropy of phasic component of DDP signals	Binary RF using LOSOCV	Accuracy:84.80%
13	Proposed KF_{LPC} model	q_{com}	Binary RF using LOSOCV	Accuracy: 90.49%
14	Proposed KF_{TS} model	q_{com}	Binary RF using LOSOCV	Accuracy: 65.39%

5.2 Future Scope of Work

The order identification technique as well as the model development procedures are proposed in this work for the specific DDP signals. However, these are applicable to most general signals, whether of human or other origins. Furthermore, since the model parameters determine the efficacy of the classification, hence it is possible to explore their use in general system condition monitoring applications.

It is to be noted in such applications that medication and hydration are crucial factors for the accurate assessment of health status since they can influence the robustness of physiological measurements in various health care settings. Hence, it is important to consider these factors carefully in such studies involving the interpretation of patient data. Suitable adjustments or corrections may be necessary to account for any physiological changes in such future work.

The DDP signals in particular can be explored for use in emotion recognition, human physiology analysis, psychiatric research, stress management, health monitoring, neuromarketing research, consumer research, gaming and virtual reality, safety and security and sleep research.

It has been shown that in case of the ACF model of these signals, the magnitude of the highest order coefficient a_n of the polynomial drops significantly from 10^{-2} to 10^{-25}) as the order increases from $n = 1$ to $n = 9$. In view of this finding, the order of the polynomial is limited to a maximum value of 9 in the algorithm proposed for determining the ACF model. It may be of interest to study the effect of the ignored higher order dynamics of the signal in various condition monitoring studies. This is particularly relevant for long duration signals since in the form of the polynomial y_{pk}, the multiplier for a_n is $(kT)^n$. This term thus increases exponentially as kT increases and hence its effective contribution may become significant as was observed from the dynamics present in the KF_{TS} model residuals in the restfulness assessment studies.

It is established that the proposed stochastic ACF-ZCL model provides physiologically interpretable tonic and phasic components of the DDP signal. It is of interest to study whether this model is useful for obtaining interpretable tonic and phasic components of general time-series signals obtained from other applications too.

Also, the restful condition studied in this work provides a useful baseline and thus these signals and their models can be used for interpreting changes in the state, whether physical or mental, due to applied stimuli. Such a study involving the signals acquired for baseline restful sitting state followed

by applied stimuli has been reported in [26]. However, there is a scope to use the ACF and/or the KF models for these signals in order to obtain additional interpretations.

The results from both the classification studies establish the utility of the 3 models in differentiating the two classes, albeit with varying significance. However, the number of parameters selected for the ACF model and in particular, the two KF models, are very low. This merits a detailed study using more parameters and also by combining parameters from all 3 models. It is to be expected that some alternative choice(s) of the classifier as well as the ranker algorithm(s) will lead to more consistent and improved interpretable results as in the case of the restfulness assessment. A judicious choice of these combinations of parameters can avoid the over fitting and subject bias.

It has been observed that the same metric parameter q_{com} for the two KF models affect the two classification results differently. Specifically, while q_{com} from KF_{LPC} model can differentiate the sitting and supine postures significantly, the same parameter from KF_{TS} model facilitates the classification of the two classes of subjects, hypertensive and normotensive, properly. The features of the ACF model perform creditably in both cases. Hence, there is a scope to combine the signal parameters stated in [1] and those of the ACF and KF_{TS} model parameters for improved and more reliable hypertensive and normotensive classification results. There is a also scope to combine signal parameters with those of the ACF and KF_{LPC} model parameters for posture change classification.

In general, the general parameters of the acquired DDP signals as well as those of the 3 models proposed in which work can be used to study, classify and interpret other physiological or external changes in state.

Appendix A

A.1 Consent form

Note to the subjects:

You are being asked to take part in the research project described below. There are no known risks associated with this research. If you do not take part in the study, there will be no penalty. If you choose to take part, you have the right to stop at any time. However, we encourage you to talk to the research group so that we know why you are leaving the project. If there are any new findings during the study that may affect you, you will be told about them. Your part in this study is confidential. None of the information will identify you by name. All records will be maintained in codes so that identity is not disclosed. The collected data and information may be used and transferred to other research studied in the Research Laboratory. Your personal information will be suitably anonymised and/or masked to prevent tracing or identifying your identity. If you decide to enrol in this project, your involvement will be required in 2 or 3 phases for the experimental portion only and this may last for about 10-12 weeks in a phase. The results of this research study may be presented at meetings or in publications; however, your identity will not be disclosed in those presentations.

ABOUT THE RESEARCH TOPIC:

Topic: "*Study and characterization of the relational aspects of human biopotentials externally acquired from multiple locations*"

Biopotential based systems like Electrocardiogram (ECG), Electroencephalogram (EEG) and Electrodermal activity (EDA) based sensors are used abundantly to characterize several infirmities or diseases in humans since the biopotentials in different parts of the human body change under various physical, mental as well as social stressors. In a recent study, biopotentials have been acquired bilaterally from human subjects using a low cost, passive, non-invasive EDA based system. A preliminary investigation of these potentials has indicated the existence of inter-relations between these bilateral signals. The aim of the present research is to perform a detailed study

of the relational aspect of these and other biopotentials acquired externally from human subjects and to derive some definitive indices. The scope of research includes a feasibility study of using these indices in an expert system to classify the overall condition of the human health inclusive of physical, mental, social and spiritual wellbeing.

The bilateral biopotentials will be acquired using two systems. The first system comprises of 4 sets of multimeters (Make: RISH Multi 18S) with compatible adapters (Make: RISH Multi SI232) for online simultaneous recording of a maximum of 4 channels of data to the PC. The second system has been built indigenously in this laboratory, which can acquire potentials simultaneously from 4 locations and send to the PC. Both systems transmit data to the PC via optically isolated transmitter-receiver module to ensure safety to the human body from electrical hazards.

I, Mr./Mrs./Dr./Ms. ---------- resident of ---------- do hereby give consent voluntarily to record my endosomatic skin potential and other data pertaining to the research work being conducted by Mr. Arindam Sarkar under the guidance of Prof. Ratna Ghosh and Prof. Bhaswati Goswami in Instrumentation and Electronics Engineering Department, Jadavpur University. I have no objection to publication of my case study details without my name in any journal or any other reviews. I am signing this consent form voluntarily, out of free will, without any pressure and in my full senses.

Date: Full Signature:

Place:

Consent form explained/witnessed by: ----------

Date: Full Signature:

Place:

A.2 Overall Questionnaire

Comprehensive questionnaire

	SUBJECT CODE : S_ _ _				
	Full name	পুরো নাম			
	Sex (Male/Female)	লিঙ্গ(ছেলে/মেয়ে)			
	Age (years)	বয়স(বছর)			
	Please Answer the following - (write Y for yes and N for no):	***নিম্নলিখিতপ্রশ্নগুলির (হ্যাঁ বা না দিয়ে) উত্তর দিন***	*Do you take any medicine prescribed by doctor?*	*Do you take any medicine on your own?*	*Remarks, if any.*
1.	Do you have any disease from the following list (any history of Chronic illness)? (Y/N)	আপনার নিচের লেখা ব্যাধিগুলি থেকে, এক বা একের বেশী কোনো ব্যাধি (দীর্ঘস্থায়ী অথবা দুরারোগ্য) আছে? (হ্যা বা না)	কোনো রোগ থাকলে, হ্যা বা না দিযে উত্তর দিন	*আপনি কি ডাক্তারের বলা কোনো ওষুধ খান?*	*আপনি কি নিয়মিত নিজে থেকে কোনো ওষুধ খান?*
	High Blood Pressure/ Hyper tension	উচ্চ রক্তচাপ			
	Diabetes	মধুমেহ			
	Kidney problem	কিডনির সমস্যা			
	High Cholesterol	কোলেস্টেরল বেশী			
	Chronic Constipation	দীর্ঘস্থায়ীকোষ্ঠ-কাঠিন্য			
	Chronic Cardiac problem	দীর্ঘস্থায়ী হার্টের সমস্যা			
	Chronic Indigestion	দীর্ঘস্থায়ী বদহজমের সমস্যা			
	Any other Chronic illness	আর অন্য কোনো দীর্ঘস্থায়ী ব্যাধি			
	Prone to cold infection	সর্দি-কাশীর ধাত			
	Asthma	হাঁপানি			
	Headache	মাথা ব্যাথা			
	Migraine	মাইগ্রেন			
	Prostrate problem	প্রস্ট্রেট সমস্যা			
	Jaundice	ন্যাবা/পাণ্ডুরোগ			
	Dengue	ডেঙ্গু			
	Chicken pox	চিকেন-পক্স			
	Menstrual problem	মাসিকের সমস্যা			
	Others (mention)	অন্য কোনো সমস্যা (থাকলে জানান)			
	Mumps	মাম্পস্			
	Toncil	টনসিল			

	Answer these questions in Yes/No	***নিচের প্রশ্নগুলির হ্যা বা না তে উত্তর দাও***	***Yes (Y)***	***No (N)***
2	Do you have any dominant or recurrent pain?	আপনার কোনো দীর্ঘস্থায়ী বা রোজকার বেথা আছে?		
	If yes (Y), please specify region.	হ্যা হলে কোথায়?		
3	Have you had surgery before?	আপনার কোনো অস্ত্রপচার হয়েছে?		
	If yes (Y), please specify	হ্যা হলে কোথায়?		
4	For females only: Any Hormone replacement (HR) therapy done?	মহিলাদের জন্য: আপনার কোনো হরমোনের চিকিত্সা হিয়েছে?		
5	Are you a consumer of alcohol?	আপনি কি মদ্যপান করেন?		
	if yes (Y) then, daily/weekly/monthly/occasional? is it more than 2 pegsor less?	হ্যা হলে, কিরোজ/সপ্তাহে/মাসে/মাঝে মধ্যে? ২ পেগের বেশি না কম?		
6	Do you smoke?	আপনি কি ধুমপান করেন?		
	if yes (Y), then what (cigarette/ biri/ cigar) ?	হ্যা হলে (সিগারেট/বিড়ি/সিগার)?		
	Is it more than 10 nos. daily or less?	দিনে ১০টার বেশি না কম?		
7	Do you take drugs?	আপনি কি মাদক সেবন করেন?		
	if yes (Y), then daily/weekly/monthly/occasional ?	হ্যা হলে, রোজ/সপ্তাহে/মাসে/মাঝে মধ্যে?		
8	Do you take Jardapaan?	আপনি কি জর্দা পান খান?		
9	Do you take Jarda masala / pan parag/ chutki /gutkha?	আপনি কি জর্দা মশলা/পান পরাগ/চুটকি/গুটখা খান?		
	If yes (Y), please tell what?, How much?	হ্যা হলে কোনটি? কতটা?		
10	Are you left handed?	আপনি কি বাম হাতে কাজ করেন?		
	Measured Physical parameters -	**শারীরিকপরিমাপ**		
	Weight in kg.	ওজন কিলোগ্রামে		
	Height in meters	উচ্চতা মিটারে		
	BMI (Body mass Index) in metric	BMI সংখ্যায়		
	Waist measure in cm.	কোমর সেন্টিমিটারে		
	Hip measure in cm.	নিতম্বসেন্টিমিটারে		
	WHR (Waist to Hip Ratio)	কোমর ও নিতম্বের অনুপাত		

A.3 Daily Questionnaire set

Strong or moderate AC effect: Amitriptyline (Elavil), Atropine, Benztropine (Cogentin), Chlorpheniramine (Actifed, Allergy & Congestion Relief, Chlor-Trimeton, Codeprex, Efidac-24 Chlorpheniramine, etc), Chlorpromazine (Thorazine), Clomipramine (Anafranil), Clozapine (Clozaril), Cyclobenzaprine (Amrix, Fexmid, Flexeril), Cyproheptadine (Periactin), Desipramine (Norpramin), Dexchlorpheniramine, Dicyclomine (Bentyl), Diphenhydramine (Advil PM, Aleve PM, Bayer PM, Benadryl, Excedrin PM, Nytol, Simply Sleep, Sominex, Tylenol PM, Unisom, etc.), Doxepin (Adapin, Silenor, Sinequan), Fesoterodine (Toviaz), Hydroxyzine (Atarax, Vistaril), Hyoscyamine (Anaspaz, Levbid, Levsin, Levsinex, NuLev), Imipramine (Tofranil), Meclizine (Antivert, Bonine), Nortriptyline (Pamelor), Orphenadrine (Norflex), Oxybutynin (Ditropan, Oxytrol), Paroxetine (Brisdelle, Paxil), Perphenazine (Trilafon), Prochlorperazine (Compazine), Promethazine (Phenergan), Protriptyline (Vivactil), Pseudoephedrine Cl/Triprolidine, HCl (Aprodine), Scopolamine (TransdermScop), Thioridazine (Mellaril), Tolterodine (Detrol), Trifluoperazine (Stelazine), Trimipramine (Surmontil)

Lesser AC effect: Alprazolam (Xanax), Fluphenazine (Prolixin), Amantadine (Symmetrel), Baclofen Carisoprodol (Soma), Cetirizine (Zyrtec), Cimetidine (Tagamet), Clorazepate (Tranxene), Codeine, Colchicine, Digoxin (Lanoxicaps, Lanoxin), Diphenoxylate (Lomotil), Furosemide (Lasix), Hydrochlorothiazide (Esidrix, Dyazide, HydroDIURIL, Maxzide & literally scores of other medications for high blood pressure), Loperamide (Imodium), Loratadine (Alavert, Claritin), Maprotiline, Nifedipine (Adalat, Procardia), Ranitidine (Zantac), Thiothixene (Navane), Tizanidine (Zanaflex)

Table 1: Daily Questionnaire set 1

Questions	DDMMYYYY	DDMMYYYY
Have you taken any drug from the list provided in table below within 24 hours?		
Please specify medicines or treatment that you are undergoing as on date, if different from that stated in comprehensive questionnaire.		
Has subject taken any anticholinergic drug knowingly or unknowingly? (If NO, then proceed for daily questionnaire 2, followed by acquisition of data, else exclude the subject for today)		

Table 2: Daily Questionnaire set 2

Questions	DDMMYYYY	DDMMYYYY
Are you stressed? Provide rating in scale of 0-5		
If yes, is the stress a) work related, b) emotional or c) any other? Please specify		
Have you taken any psychotic or antidepressant drug within 1 week?		
If yes, then provide the reason, list of medicine and time of taking medicine.		
Do you have any specific pain? Provide rating in scale of 0-5		
If yes, is the pain short term or long term?		
Is the pain in a) left half or b) right half of body or c) overall or d) not localized? If a) or b), please specify location.		
Do you have fever? Provide rating in scale of 0-5		
Do you have cough or cold? Provide rating in scale of 0-5		
Do you have any headache? Provide rating in scale of 0-5		
If yes, is the headache in a) left half or b) right half or c) overall or d) not localized? If a) or b), please specify location.		
Have you taken any cigarettes or any other intoxicating substance or caffeine containing drink within 72 hours?		

Table 3:Daily Questionnaire set 3

Question	DDMMYYYY		DDMMYYYY	
Clothing of the subject (Light, Moderate, Heavy)				
Body temperature				
Bias voltages measured for 4 electrodes with respect to 1 reference electrode.	**Tag**	**Volt**	**Tag**	**Volt**
	Rish_IL		Rish_IL	
	Rish_ML		Rish_ML	
	Rish_IR		Rish_IR	
	Rish_MR		Rish_MR	

Appendix B

B.1 Effect of additional controlled noise in Hypertensive and Normotensive Classification

In order to observe the effect of noise in the Hypertensive and Normotensive classification results, a re-classification has been done for all 3 proposed models by adding 10dB SNR random distributed Gaussian noise to the original acquired DDP signals. Using the same methodology and same numbers of 10s signals and parameters detailed in Chapter 4, a detailed study is presented step by step hereafter. Here, the two classification methods, 10FCV and LOOCV are used. This classification is henceforth referred as the 10dB noise added hypertensive-normotensive classification.

B.1.1 Selection of Attribute Evaluator

Since using InfoGainAttributeEval provided uniformly high accuracies of the classification results for all models, hence this ranker method is used in this case also. It has been observed that the accuracies are 94.97%, 94.28% and 100% using the numbers of attributes 5, 5 and 2 for the ACF, KF_{LPC} and KF_{TS} models respectively. Hence, it can be said that the overall accuracies have increased in ACF and KF_{LPC} but have decreased in case of KF_{TS}. However, the numbers of attributes have increased in this classification for all models. The rank wise selected attributes and accuracy for the 10dB noise added hypertensive-normotensive classification are stated in Table 1.

B.1.2 Selected attributes and their analysis

Applying the Supervised Attribute Selection filter in WEKA, 31 (out of total 180 parameters) attributes from ACF model, 25 (out of total 170 parameters) attributes from KF_{LPC} model and 28 (out of total 170 parameters) attributes from KF_{TS} model are selected. Using the InfoGainAttributeEval attribute

Table 1: Rank wise selected attributes and accuracy for 10dB noise added hypertensive-normotensive classification

Ranker	Number of attributes selected					
	ACF		LPC		TS	
	Number of attributes	Accuracy (%)	Number of attributes	Accuracy (%)	Number of attributes	Accuracy (%)
InfoGain AttributeEval	5	94.97	5	94.28	2	100

Table 2: List of the selected attributes using ACF, KF_{LPC}, KF_{TS} model parameters for 10dB noise added hypertensive-normotensive classification

Model	Sl No.	Selected attributes as per rank	Cross validation accuracy (%)	
ACF	1	Norkur $x_{kACF_{phasic}}$ of LH5	10FCV	LOOCV
	2	Norkur $x_{kACF_{phasic}}$ of RH1	94.97	94.92
	3	Norkur $x_{kACF_{phasic}}$ of RH3		
	4	Norkur $x_{kACF_{phasic}}$ of RH4		
	5	Norkur $x_{kACF_{phasic}}$ of RH5		
KF_{LPC}	1	q_{com} of LH1	94.28	94.39
	2	q_{com} of RH1		
	3	q_{com} of RH3		
	4	q_{com} of RH4		
	5	q_{com} of RH5		
KF_{TS}	1	q_{com} of LH1	100	99.94
	2	q_{com} of RH5		

evaluator on this subset of features, finally 5, 5 and 2 attributes respectively are selected for the 3 models as stated in Table 2. Therefore, it can be said that in this classification, Attribute Selection filter selected more attributes in ACF and KF_{TS} models and less in KF_{LPC} model as compared to the actual case (without added noise).

In order to appreciate the characteristics of all the 12 selected attributes for all of 3 models, their box plots are shown in Figures 1a, 1b, 1c, 1d and 1e for the ACF model, Figures 2a, 2b, 2c, 2d and 2e for the KF_{LPC} model and Figures 3a and 3b for the KF_{TS} model. The corresponding box plot parameters for the 3 models are listed in Tables 3, 4 and 5. The observations from these selected features and their box plots are listed hereafter.

1. As mentioned, total 12 attributes are selected for the 3 models. Of these, 3 attributes are from LH, while 9 are from RH. These are the

Table 3: Box plot parameters for 10dB noise added hypertensive normotensive classification using ACF model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
Norkur $x_{kACF_{phasic}}$ of LH5	Hypertensive	-0.13	-0.11	-0.04	0.19	0.65
	Normotensive	-1.79	-1.05	0.20	0.66	1.95
Norkur $x_{kACF_{phasic}}$ of RH1	Hypertensive	-0.13	-0.11	-0.04	0.2	0.67
	Normotensive	-1.79	-1.5	0.20	0.66	1.60
Norkur $x_{kACF_{phasic}}$ of RH3	Hypertensive	-0.53	-0.12	-0.04	0.21	0.67
	Normotensive	-1.70	-1.05	0.20	0.66	1.60
Norkur $x_{kACF_{phasic}}$ of RH4	Hypertensive	-0.59	-0.12	-0.04	0.20	0.71
	Normotensive	-1.71	-1.05	0.20	0.66	1.99
Norkur $x_{kACF_{phasic}}$ of RH5	Hypertensive	-0.38	-0.12	-0.05	0.20	0.67
	Normotensive	-1.76	-1.05	0.20	0.66	1.70

Table 4: Box plot parameters for 10dB noise added hypertensive normotensive classification using KF_{LPC} model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
q_{com} of LH1	Hypertensive	-0.19	.38	0.82	0.93	1.23
	Normotensive	0.01	0.71	1.01	1.19	1.90
q_{com} of RH1	Hypertensive	0.06	0.36	0.54	0.72	0.98
	Normotensive	-0.30	0.59	0.98	1.19	1.91
q_{com} of RH3	Hypertensive	-0.28	0.18	0.34‘	0.51	0.91
	Normotensive	-0.38	0.50	0.77	1.08	1.80
q_{com} of RH4	Hypertensive	-0.26	0.25	0.48	0.70	0.91
	Normotensive	-0.33	0.55	0.90	1.15	1.89
q_{com} of RH5	Hypertensive	-0.14	0.30	0.57	0.74	1.00
	Normotensive	-0.33	0.61	1.01	1.24	1.92

Table 5: Box plot parameters for 10dB noise added hypertensive normotensive classification using KF_{TS} model parameters

Attributes	Posture	Lower Adjacent	Lower Quartile	Median	Upper Quartile	Upper Adjacent
q_{com} of LH1	Hypertensive	-1.35	-1.29	-1.29	-1.16	-1.08
	Normotensive	-1.45	-1.30	-1.21	-1.12	-1.04
q_{com} of RH5	Hypertensive	-1.35	-1.29	-1.29	-1.16	-1.08
	Normotensive	-1.45	-1.30	-1.21	-1.12	-1.04

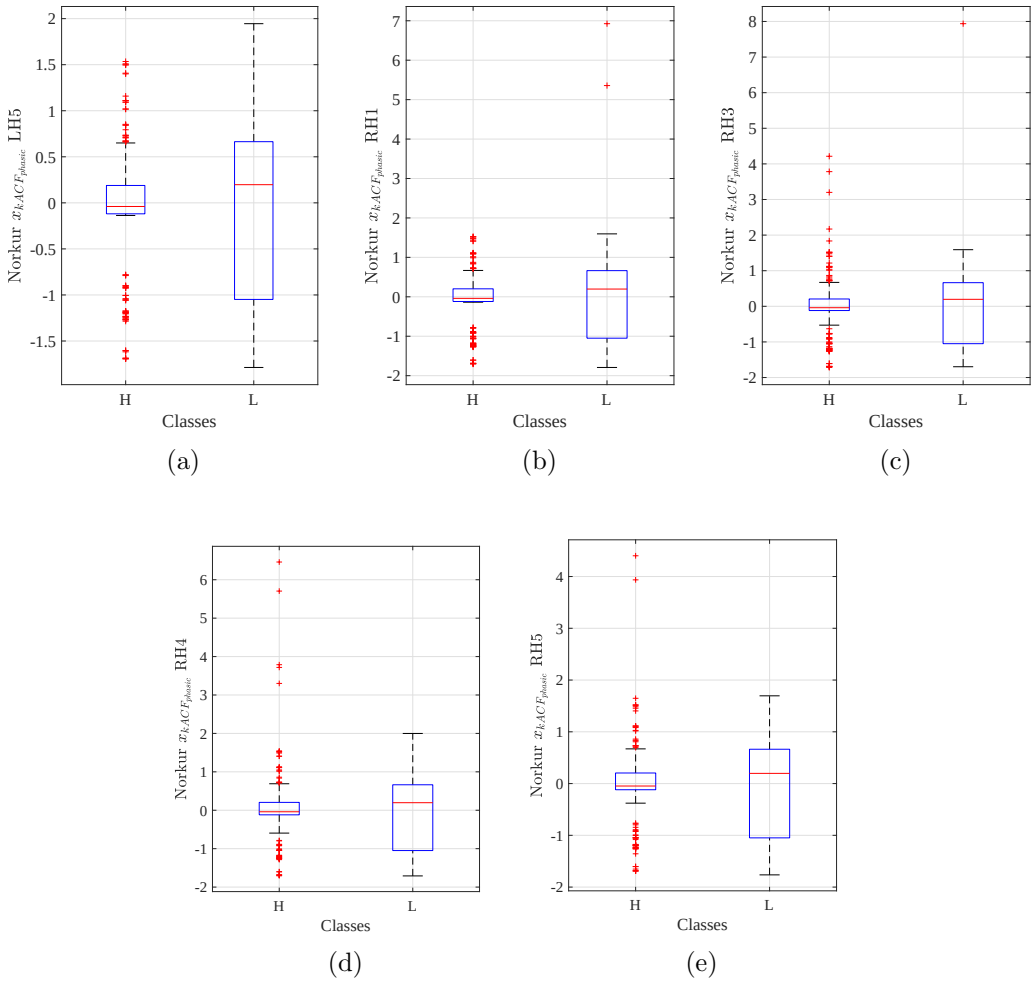


Figure 1: H: Hypertensive, N: Normotensive. Box plots of (a) Norkur $x_{kACF_{phasic}}$ of LH5 (b) Norkur $x_{kACF_{phasic}}$ of RH1(c) Norkur $x_{kACF_{phasic}}$ of RH3 (d) Norkur $x_{kACF_{phasic}}$ of RH4 (e) Norkur $x_{kACF_{phasic}}$ of RH5 for 10dB noise added hypertensive-normotensive classification.

normalized kurtosis (Norkur) of $x_{kACF_{phasic}}$ of 5 states (state 5) from LH and (states 1, 3, 4, 5) from RH for ACF model, q_{com} of 1 state from LH (state 1) and 4 states from RH (states 1, 3, 4, 5) for KF_{LPC} model and q_{com} of 1 state from LH (state 1) and 1 state from RH (state 5) for KF_{TS} model. In the previous results in Chapter 4, it has been found that the presence or absence of hypertension is most reflected in the LH parameters but the 10dB noise added classification results show most reflection in RH parameters. Hence, two results are different.

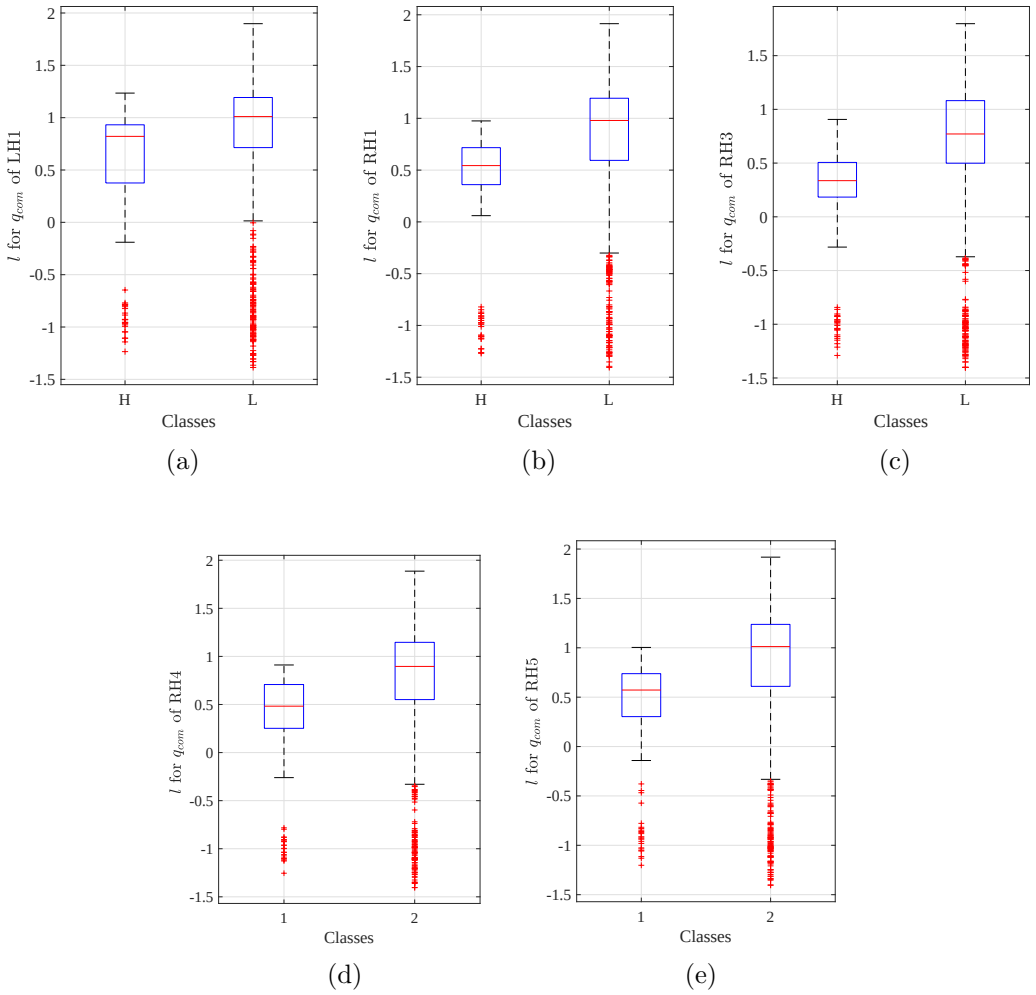


Figure 2: H: Hypertensive, N: Normotensive. Box plots of (a) l of q_{com} of LH1 (b) l of q_{com} of RH1 (c) l of q_{com} of RH3 (d) l of q_{com} of RH4 (e) l of q_{com} of RH5 for KF_{LPC} model for 10dB noise added hypertensive-normotensive classification.

2. In the ACF model, all 5 attributes are selected from only the phasic components of LH and RH signals. In the KF_{LPC} (KF_{TS}) model, 5 (2) attributes are selected from q_{com} of LH and RH. Hence, it is clear that in both classification studies (previous results and present 10dB noise added results), q_{com} is the only feature in the KF models for separation of hypertensive and normotensive classes.

3. The box plots of all 5 normalized kurtosis features selected in the ACF

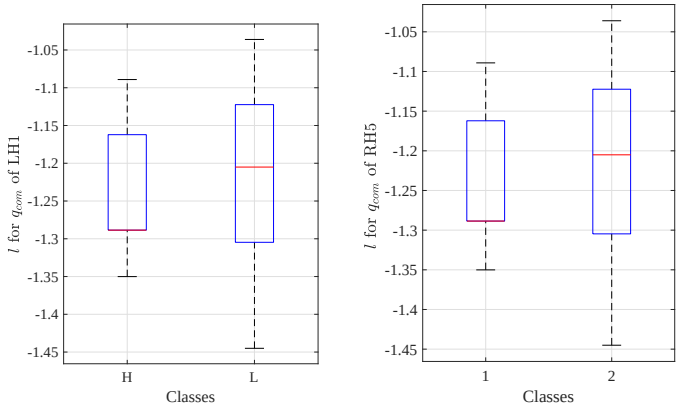


Figure 3: H: Hypertensive, N: Normotensive. Box plots of (a) l of q_{com} of LH1 (b) l of q_{com} of RH5 for KF_{TS} model for 10dB noise added hypertensive-normotensive classification.

model show that the median values are negative in hypertensive subjects and are all positive in normotensive subjects. The median values are same in all 5 attributes for both classes. Further, the lower adjacent values and upper adjacent values are quite large in normotensive subjects as compared to the hypertensive subjects.

4. From the box plots of all 5 q_{com} features selected in the KF_{LPC} model, it is observed that the median values vary from 0.34 to 0.82 in normotensive class and it varies from 0.77 to 1.01 in hypertensive class. The lower adjacent values are quite high in hypertensive subjects except q_{com} of LH1 attribute, whereas upper adjacent values are high in normotensive class. This may be due to the larger numbers of subjects in the normotensive class.

5. In case of the KF_{TS} model, only 2 features of q_{com} are selected. The nature of these two attributes in terms of the median and the lower-upper adjacent values are also almost same.

6. A comparison of these results with the previous results stated in Chapter 4 shows that due to the addition of 10dB SNR noise, the q_{com} values decrease to quite low values both in KF_{LPC} model (vary from 1.5 to 2) as well as in KF_{TS} model (vary from -1.45 to -1.05).

B.1.3 Confusion Matrices and Metrics

The confusion matrices for 10FCV and LOOCV classifications using the selected attributes are calculated and stated in Table 6. The overall results of classification using 3 model parameters are separately stated in Tables 7, 8 and 9 respectively.

Table 6: Confusion matrix (in %) of RF classifier using ACF, KF_{LPC}, KF_{TS} model parameters for 10dB noise added hypertensive-normotensive classification

CV Method	Actual	ACF		KF_{LPC}		KF_{TS}	
		HT	NT	HT	NT	HT	NT
10FCV	HT	8.01	4.17	8.07	4.11	12.18	0
	NT	0.85	86.97	1.60	86.22	0	87.82
LOOCV	HT	7.91	4.27	7.96	4.22	12.18	0
	NT	0.80	87.02	1.39	86.43	0.06	87.76
DS2 data set: N=1872, HT: Hypertensive, NT: Normotensive							

Table 7: Results for 10dB noise added Hypertensive and Normotensive classification using ACF model parameters

CV Method	Class	Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV	HT	0.904	0.658	0.990	0.761	0.902	0.790
	NT	0.954	0.990	0.658	0.972	0.902	0.978
	WA	0.948	0.950	0.698	0.946	0.902	0.955
LOOCV	HT	0.908	0.649	0.991	0.757	0.914	0.981
	NT	0.953	0.991	0.649	0.972	0.914	0.981
	WA	0.948	0.949	0.691	0.946	0.914	0.959

From the confusion matrices, it is observed that the matrix structures are similar for the 3 models using the 2 different cross-validation techniques. Hence, it is expected that for all the models, the WA of the matrices elements (precision, Recall, Specificity, F1-Score, ROC, PRC) will also be similar using both techniques. This type of identical results were also observed when classification was done using the acquired signal parameters in Chapter 4.

1. Except the WA of specificity, all the WA values of the confusion matrix elements are greater than 90% in both techniques for ACF model stated in Table 7. From Table 4.9 of Chapter 4, it was observed that the WA values of these matrices elements were greater than 98%. In both

Table 8: Results for 10dB noise added Hypertensive and Normotensive classification using KF_{LPC} model parameters

CV Method	Class	Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV	HT	0.834	0.662	0.982	0.738	0.964	0.818
	NT	0.954	0.982	0.662	0.968	0.964	0.994
	WA	0.940	0.943	0.701	0.940	0.964	0.972
LOOCV	HT	0.851	0.654	0.984	0.739	0.971	0.841
	NT	0.984	0.969	0.047	0.969	0.971	0.995
	WA	0.941	0.944	0.694	0.941	0.971	0.976

Table 9: Results for 10dB noise added Hypertensive and Normotensive classification using KF_{TS} model parameters

CV Method	Class	Precision	Recall	Specificity	F1-Score	ROC Area	PRC Area
10FCV	HT	1.00	1.00	1.00	1.00	1.00	1.00
	NT	1.00	1.00	1.00	1.00	1.00	1.00
	WA	1.00	1.00	1.00	1.00	1.00	1.00
LOOCV	HT	0.996	1.00	0.999	0.998	1.00	1.00
	NT	1.00	0.999	1.00	1.00	1.00	1.00
	WA	0.999	0.999	1.00	0.999	1.00	1.00

classifications (with and without added 10dB noise), the specificity values are low.

2. In Table 8 for the KF_{LPC} model, the specificity value is almost 69% while other parameters are greater than 94% using 10FCV and LOOCV. As in Table 4.10, the low specificity value is also maintained here. From Table 4.10, it was observed that using LOOCV, the values of matrix elements are 100% but these values decrease by almost 6% on adding the 10dB noise.

3. From Table 9 for the KF_{TS} model, it is observed that the values of all elements are almost 100%. Using the KF_{TS} model parameters, 2 attributes were finally selected and from Figure 3a and 3b, it is observed that both classes are distinctly separable. Hence, the 100% results, stated in Table 1, are to be expected.

4. Comparing the two KF models, it can be said that on adding noise, all the values in the confusion matrix are less in case of the KF_{LPC} model although the number of selected attributes are more. Caomparing these

results with the original results in Table 4.10 and Table 4.11, it is found that the results were opposite with the confusion matrix values being 100% in case of KF_{LPC} model.

Thus in this hypertensive-normotensive classification study, where 10dB noise is added to observe the changes of overall results using the parameters of the 3 proposed model parameters, it is found that in general more numbers of attributes are selected in this case. In case of the ACF model, a different set of attributes is selected, specifically normalized kurtosis in place of normalized variance. However, the same attributes are selected in both KF models, albeit for different states. The overall accuracies are marginally low (high) in ACF, KF_{LPC} (in KF_{TS}) when compared to the results of Chapter 4. Therefore, it can be said that of the 3 models, the KF_{TS} model is the most suited for this classification although the other models are also quite comparable and robust.

Bibliography

[1] A. Sarkar, "Differential Dermal Potential: Acquisition, Validation and Applications," Ph.D. dissertation, IEE, JU, Kolkata, 2023.

[2] L. Cromwell, F. J. Weibell, E. A. Pfeiffer, and L. B. Usselman, *Biomedical Instrumentation and Measurements*, 2nd ed. New Jersey, U.S.A: Prentice-Hall Inc, 1980.

[3] N. V. Thakor, *Biopotentials and Electrophysiology Measurement*, 2nd ed. Boca Raton: CRC Press, 31 January, 2017.

[4] A. Akinin, A. Paul, J. Wang, A. Buccino, and G. Cauwenberghs, *Biopotential Measurements and Electrodes*, 3rd ed., B. He, Ed. Switzerland: Springer, 2020.

[5] K. U. Menon, D. Jose, and M. V. Ramesh, "SleepGaze: A Wireless System for Monitoring and Detection of Sleep Disorders," in *Proceedings of the International Conference on Bioinformatics & Computational Biology (BIOCOMP)*. The Steering Committee of The World Congress in Computer Science, Computer, 2011, p. 1.

[6] O. Razumnikova and M. Bakaev, "Ontology of Frequency-Spatial Organization of Brain Activity Reflecting the Cognitive Reserves," in *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. IEEE, 2019, pp. 0950–0954.

[7] S. K. Berkaya, A. K. Uysal, E. S. Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, "A survey on ECG analysis," *Biomedical Signal Processing and Control*, vol. 43, pp. 216–235, May, 2018.

[8] Z. Sankari and H. Adeli, "HeartSaver: A mobile cardiac monitoring system for auto-detection of atrial fibrillation, myocardial infarction, and atrio-ventricular block," *Computers in Biology and Medicine*, vol. 41, no. 4, pp. 211–220, 2011.

[9] C. R. Gloschat, "Development of High Resolution Tools for Investigating Cardiac Arrhythmia Dynamics," Ph.D. dissertation, Washington University in St. Louis, 2017.

[10] M. Teplan *et al.*, "Fundamentals of EEG measurement," *Measurement Science Review*, vol. 2, no. 2, pp. 1–11, 2002.

[11] P. Konrad, "The ACF of EMG," *A Practical Introduction to Kinesiological Electromyography*, vol. 1, no. 2005, pp. 30–5, 2005.

[12] C.-T. Lin, J.-T. King, P. Bharadwaj, C.-H. Chen, A. Gupta, W. Ding, and M. Prasad, "EOG-Based Eye Movement Classification and Application on HCI Baseball Game," *IEEE Access*, vol. 7, pp. 96 166–96 176, 2019.

[13] L. Y. Deng, C.-L. Hsu, T.-C. Lin, J.-S. Tuan, and S.-M. Chang, "EOG-Based Human–Computer Interface System Development," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3337–3343, 2010.

[14] I. Perlman, "The Electroretinogram: ERG," *Webvision: The Organization of the Retina and Visual System [Internet]*, 2007.

[15] W. Boucsein, *Electrodermal Activity*, 2nd ed. Springer Science & Business Media, Germany, 2012.

[16] S. Cecchi, A. Piersanti, A. Poli, and S. Spinsante, "Physical Stimuli and Emotions: EDA Features Analysis from a Wrist-Worn Measurement Sensor," in *2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. IEEE, 2020, pp. 1–6.

[17] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, "A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments," *Psychophysiology*, vol. 49, no. 1, pp. 1017–1034, 2013.

[18] M. Lajante, O. Droulers, T. Dondaine, and D. Amarantini, "Opening the "black box" of electrodermal activity in consumer neuroscience research." *Journal of Neuroscience, Psychology, and Economics*, vol. 5, no. 4, p. 238, 2012.

[19] B. U. Cowley and J. Torniainen, "A short review and primer on electrodermal activity in human computer interaction applications," *arXiv preprint arXiv:1608.06986*, 2016.

[20] J. Shukla, M. Barreda-Angeles, J. Oliver, G. C. Nandi, and D. Puig, "Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity," *IEEE Transactions on Affective Computing*, vol. 12, no. 4, pp. 857–869, 2019.

[21] A. Bhattacharya, D. Basu, B. Goswami, and R. Ghosh, "Measuring human biopotentials to ascertain parameters for health," in *2011 International Conference on Communication and Industrial Application*. IEEE, 2011, pp. 1–4.

[22] ——, "A reliable signal conditioning circuit to acquire human biopotentials," *International Journal on Smart Sensing and Intelligent Systems*, vol. 6, no. 3, pp. 1116–1132, 2013.

[23] A. Bhattacharya, "Study and characterization of non-organ specific biopotential signals acquired from the human system," Ph.D. dissertation, IEE, JU, Kolkata, India, 2019.

[24] A. Sarkar, B. Goswami, and R. Ghosh, "Validating Differential Dermal Potentials for use in Real-time Human Condition Monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–9, 2022.

[25] A. Sarkar, A. Bhattacharya, R. Ghosh, and B. Goswami, "A Comparative Study of Biopotentials Acquired from Left and Right Hands of Human Subjects," in *Advanced Computational and Communication Paradigms*. Springer, 2018, pp. 110–117.

[26] D. Jaiswal, M. Moulick, D. Chatterjee, R. Ranjan, R. K. Ramakrishnan, A. Pal, and R. Ghosh, "Assessment of Cognitive Load from Biopotentials Measured using Wearable Endosomatic Device," in *Proceedings of the 6th ACM Workshop on Wearable Systems and Applications*, 2020, pp. 13–18.

[27] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. John Wiley & Sons, New Jersey, 2015.

[28] S. R. Devasahayam, *Signals and Systems in Biomedical Engineering: Signal Processing and Physiological Systems Modeling*, 2nd ed. New York, Heidelberg Dordrecht London: Springer Science & Business Media, 2012.

[29] R. W. J. Stern Robert M and Q. K. S., *Psychophysiological recording*, 2nd ed. Oxford University Press, USA, 2001.

[30] M. Tlili, A. Maalej, M. B. Romdhane, F. Rivet, D. Dallet, and C. Rebai, "Mathematical modeling of clean and noisy ECG signals in a level-crossing sampling context," in *2016 International Symposium on Signal, Image, Video and Communications (ISIVC)*. IEEE, 2016, pp. 359–363.

[31] M. Sidorova and N. Serzhantova, "Method for Forming Mathematical Models of Measured Electrophysiological Signals," in *Moscow Workshop on Electronic and Networking Technologies (MWENT)*. IEEE, 2020, pp. 1–5.

[32] H. Limaye and V. Deshmukh, "ECG Noise Sources and Various Noise Removal Techniques: A Survey," *International Journal of Application or Innovation in Engineering & Management*, vol. 5, no. 2, pp. 86–92, 2016.

[33] K. Takahashi, "Remarks on SVM-based emotion recognition from multi-modal bio-potential signals," in *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE, 2004, pp. 95–100.

[34] C. K. Ho and M. Sasaki, "Brain-wave bio potentials based mobile robot control: wavelet-neural network pattern recognition approach," in *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*, vol. 1. IEEE, 2001, pp. 322–328.

[35] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clinical Neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007.

[36] E. García-Cossio, M. Severens, B. Nienhuis, J. Duysens, P. Desain, N. Keijsers, and J. Farquhar, "Decoding Sensorimotor Rhythms During Robotic-Assisted Treadmill Walking for Brain Computer Interface (BCI) Applications," *PloS One*, vol. 10, no. 12, p. e0137910, 2015.

[37] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, T. M. Vaughan *et al.*, "Brain-Computer Interface Technology: A Review of the First International Meeting," *IEEE Transactions On Rehabilitation Engineering*, vol. 8, no. 2, pp. 164–173, 2000.

[38] R. H. Shumway, D. S. Stoffer, and D. S. Stoffer, *Time Series Analysis and Its Applications*, ser. 1431-875X. USA: Springer, 2000, vol. 3.

[39] I. N. Soyiri and D. D. Reidpath, "An overview of health forecasting," *Environmental Health and Preventive Medicine*, vol. 18, no. 1, pp. 1–9, 2013.

[40] M. Villani, A. Earnest, N. Nanayakkara, K. Smith, B. De Courten, and S. Zoungas, "Time series modelling to forecast prehospital EMS demand for diabetic emergencies," *BMC Health Services Research*, vol. 17, no. 1, pp. 1–9, 2017.

[41] A. Amei, W. Fu, and C.-H. Ho, "Time Series Analysis for Predicting the Occurrences of Large Scale Earthquakes," *International Journal of Applied Science and Technology*, vol. 2, no. 7, 2012.

[42] R. S. J. Sparks, "Forecasting Volcanic Eruptions," *Earth and Planetary Science Letters*, vol. 210, no. 1-2, pp. 1–15, 2003.

[43] C. Bretó, D. He, E. L. Ionides, and A. A. King, "Time Series Analysis via Mechanistic Models," *The Annals of Applied Statistics*, pp. 319–348, 2009.

[44] A. Hastings, K. C. Abbott, K. Cuddington, T. B. Francis, Y.-C. Lai, A. Morozov, S. Petrovskii, and M. L. Zeeman, "Effects of stochasticity on the length and behaviour of ecological transients," *Journal of the Royal Society Interface*, vol. 18, no. 180, p. 20210257, 2021.

[45] M. Fujiwara and T. Takada, "Environmental stochasticity," *Encyclopedia of Life Sciences*, pp. 1–8, 2017.

[46] H. F. Posada-Quintero and K. H. Chon, "Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review," *Sensors*, vol. 20, no. 2, p. 479, 2020.

[47] C. Chatfield and H. Xing, *The Analysis of Time Series: An Introduction with R*, 7th ed. Boca Raton: CRC Press, 2019.

[48] G. R. Cooper and C. D. McGillem, *Probabilistic Methods of Signal and System Analysis*, 3rd ed. Oxford University Press, New Delhi, 1986.

[49] R. Shiavi, *Introduction to Applied Statistical Signal Analysis: Guide to Biomedical and Electrical Engineering Applications*, 3rd ed. Elsevier, USA, 2010.

[50] T. F. Quatieri, *Discrete-Time Speech Signal Processing: Principles and Practice*. Pearson Education India, 2006.

[51] C. Chatfield, *The Analysis of Time Series: An Introduction*, 5th ed. New York, NY.: Chapman and Hall/CRC, 1996.

[52] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Prentice Hall PTR, New Jersey, 1992.

[53] F. Huang, T. Qin, L. Wang, H. Wan, and J. Ren, "An ECG Signal Prediction Method Based on ARIMA Model and DWT," in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, vol. 1. IEEE, 2019, pp. 1298–1304.

[54] H. D. Critchley, "Electrodermal Responses: What Happens in the Brain," *The Neuroscientist*, vol. 8, no. 2, pp. 132–142, 2002.

[55] L. Ljung, "System Identification," in *Signal Analysis and Prediction*. Boston, MA: Birkhäuser Boston, 1998, pp. 163–173.

[56] L. Zhu, Y. Wang, and Q. Fan, "MODWT-ARMA model for time series prediction," *Applied Mathematical Modelling*, vol. 38, no. 5-6, pp. 1859–1865, 2014.

[57] M. S. Aleksić and M. M. Ristić, "A geometric minification integer-valued autoregressive model," *Applied Mathematical Modelling*, vol. 90, pp. 265–280, 2021.

[58] J. Garza-Ulloa, H. Yu, and T. Sarkodie-Gyan, "A mathematical model for the validation of the ground reaction force sensor in human gait analysis," *Measurement*, vol. 45, no. 4, pp. 755–762, 2012.

[59] S. Lu, K. H. Ju, and K. H. Chon, "A new algorithm for linear and nonlinear ARMA model parameter estimation using affine geometry [and application to blood flow/pressure data]," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 10, pp. 1116–1124, 2001.

[60] D. J. Christini, F. Bennett, K. Lutchen, H. Ahmed, J. Hausdorff, and N. Oriol, "Application of linear and nonlinear time series modeling to heart rate dynamics analysis," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 4, pp. 411–415, 1995.

[61] T. Bermudez, D. Lowe, and A.-M. Arlaud-Lamborelle, "EEG/ECG information fusion for epileptic event detection," in *2009 16th International Conference on Digital Signal Processing*. IEEE, 2009, pp. 1–8.

[62] B. V. Kini and C. C. Sekhar, "Large margin mixture of AR models for time series classification," *Applied Soft Computing*, vol. 13, no. 1, pp. 361–371, 2013.

[63] S. Khalighi, T. Sousa, G. Pires, and U. Nunes, "Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnographic channels," *Expert Systems with Applications*, vol. 40, no. 17, pp. 7046–7059, 2013.

[64] S. Pourmohammadi and A. Maleki, "Stress detection using ECG and EMG signals: A comprehensive study," *Computer Methods and Programs in Biomedicine*, vol. 193, p. 105482, 2020.

[65] J. Yu, S. Park, S.-H. Kwon, C. M. B. Ho, C.-S. Pyo, and H. Lee, "AI-Based Stroke Disease Prediction System Using Real-Time Electromyography Signals," *Applied Sciences*, vol. 10, no. 19, p. 6791, 2020.

[66] M. Arnold, X. Milner, H. Witte, R. Bauer, and C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 5, pp. 553–562, 1998.

[67] L. Lu, J. Mao, W. Wang, G. Ding, and Z. Zhang, "A Study of Personal Recognition Method Based on EMG Signal," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 4, pp. 681–691, 2020.

[68] G. Benrhmach, K. Namir, A. Namir, and J. Bouyaghroumni, "Nonlinear Autoregressive Neural Network and Extended Kalman Filters for Prediction of Financial Time Series," *Journal of Applied Mathematics*, vol. 2020, pp. 1–6, 2020.

[69] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, 1st ed. United Kingdom: Cambridge University Press, 1990.

[70] V. M. Moreno and A. Pigazo, *Kalman Filter: Recent Advances and Applications*. Vienna, Austria.: InTech, April 2009.

[71] O. Das, B. Goswami, and R. Ghosh, "Application of the Tuned Kalman Filter in Speech Enhancement," in *2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI)*. IEEE, 2016, pp. 62–66.

[72] S. So, A. E. George, R. Ghosh, and K. K. Paliwal, "A Non-Iterative Kalman Filtering Algorithm with Dynamic Gain Adjustment for Single-Channel Speech Enhancement," *International Journal of Signal Processing Systems*, vol. 4, no. 4, pp. 263–268, 2016.

[73] A. E. George, S. So, R. Ghosh, and K. K. Paliwal, "Robustness metric-based tuning of the augmented Kalman filter for the enhancement of speech corrupted with coloured noise," *Speech Communication*, vol. 105, pp. 62–76, 2018.

[74] N. Dionelis and M. Brookes, "Phase-Aware Single-Channel Speech Enhancement With Modulation-Domain Kalman Filtering," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 5, pp. 937–950, 2018.

[75] S. So, A. E. George, R. Ghosh, and K. K. Paliwal, "Kalman Filter with Sensitivity Tuning for Improved Noise Reduction in Speech," *Circuits, Systems, and Signal Processing*, vol. 36, no. 4, pp. 1476–1492, 2017.

[76] N. Dionelis and M. Brookes, "Modulation-domain speech enhancement using a Kalman filter with a Bayesian update of speech and noise in the log-spectral domain," in *2017 Hands-free Speech Communications and Microphone Arrays (HSCMA)*. IEEE, 2017, pp. 111–115.

[77] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A Nonlinear Bayesian Filtering Framework for ECG Denoising," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2172–2185, 2007.

[78] R. Sameni, M. B. Shamsollahi, and C. Jutten, "Model-based Bayesian filtering of cardiac contaminants from biomedical recordings," *Physiological Measurement*, vol. 29, no. 5, p. 595, 2008.

[79] M. Rohál'ová, P. Sykacek, M. Koskaand, and G. Dorffner, "Detection of the EEG Artifacts by the Means of the (Extended) Kalman Filter," *Measurement Science Review*, vol. 1, no. 1, pp. 59–62, 2001.

[80] S.-H. Kim, H.-J. Yang, N. A. T. Nguyen, R. M. Mehmood, and S.-W. Lee, "Parameter Estimation using Unscented Kalman Filter on the

Gray-Box Model for Dynamic EEG System Modeling," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6175–6185, 2020.

[81] V. P. Oikonomou, A. T. Tzallas, and D. I. Fotiadis, "A Kalman filter based methodology for EEG spike enhancement," *Computer Methods and Programs in Biomedicine*, vol. 85, no. 2, pp. 101–108, 2007.

[82] A. Lawrance and P. Lewis, "Modelling and Residual Analysis of Nonlinear Autoregressive Time Series in Exponential Variables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 47, no. 2, pp. 165–183, 1985.

[83] C. F. Baum, "Residual Diagnostics for Cross-Section Time Series Regression Models," *The Stata Journal*, vol. 1, no. 1, pp. 101–104, 2001.

[84] S. Eyheramendy, F. Elorrieta, and W. Palma, "An irregular discrete time series model to identify residuals with autocorrelation in astronomical light curves," *Monthly Notices of the Royal Astronomical Society*, vol. 481, no. 4, pp. 4311–4322, 2018.

[85] M. E. Mann and J. M. Lees, "Robust estimation of background noise and signal detection in climatic time series," *Climatic Change*, vol. 33, no. 3, pp. 409–445, 1996.

[86] R. D. Prokasy William, *Electrodermal activity in psychological research.* New York and London: Academic Press, 1973.

[87] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80–91, 2010.

[88] H. F. Posada-Quintero and K. H. Chon, "Phasic Component of Electrodermal Activity is more Correlated to Brain Activity than Tonic Component," in *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI).* IEEE, 2019, pp. 1–4.

[89] M. J. Christie and P. H. Venables, "Characteristics of palmar skin potential and conductance in relaxed human subjects," *Psychophysiology*, vol. 8, no. 4, pp. 525–532, 1971.

[90] D. Lykken, R. Rose, B. Luther, and M. Maley, "Correcting psychophysiological measures for individual differences in range." *Psychological Bulletin*, vol. 66, no. 6, p. 481, 1966.

[91] S. Patidar, D. P. Jenkins, and S. A. Simpson, "Stochastic modelling techniques for generating synthetic energy demand profiles," *International Journal of Energy and Statistics*, vol. 4, no. 03, p. 1650014, 2016.

[92] P. J. Bota, C. Wang, A. L. Fred, and H. P. Da Silva, "A review, current challenges, and Future Possibilities on Emotion Recognition Using Machine Learning and PhysioSignals," *IEEE Access*, vol. 7, pp. 140 990–141 020, 2019.

[93] A. Greco, G. Valenza, J. Lazaro, J. M. Garzon-Rey, J. Aguilo, C. De-la Camara, R. Bailon, and E. P. Scilingo, "Acute Stress State Classification Based on Electrodermal Activity Modeling," *IEEE Transactions on Affective Computing*, 2021.

[94] C. Tronstad, O. M. Staal, S. Sælid, and Ø. G. Martinsen, "Model-based filtering for artifact and noise suppression with state estimation for electrodermal activity measurements in real time," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 2750–2753.

[95] A. Greco, G. Valenza, E. P. Scilingo, A. Greco, G. Valenza, and E. P. Scilingo, "Modeling for the Analysis of the EDA," *Advances in Electrodermal Activity Processing with Applications for Mental Health: From Heuristic Methods to Convex Optimization*, pp. 19–33, 2016.

[96] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating Stress from Cognitive Load using a Wearable EDA Device," *IEEE Transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 410–417, 2009.

[97] S. Fisher, "Body image and asymmetry of body reactivity." *The Journal of Abnormal and Social Psychology*, vol. 57, no. 3, p. 292, 1958.

[98] P. A. Obrist, "Skin resistance levels and galvanic skin response: Unilateral differences," *Science*, vol. 139, no. 3551, pp. 227–228, 1963.

[99] D. T. Lykken, W. G. Iacono, K. Haroian, and T. Bouchard Jr, "Habituation of the skin conductance response to strong stimuli: A twin study," *Psychophysiology*, vol. 25, no. 1, pp. 4–15, 1988.

[100] M. Fredrikson, "Orienting and defensive reactions to phobic and conditioned fear stimuli in phobics and normals," *Psychophysiology*, vol. 18, no. 4, pp. 456–465, 1981.

[101] E. Duffy, "The concept of energy mobilization." *Psychological Review*, vol. 58, no. 1, p. 30, 1951.

[102] R. W. Picard, S. Fedor, and Y. Ayzenberg, "Multiple arousal theory and daily-life electrodermal activity asymmetry," *Emotion Review*, vol. 8, no. 1, pp. 62–75, 2016.

[103] M. E. Dawson, A. M. Schell, and D. L. Filion, *Handbook Of Psychophysiology*, 3rd ed. Cambridge University Press, 2007, ch. The electrodermal system., pp. 159–181.

[104] E. Binboga, S. Guven, F. Çatıkkaş, O. Bayazıt, and S. Tok, "Psychophysiological responses to competition and the big five personality traits," *Journal of Human Kinetics*, vol. 33, pp. 187–194, 2012.

[105] R. Bhoja, O. T. Guttman, A. A. Fox, E. Melikman, M. Kosemund, and K. J. Gingrich, "Psychophysiological Stress Indicators of Heart Hate Variability and Electrodermal Activity with Application in Healthcare Simulation Research," *Simulation in Healthcare*, vol. 15, no. 1, pp. 39–45, 2020.

[106] L. Silvert, S. Delplanque, H. Bouwalerh, C. Verpoort, and H. Sequeira, "Autonomic responding to aversive words without conscious valence discrimination," *International Journal of Psychophysiology*, vol. 53, no. 2, pp. 135–145, 2004.

[107] M. Page and R. C. A. Robson, "Galvanic skin responses from asking stressful questions," *British Journal of Nursing*, vol. 16, no. 10, pp. 622–627, 2007.

[108] R. Zangroniz, A. Martinez Rodrigo, J. M. Pastor García, M. López Bonal, and A. Fernández-Caballero, "Electrodermal activity sensor for classification of calm/distress condition," *Sensors*, vol. 17, p. 2324, 10 2017.

[109] M. Lader and L. Wing, "Habituation of the psycho-galvanic reflex in patients with anxiety states and in normal subjects," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 27, no. 3, p. 210, 1964.

[110] M. H. Lader and L. Wing, *Physiological measures, sedative drugs, and morbid anxiety.* Oxford UP, 1966, vol. 14.

[111] J. Kritikos, G. Tzannetos, C. Zoitaki, S. Poulopoulou, and D. Koutsouris, "Anxiety detection from electrodermal activity sensor with

movement & interaction during virtual reality simulation," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 571–576.

[112] M. Sarchiapone, C. Gramaglia, M. Iosue, V. Carli, L. Mandelli, A. Serretti, D. Marangon, and P. Zeppegno, "The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis," *BMC Psychiatry*, vol. 18, no. 1, p. 22, 2018.

[113] W. G. Iacono, D. T. Lykken, L. J. Peloquin, A. E. Lumry, R. H. Valentine, and V. B. Tuason, "Electrodermal activity in euthymic unipolar and bipolar affective disorders: A possible marker for depression," *Archives of General Psychiatry*, vol. 40, no. 5, pp. 557–565, 1983.

[114] W. G. Iacono, L. J. Peloquin, D. T. Lykken, K. P. Haroian, R. H. Valentine, and V. B. Tuason, "Electrodermal activity in euthymic patients with affective disorders: One-year retest stability and the effects of stimulus intensity and significance." *Journal of Abnormal Psychology*, vol. 93, no. 3, p. 304, 1984.

[115] J. Wendt, M. Lotze, A. I. Weike, N. Hosten, and A. O. Hamm, "Brain activation and defensive response mobilization during sustained exposure to phobia-related and other affective pictures in spider phobia," *Psychophysiology*, vol. 45, no. 2, pp. 205–215, 2008.

[116] M. C. Corballis and I. L. Beale, "Bilateral symmetry and behavior," *Psychological Review*, vol. 77, no. 5, pp. 451–464, 1970.

[117] K. Amunts, L. Jäncke, H. Mohlberg, H. Steinmetz, and K. Zilles, "Interhemispheric asymmetry of the human motor cortex related to handedness and gender," *Neuropsychologia*, vol. 38, no. 3, pp. 304–312, 2000.

[118] C. D. Good, I. Johnsrude, J. Ashburner, R. N. Henson, K. J. Friston, and R. S. Frackowiak, "Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains," *Neuroimage*, vol. 14, no. 3, pp. 685–700, 2001.

[119] N. TaheriNejad and D. Pollreisz, "Assessment of Physiological Signals During Happiness, Sadness, Pain or Anger," in *Wireless Mobile Communication and Healthcare: 6th International Conference, MobiHealth 2016, Milan, Italy, November 14-16, 2016, Proceedings 6*. Springer, 2017, pp. 107–114.

[120] R. T. Faghih, P. A. Stokes, M.-F. Marin, R. G. Zsido, S. Zorowitz, B. L. Rosenbaum, H. Song, M. R. Milad, D. D. Dougherty, E. N. Eskandar *et al.*, "Characterization of Fear Conditioning and Fear Extinction by Analysis of Electrodermal Activity," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 7814–7818.

[121] A. Momin, S. Bhattacharya, S. Sanyal, and P. Chakraborty, "Visual Attention, Mental Stress and Gender: a Study Using Physiological Signals," *IEEE Access*, vol. 8, pp. 165 973–165 988, 2020.

[122] U. Desai and A. D. Shetty, "Electrodermal activity (EDA) for Treatment of Neurological and Psychiatric Disorder Patients: a Review," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1. IEEE, 2021, pp. 1424–1430.

[123] O. Adeola, O. Evans, J. Ndubuisi Edeh, and I. Adisa, "The Future of Marketing: Artificial Intelligence, Virtual Reality, and Neuromarketing," *Marketing communications and brand development in emerging economies Volume I: Contemporary and future perspectives*, pp. 253–280, 2022.

[124] A. O. Johansen, J. Mølgaard, S. S. Rasmussen, Y. Gu, K. K. Grønbæk, H. B. Sørensen, E. K. Aasvang, and C. S. Meyhoff, "Deviations in Continuously Monitored Electrodermal Activity Before Severe Clinical Complications: A Clinical Prospective Observational Explorative Cohort Study," *Journal of Clinical Monitoring and Computing*, vol. 37, no. 6, pp. 1573–1584, 2023.

[125] A. Sano, R. W. Picard, and R. Stickgold, "Quantitative Analysis of Wrist Electrodermal Activity During Sleep," *International Journal of Psychophysiology*, vol. 94, no. 3, pp. 382–389, 2014.

[126] A. Herlan, J. Ottenbacher, J. Schneider, D. Riemann, and B. Feige, "Electrodermal Activity Patterns in Sleep Stages and Their Utility for sleep Versus Wake Classification," *Journal of sleep research*, vol. 28, no. 2, p. e12694, 2019.

[127] L. D. Bolaños, J. M. Vicente-Samper, D. Z. Vinaroz, O. A. Vivas, V. E. Sala, and J. M. Sabater-Navarro, "Low-cost EDA Device for Screening Diabetic Neuropathy," in *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2019, pp. 253–258.

[128] S. Singaram, K. Ramakrishnan, and S. Periyasamy, "Electrodermal Signal Analysis using Continuous Wavelet Transform as a Tool for Quantification of Sweat Gland Activity in Diabetic Kidney Disease," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 237, no. 8, pp. 919–927, 2023.

[129] M. Žnidarič, D. Škrinjar, and A. Kapel, "Electrodermal Activity and Heart Rate Variability for Detection of Peripheral Abnormalities in Type 2 Diabetes: A review," *Biomolecules and Biomedicine*, vol. 23, no. 5, p. 740, 2023.

[130] S. G. Caliskan, M. D. Bilgin, and M. Polatli, "Nonlinear Analysis of Electrodermal Activity Signals for Healthy Subjects and Patients with Chronic Obstructive Pulmonary Disease," *Australasian Physical & Engineering Sciences in Medicine*, vol. 41, pp. 487–494, 2018.

[131] D. N. O'Connell and M. T. Orne, "Endosomatic electrodermal correlates of hypnotic depth and susceptibility," *Journal of Psychiatric Research*, vol. 6, no. 1, pp. 1–12, 1968.

[132] A. Sarkar, S. Biswas, G. Datwal, S. Debnath, B. Goswami, and R. Ghosh, "Design and calibration of a multi-channel low voltage data acquisition system," in *2018 IEEE Applied Signal Processing Conference (ASPCON)*, 2018, pp. 119–123.

[133] R. Ghosh, A. Bhattacharya, D. Mondal, S. Biswas, S. Nandy, and A. Sarkar, ""An expert system to determine the human condition using bilateral endosomatic EDA and the method thereof," ," Indian patent 201 931 038 304, Oct. 23, 2022, grant No. 402187. [Online]. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=IN309729256&_cid=P10-L776DE-43223-1

[134] Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. Roth, M. E. Dawson, and D. L. Filion, "Publication Recommendations for Electrodermal Measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, 2012.

[135] M. Calzavara, A. Persona, F. Sgarbossa, and V. Visentin, "A model for rest allowance estimation to improve tasks assignment to operators," *International Journal of Production Research*, vol. 57, no. 3, pp. 948–962, 2019.

[136] H. Kaneko, H. Itoh, H. Kiriyama, T. Kamon, K. Fujiu, K. Morita, N. Michihata, T. Jo, N. Takeda, H. Morita *et al.*, "Restfulness from sleep and subsequent cardiovascular disease in the general population," *Scientific Reports*, vol. 10, no. 1, p. 19674, 2020.

[137] N. Kleitman, *Sleep and wakefulness*. Chicago and London: University of Chicago Press, 1987.

[138] K. Cherry, "The 4 Stages of Sleep: What happens in the brain and body during NREM and REM sleep," Online, Aug. 2022. [Online]. Available: https://www.verywellhealth.com/the-four-stages-of-sleep-2795920#citation-12

[139] B. M. Altevogt, H. R. Colten *et al.*, *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*. Washington, DC: National Academies Press, 2006.

[140] J. Wörle, B. Metz, C. Thiele, and G. Weller, "Detecting sleep in drivers during highly automated driving: The potential of physiological parameters," *IET Intelligent Transport Systems*, vol. 13, no. 8, pp. 1241–1248, 2019.

[141] M. Misbhauddin, A. AlMutlaq, A. Almithn, N. Alshukr, and M. Aleesa, "Real-time driver drowsiness detection using wearable technology," in *Proceedings of the 4th international conference on smart city applications*, 2019, pp. 1–6.

[142] A. Němcová, V. Svozilová, K. Bucsuházy, R. Smíšek, M. Mézl, B. Hesko, M. Belák, M. Bilík, P. Maxera, M. Seitl *et al.*, "Multimodal Features for Detection of Driver Stress and Fatigue," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3214–3233, 2020.

[143] R. A. Hazarika, A. Abraham, S. N. Sur, A. K. Maji, and D. Kandar, "Different techniques for Alzheimer's disease classification using brain images: a study," *International Journal of Multimedia Information Retrieval*, pp. 1–20, 2021.

[144] V. V. Kale, S. T. Hamde, and R. S. Holambe, "Brain disease diagnosis using local binary pattern and steerable pyramid," *International Journal of Multimedia Information Retrieval*, vol. 8, no. 3, pp. 155–165, 2019.

[145] V. Jalali and D. Kaur, "A study of classification and feature extraction techniques for brain tumor detection," *International Journal of Multimedia Information Retrieval*, pp. 1–20, 2020.

[146] S. Saman and S. J. Narayanan, "Survey on brain tumor segmentation and feature extraction of MR images," *International Journal of Multimedia Information Retrieval*, vol. 8, no. 2, pp. 79–99, 2019.

[147] A. Hazra, "A comprehensive survey on chest diseases analysis: technique, challenges and future research directions," *International Journal of Multimedia Information Retrieval*, pp. 1–28, 2021.

[148] A. Hazarika, L. Dutta, M. Boro, M. Barthakur, and M. Bhuyan, "An automatic feature extraction and fusion model: application to electromyogram (EMG) signal classification," *International Journal of Multimedia Information Retrieval*, vol. 7, no. 3, pp. 173–186, 2018.

[149] K. El Asnaoui, "Design ensemble deep learning model for pneumonia disease classification," *International Journal of Multimedia Information Retrieval*, vol. 10, no. 1, pp. 55–68, 2021.

[150] A. K. Jaiswal and H. Banka, "Epileptic seizure detection in EEG signal using machine learning techniques," *Australasian Physical & Engineering Sciences in Medicine*, vol. 41, no. 1, pp. 81–94, 2018.

[151] T. Khatibi and N. Rabinezhadsadatmahaleh, "Proposing feature engineering method based on deep learning and K-NNs for ECG beat classification and arrhythmia detection," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 1, pp. 49–68, 2020.

[152] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, 2020.

[153] Y. Liang, Z. Chen, R. Ward, and M. Elgendi, "Hypertension assessment via ECG and PPG signals: An evaluation using MIMIC database," *Diagnostics*, vol. 8, no. 3, p. 65, 2018.

[154] J. S. Rajput, M. Sharma, and U. R. Acharya, "Hypertension diagnosis index for discrimination of high-risk hypertension ECG signals using optimal orthogonal wavelet filter bank," *International Journal of Environmental Research and Public Health*, vol. 16, no. 21, p. 4068, 2019.

[155] J. S. Rajput, M. Sharma, R. San Tan, and U. R. Acharya, "Automated detection of severity of hypertension ECG signals using an optimal bi-orthogonal wavelet filter bank," *Computers in Biology and Medicine*, vol. 123, p. 103924, 2020.

[156] G. Kovacs, A. Avian, V. Foris, M. Tscherner, X. Kqiku, P. Douschan, G. Bachmaier, A. Olschewski, M. Matucci-Cerinic, and H. Olschewski, "Use of ECG and Other Simple Non-invasive Tools to Assess Pulmonary Hypertension," *PloS One*, vol. 11, no. 12, p. e0168706, 2016.

[157] Y. Luo, Y. Li, Y. Lu, S. Lin, and X. Liu, "The Prediction of Hypertension Based on Convolution Neural Network," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*. IEEE, 2018, pp. 2122–2127.

[158] I. Marin and N. Goga, "Hypertension detection based on machine learning," in *Proceedings of the 6th Conference on the Engineering of Computer Based Systems*, 2019, pp. 1–4.

[159] W. Chang, Y. Liu, Y. Xiao, X. Yuan, X. Xu, S. Zhang, and S. Zhou, "A machine-learning-based prediction method for hypertension outcomes based on medical data," *Diagnostics*, vol. 9, no. 4, p. 178, 2019.

[160] J. Fahrenberg and M. Myrtek, *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Washington: Hogrefe & Huber Publishers, 1996.

[161] A. R. Zapater, D. M. Silveira, A. d. Vitta, C. R. Padovani, and J. C. P. d. Silva, "Seat posture: the efficiency of an educational program for scholars," *Ciência & Saúde Coletiva*, vol. 9, no. 1, pp. 191–199, 2004.

[162] K. Nomura, M. Iwata, O. Augereau, and K. Kise, "Estimation of student's engagement based on the posture," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, 2019, pp. 164–167.

[163] T. Bhattacharya, S. Biswas, N. Aslam, and S. Chattopadhyay, "Posture detection using WBAN and its application in remote healthcare monitoring," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE, 2016, pp. 2027–2034.

[164] J. Jeffin Gracewell and S. Pavalarajan, "Fall detection based on posture classification for smart home environment," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3581–3588, 2021.

[165] B. Lewandowski, J. Liebner, T. Wengefeld, S. Müller, and H.-M. Gross, "Fast and robust 3D person detector and posture estimator for mobile robotic applications," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 4869–4875.

[166] N. N. Vo and H. Yu, "A Wearable Posture Detection Device for Inpatient Healthcare," *Journal of Medical Devices-Transaction of The ASME*, vol. 8, no. 3, 2014.

[167] K. Peterson, E. T. Ozawa, G. M. Pantalos, and M. K. Sharp, "Numerical Simulation of the Influence of Gravity and Posture on Cardiac Performance," *Annals of Biomedical Engineering*, vol. 30, no. 2, pp. 247–259, 2002.

[168] P. Van Leeuwen and H. C. Kümmell, "Effects of Posture on Cardiac Time Intervals," *American Journal of Noninvasive Cardiology*, vol. 5, pp. 125–128, 1991.

[169] S. S. Thomas, V. Nathan, C. Zong, K. Soundarapandian, X. Shi, and R. Jafari, "BioWatch: A Noninvasive Wrist-Based Blood Pressure Monitor that Incorporates Training Techniques for Posture and Subject Variability," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 5, pp. 1291–1300, 2015.

[170] Y. Yoshida, E. Yuda, K. Yamamoto, Y. Miura, and J. Hayano, "Machine-learning estimation of body posture and physical activity by wearable acceleration and heartbeat sensors," *Signal & Image Processing: An International Journal (SIPIJ)*, vol. 10, no. 3, pp. 1–9, 2019.

[171] D. Sasakawa, N. Honma, T. Nakayama, and S. Iizuka, "Human posture identification using a MIMO array," *Electronics*, vol. 7, no. 3, p. 37, 2018.

[172] *RISH Multi 18S*, Rishabh Instruments Pvt. Ltd. [Online]. Available: http://rishabh.co.in/uploads/product/15030919_RISH_multi_18S_manual_with_back_new.pdf

[173] *RISH Multi SI 232*, Rishabh Instruments Pvt. Ltd. [Online]. Available: http://rishabh.co.in/uploads/product/SI_232_MANUAL_10-2011_NEW-1.pdf

[174] M. Saha, R. Ghosh, and B. Goswami, "Robustness and Sensitivity Metrics for Tuning the Extended Kalman Filter," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 4, pp. 964–971, 2013.

[175] C. P. SP and M. A. V, *Introductory time series with R*. London, New York: Springer Science & Business Media, 2009.

[176] J. P. Bentley, *Principles of Measurement Systems*, 4th ed. Pearson Prentice-Hall, England, 2005.

[177] S. Souissi, O. B. Rhouma, and C. Rebai, "Statistical modeling of mains zero crossing variation in powerline communication," *Measurement*, vol. 90, pp. 158–167, 2016.

[178] K. Wang, Y. Wu, Y. Gao, and Y. Li, "New methods to estimate the observed noise variance for an ARMA model," *Measurement*, vol. 99, pp. 164–170, 2017.

[179] G. Strang, *Introduction to Linear Algebra*, 4th ed. Wellesley-Cambridge Press Wellesley, MA, USA, 2009.

[180] A. Koskinas, E. Zaharopoulou, G. Pouliasis, I. Deligiannis, P. Dimitriadis, T. Iliopoulou, N. Mamassis, and D. Koutsoyiannis, "Estimating the Statistical Significance of Cross–Correlations between Hydroclimatic Processes in the Presence of Long–Range Dependence," *Earth*, vol. 3, no. 3, pp. 1027–1041, 2022.

[181] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions, Volume 2*. John Wiley & Sons, 1995, vol. 289.

[182] M. Borowska, "Entropy-based algorithms in the analysis of biomedical signals," *Studies in Logic, Grammar and Rhetoric*, vol. 43, no. 1, pp. 21–32, 2015.

[183] İ. Eşer, L. Khorshid, Ü. Yapucu Güneş, and Y. Demir, "The effect of different body positions on blood pressure," *Journal of Clinical Nursing*, vol. 16, no. 1, pp. 137–140, 2007.

[184] R. Arnheim, *Entropy and art: An essay on disorder and order*. Univ of California Press, 1974.

[185] M. Mallick and B. La Scala, "Comparison of Single-point and Two-point Difference Track Initiation Algorithms using Position Measurements," *Acta Automatica Sinica*, vol. 34, no. 3, pp. 258–265, 2008.

[186] G. Powell and I. Percival, "A spectral entropy method for distinguishing regular and irregular motion of Hamiltonian systems," *Journal of Physics A: Mathematical and General*, vol. 12, no. 11, p. 2053, 1979.

[187] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. Bhat, and J. E. Koh, "Application of entropies for automated diagnosis of epilepsy using EEG signals: A review," *Knowledge-Based Systems*, vol. 88, pp. 85–96, 2015.

[188] A. Howedi, A. Lotfi, and A. Pourabdollah, "An Entropy-Based Bpproach for Anomaly Detection in Activities of Daily Living in the Presence of a Visitor," *Entropy*, vol. 22, no. 8, p. 845, 2020.

[189] A. Sarkar, B. Goswami, and R. Ghosh, "Classification of Hypertensive and Normotensive Subjects Using Bilateral Differential Biopotential Signals," *Research Square*, 2021, doi https://doi.org/10.21203/rs.3.rs-813015/v1.

[190] G. Holmes, A. Donkin, and I. H. Witten, "WEKA: A machine learning workbench," in *Proceedings of ANZIIS'94-Australian New Zealnd Intelligent Information Systems Conference*. IEEE, 1994, pp. 357–361.

[191] B. Çığşar, D. Ünal *et al.*, "Comparison of Data Mining Classification Algorithms Determining the Default Risk," *Scientific Programming*, vol. 2019, 2019.

[192] C. A. Kumar, M. Sooraj, and S. Ramakrishnan, "A Comparative Performance Evaluation of Supervised Feature Selection Algorithms on Microarray Datasets," *Procedia Computer Science*, vol. 115, pp. 209–217, 2017.

[193] C. S. Kumar and R. Sree, "Application of Ranking Based Attribute Selection Filters to Perform Automated Evaluation of Descriptive Answers Through Sequential Minimal Optimization Models." *ICTACT Journal on Soft Computing*, vol. 5, no. 1, 2014.

[194] S. Gnanambal, M. Thangaraj, V. Meenatchi, and V. Gayathri, "Classification Algorithms with Attribute Selection: an evaluation study using WEKA," *International Journal of Advanced Networking and Applications*, vol. 9, no. 6, pp. 3640–3644, 2018.

[195] S. F. Rosario, K. Thangadurai *et al.*, "RELIEF: Feature Selection Approach," *International Journal of Innovative Research and Development*, vol. 4, no. 11, pp. 218–224, 2015.

[196] M. Gillberg, G. Kecklund, J. Axelsson, and T. Åkerstedt, "The Effects of a Short Daytime Nap After Restricted Night Sleep," *Sleep*, vol. 19, no. 7, pp. 570–575, 1996.

[197] J. A. Horne and L. A. Reyner, "Counteracting Driver Sleepiness: Effects of Napping, Caffeine, and Placebo," *Psychophysiology*, vol. 33, no. 3, pp. 306–309, 1996.

[198] M. T. DMSc and H. A. DMSc, "Maintenance of Alertness and Performance by a Brief Nap After Lunch Under Prior Sleep Deficit," *Sleep*, vol. 23, no. 6, p. 813, 2000.

[199] A. J. Tietzel and L. C. Lack, "The Short-term Benefits of Brief and Long Naps Following Nocturnal Sleep Restriction," *Sleep*, vol. 24, no. 3, pp. 293–300, 2001.

[200] ——, "The Recuperative Value of Brief and Ultra-brief Naps on Alertness and Cognitive Performance," *Journal of sleep research*, vol. 11, no. 3, pp. 213–218, 2002.

[201] M. Hayashi, N. Motoyoshi, and T. Hori, "Recuperative Power of a Short Daytime Nap With or Without Stage 2 Sleep," *Sleep*, vol. 28, no. 7, pp. 829–836, 2005.

[202] M. Hayashi, S. Ito, and T. Hori, "The Effects of a 20-min Nap at Noon on Sleepiness, Performance and EEG Activity," *International Journal of Psychophysiology*, vol. 32, no. 2, pp. 173–180, 1999.

[203] M. Hayashi, M. Watanabe, and T. Hori, "The Effects of a 20 min Nap in the Mid-afternoon on Mood, Performance and EEG Activity," *Clinical Neurophysiology*, vol. 110, no. 2, pp. 272–279, 1999.

[204] M. Hayashi, A. Masuda, and T. Hori, "The Alerting Effects of Caffeine, Bright Light and Face Washing After a Short Daytime Nap," *Clinical Neurophysiology*, vol. 114, no. 12, pp. 2268–2278, 2003.

[205] M. Takahashi, H. Fukuda, and H. Arito, "Brief Naps During Postlunch Rest: Effects on Alertness, Performance, and Autonomic Balance," *European journal of applied physiology and occupational physiology*, vol. 78, no. 2, pp. 93–98, 1998.

[206] A. Brooks and L. Lack, "A Brief Afternoon Nap Following Nocturnal Sleep Restriction: Which Nap Duration is Most Recuperative?" *Sleep*, vol. 29, no. 6, pp. 831–840, 2006.

[207] N. Al-Busaidi, S. Morrison, B. Hadley, A. McAlister, and M. Brown, "Subjective Benefit Versus Sleep Inertia After Short Post Lunch Nap," *INTERNATIONAL JOURNAL OF ADVANCED MEDICINE*, vol. 2, no. 4, pp. 6–10, 2018.

[208] C. J. Hilditch, S. A. Centofanti, J. Dorrian, and S. Banks, "A 30-minute, but not a 10-minute Nighttime Nap is Associated with Sleep Inertia," *Sleep*, vol. 39, no. 3, pp. 675–685, 2016.

[209] M. Romdhani, N. Souissi, Y. Chaabouni, K. Mahdouani, T. Driss, K. Chamari, and O. Hammouda, "Improved Physical Performance and Decreased Muscular and Oxidative Damage with Postlunch Napping After Partial Sleep Deprivation in Athletes," *International journal of sports physiology and performance*, vol. 15, no. 6, pp. 874–883, 2020.

[210] A. Who, "Global Brief on Hypertension," *World Health Organization*, 2013.

[211] "Cardiovascular Diseases (CVDs)," *World Health Organization*, 11 June 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[212] WHO, "Hypertension in India - WHO," *World Health Organization*, 2021. [Online]. Available: https://www.who.int/india/health-topics/hypertension

[213] H. Ni, Y. Wang, G. Xu, Z. Shao, W. Zhang, and X. Zhou, "Multi-scale Fine-Grained Heart Rate Variability Analysis for Recognizing the Severity of Hypertension," *Computational and Mathematical Methods in Medicine 2019*, 2019.

[214] M. Simjanoska, M. Gjoreski, M. Gams, and A. Madevska Bogdanova, "Non-invasive blood pressure estimation from ECG using machine learning techniques," *Sensors*, vol. 18, no. 4, p. 1160, 2018.

[215] M. Poddar, A. C. Birajdar, J. Virmani *et al.*, "Automated Classification of Hypertension and Coronary Artery Disease Patients by PNN, KNN, and SVM Classifiers Using HRV Analysis," in *Machine Learning in Bio-signal Analysis and Diagnostic Imaging*. Elsevier, 2019, pp. 99–125.

[216] T. Seidler, K. Hellenkamp, B. Unsoeld, S. Mushemi-Blake, A. Shah, G. Hasenfuss, and A. Leha, "A machine learning approach for the prediction of pulmonary hypertension," *Journal of the American College of Cardiology*, vol. 73, no. 9S1, pp. 1589–1589, 2019.

[217] N. Foubert, A. M. McKee, R. A. Goubran, and F. Knoefel, "Lying and Sitting Posture Recognition and Transition Detection using a Pressure Sensor Array," in *2012 IEEE International Symposium on Medical Measurements and Applications Proceedings*. IEEE, 2012, pp. 1–6.

[218] H. Jeong and W. Park, "Developing and Evaluating a Mixed Sensor Smart Chair System for Real-time Posture Classification: Combining Pressure and Distance Sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1805–1813, 2020.

[219] T. E. Lockhart, R. Soangra, J. Zhang, and X. Wu, "Wavelet Based Automated Postural Event Detection and Aactivity Classification with Single IMU (TEMPO)," *Biomedical sciences instrumentation*, vol. 49, p. 224, 2013.

[220] K. Lyden, D. John, P. Dall, M. H. Granat *et al.*, "Differentiating Sitting and Lying using a Thigh-worn Accelerometer," *Med Sci Sports Exerc*, vol. 48, no. 4, pp. 742–747, 2016.

[221] D. R. Bassett Jr, D. John, S. A. Conger, B. C. Rider, R. M. Passmore, and J. M. Clark, "Detection of Lying down, Sitting, Standing, and Stepping using Two ActivPAL Monitors." *Medicine and science in sports and exercise*, vol. 46, no. 10, pp. 2025–2029, 2014.

[222] F. Foerster, M. Smeja, and J. Fahrenberg, "Detection of Posture and Motion by Accelerometry: A Validation Study in Ambulatory Monitoring," *Computers in human behavior*, vol. 15, no. 5, pp. 571–583, 1999.

[223] S. Wahabi, *Variability in ECG Biometrics: State of the Art and Subspace Methods*. University of Toronto (Canada), 2015.

[224] A. D. Ruiz, J. S. Mejía, J. M. López, and B. F. Giraldo, "Characterization of Cardiac and Respiratory System of Healthy Subjects in Supine and Sitting Position," in *Iberian Conference on Pattern Recognition and Image Analysis*. Springer, 2019, pp. 367–377.

[225] S. Wahabi, S. Pouryayevali, and D. Hatzinakos, "Posture-Invariant ECG Recognition with Posture Detection," in *2015 IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2015, pp. 1812–1816.

[226] N. Rikatsih and A. A. Supianto, "Classification of Posture Reconstruction with Univariate Time Series Data Type," in *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*. IEEE, 2018, pp. 322–325.