Understanding Peripheral Blood Pressure Signals: A Statistical Learning Approach

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Engineering with a concentration in Electrical Engineering

by

Md Abul Hayat
Bangladesh University of Engineering and Technology
Bachelor of Science in Electrical and Electronic Engineering, 2015
University of Arkansas
Master of Science in Statistics and Analytics, 2021

August 2023
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Jingxian Wu, Ph.D.
Dissertation Director

Morten O. Jensen, Ph.D.
Committee Member

Roy A. McCann, Ph.D.
Committee Member

Yue Zhao, Ph.D.
Committee Member

Abstract

Proper estimation of body fluid status for human or animal subjects has always been a challenging problem. Accurate and timely estimate of body fluid can prevent life threatening conditions under trauma and severe dehydration. The main objective of this research is the estimation, classification and detection of dehydration in human and animal subjects using peripheral blood pressure (PBP) signals. Peripheral venous pressure (PVP) and peripheral arterial pressure (PAP) signals have been investigated in this research. Both PVP and PAP signals are PBP signals. A dataset of PVP signals was collected using standard peripheral intravenous catheters from human subjects suffering from hypertrophic pyloric stenosis. Using this dataset, we successfully classified dehydrated subjects from hydrated subjects using regularized logistic regression on frequency domain data of the PVP signals. During the data acquisition process, the PVP signals was corrupted by noise and blood clot. So, we developed an unsupervised anomaly detection algorithm for PVP signals using hidden Markov model and Kalman filter. This anomaly detection algorithm removed the human bias in data-preprocessing. Another dataset of PAP and PVP signals was collected from pigs under anesthesia using the Millar catheter. We proposed a integral pulse frequency modulation (IPFM) based signal model for both PAP and PVP signals. The proposed model-synthesized signal is highly correlated with the experimental data. The model-synthesized signals also performs similar to experimental signals under classification tasks. We also examine the model estimated parameters both qualitatively and quantitatively. This model can also quantify the effect of respiratory rate on heart rate variability. Increasing doses of anesthesia has similar effect of getting hydrated from dehydration.

Acknowledgements

First of all, I am immensely grateful to Dr. Jingxian Wu for giving me the opportunity to pursue Ph.D. studies. I have learned a so much from his calm and composed way of life and thinking process. Without his help and mentorship, it would be impossible. He never said 'No' to any of my requests. I want to thank him for trusting and keeping faith on me with patience. I cannot thank him enough to help me get the M.S. in Statistics while being a Ph.D. student in Electrical Engineering.

My sincere gratitude towards Dr. Morten O. Jensen and Dr. Hanna K. Jensen for their constant support, motivation and inspiration from the very beginning. Special thanks to Ali Al-Alawi, Mugisha Nitunga, Kaylee Henry, Lauren Crimmins-Pierce and Gabriel Bonvillain for helping and working with me. I would also like to express gratitude towards our collaborators from the UAMS - Dr. Kevin W. Sexton, Dr. Joseph A. Sanford Jr., and Dr. Patrick C. Bonasso II for their constant support and time amid busy schedule. It has been a pleasure to work with all of you. Sincere thanks to Dr. Roy McCann and Dr. Yue Zhao for being members in my dissertation committee and thoughtful suggestions.

I have thoroughly enjoyed the Statistics courses offered at the university. The statistical knowledge I gained helped me a lot in my research. I want to thank Dr. Giovanni Petris and Dr. Jyotishka Datta for introducing me to Bayesian statistics and inspiring me to pursue M.S. in Statistics and Analytics.

I want to especially thank Samrat Nath for being my friend and mentor from the very day I landed in Fayetteville, AR. My heartiest gratitude is towards the group members - Zuoen Wang, Israel Akingeneye, Tanny Chavez Esparza, Yaze Li, and Andong Zhou.

I want to thank the faculties and staffs of the Electrical Engineering department for being a part of my journey. Special thanks to Mr. Robert Saunders for giving me the opportunity to teach Electric Circuits I (ELEG 2103) in Spring and Summer semesters of 2023. I am grateful to you for so many reasons. I want to express my gratitude to my friends and well-wishers for all the support and love over years. We thoroughly enjoyed our stay in

Northwest Arkansas.

The story of my Ph.D. would be incomplete if I do not mention about Mustafa Mustafa. Mustafa was my mentor at Lawrence Berkeley National Lab summer internship (2020). I cannot thank him enough for giving me the opportunity to work on the exciting self-supervised contrastive learning project. I have learned so much from him and the team-members there. I also want to thank William Sean Kennedy and Yassar Azam for trusting me for the internship opportunity at Nokia Bell Labs (2019) and Amazon (2021) respectively.

My father, Md Abu Bakkar Mridha, and mother, Samirun Nessa always supported my academic journey. My mother's role has been instrumental in my educational journey. Without her support, morally high upbringing, and love for education, it would have been impossible to secure whatever I have done today. If I look back, it feels amazing how I ended up where I am today. My brother, Mohammad Abul Bashar, has played a very important role in my upbringing. I am using this as an opportunity to thank him for all his love and support since my early childhood.

I am grateful to my wife, Zin-an Arshad, for sacrificing her most valuable six years in career for me and our kids. She has been my rock and solace during this journey. No matter how much words I use, it is not going to be enough.

Dedication

Samirun Nessa – *My Mother*

Zin-an Arshad – *My Wife*

Zeeana Hayat, Zaira Hayat – *My Daughters*

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation and Background

Estimating the amount of body fluid in patients has always been a challenge under critical health monitoring scenarios. Accurate and timely detection of body fluid or haemodynamic status can be a very useful tool during severe trauma and other life-threatening scenarios. In the United States alone, dehydration affects 30 million children annually and accounts for 400,000 pediatric emergency room visits [2, 3]. Several existing studies have attempted to define level of dehydration [4, 5, 6, 7, 8]. In children, there is no standard of care for volume status measurement and it is still a diagnostic challenge. Unrecognized dehydration or incorrect estimation of its degree can adversely affect the morbidity and mortality of a dehydrated infant [9, 10]. With increasing amount of data being generated in the healthcare industry in recent time, it is quite obvious that future health monitoring system is going to be heavily dependent on data. This research was aimed at answering this dehydration detection problem using a data driven approach.

The blood pressure signals collected from the peripheral vein and artery are defined as peripheral blood pressure signals (PBP). The two PBP signals being discussed in this dissertation are peripheral venous pressure (PVP) signals and peripheral arterial pressure (PAP) signals. The main goal of this dissertation is to quantitatively model, analyze, and understand the relationship between PVP signal and human volume status due to various medical conditions, such as blood loss or dehydration. Understanding the relationship between PVP and body fluid status can immensely contribute to the understanding of haemodynamic status of patients. Apart from that, PVP signal can be easily collected using the ubiquitously available peripheral catheters. This makes the PVP signal a very useful and available choice for haemodynamic status monitoring.

Compared to other widely used biomedical signals such as Electrocardiogram (ECG), Electroencephalogram (EEG) and Photoplethysmogram (PPG), PVP signals have received considerably less attention from the scientific community. Existing works on PVP signals are

mostly qualitative and does not incorporate advanced techniques such as statistical signal processing or artificial intelligence. The recent success of data driven approaches in various fields has made it a promising technique in the biomedical and healthcare areas.

In this dissertation, we aim to understand the utility of PVP signal and finding its relation to other physiological and haemodynamic conditions. Following subsection include a comprehensive review of all previous works done on venous pressure and related signals. Along with the literature review, this chapter also introduces the very basics of heart physiology to help understand the following chapters.

1.2 Literature Review

Peripheral venous pressure (PVP) signal has been a subject of research primarily among the physicians. The works can be divided into three different broad groups which are discussed here. Among review papers of venous signals, [11] is probably the earliest. Among others, [12] is a accessible review on venous blood pressure and vein physiology. A recent and more extensive review on PVP signal is [13].

1.2.1 PVP Signals and Breathing Pattern

Early analytical works related to PVP signals include [14]. Using three patients, the authors demonstrated the similarity of PPG, PVP (collected from the dorsum of the hand) and CVP signals. Works on impact of breathing on PVP signals has been presented in [15, 16, 17]. [15] mainly focuses on the relation between PPG and different breathing patterns along with PVP signals. [16] discusses impact of three different breathing patterns (Muller, Valsalva and Incentive Spirometry) on the PVP signals collected from the hand. Another recent study [17] discusses the effect of respiratory modulation of peripheral vesoconstriction from signal modelling perspective.

1.2.2 Relationship Between PVP and CVP Signals

Blood pressure measured at the superior vena cava (Fig. 1.1) is referred as Central venous pressure (CVP). The measurement of CVP requires the insertion and maintenance of a catheter in the thorax. For more than a decade, most of the publications on PVP signals were based on finding its relation with CVP signals using simple statistical methods. Interested reader can refer to [18], this paper discusses the usefulness and limitations of using CVP as a guide to clinical fluid management.

To the best of authors knowledge, [19] is the first paper to demonstrate high correlation ($r = 0.82$) between CVP and PVP signals based on 15 neurological patients and 1026 paired measurements. This paper also claims that PVP approximates mean systemic pressure, which is relatively independent of cardiac influences, and reflects volume status alone. A similar study [20] of 150 patients without cardiac dysfunction undergoing major elective non-cardiac surgery also demonstrated CVP and PVP correlation ($r = 0.86$). Another study [21] demonstrated strong correlation ($r = 0.89$) between CVP and PVP preformed on 500 patients. It also provides linear relationship between PVP and CVP in different patient positions, different catheter diameters, and catheterization sites. Similarly in [22], 188 measurements in 19 patients undergoing cardiac surgery demonstrated CVP and PVP values to be correlated with a bias of the PVP between -0.72 and 0 mmHg compared to the CVP. Patients aged 3-9 years also demonstrated correlation of $r = 0.93$ between PVP and CVP signals [23]. Patients undergoing kidney surgery also demonstrated high correlation in [24, 25, 26]. Interestingly, [27] demonstrated that the difference between PVP and CVP decreases with increasing CVP. Other studies demonstrating CVP and PVP correlation include patients under infants and children heart disease [28, 29], right heart catheterization [30], laparoscopic colorectal surgery [31] and Fontan operation [32, 33]. In [33], the researchers have proposed a liner relationship between CVP and VEN signal for adult Fontan patients. PVP measurements in patients With acute decompensated heart failure are also correlated with CVP signals in [34].

1.2.3 Recent PVP Research

Papers discussing other than the relation between CVP and PVP signals started to emerge from last decade. Impact of lower body negative pressure (LBNP) hypovolemia (dehydration) on PVP signal has been discussed in [35, 36]. This study [35] sought to examine the impact of central hypovolemia on PVP in spontaneously breathing subjects. In response to mild hypovolemia induced by LBNP, changes in cardiac modulation and other PVP waveform parameters identified hypovolemia before detectable hemodynamic changes. Advanced time-frequency based techniques like synchrosqueezing transform have been applied on PVP signal to extract meaningful phase information under noise [37, 36]. Another research [38] on pigs demonstrated that the strength of heart rate frequency component gradually increases from hypovolemia (dehydration), euvolemia (hydrated stage) to hypervolemia (fluid overload). Another similar research [39] on patients under haemodialysis proposed that peripheral i.v. analysis is a method for analysing the peripheral venous waveform that has been used to monitor volume status. It was a proof-of-concept study for evaluating the efficacy of PIVA in detecting changes in fluid volume. Another interesting pilot study [40] on hospitalized patients with acute decompensated heart failure demonstrated PVP signal on admission correlated with the amount of volume removed during diuresis (increased production of urine). Some recent publications [41, 42] also discusses experimental variations and challenges in PVP signals acquisition. Interesting findings from start-ups like VoluMetrix [43] published studies [44] on heart rate variability monitoring utilizing a non-invasive venous pressure signal. This is a different signal from the one (minimally invasive) that has been discussed in the following chapters of the document.

Recent researches include craniosynostosis (a kind of skull surgery) patients under the influence to two different anesthetics, namely - propofol and isoflurane [45] [46]. Using MANOVA, this study has found that anesthetics influence the PVP signal. A simple k-nearest neighbor algorithm can classify different isoflurane stages. Also provides evidence of cross-talk between arterial and venous blood pressure. Another study [47] on pigs showed

that Pearson correlation coefficients between venous and arterial signals decreased as anesthetic dosage increased. In an opposing fashion, the correlation coefficients increased as hemorrhage occurred. This study claims that anesthetics can influence PVP signals and hence needs more in-depth study.

1.3 Cardiac Physiology

A brief idea on cardiac physiology is necessary to understand the underlying dynamics of the signals generated. The blood flow of the human body is generated by heart. The heart is divided into four chambers (Fig. 1.1). The upper chambers where blood come into the heart are called atria (singular: atrium). The lower chambers that drive blood out of the heart are called ventricles (singular: ventricle). Veins and arteries are responsible to bring blood in and out from the heart respectively. Arteries deliver the blood to the whole body and veins collects the blood and brings back to the heart. Superior and inferior vena cava collects deoxygenated blood from rest of the body and end at right atrium (RA). On the other hand, oxygenated bloods are supplied to left atrium (LA) by pulmonary veins (from lungs). Right ventricle (RV) pushes the deoxygenated blood to lungs through pulmonary artery. On the other hand, left ventricle (LV) pushes the oxygenated blood aorta. These four compartments are separated by valves. The valve between RA and RV is called tricuspid or right atrioventricular valve. Whereas, LA and LV are seperated by bicuspid or Mitral valve. The semi-lunar valve between RV and pulmonary artery is called pulmonary valve. Similarly, the semi-lunar valve between LV and aorta is called aortic valve.

The whole arterial and venous signal generation process can be understood from the Wiggers diagram. The interested readers can refer to [48]. The arterial and venous signals are related to each other and there is increasing amount of interest among researchers to find out the relationship between the two.

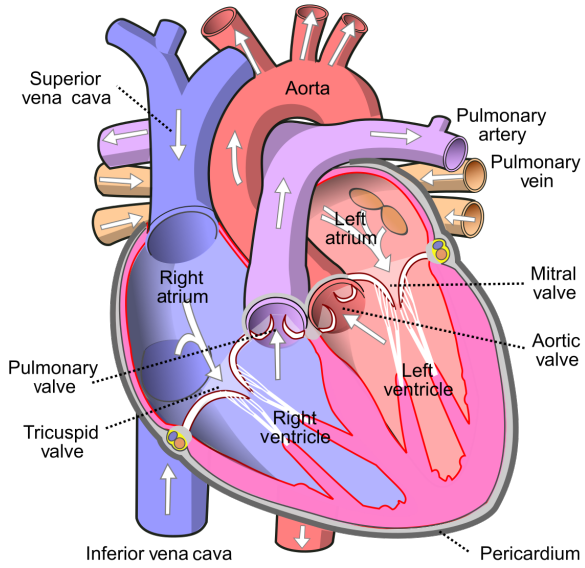

Figure 1.1: Diagram of the human heart [1].

1.4 Dissertation Outline

The outline of the rest of this dissertation is as follows.

Chapter 2: This chapter describes the primary result obtained using 18 human subjects suffering from hypertrophic pyloric stenosis. Among these 18 patients, 10 were resuscitated or hydrated and 8 were hypovolemic or dehydrated patients. PVP signal was collected using standard peripheral intravenous (PIV) catheter. This study proves that it is possible to separate these two groups of patients with sensitivity $> 96\%$ and specificity $> 92\%$ using regularized logistic regression on frequency domain data. Using statistical methods, it proves that signal power at 2.4 Hz (heart rate) is a crucial indicator to distinguish the two different classes. Also, there are other statistical results that prove the utility of the PVP signal in distinguishing the hydrated subjects from dehydrated subjects.

Chapter 3: While collecting the signals using PIV catheters the main challenge of the data collection was clots or air in arterial tubing. These clots or air corrupted the signal. As these PVP signals were collected from babies and from peripheral veins, the signals were of very low frequency. So eliminating anomaly from the signal was very important.

In Chapter 2, anomalous part of the data was manually cleaned. The manual cleaning process introduces subjective bias in the result. This is why we proposed a hidden Markov model based anomaly detection algorithm with Kalman filter to in this chapter to remove anomalies in objective way. This chapter is continuation of the research in Chapter 2 and similar classification results are discussed here. It should be noted that data used in Chapter 2 and 3 are collected from human subjects.

Chapter 4: The data used in this chapter is collected from 4 female pigs under anesthesia from brachial artery and brachial vein. Peripheral artery pressure (PAP) and peripheral veinous pressure (PVP) signals are the main focus of this chapter. In this case, we have used Millar catheters that are different from PIV catheters. Signal collected using Millar catheters have high fidelity and does not suffer from the blood clotting issue. This chapter proposes a integral pulse frequency modulation (IPFM) based signal model for PAP and PVP signals. This is first ever PVP or PAP signal modeling approach. Proposed model-synthesized signal is highly correlated with the experimental signal correlation coefficient $\rho > 0.99$ for PAP and $\rho > 0.90$ PVP signals. In classification problems the model-synthesized signal and the experimental signal also performed similar which proves that our model estimated signal is retaining class discriminatory features. Using IPFM-based model, we can also examine the relationship of heart rate variability (HRV) and respiratory rate. We also presented some qualitative and quantitative results based on this proposed model.

Chapter 5: This chapter concludes this dissertation with a brief summary and future research directions.

2 Venous Physiology Predicts Dehydration in the Pediatric Population

2.1 Abstract

There does not exist any standard dehydration monitoring method for children. This study attempts to determine the utility of Fast Fourier Transform (FFT) based regression method of a peripheral venous pressure (PVP) waveform to predict dehydration.

PVP waveforms were collected from 18 patients. Groups were defined as resuscitated or hydrated (serum chloride $\geq$ 100 mmol/L) and hypovolemic or dehydrated (serum chloride $<$ 100 mmol/L). Data were collected on emergency department admission and after a 20 cc/kg fluid bolus. The MATLAB (MathWorks) software analyzed nonoverlapping 10 s window signals. Based on the analysis, 2.4 Hz (144 bps) was the most demonstrative frequency to compare the PVP signal power (mmHg).

Frequency components of admitted subjects were compared between 10 (56%) resuscitated and 8 (44%) hypovolemic patients. The PVP signal power was higher in resuscitated patients (median 0.174 mmHg, IQR: 0.079-0.374 mmHg) than in hypovolemic patients (median 0.026 mmHg, IQR: 0.001-0.057 mmHg), ($p < 0.001$) at 2.4 Hz. Fourteen patients received a bolus regardless of laboratory values: 6 (43%) resuscitated and 8 (57%) hypovolemic. In resuscitated patients, the signal power did not change significantly after the fluid bolus at 2.4 Hz (median 0.142 mmHg, IQR: 0.032-0.383 mmHg) ($p = 0.019$), whereas significantly increased signal power (median 0.0474 mmHg, IQR: 0.019-0.110 mmHg) was observed in the hypovolemic patients after a fluid bolus at 2.4 Hz ($p < 0.001$). The algorithm predicted dehydration for window-level analysis (sensitivity 97.95%, specificity 93.07%). The algorithm also predicted dehydration for patient-level analysis (sensitivity 100%, specificity 100%).

PVP waveforms can predict dehydration in hypertrophic pyloric stenosis. Further work is needed to determine the utility of PVP analysis to guide fluid resuscitation status in other pediatric populations.

2.2 Introduction

Dehydration is common and sometimes life-threatening in children. Clinicians use vital signs, urine output, physical examination findings, and laboratory values to estimate the level of dehydration and intravascular volume status; however, determining level of dehydration is still difficult as there is no standard dehydration monitor in children [2]. Scales have been developed to classify dehydration; however, they are not accurate, and more precise predictors are still needed [49, 50]. The lack of a suitable gold standard has been a barrier to clinical research and may have hampered good clinical care.

Analysis of peripheral venous pressure (PVP) waveforms is a novel method of monitoring intravascular volume, which may provide earlier sensitivity in detection of volume depletion [38]. Fast Fourier Transform (FFT) is a powerful technique, which facilitates analysis of hemodynamic signals in the frequency domain [51]. Previous work in a young porcine model has demonstrated that FFT of a PVP waveform correlated with volume status more sensitively than standard vital sign monitoring [38]. The use of PVP waveforms could represent a significant advance in the field, particularly, as most children with severe dehydration often require intravenous rehydration.

Hypertrophic pyloric stenosis (HPS) is a common pediatric condition in which patients encounter varying degrees of dehydration and marked biochemical derangement [52, 53, 54]. Electrolytes are used as a marker of resuscitation. Fluid replacement corrects the hypochloremic, hypokalemic metabolic alkalosis to avoid postoperative apnea [55]. Typically, if electrolyte abnormalities ($Cl^- < 100$ mmol/L, $HCO_3 \geq 30$ mmol/L) are present on diagnosis of HPS, aggressive intravenous fluid (IVF) resuscitation is given [56, 57, 58]. In [56], the authors found chloride to be the most sensitive and specific indicator of the need for multiple saline boluses [56]. Identifying a noninvasive monitor via a PVP waveform that correlates with hydration status and the presently used chloride values could be beneficial in determining adequate resuscitation.

In this study, we use PVP waveform analysis via a standard peripheral intravenous

(PIV) catheter to study dehydration in children. We hypothesized that FFT could predict dehydration in patients with HPS.

2.3 Materials & Methods

This study was performed in accordance with the University of Arkansas for Medical Sciences Institutional Review Board (IRB: 206193). After obtaining informed parental consent, PVP waveforms were collected from 32 patients with ultrasound-proven HPS.

Patients with cardiac defects, patients who were previously admitted to an outside hospital and received IVF, and patients receiving mechanical ventilation were excluded from the study. Among 32 patients, 29 (91%) had a 24-gauge Insyte-N Autoguard PIV catheter (Becton Dickinson Infusion Therapy Systems, Sandy, UT) placed in the upper or lower extremity at time of emergency department admission. The PIV was connected to a Deltran II pressure transducer (ADInstruments, Colorado Springs, CO) interfaced with a PowerLab (ADInstruments) data acquisition system via 48-inch arterial pressuring tubing (Smiths Medical, Dublin, OH). A standard algorithm for fluid resuscitation was used based on the initial chloride and carbon dioxide laboratory values [56]. Groups were defined as resuscitated (i.e., ready for the operating room) when serum chloride was $\geq$ 100 mmol/L or bicarbonate $<$ 30 mmol/L and hypovolemic when serum chloride $<$ 100 mmol/L or bicarbonate $\geq$ 30 mmol/L. Data were collected on emergency department admission and after a 20 cc/kg fluid bolus.

The data collected before and after bolus were analyzed by using signal processing algorithms developed in the MATLAB (MathWorks, Natick, MA) software. The objective of this analysis was to identify a quantitative relationship between the PVP waveforms and the intravascular volume status of patients.

2.4 Signal Processing Algorithm

2.4.1 Sampling

For each patient, the entire PVP waveform was divided into nonoverlapping windows of 10 s, with a sampling rate of 1000 Hz. There are total 329 windows from all hypovolemic patients, and 343 from all resuscitated patients. FFT was performed over the time domain signals in each window to obtain the frequency domain signal representation of components between 0 and 20 Hz [59]. A window size of 10 s yields a frequency domain resolution of $1/10 = 0.1$ Hz, meaning the space between two adjacent frequency domain samples is 0.1 Hz. Thus, there are a total of 200 frequency domain samples in each window between 0 and 19.9 Hz.

2.4.2 Logistic Regression

Logistic regression with LASSO (Least Absolute Shrinkage & Selection Operator), Ridge, and Elastic-net are performed over the frequency domain data to establish the correlation between the frequency domain PVP waveform signals and the hypovolemic or resuscitated status of a patient [60, 61, 3]. Denote the frequency domain samples in the i-th window as a 201-dimension vector X_i, with the first element being 1 and the other 200 elements corresponding to the frequency domain samples between 0 and 19.9 Hz. Define a binary variable Y_i to represent the hypovolemic or resuscitated status of the i-th window, with $Y_i = 1$ for hypovolemic and $Y_i = 0$ for resuscitated. For each window, the logistic regression model can be used to calculate its probability of being in one of the two categories, as

$$P(Y_i = 1) = \frac{1}{1 + e^{-\beta^T X_i}}$$
$$P(Y_i = 0) = \frac{e^{-\beta^T X_i}}{1 + e^{-\beta^T X_i}},$$

where β is a 201-dimension regression coefficient vector to be determined through training and β^T denotes the transpose of β. The coefficient vector β can be calculated by minimizing the following negative log likelihood function $l(\beta)$ during the training process

$$l(\beta) = \sum_{i=1}^{n}[(1 - y_i)\beta^T X_i + \log(1 + e^{-\beta^T X_i})] + \lambda P_\alpha(\beta),$$

where,

$$P_\alpha(\beta) = \frac{1-\alpha}{2}\|\beta\|_2^2 + \alpha\|\beta\|_1.$$

Here, n is the total number of training windows, $\|\beta\|_1$ and $\|\beta\|_2$ are the L_1 and L_2 norms of β, respectively, and λ is a tuning parameter that can be selected using cross-validation (CV). If $\alpha = 1$, this regression becomes LASSO; if $\alpha = 0$, this becomes Ridge. Both LASSO and Ridge regressions are special cases of Elastic-net regularization that basically mixes LASSO and Ridge using the parameter $\alpha \in [0, 1]$.

Among all the windows, 70% from each category were randomly chosen as the training data, and the remaining 30% were used for testing. Specifically, in our experiment, 232 windows with $Y = 1$ and 242 windows with $Y = 0$ were used for training. We performed 5-fold CV with these training data to tune the parameter λ. Within each CV, 80% of the training data have been used for training purpose and 20% have been used to find the deviance for selecting proper λ. For our case, $\lambda = 0.0055$ gives the minimum deviance during CV, and it is used during the testing process for logistic regression with LASSO (Fig. 2.1). The optimum λ forced 158 of the 201 coefficients of β to be zero [59].

2.4.3 Window-level Classification

Once the coefficient vector β was estimated through the training process, the probability $P(Y = 1)$ for each window in the testing data was then calculated, and the calculated probabilities were thresholded at 0.5, that is, the window was classified as dehydrated if

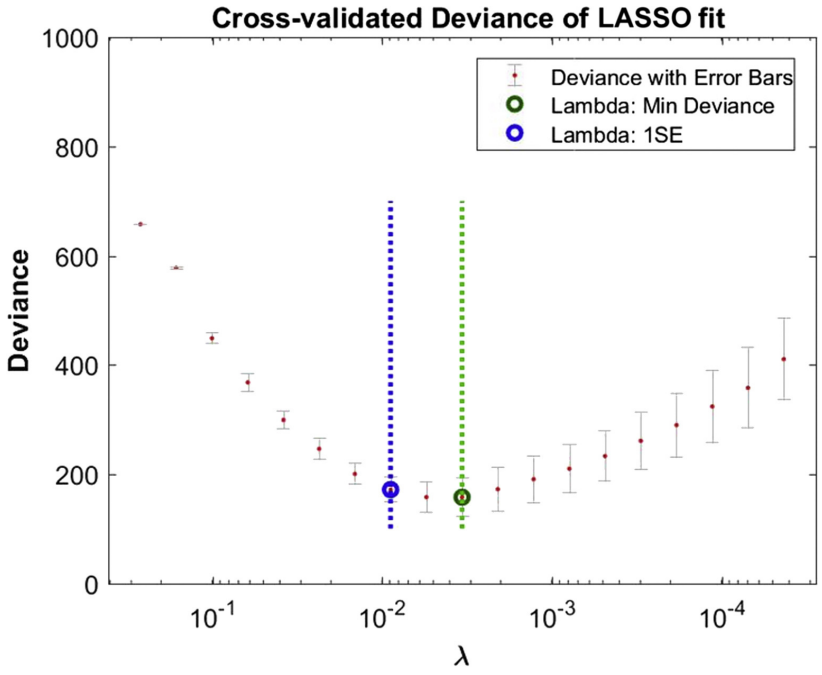

Figure 2.1: Cross-validation for optimizing the tuning parameter λ. The minimum deviance is obtained at $\lambda = 0.0055$.

$P(Y_i = 1) > 0.5$ and hydrated otherwise. We evaluated 101 windows from resuscitated patients and 97 windows from hypovolemic patients. The classification results were compared with their true values for each window.

2.4.4 Patient-level Classification

Patient-level classification was determined as whether a patient is hypovolemic or resuscitated by using a collection of multiple PIV windows from the same patient. The patient level classification was performed with a majority decision rule, which classifies a patient as hypovolemic if this patient has more windows classified as hypovolemic than resuscitated, and vice versa.

2.5 Comparison of Study Groups

The Kolmogorov-Smirnov (KS) two-sample test has been used to test the significance of the difference in frequency domain PVP data from hypovolemic or resuscitated patients

[62]. The KS two-sample test is a nonparametric test that quantifies the difference between the empirical distribution functions of two group of data, which are the frequency domain PVP data from hypovolemic and resuscitated patients, respectively. To study the impact of the signals at different frequencies, the KS twosample test is performed for signals at each frequency separately.

The setup of this hypothesis testing is - H_0: hypovolemic and resuscitated samples have the same distribution at a frequency f Hz; H_1: hypovolemic and resuscitated samples have different distributions at f Hz. The KS two-sample test has been performed over the frequency range from 1.5 Hz to 5.5 Hz.

2.6 Results

2.6.1 Clinical Demographics

A total number of 34 patients were enrolled. Among these, 16 patients were excluded; 5 for incorrect PIV type, 9 for incorrect data collection times, and 2 for incorrect PIV location (scalp or hand).

Eighteen patients were compared in which 89% (16/18) were male and 39% (7/18) the first-born son. There was family history of HPS in 28%. The mean gestational age was 38 4/7 weeks, mean age at study enrollment was 37.2 days, and mean admission weight was 3.93 kg. The mean days of projectile vomiting symptoms was 4.28 days. Average ultrasound measurements for width and length were 4.4 mm and 18.9 mm, respectively. No PIV infiltration occurred during data collection, which was monitored by the authors during each data point collection. No PIVs were lost during system setup.

2.6.2 Classification

Window-level Classification

When the classification was performed on the window level, the logistic regression LASSO algorithm yielded a 97.95% sensitivity and a 93.07% specificity. Other models such

Table 2.1: Summary of results for Elastic-net, LASSO, and Ridge regression.

α	Train. Sen., (%)	Train. Spec., (%)	Test. Sen., (%)	Test Spec., (%)	Non-zero Coeff.
0.0001	94.40	95.87	97.94	93.07	201
0.5	99.57	99.59	96.91	93.07	73
0.75	99.57	100	96.91	92.08	64
1	99.57	99.17	97.95	93.07	43

as Ridge and Elastic-net shows almost similar performances (Table 2.1). The LASSO model is more significant than others because it offers a sparse solution with the coefficients at many frequencies to be 0; thus, it helps us understand the impact of the PVP signals on the dehydration level at different frequencies.

Patient-level Classification

Patient-level classification was determined as whether a patient is hypovolemic or resuscitated by using a collection of multiple PIV windows from the same patient. The patient level classification was performed with a majority decision rule, which classifies a patient as hypovolemic if this patient has more windows classified as hypovolemic than resuscitated, and vice versa.

2.6.3 Peripheral Venous Waveform Signal Analysis

The KS two-sample test was performed over the frequency domain data from all patients to test the difference between the PVP waveforms of the hypovolemic and resuscitated patients at different frequencies. The results of the KS twosample test are depicted in Fig. 2.2, where the logarithm of the p-value is shown as a function of the frequency (Fig. 2.2). The p-value is defined as the probability that the null hypothesis H_0 is true; thus, a small p-value means a bigger difference between the two distributions. It is clear from this figure that the two samples have different distributions (reject H_0). Specifically, the smallest p-values are observed around the f_1 frequency at 2.4 Hz, which corresponds to the heart rate of 144 bps of pediatric patients. The corresponding empirical cumulative distribution

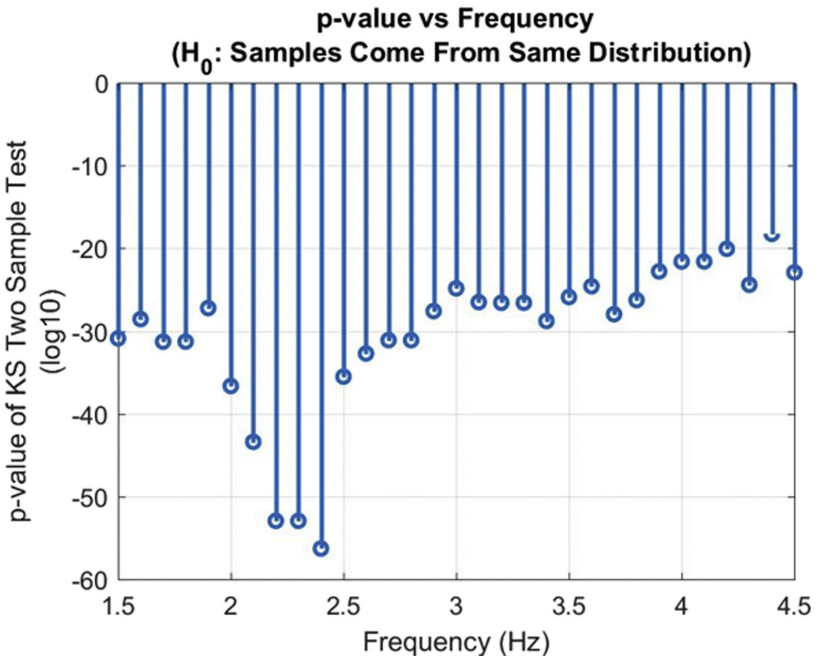

Figure 2.2: p-value of the Kolmogorov-Smirnov two-sample test over frequency range 1.5 to 4.5 Hz.

functions of the two groups of data at 2.4 Hz are shown in Fig. 2.3, which again clearly shows the difference in distributions between the two groups of data (Fig. 2.3).

A similar KS two-sample test with resuscitated patients was performed at the frequencies $f_2 = 4.8$ Hz and $f_3 = 7.2$ Hz, respectively. These two frequencies are the second and third harmonics of f_1. The KS two-sample test was performed between signals collected from resuscitated patients before and after bolus. This KS two-sample test yields a p-value of 0.0053 at f_2, and a p-value of 0.5154 at f_3. For comparison, the p-value obtained from the same test at f_1 is 0. With a significance level of 0.5%, it can be concluded that there is no significant relation between the volume status of resuscitated patients and the signals at f_2 and f_3. On the other hand, strong relation is observed between volume status and the PVP signals at the heart rate frequency f_1.

To further highlight the difference in PVP signals, the admission PVP signals were compared between 10 (56%) children and 8 (46%) hypovolemic patients. The PVP signal power at 2.4 Hz was higher in resuscitated patients (median 0.174 mmHg, IQR: 0.079-0.374

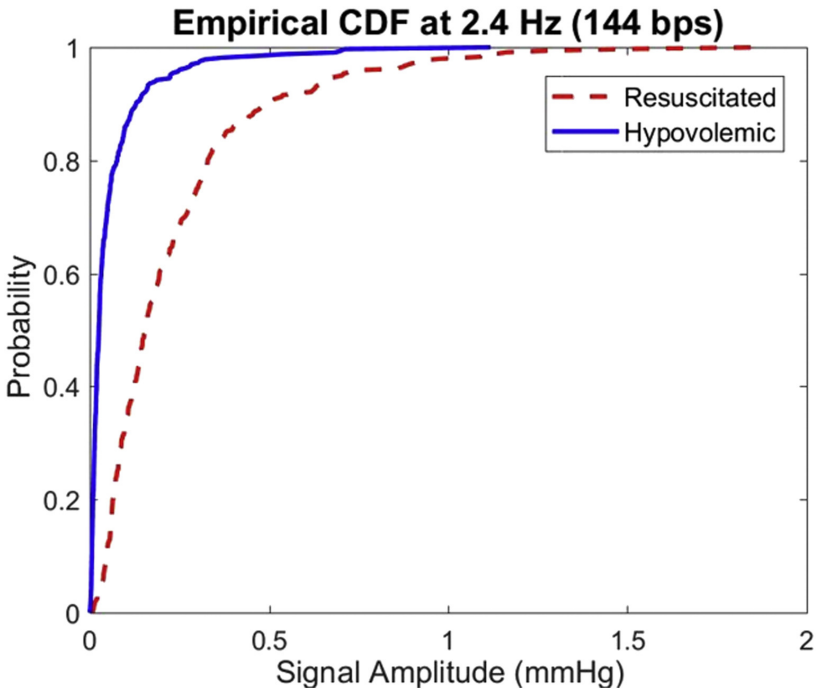

Figure 2.3: The empirical cumulative distribution functions (CDFs) of data from hypovolemic or resuscitated patients at 2.4 Hz.

mmHg) than in hypovolemic patients (median 0.026 mmHg, IQR: 0.001-0.057 mmHg), ($p <$ 0.001).

A total of 14 patients received a bolus regardless of laboratory values: 6 (43%) resuscitated and 8 (57%) hypovolemic. The box plots of the PVP signal power at 2.4 Hz for resuscitated and hypovolemic patients are shown in Figs. 2.4 and 2.5, respectively. In resuscitated patients, the signal power did not change significantly after the fluid bolus (median 0.142 mmHg, IQR: 0.032-0.383 mmHg) ($p = 0.019$), whereas significantly increased signal power (median 0.0474 mmHg, IQR: 0.019-0.110 mmHg) was observed in the hypovolemic patients after a fluid bolus. ($p < 0.001$) (Figs. 2.4 and 2.5).

2.7 Discussion

In this proof of concept study, PVP waveform analysis determined volume status by predicting dehydration. We found that the PVP waveforms have higher amplitudes when

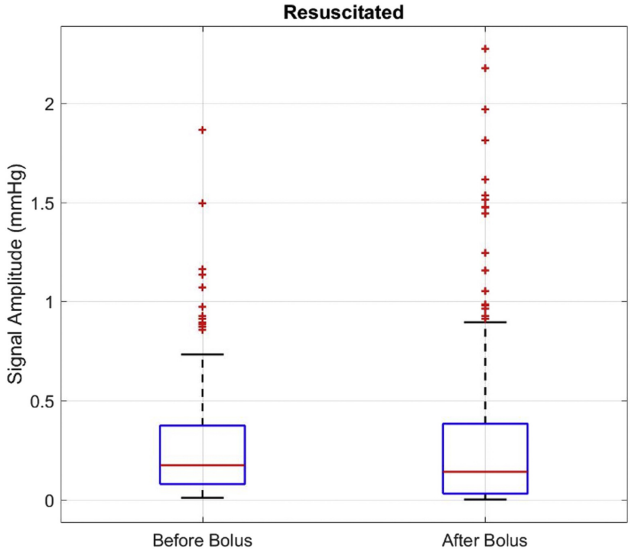

Figure 2.4: Box plot of peripheral venous pressure (PVP) signal amplitude at 2.4 Hz for resuscitated patients before and after bolus. The signal amplitude does not change significantly. The outliers are plotted using "+" symbol in the figure.

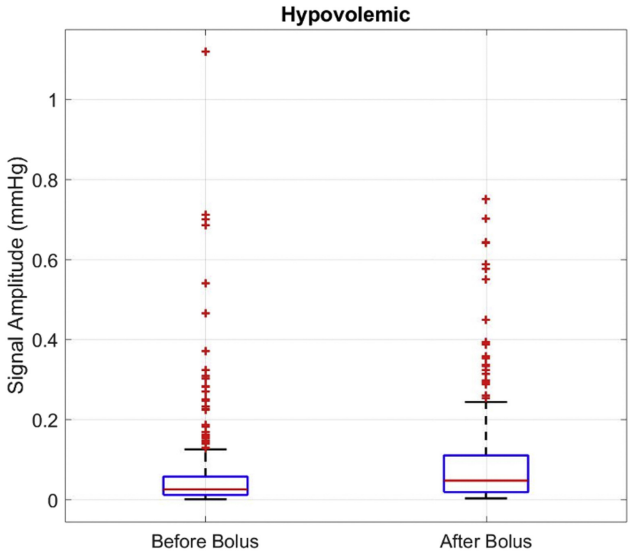

Figure 2.5: Box plot of peripheral venous pressure (PVP) signal amplitude at 2.4 Hz for hypovolemic patients before and after bolus. The signal amplitude changes significantly due to bolus. The outliers are plotted using "+" symbol in the figure.

comparing resuscitated patients with hypovolemic patients, which associated to above and below a chloride value of 100 mmol/L. In addition, we found the PVP signal power changed significantly after IVF bolus in hypovolemic patients. These findings may guide physicians using a noninvasive venous waveform to assess volume status and direct resuscitation in an awake pediatric patient via a PIV.

In the United States alone, dehydration affects 30 million children annually and accounts for 400,000 pediatric emergency room visits [2, 3]. Several existing studies have attempted to define level of dehydration [4, 5, 6, 7, 8]. In children, there is no standard of care for volume status measurement and it is still a diagnostic challenge. Unrecognized dehydration or incorrect estimation of its degree can adversely affect the morbidity and mortality of a dehydrated infant [9, 10]. On presentation, the degree of dehydration was traditionally assessed based on history and physical examination findings alone; however, this practice was called into question as being too subjective [4]. Measurements of body weight, use of clinical scales, physical examination, urine output, and laboratory values have been studied. The difference in body weight between the ill weight and the pre-illness or baseline weight of the patient is used; however, these weights may not always be available to the clinician and variation among different weight scales used can make it challenging to apply clinically [5]. In response, more objective clinical scales were developed and patients were stratified into severity groups [49, 50]. Based on this stratification, many patients are resuscitated with IVF; however, endpoints of resuscitation remain subjective. Some clinicians rely on physical examination findings such as skin turgor and capillary refill. Unfortunately, these are less precise and cannot be monitored continuously. Urine output can be monitored in real time but requires invasive catheter placement into the bladder or weighing of diapers. Many clinicians also use laboratory values including sodium, bicarbonate, lactate, and urine specific gravity to assess adequacy of resuscitation [6, 7, 8]; however, reliance on these values have several drawbacks: invasive venipuncture is required, the dehydrated state can lead to difficulty accessing small caliber veins for blood draws, and specimens can be hemolyzed. In addition, serum electrolytes can vary depending on the disease state, which caused the

dehydration and may be of limited value for detecting the degree of dehydration [6]. In selected cases, electrolyte abnormalities may exist including derangements in sodium levels or acidosis characterized by low bicarbonate levels or elevated lactate levels. Urine specific gravity and presence of ketones can assist in the evaluation of dehydration [7]. End-tidal carbon dioxide measurements have been studied in an attempt to assess degrees of dehydration in children but as of now have not proven to be an effective tool in determining the level of dehydration [8]. A gold standard assessment of dehydration is still lacking [4, 5, 6, 7, 8].

Recent attempts have been made to use existing technology as an objective, noninvasive means of assessing intravascular volume status in real time. Assessment by nearinfrared spectroscopy has been studied using regional oxygen saturation as a surrogate for intravascular volume status [10]. However, there are no accepted normal standards for children of varying ages and body mass indices. Use of point-of-care bedside ultrasound has shown the most promise. Although assessment of inferior vena cava (IVC) collapsibility during inspiration has been less reliable, measurement of the IVC to aortic ratio (IVC/Ao) using this modality allows the user to more objectively assess adequacy of resuscitation [7, 63]. In a previous study of HPS patients, [64] found that as hydration status improved, IVC/Ao ratio approached 1. Specifically, a ratio of 0.75 correlated to a serum bicarbonate level of 30 mmol/L, an accepted endpoint of resuscitation for this disease process [63]. This ratio is similar to values found by other groups when ultrasound was used for assessment of dehydration from acute gastroenteritis [63, 65, 64, 66]. The search for a rapid, noninvasive, and accurate method for assessing the degree of hydration in infants is mandatory especially in countries with limited resources [66]. Analysis of PVP waveforms could represent a modality to assess volume status via noninvasive means.

The venous system is a highly compliant system that can accommodate large changes in volume with minimal changes in pressure. Compensatory venous vasoconstriction to volume depletion diverts blood from the periphery to the central vasculature to maintain cardiac output and ultimately endorgan perfusion [67]. This may explain the insensitivity of arterial blood pressure, pulse rate, and PVP to initial volume depletion. PVP waveform

analysis of the venous system was not rigorously examined until appropriate sensing and amplifying technologies became available [67, 68]. PVP is measured after PIV catheter insertion, the most common procedure performed in the United States [69]. Detection of these small changes in PVP became possible through analysis of FFTs, which are highly sensitive to changes in venous compliances due to volume [70, 35]. PVP waveform analysis of compensatory venous changes represents a significant paradigm shift from dynamic arterial-based measurements [39]. Venous waves are generated by the cardiac cycle and propagated as harmonics [70]. The f_1 waveform that correlates with the heart rate is affected by mild hypovolemia [35]. The authors in [35] proposed that f_1 was associated with a decreased venous blood volume so that the venous waveform is no longer transmitted back from the right atrium.

For hydrated patients, a strong correlation was found at the frequency near the heart rate of the patients (f_1). Our analysis proposes an additional mechanism of hydromechanical interaction to explain the change in f_1 that correlates the heart rate with hypovolemia. In the hydrated patient, the signal from arterial pulse pressure waveform crossover to the venous side of the systemic circulation becomes stronger (Fig. 2.6). The hydromechanical interaction between the arterial and the venous side of the circulation depends on the diameter of the veins and arteries, which in the dehydrated state is lower than in the hydrated state. Hence, in the hydrated state, the direct physical interaction or at least the increase in proximity between the veins and arteries is responsible for the crossover of the pulse pressure waveform. The most demonstrative frequency was 2.4 Hz, equaling 144 beats per minute, which is a very realistic heart rate for dehydrated pediatric patients. This hydromechanical crossover interaction may be impacting the change in f_1 in a combination with the theory proposed by [35] described previously.

There were study limitations. The study was of a limited number of patients with HPS who were categorized as resuscitated or hypovolemic based on laboratory values. The laboratory values provided an objective means of comparison; however, no absolute values of volume status could be determined. Chloride has not been a validated marker of dehydration

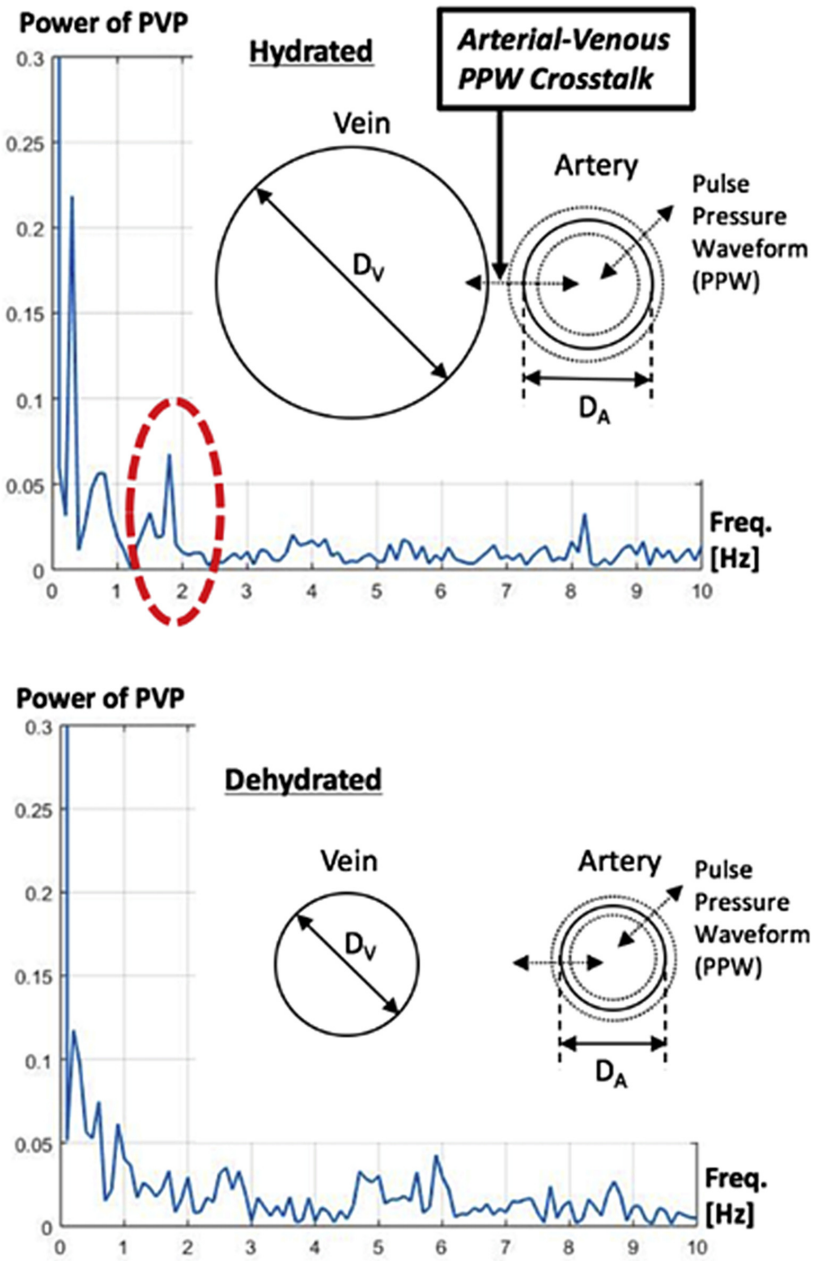

Figure 2.6: The power spectral density (PSD) of peripheral venous pressure (PVP) for the hydrated patient (top) identifies a peak at frequencies around the heart rate (red dotted line). In this example, the peak is shown at approximately 1.8 Hz = 108 bpm. In the same patient during dehydration (bottom), this phenomenon does not exist. The vein diameter (D_V) is significantly larger in the hydrated state. The arterial diameter (D_A) changes slightly between hydration and dehydration. Combined, this causes the hydro-mechanical interaction of pressure signals between the arterial and the venous side of the circulation to be stronger when the patient is hydrated.

but is an accepted endpoint of resuscitation in HPS, and its use in more critically ill children with other causes of dehydration is unknown [56, 58]. The study only assessed a change in the PVP waveform signal from baseline to after fluid bolus. This limitation is documented in other PVP waveform analysis articles; however, finding a strong correlation between fluid bolus given and change in the PVP waveform signal is evident [39]. Another limitation of the study was waveform variation based on type of PIV, PIV location, clots or air in arterial tubing, and inability to collect proper waveforms [41]. One physician was responsible for accurate waveform collection, and those patients were excluded from the study; however, some variation may still exist.

HPS was chosen as the model because laboratory values have been used as a predictive marker for the level of dehydration and subsequent volume needed for resuscitation. All HPS patients are hypovolemic and received some level of IVF resuscitation. This allows us to compare the dehydrated state with a more hydrated state within an individual and determine if there are changes in the PVP signal that reflect the amount of rehydration. Future models considered are dehydrated patients secondary to diabetic ketoacidosis or severe gastroenteritis. Our fundamental hypothesis is that the absolute pressure is not as relevant as the signal from venous augmentation during circulation. Static measures of fluid responsiveness such as central venous pressure (CVP) may not be the most appropriate and may be less accurate physiologically than dynamic measures [71, 72]. No age-specific CVP norms have been defined, and the comparison of CVP to FFT of the PVP waveforms has not been performed in the pediatric population. Moreover, CVP may be less accurate in the infant population specifically those that require central venous access for resuscitation, that is, patients with complex congenital heart disease or chronic lung disease. Those patients with abnormal cardiopulmonary systems would introduce more variability into the experimental model.

This clinical model represented a homogenous group of awake patients without comorbidities and were dehydrated. Results from this study shows that PVP waveforms can be performed and predict dehydration. The ideal modality to determine dehydration in the

pediatric age group would be reproducible, noninvasive, user-independent, and provide a real-time assessment of volume status which this study supports. PVP waveform analysis has the potential to monitor volume status in other clinical conditions such as perioperative goal-directed fluid therapy, congestive heart failure optimization, acute kidney injury, trauma, and hemorrhage [39]. This description and utilization of PVP waveforms in pediatric patients could significantly change the evaluation and management of volume status in children using a future technology that can adjust the rate and volume of fluid administration based on continuous PVP waveform analysis. This novel approach to a more objective assessment of dehydration in the pediatric population serves as an interesting pilot study for a hypothesis generating data in future studies and more general populations.

2.8 Conclusion

FFT of PVP waveforms is a novel modality that can assess volume status and predict dehydration in HPS. PVP waveform analysis can predict dehydration with a 97.75% sensitivity and a 93.0% specificity. We showed that fluid bolus administration in an awake pediatric patient can be performed possibly offering a significant advantage over dynamic monitoring modalities. Further work is needed to determine the utility of PVP analysis to guide fluid resuscitation status in children, which could profoundly impact treatment of dehydration in children.

3 Unsupervised Anomaly Detection in Peripheral Venous Pressure Signals with Hidden Markov Models

3.1 Abstract

This chapter proposes an automatic anomaly detection and removal algorithm for peripheral venous pressure (PVP) signals, which can be used to predict intravascular volume loss in humans. PVP signal collection is a minimally invasive procedure that can be performed by using a standard peripheral intravenous (PIV) catheter and a commercial pressure-monitoring transducer. PVP signals are highly susceptible to motion and noise artifacts such as patient movements or unintended manipulation of PIV lines. Anomalies in PVP signals can corrupt useful information and seriously affect the integrity of PVP signal analysis. We propose to detect and remove such anomalies by exploiting the properties of PVP signals. Specifically, a dynamic linear model (DLM) with a Kalman filter is used to track and predict the time-domain evolution of PVP signals. The prediction residuals of the Kalman filter are then modeled with a hidden Markov model (HMM), with the normal and anomalous status of the signal modeled by using binary states of a hidden Markov chain. The HMM parameters along with the hidden states are iteratively estimated by using an unsupervised learning algorithm with a modified Baum-Welch method. The anomaly detection algorithm is applied to clinical data from a cohort of 24 pediatric patients with hypertrophic pyloric stenosis. Experimental results demonstrate that the proposed unsupervised anomaly detection algorithm can efficiently remove anomalies in PVP signals without the need of a training phase. The algorithm can also be applied to other time series signals, such as Electrocardiography (ECG) and Photoplethysmogram (PPG) signals.

3.2 Introduction

Dehydration or loss of intravascular blood volume is a common and potentially life-threating condition. Dehydration affects 30 million children annually and accounts for 400,000 pediatric emergency room visits in the United States [2, 4]. Often severe dehy-

dration (hypovolemia) or occult bleeding is not recognized by monitoring vital signs until the stage where end-organ damage occurs [73, 74]. However, there is no standardized measurement for intravascular volume in adults or children. This necessitates the development of technologies that would accurately assess the volume status of a patient to guide resuscitation and treatment. Analysis of peripheral venous pressure (PVP) waveforms is a novel method of monitoring intravascular volume and may provide valuable information in earlier detection of volume depletion [75, 76]. Previous studies have demonstrated that PVP signals strongly correlate with central venous pressure under different conditions [20, 22, 19, 25, 24]. The shapes of PVP waveforms could be affected by a number of physiological phenomena, such as respiration, heart rate (HR), and/or systolic blood pressure (SBP). With porcine models, it has been shown in [75] that PVP signals can be a good indicator of intravascular volume depletion due to hemorrhage. Similarly, the HR frequency components of PVP signals have been used to quantify blood loss with porcine models in [38].

Pilot studies have shown evidence that PVP signal can be used for early hemorrhage detection in humans [70]. It has been demonstrated in previous studies that PVP signals can be used to predict dehydration in pediatric patients [76, 77, 78]. PVP signal collection is a minimally invasive technology consisting of a standard peripheral intravenous (PIV) catheter and a commercial pressure-monitoring transducer. However, PVP signals are highly susceptible to non-stationary noise, such as random patient movements or unintended physical manipulation of the PIV catheter during the measurement process. The PVP signals will be corrupted under such events, which results in anomalies in the collected data. The presence of anomalies will seriously undermine the integrity of the analysis results. Thus it is necessary to detect and remove such anomalies before any further processing.

Anomaly detection has a wide range of applications, and it is critical for biomedical signal processing [79]. One of the main sources of anomalies in time series biomedical signals are motion and noise artifacts (MNA), which impose serious challenges in the accurate collection, modeling, and analysis of signals such as PVP [77], Electrocardiography (ECG) [80], and Photoplethysmography (PPG) [81]. Anomalies in PVP signals are detected and

removed manually through visual inspections by clinical staffs in [76] and [78], where sections of PVP signals with large variations are removed before further processing. A simple automatic anomaly detection method is adopted in [77], where signals with excessively large variances are labeled as anomalies. An automatic MNA detection method for ECG signals is developed in [80] by using empirical mode decomposition, and the cleaned signals are used to detect diseases like atrial fibrillation (AF). An adaptive step-size least mean squares (AS-LMS) filter is used to reduce motion artifacts in PPG signals in [81]. Anomaly detections in time series can also be performed by using deep learning algorithms, such as recurrent neural networks with long short-term memory (LSTM) [82, 83].

The objective of this chapter is to develop an efficient and unsupervised anomaly detection algorithm by exploiting the features of PVP signals. To model the correlation of consecutive samples of the PVP signal, we propose to use a dynamic linear model (DLM) [84] with Kalman filter to track and predict the evolution of the PVP signal in the time domain. It is expected that the model residual, which is the difference between the actual signal and the one predicted by the Kalman filter, is a stationary random process with small variances for normal PVP signals. On the other hand, since anomalies or bad data violate the normal dynamics of PVP signals, the DLM residual will be large for anomalies. Motivated by this fact, we propose to detect and remove PVP anomalies by analyzing the differences in statistical behaviors of Kalman filter residuals.

Specifically, the Kalman filter prediction residuals are modeled by using a hidden Markov model (HMM). HMM is a powerful statistical tool that has been widely used in many different disciplines for various applications [85]. For example, it has been used for protein structure prediction [86], genome sequencing [87], speech processing [85, 88], network intrusion detection [89], traffic sensing [90], etc. In this chapter, the status of the PVP data is modeled by using a Markov chain with two hidden states: normal or anomalous states. The distributions of the prediction residuals depend on the two hidden states of the PVP signals. The unknown parameters of the statistical model are iteratively learned by using a modified Baum-Welch algorithm [91]. The original Baum-Welch algorithm [91] suffers from

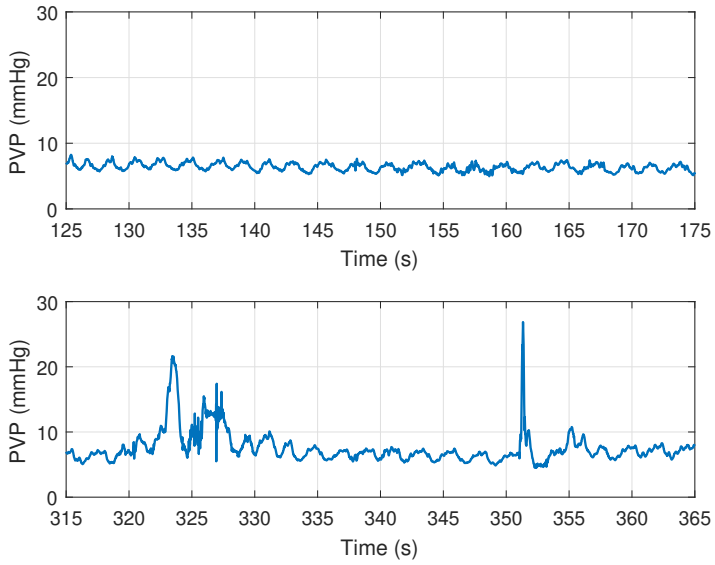

Figure 3.1: Exemplary PVP signal (a) without anomaly and (b) with anomaly.

mathematical underflow due to limited computer precision. A variable scaling approach [85] is adopted in this chapter to ensure the numerical stability of the iterative algorithm. In addition, we provide detailed explanations on physical meanings of the scaled probability variables which to our understanding are not available in the literature. The results of the modified Baum-Welch algorithm are used to detect and remove anomalies in PVP signals. The algorithm has been applied to clinical data from a cohort of 24 pediatric patients. Experiment results demonstrate that the algorithm can efficiently remove anomalies in PVP signals without a training phase. The proposed anomaly detection algorithm can also be applied to other time-series biomedical signals, such as ECG and PPG signals.

The remainder of this chapter is organized as follows. The data acquisition process is described in Section 2. Section 3 presents the representation of PVP signals with the DLM and HMM. The anomaly detection is performed with a modified Baum-Welch algorithm, which is also described. Experimental results are presented in Section 4, and Section 5 concludes the chapter.

3.3 Methods

PVP waveform data was collected in a closed system via a 22-gauge Insyte-N Autoguard PIV catheter (Becton Dickinson Infusion Therapy Systems, Sandy, Utah, USA) connected to a 48-inch arterial pressuring tubing (Smiths Medical, Dublin, OH, USA). The arterial pressure tubing was connected to a Deltran II pressure transducer (ADInstruments, Colorado Springs, CO, USA) interfaced with a PowerLab data acquisition system (ADInstruments). A similar data collection procedure was used in [77] and [41]. Primary data manipulation and visualization have been done using LabChart v8 (ADInstruments) software of PowerLab. Acquired data was saved in .adicht and .mat formats.

Patients were defined as euvolemic or hydrated when the concentration of serum chloride was $\geq$ 100 mmol/L or bicarbonate $<$ 30 mmol/L and hypovolemic or dehydrated when serum chloride $<$ 100 mmol/L or bicarbonate $\geq$ 30 mmol/L. Data were collected on emergency department admission of patients suffering from Hypertropic pyloric stenosis.

The study was approved by the University of Arkansas for Medical Sciences (UAMS) Institutional Review Board (IRB) (Protocols 206193 & 207085). All procedures performed on the study participants were in accordance with the ethical standards of the IRB and with the 1964 Helsinki declaration and its later amendments. Informed consent was obtained from the legal guardians of the children participating in the study.

3.4 Theory

In this section, the PVP signals are represented and modeled by using dynamic linear models (DLM) [84]. The results will be used to detect and remove anomalies in PVP signals. The PVP signals are collected by using PIV catheters from a peripheral part of the body [77]. As a result, the signal is highly susceptible to non-stationary noise caused by events such as the random movement of the patients or an unintended touch of the catheter during the measurement process. The PVP signals will be corrupted under such events, which result in anomalies in the collected data. Fig. 3.1 shows a normal PVP signal and a corrupted PVP

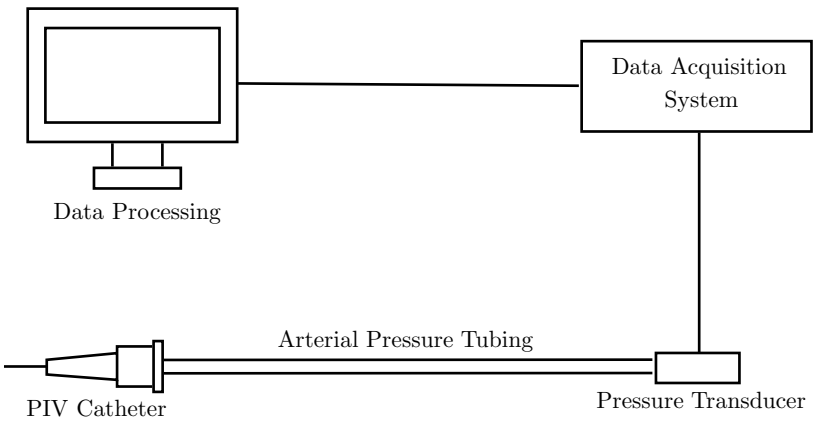

Figure 3.2: Schematic diagram of the data acquisition system. Peripheral intravenous (PIV) catheter is inserted into peripheral vein of a patient. Arterial pressure tubing connects the PIV catheter to pressure transducer that converts pressure to equivalent electrical signal. This signal is acquired and processed afterwards.

signal. It is apparent that the statistical behaviors of the PVP signals with anomalies are considerably different from its normal counterpart. The presence of anomalies will seriously undermine the integrity of the analysis results. Thus it is necessary to detect and remove such anomalies before any further processing.

We propose to capture the dynamic behaviors of PVP signals by using a DLM [84], which is a special case of the Gaussian state space model and commonly used for time series modeling and forecast. The DLM can be used in combination with the Kalman filter to track and predict the dynamic evolution of the PVP signal in the time domain. Under normal conditions, it is expected that Kalman filter with the DLM can accurately predict normal PVP signal values, and this will result in very small prediction residuals (prediction errors). On the other hand, abnormal signals will lead to large prediction residuals due to the unpredicted nature of abnormal events. Thus anomalies can be detected by analyzing the statistical properties of prediction residuals of Kalman filter.

3.4.1 Dynamic Linear Model with Kalman Filter

The discrete-time PVP signal can be modeled as a time series y_t, for $t = 1, 2, \cdots, T$. Define $\mathbf{y}_{1:T} = [y_1, y_2, \cdots, y_T]$ to simplify notation. In this DLM, it is assumed that each

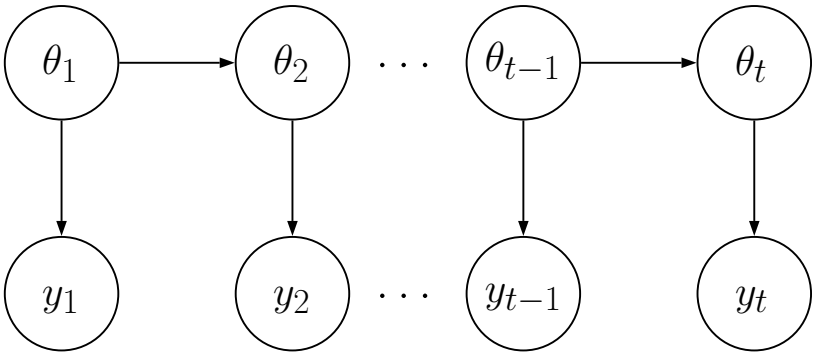

Figure 3.3: Dependence structure of dynamic linear model. Here, θ_i's are forming a first-order Markov chain. Also, θ_{i+1} and y_i follow a Gaussian distribution depending on θ_i under a linear relationship. $\{\theta_i\}$ and $\{y_i\}$ are continuous random variables.

sample of the PVP signal, y_t, is associated with a time-varying state, θ_t, as shown in Fig. 3.3. The DLM can be represented as [84]

$$y_t = F\theta_t + v_t; \tag{3.1a}$$

$$\theta_t = G\theta_{t-1} + w_t. \tag{3.1b}$$

where (3.1a) is the observation model describing the relationship between state variable θ_t and observation variable y_t, with v_t being the observation noise, and (3.1b) is the state transition model with w_t being model uncertainties. Here, the coeffcients F and G are constants for a given patient, and it is assumed that both v_t and w_t are independent zero-mean Gaussian distributed with variances σ_v^2 and σ_w^2, respectively, that is $v_t \sim \mathcal{N}(0, \sigma_v^2)$, $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $v_t \perp w_t$.

With the DLM given in (3.1), we can apply the Kalman filter to iteratively predict the values of the PVP signal y_t and the state variable θ_t by using all previous observations $\mathbf{y}_{1:t-1}$. In the implementation of the Kalman filter, denote $\hat{\theta}_{t|\tau}$ as the estimation of θ_t by using $\mathbf{y}_{1:\tau}$ with $\tau = t-1$ or t, $R_{t|\tau} = \mathbb{E}[|\hat{\theta}_{t|\tau} - \theta_{t|\tau}|^2]$ as the corresponding error variance; denote $\hat{y}_{t|\tau}$ as the estimation of y_t by using $\mathbf{y}_{1:\tau}$ with $\tau = t-1$, with error variance $Q_{t|\tau} = \mathbb{E}[|\hat{y}_{t|\tau} - y|^2]$. The values of $\hat{\theta}_{t|\tau}$, $R_{t|\tau}$, $\hat{y}_{t|\tau}$, and $Q_{t|\tau}$ can be iteratively updated by using the Kalman filter,

and details are given in Algorithm 1.

The DLM is completely represented by the initial parameter set, $\{F, G, \sigma_v^2, \sigma_w^2, \hat{\theta}_{0|0}, R_{0|0}\}$. The values of these parameters for a given patient can be estimated by applying maximum likelihood estimation on the data of that patient. The parameters are estimated by using the maximum likelihood estimation algorithm L-BFGS-B [92]. Once the parameters are estimated, Kalman filter can then be applied to track and estimate the dynamic evolution of the data in the time domain.

Algorithm 1 : Kalman Filter

1: **Input:** Discrete time dataset $\mathbf{y}_{1:T}$;
2: Estimate $\{F, G, \sigma_v^2, \sigma_w^2\, \hat{\theta}_{0|0}, R_{0|0}\}$ by applying the L-BFGS-B algorithm on $\mathbf{y}_{1:T}$; initialize $t = 0$.
3: **do**
4: $t \leftarrow t + 1$;
5: Prediction of θ_t:

$$\hat{\theta}_{t|t-1} = G\hat{\theta}_{t-1|t-1} \tag{3.2}$$
$$R_{t|t-1} = G^2 R_{t-1|t-1} + \sigma_w^2 \tag{3.3}$$

6: Prediction of y_t:

$$\hat{y}_{t|t-1} = F\hat{\theta}_{t|t-1} \tag{3.4}$$
$$Q_{t|t-1} = F^2 R_{t|t-1} + \sigma_v^2 \tag{3.5}$$

7: Calculate prediction residual: $x_t = y_t - \hat{y}_{t|t-1}$;
8: Update estimation of θ_t:

$$\hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + FR_{t|t-1}Q_{t|t-1}^{-1}x_t \tag{3.6}$$
$$R_{t|t} = R_{t|t-1} - F^2 R_{t|t-1}^2 Q_{t|t-1}^{-1} \tag{3.7}$$

9: **while** $t \leq T$
10: **Output:** Prediction residual $\mathbf{x}_{1:T}$.

The Kalman filter described in Algorithm 1 iteratively tracks and predicts the state and observation variables using the DLM. The output of the Kalman filter is the prediction residual, $x_t = y_t - \hat{y}_{t|t-1}$, which is the difference between the predicted and actual value of y_t. It is expected that the prediction residuals under normal and abnormal conditions will exhibit different statistical properties, thus they can be used for anomaly detection.

3.4.2 Modeling PVP Prediction Residuals with HMM

The prediction residuals at the output of the Kalman filter are modeled by using HMM. Each residual x_t is associated with a binary hidden state $s_t \in \{0, 1\}$, where $s_t = 0$ indicates normal data, and $s_t = 1$ indicates anomalies. The HMM is illustrated in Fig. 3.4. Under the HMM, the residuals are conditionally independent with each other conditioned on their respective hidden states. The dependence of the residuals are represented through the hidden states. To simplify notation, define $\mathbf{x}_{1:T} = [x_1, x_2, \cdots, x_T]$.

The hidden states form a first-order Markov chain with transition probability, $a_{ij} = \Pr(s_{t+1} = j | s_t = i)$, for $i, j \in \{0, 1\}$. Define the transition probability matrix as

$$\mathbf{A} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \tag{3.8}$$

The initial distribution of the two states in the Markov chain is assumed to be $\pi_k = \Pr(s_1 = k)$ for $k \in \{0, 1\}$. Define the initial probability vector as $\boldsymbol{\pi} = [\pi_0, \pi_1]^T$, where the operator $(\cdot)^T$ represents matrix transpose.

The statistical distribution of each prediction residual depends on its hidden state. For state $s_t = k$, it is assumed that the prediction residuals follows a Gaussian distribution with mean μ_k and variance σ_k^2, that is,

$$x_t | (s_t = k) \sim \mathcal{N}(\mu_k, \sigma_k^2), \quad \text{for} \quad k \in \{0, 1\} \tag{3.9}$$

It is assumed that $\sigma_1^2 > \sigma_0^2$ due to the presence of noise or interference in the corrupted signals. Define the sample distribution parameter set as $\mathcal{B} = \{\mu_0, \mu_1, \sigma_0^2, \sigma_1^2\}$.

Based on the above definitions, the HMM can be represented by using the parameter set $\boldsymbol{\lambda} = \{\mathbf{A}, \boldsymbol{\pi}, \mathcal{B}\}$. The HMM is assumed to be time-homogeneous, that is, the model parameters do not change over time.

The objective of anomaly detection is to identify the state $\{s_t\}$, by using the pre-

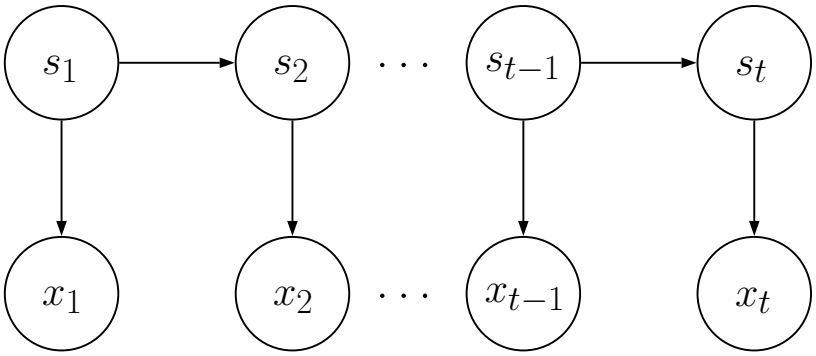

Figure 3.4: Dependence structure of first-order hidden Markov model. Unlike Fig. 3.3, here $\{s_i\}$ are discrete random variables and x_i follows a Gaussian distribution depending on s_i.

diction residuals $\{x_t\}$. The identification of the state variables require the knowledge of the parameter set $\boldsymbol{\lambda}$, which can be directly learned from $\{x_t\}$ by using the Baum-Welch algorithm.

Define the posterior probability of the hidden state variable as

$$\gamma_t(i) = \Pr(s_t = i|\mathbf{x}_{1:T}, \boldsymbol{\lambda}). \tag{3.10}$$

If the HMM parameter set $\boldsymbol{\lambda}$ is known, then the anomaly detection algorithm can be formulated as

$$\hat{s}_t = \underset{i\in\{0,1\}}{\operatorname{argmax}} \quad \gamma_t(i). \tag{3.11}$$

The parameter set $\boldsymbol{\lambda}$ is unknown, and all parameters can be estimated from prediction residual $\{x_t\}$ by using the Baum-Welch algorithm.

3.4.3 Unsupervised Learning with a Modified Baum-Welch Algorithm

In this section, unsupervised learning is performed to detect the anomalies in PVP signals with the help of a modified Baum-Welch algorithm [91]. The Baum-Welch algorithm is an expectation-maximization algorithm [93] that can iteratively learn the values of the

unknown parameter set, $\boldsymbol{\lambda}$, by maximizing the log-likelihood function

$$\ell = \log p(\mathbf{x}_{1:T}|\boldsymbol{\lambda}) \tag{3.12}$$

The unsupervised learning process is carried out by using the observations, $\mathbf{x}_{1:T}$, without the need of a training phase.

The implementation of the Baum-Welch algorithm requires the iterative calculation of a forward probability, $\alpha_t(i)$, and a backward probability, $\beta_t(i)$, which are defined, respectively, as

$$\alpha_t(i) = \begin{cases} \pi_i p(x_1|s_1 = i, \boldsymbol{\lambda}), & t = 1, \\ \Pr(\mathbf{x}_{1:t}, s_t = i|\boldsymbol{\lambda}), & 2 \leq t \leq T \end{cases}$$

$$\beta_t(i) = \begin{cases} p(\mathbf{x}_{(t+1):T}|s_t = i, \boldsymbol{\lambda}), & 1 \leq t \leq T-1, \\ 1, & t = T. \end{cases}$$

Based on the above definitions and the Markov property of the hidden states, the forward and backward probabilities can be iteratively updated as

$$\alpha_t(i) = \sum_{j \in \{0,1\}} \alpha_{t-1}(j) a_{ji} p(x_t|s_t = i, \boldsymbol{\lambda}), \quad 2 \leq t \leq T$$

$$\beta_t(i) = \sum_{j \in \{0,1\}} a_{ij} p(x_{t+1}|s_{t+1} = j, \boldsymbol{\lambda}) \beta_{t+1}(j),$$

$$\text{for} \quad 1 \leq t \leq T-1.$$

As the number of samples, T, increases, the values of $\alpha_t(i)$ and $\beta_t(i)$ get very small and technically become zero due to limited computer precision. To solve the problem of mathematical underflow, the 'log-sum-exponent' trick is used in many cases. However, the 'log-sum-exponent' trick is an approximate solution and it introduces additional computation time. To achieve mathematically accurate and robust estimation, we adopt the variable scaling approach as proposed in [85]. The notation is ambiguous while defining the scaled

probabilities in [85]. This chapter tries to give the physical meaning and better understanding of the scaled variables.

3.4.4 Modified Baum-Welch Algorithm

Define a modified forward probability variable, $\tilde{\alpha}_t(i)$, as

$$\tilde{\alpha}_t(i) = \begin{cases} \Pr(x_1, s_1 = i|\boldsymbol{\lambda}), & t = 1 \\ \Pr(x_t, s_t = i|\mathbf{x}_{1:t-1}, \boldsymbol{\lambda}), & t > 1 \end{cases} \tag{3.13}$$

The variable $\tilde{\alpha}_t(i)$ can be interpreted as a scaled version of $\alpha_t(i)$ as

$$\tilde{\alpha}_t(i) = C_{t-1}\alpha_t(i) \tag{3.14}$$

where the scaling variable C_t is defined as follows

$$C_t^{-1} = \begin{cases} 1, & t = 0, \\ p(\mathbf{x}_{1:t}|\boldsymbol{\lambda}) = \alpha_t(0) + \alpha_t(1), & t \geq 1 \end{cases}$$

Lemma 1 *The scaled forward probability variable, $\tilde{\alpha}_t(i)$, can be calculated in an iterative manner as*

$$\tilde{\alpha}_t(i) = c_{t-1} \sum_{j \in \{0,1\}} \tilde{\alpha}_{t-1}(j) \cdot a_{ji} \cdot p(x_t|s_t = i, \lambda), \tag{3.15}$$

for $2 \leq t \leq T$ and $\tilde{\alpha}_1(i) = \alpha_1(i)$, and c_t is defined as

$$c_t^{-1} = p(x_t|\boldsymbol{x}_{1:t-1}, \boldsymbol{\lambda}) = \tilde{\alpha}_t(0) + \tilde{\alpha}_t(1) \tag{3.16}$$

Proof:

$$\begin{aligned}\tilde{\alpha}(i) &= \Pr(x_t, s_t = i|\mathbf{x}_{1:t-1}, \boldsymbol{\lambda}) \\ &= \frac{\Pr(\mathbf{x}_{1:t}, s_t = i|\boldsymbol{\lambda})}{p(\mathbf{x}_{1:t-1}|\boldsymbol{\lambda})} \\ &= \frac{\Pr(\mathbf{x}_{1:t}, s_t = i|\boldsymbol{\lambda})}{p(x_{t-1}|\mathbf{x}_{1:t-2}, \boldsymbol{\lambda}) \cdot p(\mathbf{x}_{1:t-2}|\boldsymbol{\lambda})},\end{aligned} \tag{3.17}$$

The numerator in the above expression can be calculated as

$$\begin{aligned}&\Pr(\mathbf{x}_{1:t}, s_t = i|\boldsymbol{\lambda}) \\ &= \sum_{j\in\{0,1\}} \Pr(\mathbf{x}_{1:t}, s_{t-1} = j, s_t = i|\boldsymbol{\lambda}) \\ &= \sum_{j\in\{0,1\}} \Pr(\mathbf{x}_{1:t-1}, s_{t-1} = j|\boldsymbol{\lambda}) \cdot a_{ji} \cdot p(x_t|s_t = i, \boldsymbol{\lambda}) \\ &= p(\mathbf{x}_{1:t-2}|\boldsymbol{\lambda}) \cdot p(x_t|s_t = i, \boldsymbol{\lambda}) \\ &\times \sum_{j\in\{0,1\}} \left[\Pr(x_{t-1}, s_{t-1} = j|\mathbf{x}_{1:t-2}, \boldsymbol{\lambda}) \cdot a_{ji}\right]\end{aligned} \tag{3.18}$$

Substituting (3.17) into (3.18) results in

$$\tilde{\alpha}(i) = c_{t-1} \sum_{j\in\{0,1\}} \tilde{\alpha}_{t-1}(j) \cdot a_{ji} \cdot p(x_t|s_t = i, \boldsymbol{\lambda}).$$

This completes the proof. ■

Similarly, define a scaled backward probability variable, $\tilde{\beta}_t(i)$, as

$$\tilde{\beta}_t(i) = \begin{cases} \frac{\Pr(\mathbf{x}_{t+1:T}|s_t=i,\boldsymbol{\lambda})}{\Pr(\mathbf{x}_{t+1:T}|\mathbf{x}_{1:t}=i,\boldsymbol{\lambda})}, & 1 \le t \le T-1 \\ 1, & t = 1. \end{cases} \tag{3.19}$$

Also,

$$\tilde{\beta}_t(i) = D_{t+1}\beta_t(i), \tag{3.20}$$

where the scaling variable D_t is defined as

$$D_t^{-1} = \begin{cases} \Pr(\mathbf{x}_{t:T}|\mathbf{x}_{1:t-1}, \boldsymbol{\lambda}), & 1 \leq t \leq T, \\ 1, & t = T+1. \end{cases} \tag{3.21}$$

Lemma 2 *The scaled backward probability variable, $\tilde{\beta}_t(i)$, can be calculated in an iterative manner as*

$$\tilde{\beta}_t(i) = c_{t+1} \sum_{j \in \{0,1\}} a_{ij} p(x_{t+1}|s_{t+1} = j, \lambda) \tilde{\beta}_{t+1}(j) \tag{3.22}$$

where $1 \leq t \leq T-1$ and $\tilde{\beta}_T(i) = \beta_T(i) = 1$.

Proof:

$$\begin{aligned} \tilde{\beta}_t(i) &= \frac{\Pr(\mathbf{x}_{t+1:T}|s_t = i, \boldsymbol{\lambda})}{p(\mathbf{x}_{t+1:T}|\mathbf{x}_{1:t}, \boldsymbol{\lambda})} \\ &= \frac{\Pr(\mathbf{x}_{t+1:T}|s_t = i, \boldsymbol{\lambda})}{p(x_{t+1}|\mathbf{x}_{1:t}, \boldsymbol{\lambda}) \cdot p(\mathbf{x}_{t+2:T}|\mathbf{x}_{1:t+1}, \boldsymbol{\lambda})} \\ &= c_{t+1} \sum_{j \in \{0,1\}} a_{ij} p(x_{t+1}|s_{t+1} = j, \boldsymbol{\lambda}) \tilde{\beta}_{t+1}(j). \end{aligned}$$

The last equality in the above equation is based on the fact that

$$\begin{aligned} &\Pr(\mathbf{x}_{t+1:T}|s_t = i, \boldsymbol{\lambda}) \\ &= \sum_{j \in \{0,1\}} a_{ij} p(x_{t+1}|s_{t+1} = j, \boldsymbol{\lambda}) \Pr(\mathbf{x}_{t+2:T}|s_{t+1} = j, \boldsymbol{\lambda}). \end{aligned}$$

This completes the proof. ■

Based on the definitions of C_t, D_t, and c_t, we have the following relationship among

them

$$C_t = c_1 c_2 \cdots c_t = \prod_{i=1}^{t} c_i = c_t C_{t-1}$$
$$D_t = c_t c_{t+1} \cdots c_T = \prod_{i=t}^{T} c_i = c_t D_{t+1}$$

The likelihood function can then be calculated as

$$p(\mathbf{x}_{1:T}|\boldsymbol{\lambda}) = C_T^{-1} = \frac{c_t}{C_t D_t} = C_t^{-1} D_{t+1}^{-1} = c_t^{-1} C_{t-1}^{-1} D_{t+1}^{-1} \tag{3.23}$$

The posterior probability $\gamma_t(i)$ defined in (3.10) can be calculated as

$$\begin{aligned}
\gamma_t(i) &= \frac{p(\mathbf{x}_{1:T}, s_t = i|\boldsymbol{\lambda})}{p(\mathbf{x}_{1:T}|\boldsymbol{\lambda})} \\
&= \frac{\Pr(\mathbf{x}_{1:t}, s_t = i|\boldsymbol{\lambda}) \cdot p(\mathbf{x}_{(t+1):T}|s_t = i, \boldsymbol{\lambda})}{p(\mathbf{x}_{1:T}|\boldsymbol{\lambda})} \\
&= \alpha_t(i)\beta_t(i) C_{t-1} D_{t+1} c_t \\
&= c_t \tilde{\alpha}_t(i) \tilde{\beta}_t(i)
\end{aligned} \tag{3.24}$$

During implementation, the relationship $\gamma_t(0) + \gamma_t(1) = 1$ can be used for sanity check. Due to the structure of HMM, $\mathbf{x}_{1:t}$ and $\mathbf{x}_{t+1:T}$ are conditionally independent given s_t.

With $\tilde{\alpha}_t(i)$ and $\tilde{\beta}_t(i)$, we can calculate the posterior probability of transitioning from state i at time t to state j at time $t+1$ as the implementation of the Baum-Welch algorithm requires the definition of the following probability

$$\begin{aligned}
\xi_t(i,j) &= \Pr(s_t = i, s_{t+1} = j|\mathbf{x}_{1:T}, \boldsymbol{\lambda}) \\
&= \frac{\alpha_t(i)\beta_{t+1}(j) a_{ij} p(x_{t+1}|s_{t+1} = j)}{p(\mathbf{x}_{1:T}|\boldsymbol{\lambda})} \\
&= c_t a_{ij} \tilde{\alpha}_t(i) \tilde{\beta}_t(j) p(x_{t+1}|s_{t+1} = j)
\end{aligned} \tag{3.25}$$

The Baum-Welch algorithm updates the values of the parameter set $\boldsymbol{\lambda}$ in an iterative

manner. Each iteration contains an expectation step (E-step) and a maximization step (M-step).

In the E-step, the probabilities, $\tilde{\alpha}_t(i)$, $\tilde{\beta}_t(i)$, $\gamma_t(i)$, and $\xi_t(i,j)$, are calculated by using (3.15), (3.22), (3.24) and (3.25) with the parameter set $\boldsymbol{\lambda}$ from the previous iteration.

In the M-step, the parameter set $\boldsymbol{\lambda}$ are updated by using the probabilities $\gamma_t(i)$ and $\xi_t(i,j)$ obtained in the E step. The parameter estimations [94] performed at the M-step in each iteration are

$$\pi_i = \gamma_1(i); \tag{3.26}$$

$$\mu_i = \frac{\sum_{t=1}^{T} \gamma_t(i) x_t}{\sum_{t=1}^{T} \gamma_t(i)} \tag{3.27}$$

$$\sigma_i^2 = \frac{\sum_{t=1}^{T} \gamma_t(i)(x_t - \mu_i)^2}{\sum_{t=1}^{T} \gamma_t(i)} \tag{3.28}$$

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{3.29}$$

for $i, j \in \{0, 1\}$. It is important to note that the estimated parameters are calculated using $\tilde{\alpha}_t(i)$ and $\tilde{\beta}_t(i)$ instead of using α_t and β_t.

In the k-th iteration, we can update the log-likelihood function as

$$\ell(k) = -\log C_T. \tag{3.30}$$

The iterative procedure is terminated if $\ell(k+1) - \ell(k)$ is less than a predefined threshold ϵ.

At the beginning of the iteration process, the parameters are initialized as

$$\pi_0 = \pi_1 = 0.5 \tag{3.31a}$$

$$\mathbf{A} = \tfrac{1}{2}\mathbf{I}_2 \tag{3.31b}$$

$$\mu_0 = 0 \tag{3.31c}$$

$$\mu_1 = 0 \tag{3.31d}$$

$$\sigma_0 = s \tag{3.31e}$$

$$\sigma_1 = \frac{1}{2}\max(|\mathbf{x}_{1:T}|). \tag{3.31f}$$

Here, $\bar{x} = \sum_{i=1}^{T} x_i$ and $s^2 = \frac{1}{T-1}\sum_{i=1}^{T}(x_i - \bar{x})^2$ are sample mean and sample variance of $\mathbf{x}_{1:T}$, respectively, and $\mathbf{I}_2$ is a size 2 identity matrix. Also, $|.|$ is the absolute value operator, operating individually on each element of the vector. Proper care should be taken while selecting values of σ_0 and σ_1, if these values are too big or small then then likelihood $\Pr(x_t|s_t = i)$ can be very small which will force subsequent forward and backward probabilities to zero.

The modified Baum-Welch algorithm is summarized in Algorithm 2.

Algorithm 2 : Baum-Welch Algorithm

1: **Input:** Discrete time dataset $\mathbf{x}_{1:T}$;
2: Initialize $\boldsymbol{\lambda}$ with (3.31), set $k = 0$ and $\ell(0) = -\infty$;
3: **do**
4: $k \leftarrow k + 1$;
5: Calculate $\tilde{\alpha}_t(i)$ and c_t using (3.15) and (3.16);
6: Calculate $\tilde{\beta}_t(i)$ using (3.22);
7: Calculate $\gamma_t(i)$ and $\xi_t(i, j)$ using (3.24) and (3.25);
8: Update $\pi_i, \mu_i, \sigma_i^2, \mathbf{A}$ using (3.26)-(3.29), here $\sigma_1^2 > \sigma_0^2$;
9: Calculate $\ell(k)$ using (3.30);
10: **while** $\ell(k) - \ell(k-1) > \epsilon$
11: **Output:** $\boldsymbol{\lambda}$, $\gamma_t(i)$, for $t = 1, \cdots, T$ and $i \in \{0, 1\}$.

3.4.5 Anomaly Removals

Based on the output of the Baum-Welch algorithm, the hidden state $\hat{s}_t$ of each sample can be detected by applying $\gamma_t(i)$ in (3.11). Samples with estimated hidden state $\hat{s}_t = 1$ are

labeled as anomalies.

Since PVP signals are usually analyzed in the frequency domain, we need to ensure the continuity of the signal in the time domain after the removal of the anomalous samples. To achieve this goal, the PVP signals are first divided into non-overlapping windows with w samples per window. If the percent of corresponding anomalous residuals within a window exceeds a certain threshold ζ (e.g. $\zeta = 15\%$), then all samples within this window are discarded. Such a window-based anomaly removal ensures the time continuity of samples within the same window. The remaining windows are considered as normal windows with valid data. Fast Fourier transform (FFT) can then be applied to each normal window to obtain the frequency domain representation of the PVP signal. With a window size of w samples with sampling period T_0, the frequency resolution of the frequency domain signal is $\frac{1}{wT_0}$.

The window-based anomaly removal approach will remove both normal and anomalous sample inside a window if the percentage of anomalous samples inside that window exceeds a certain threshold. Thus using a smaller window size can effectively reduce the amount of normal samples being removed. On the other hand, a smaller window will results in a lower resolution in the frequency domain. Thus the window size determines a trade-off between the precision of anomaly removal and frequency domain resolution. In this chapter, we pick a window size of 10,000 samples, which corresponds to a duration of 10 seconds with a sampling rate of 10 kHz, to ensure a frequency domain resolution of 0.1 Hz. Algorithm 3 summarizes the procedure of window-based anomaly removal.

3.5 Results

The proposed anomaly detection and removal algorithm is applied to clinical data to verify its effectiveness. The data used in the experiments are PVP signals collected from a cohort of 24 pediatric patients suffering from Pyloric stenosis during emergency department admissions. The characteristics of the patient population are summarized in Table 3.2. Based on the concentration of serum chloride or bicarbonate of each patient, patients were classified

Algorithm 3 : Window-based Anomaly Removal

1: **Input:** PVP signal $\mathbf{y}_{1:T}$, window size w, threshold ζ;
2: Infer $\hat{\mathbf{s}}_{1:T}$ using Algorithms 1 and 2.
3: **for** $n = 1$ to $\lfloor T/w \rfloor$ **do**
4: Calculate the percentage of anomaly samples in the n-th window

$$\tau_n = \frac{1}{w} \sum_{i=1}^{w} \hat{s}_{(n-1)w+i}$$

5: If $w \geq \zeta$, discard the n-th window;
6: **end for**
7: **Output:** Normal windows.

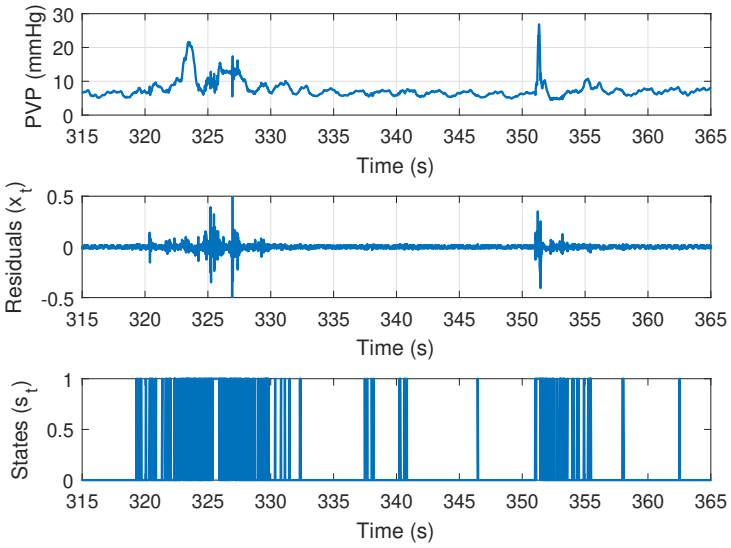

Figure 3.5: (Top) Exemplary PVP signal from Fig. 3.1; (middle) Prediction residual of the Kalman filter; and (bottom) estimated hidden states ($s_t = 0$: normal sample; $s_t = 1$: anomalous sample).

Table 3.1: Number of normal and anomalous windows for each patient ($\zeta = 15\%$).

Patient	Status	Normal	Anomaly	Anomaly (%)
5	Hypo.	55	9	14.06
6	Hypo.	65	2	02.99
7	Hypo.	14	34	70.83
9	Euvo.	55	1	01.79
10	Euvo.	42	24	36.36
12	Euvo.	35	23	39.66
18	Hypo.	27	17	38.64
20	Euvo.	36	5	12.20
22	Hypo.	42	5	10.64
23	Euvo.	26	5	16.13
24	Euvo.	26	7	21.21
25	Hypo.	49	9	15.52
26	Hypo.	45	4	08.16
27	Euvo.	22	11	33.33
28	Euvo.	45	2	04.26
29	Hypo.	43	3	06.52
30	Euvo.	37	9	19.57
31	Euvo.	52	9	14.75
32	Hypo.	38	9	19.15
33	Euvo.	30	14	31.82
34	Hypo.	30	15	23.33
35	Euvo.	50	7	12.28
37	Euvo.	40	13	24.53
39	Euvo.	40	9	18.37

Table 3.2: Patient characteristics

	Euvo.	Hypo.
Patients	14	10
Average weight (kg)	4.17	3.89
Minimum weight (kg)	2.76	2.72
Maximum weight (kg)	5.82	4.72
Standard deviation (kg)	0.78	0.70
Mean age (days)	38.1	40.1

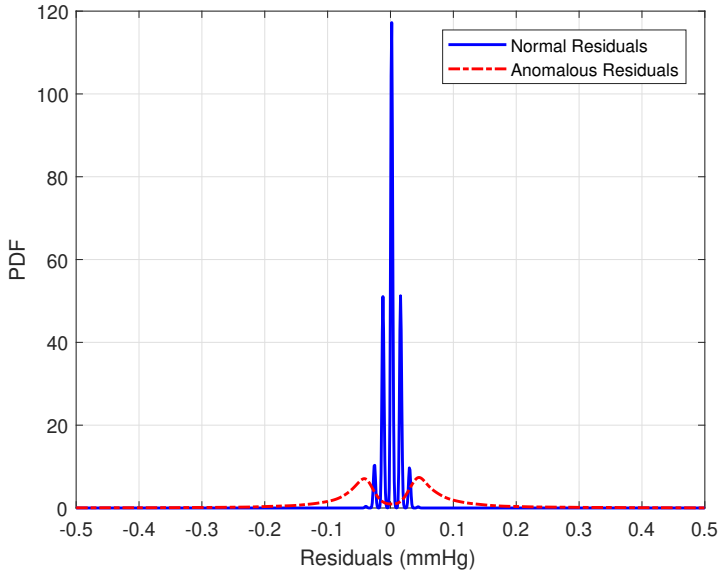

Figure 3.6: Empirical probability density function of prediction residuals (x_t) of patient 10.

Table 3.3: Number of Windows in Dataset

	Training Set			Testing Set		
	Euvo.	Hypo.	Total	Euvo.	Hypo.	Total
Raw Data	466	354	820	209	161	370
Manual	333	273	606	155	124	279
Algorithm in [77]	372	306	678	172	139	311
Proposed Algorithm	371	281	652	165	127	292

Table 3.4: Testing Classification Results

Parameter	Raw Data	Manual	Algorithm in [77]	Proposed Algorithm
True Positive Rate	45.96%	69.35%	63.31%	71.65%
True Negative Rate	76.08%	77.42%	79.65%	81.21%
Precision	59.68%	73.83%	71.54%	74.60%
F_1 Score	52.00%	75.41%	67.18%	73.09%
Accuracy	62.97%	71.07%	72.35%	77.05%
Windows used	100%	70.20%	84.05%	78.92%

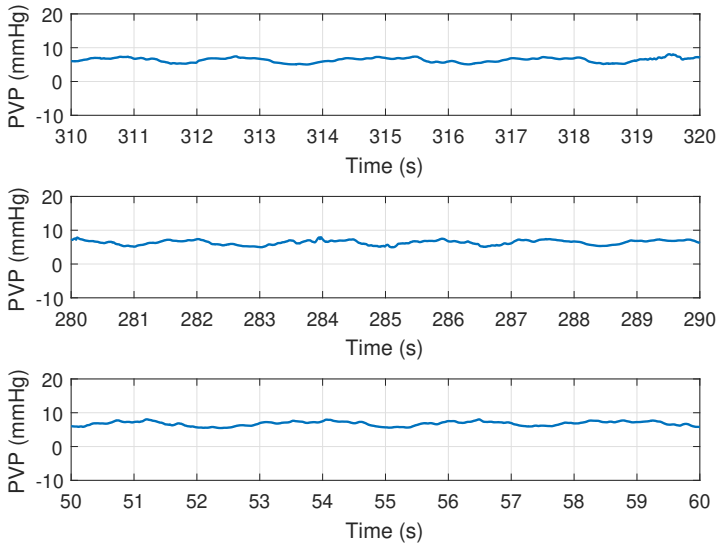

Figure 3.7: Example of normal windows inferred by the proposed model. Windows have a periodic structure and the amplitude does not change abruptly.

as either euvolemic or hypovolemic as described in Section 2. Among the 24 patients, 14 were labeled as euvolemic and 10 were labeled as hypovolemic. The signals for each patient are collected over a continuous period of time with a sampling rate of $f_s = 1$ kHz. The PVP signal is also affected by heart rate in the range $1.9 - 2.5$ Hz ($114 - 150$ bps) and respiratory rate at $0.25 - 0.45$ Hz ($15 - 27$ bps) and their higher order harmonics. Amplitude of the signal varies from $5 - 40$ mmHg. The signal duration collected from each patient ranges from 310 to 670 seconds.

For each patient, the unsupervised anomaly detection algorithm is applied to the entire data sequence collected from that patient. As an example, Fig. 3.5 shows the PVP signal along with the estimated hidden states from patient 20 for a duration of 50 seconds. Hidden states 1 and 0 correspond to anomalous samples or normal samples, respectively. The probability density functions of anomalous and normal data residuals for patient 10 are shown in Fig. 3.6. As can be seen from the results, all large fluctuations in the signals are labeled as anomalous by the proposed algorithm. The large fluctuations are mainly caused by sudden movements of the pediatric patients during the data collection process, and they

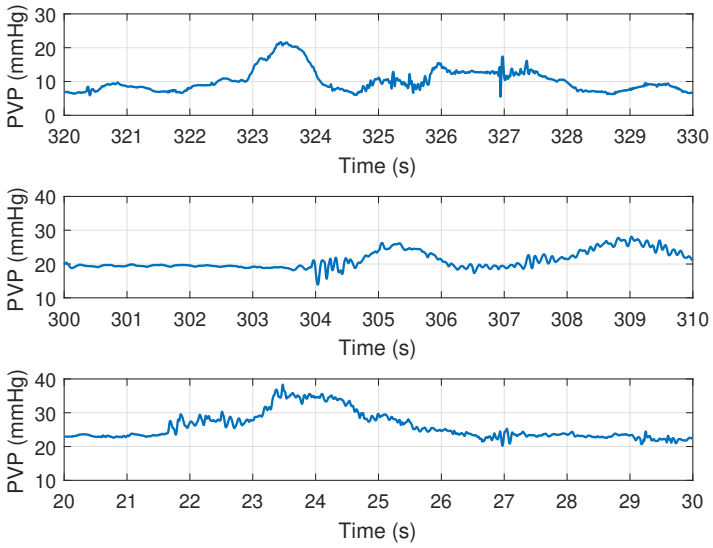

Figure 3.8: Example of anomalous windows inferred by the proposed model. In theses windows, amplitude changes abruptly (10-15 mmHg higher than average) indicating random MNAs.

deviate significantly from normal PVP signals. The results in Fig. 3.6 demonstrate that the distributions of residuals from anomalous data are quite different from their normal counterpart, thus the algorithm can effectively distinguish between the two by analyzing their statistical distributions.

Among all PVP signal windows, 20.67% are detected as anomalous by the HMM-based anomaly detection algorithm. Here, $\epsilon = 10^{-5}$ has been used. A window is declared as anomalous if it contains more than $\zeta = 15\%$ of anomalous residual samples. Figs. 3.7 and 3.8 show a few examples of the PVP signals in normal and anomalous windows, respectively. Table 3.1 lists the number of normal and anomalous windows for each patient.

The direct quantitative evaluation of the anomaly detection algorithm requires labeled data with each sample classified as either normal or anomalous. However, due to the unexpected nature of anomaly events, no labeled data is available for testing of the HMM-based anomaly detection algorithm. Manually labeling the data might introduce subjective bias, which will affect the objectiveness of the evaluation results. The only objective labels that are available to the data are the euvolemic or hypovolemic status of the patients. It has

been shown in [76] that there is a strong correlation between PVP signals and the volume status of pediatric patients. That is, the dehydration status of a pediatric patient can be accurately predicted by analyzing PVP signals. Motivated by this result, we propose to indirectly evaluate the effectiveness of the HMM-based anomaly detection algorithm through volume status prediction.

Based on the results in [76], our hypothesis is that removing anomaly in PVP signals can improve the prediction accuracy of dehydration in pediatric patients. To verify the hypothesis, the experiment is designed as follows. First, the original raw PVP signals without cleaning are used to train and detect the volume status of the patients. The accuracy of the prediction can be quantitatively evaluated given that all patients are objectively labeled as euvolemic or hypovolemic. Second, apply the HMM-based anomaly detection algorithm to the PVP signals, and remove all anomalous windows. Then perform volume status prediction with the cleaned PVP signal after anomaly removal. It is expected that the cleaned signal can achieve a higher accuracy of volume status prediction compared to its uncleaned counterpart.

The volume status prediction is performed by using regularized logistic regression as described in [76]. The regularized logistic regression is a supervised learning algorithm for classification, thus it requires a training phase to build the classification model. It should be noted the training is only used for volume status classification, and no training is needed for the proposed anomaly detection algorithm.

During the analysis, the data are divided into non-overlapping windows with $w = 10,000$ samples per window. Thus the duration of each window is $T_w = 10$ seconds. Signals in each window are converted to the frequency domain by using FFT. The frequency resolution of the frequency domain signals is $f_0 = \frac{1}{T_w} = 0.1$ Hz, with the highest frequency being $\frac{1}{2}f_s = 500$ Hz. Since PVP signals are mainly in the low frequency range, only signal components below 20 Hz are used in the analysis. This results in a frequency domain vector of length 200 for each window. For volume status prediction with regularized logistic regression, the training-testing data was split with a $70\% - 30\%$ ratio. During training, a 5-fold cross validation was used to tune the regularization parameter. The number of windows used in

both cases are presented in Table 3.3.

The classification results obtained from raw data, manually cleaned data, data cleaned by a simple anomaly removal method in [77], and data cleaned by the proposed algorithm are summarized in Table 3.4. For the algorithm in [77], the model parameters are set as $a = 4$, $b = -0.1$ and $w_s = 100$. Even though the denominator used by different algorithms are different in Tables 3.3 and 3.4, the difference is relatively small. For example, the numbers of testing samples for manual cleaning, algorithm in [77], and proposed algorithm are 279, 311, and 292, respectively, which are within 5.7% of their mean value. As can be seen from the results, after removing anomaly with the proposed anomaly detection algorithm, the accuracy of volume status prediction is considerably improved compared to the results obtained from the uncleaned raw data. The proposed algorithm also outperforms manual cleaning or the cleaning algorithm in [77]. These results demonstrate that the HMM-based anomaly detection algorithm can effectively remove the anomalous data from the collected PVP signals, thus more accurate volume status prediction is achieved by using the cleaned PVP signals.

3.6 Conclusion

An unsupervised anomaly detection algorithm has been proposed to remove MNA in PVP signals. The PVP signals were modeled by using a DLM to capture correlation of neighboring samples of the PVP signals. Kalman filter was applied to the DLM to track and predict the PVP signal in the time domain. The prediction residuals were then modeled by using an HMM, where the normal and anomalous states of the PVP signals were represented by using the binary states of the hidden Markov chain. The HMM parameters along with the hidden states were iteratively estimated with a modified Baum-Welch algorithm to ensure the numerical stability. The proposed algorithm was verified by comparing the hypovolemic prediction accuracy of PVP signals before and after the anomaly processing. Experimental results with clinical data of a cohort of 24 pediatric patients demonstrated that the algorithm can efficiently remove anomalies in PVP signals. The algorithm can also be used for anomaly

detection in other time series signals, such as ECG and PPG signals. All supplementary materials and codes of this chapter can be found at this github repository:

`https://github.com/i2pt/UAD_PVPS_HMM`

4 Modeling Peripheral Arterial and Venous Pressure Signals with Integral Pulse Frequency Modulation

4.1 Abstract

This chapter proposes a novel mathematical modeling framework for peripheral arterial and peripheral venous blood pressure signals from porcine experiments. Peripheral blood pressure signals can be acquired using regular catheters during standard patient treatment. The minimally invasive nature and ubiquitous availability of catheters render it an ideal candidate for a wide range of applications. However, there is no analytical model in the literature for peripheral blood pressure signals. We address this issue by proposing a model for these signals under the integral pulse frequency modulation (IPFM) framework. The model incorporates the impacts of physiological phenomena, such as the heartbeat pulse shape variation, heart rate variability, respiratory rate, etc. The model parameters are obtained by applying the IPFM model to experimental data collected from four pigs under different anesthetic dosages. The proposed model can fit the experimental data with Pearson correlation coefficients greater than 0.99 and 0.90 for arterial and venous blood pressure signals, respectively. The performance of model-synthesized data on the classification of two different anesthesia is comparable with experimental data. Parameters like pulse shape and duration can also work as distinguishable features under different anesthesia. We also proposed a way to distinguish respiratory-induced heart rate variability from other causes. Increasing doses vasodilating anesthesia is similar to the scenario of going from dehydration to hydration. Thus the results obtained in this study can be extended in distinguishing hydrated and dehydrated subjects. This model can be extended to similar biomedical signals like photoplethysmography, cerebral blood flow velocity, and Doppler waveforms.

4.2 Introduction

Peripheral arterial pressure (PAP) and peripheral venous pressure (PVP) signal properties can provide valuable information regarding the clinical conditions of patients. For ex-

ample, the PVP signal is strongly correlated with the central venous pressure (CVP) signal [31, 18]. CVP is the pressure in the first and second-order branches of veins that drain into the right atrium of the heart. However, the collection of CVP relies on the invasive placement of a central monitor and contains significant risk and time to place correctly. PAP or PVP signals, on the other hand, can be collected with regular intravenous and intra-arterial catheters that are widely available and routinely used. The minimally invasive nature combined with the ubiquitous availability of catheters renders it as an ideal candidate for a wide range of applications in patient care. For example, PVP signals have been used in a wide range of medical scenarios, such as pediatric surgery [23], laparoscopic colorectal surgery [31], craniotomy or complex spine surgery [19], patients with acute decompensated heart failure [34], etc.

The study and analysis of PVP signals have attracted considerable attention recently [70, 41, 78, 77, 95, 47, 46]. It is shown that PVP signals can be used to predict dehydration in the pediatric population [76]. PVP signals are used to assess the dehydration level of infants suffering from pyloric stenosis and hemorrhage during craniosynostosis repair [45]. PVP signals are also related to other peripheral blood pressure signals. A recent study shows that the PVP signal is strongly correlated with peripheral arterial pressure (PAP) [96] signal. The Photoplethysmogram (PPG) signal is similar in nature to PVP signal in that both are directly related to blood flow, and the relationship between PPG and PVP signals is studied in [97]. However, there is a lack of research on the modeling of PAP and PVP signals. Parametric modeling of these signals can help us understand the generation process behind the signals, thus providing insights into their relations to various physiological phenomena.

Signal modeling is an important tool that can be used to understand the physiology and generative process behind a biomedical signals. It also helps to quantify and predict the changes in the signal under different internal and external impact circumstances. This can significantly improve treatment outcomes in many situations, which is the overall goal of this work. Several models are developed to represent the behaviors of PPG signals. For example, in [98], the PPG signal is modeled by using a synthesis-by-analysis technique. Synthetic PPG

models are developed to generate PPG waveforms with two combined Gaussian functions, and they can model both regular and irregular heartbeats [99] and [100]. A modulation-based PPG model is proposed in [101] to extract vital signs from PPG signals. There are a wide range of models for ECG signals. In [102], a modified Zeeman model with an artificial neural network is developed to generate ECG signals. In [103], 12-lead ECG signals are modeled by using bidomain equations coupled to a phenomenological ionic model in the heart and a generalized Laplace equation in the torso. An ECG model that simulates the main pacemakers in the heart by using a set of three nonlinear oscillators produced by discretized reaction-diffusion system is proposed in [104]. In [105], Gaussian functions along with hybrid optimization methods have been proposed to model ECG morphology in different cardiac dysrhythmias. To the best of our knowledge, no analytical model is available for peripheral blood pressure signals.

In this chapter, we propose to build an analytical model for PAP and PVP signals with the framework of integral pulse frequency modulation (IPFM). IPFM was first introduced in [106] for analyzing nervous systems. It was later applied to model heart rate variability (HRV) studies [107, 108]. IPFM uses irregularly spaced spikes to imitate the physical properties behind HRV [109]. Improved HRV modeling and analysis with IPFM are presented in [110, 111, 112]. In all existing works, the IPFM model has only been used to model the onset of pulses, and none can be found to model the signal itself. To the best of our knowledge, the IPFM model was never applied to PAP and PVP signals before.

In recognition that the peripheral blood pressure signals are affected by HRV through frequency modulation (FM) and respiration signals through amplitude modulation (AM), we propose to capture these physiological phenomena by using IPFM. The utilization of IPFM allows us to focus on the physiology of peripheral blood pressure signals without dealing with complex spectrum analyses of AM-FM signals. Unlike existing IPFM models that focus solely on the onset of pulses, the proposed work combines IPFM with convolution analysis, which allows us to model the shape of the peripheral waveforms.

Under the IPFM framework, the proposed model is developed by analyzing the prop-

erties of PAP and PVP signals collected from pig experiments. The newly proposed model is more general than the modulation model in [101]. The model parameters are learned by fitting the developed IPFM model to the experimental data. The accuracy and efficacy of the model are evaluated by comparing model-synthesized data with experimental data. It is demonstrated that the proposed model matches very well with the experimental results.

The rest of the chapter is organized as follows. Section 4.3 discusses the experimental setup and data acquisition process. The proposed IPFM modeling framework is introduced in Section 4.4. In Section 4.5, a numerical modeling fitting method is proposed to estimate the model parameters by fitting the model to the experimental data. The performance of the model is evaluated in Section 4.6. Section 4.7 discusses some qualitative properties of some of the model parameters, and Section 4.9 concludes the chapter.

4.3 Experiment Setup and Data Acquisition

This section describes the setup of the experiment, the demographics of the animal subjects, the data acquisition system, and the data collection workflow at different stages of anesthesia, and the overall design of the experiment.

4.3.1 Pig Demographics

In this study, four female pigs were used. All pigs were between ages 16-17 weeks, at a weight of 70-74 kg at the time of the experiment. The detailed demographics are listed in Table 4.1. During the experiment, the pigs were anesthetized and experienced controlled blood loss. The average blood-weight ratio is 13.97 ml/kg. The study was performed following the approval of the University of Arkansas for Medical Sciences (UAMS) institutional animal care and use committee.

4.3.2 Data Acquisition System

Millar Mikro-Tip® solid-state catheters with a pressure sensor(s) provide high-fidelity cardiovascular pressures (the heart and vascular system) at the source in small animals [113].

Table 4.1: Pig subjects used in this study

Subject	Weight (kg)	Age (weeks)	Controlled Blood Loss (ml)	Blood-Weight Ratio (ml/kg)
Pig1	70	16	1270	18.14
Pig2	74	17	923	12.47
Pig3	74	17	963	13.01
Pig4	73	17	910	12.47

Two Millar pressure catheters are used at the brachial artery and brachial vein to collect arterial (PAP) and venous (PVP) pressure signals, respectively. Fig. 4.1 is the schematic diagram of the data acquisition system. For Pigs 1 and 2, Millar catheters of size 5F (Millar SPR-350S) were used. Whereas, Millar catheters of size 2F (Millar SPR-320) were used for Pigs 3 and 4. Here, F indicates the French catheter scale (1F corresponds to an external diameter of 1/3 mm). Catheters were then connected to a pressure control unit (Millar PCU-2000) via catheter interface cables (ADInstruments PEC-10D). The output of the pressure control unit is then connected to a National Instruments (NI) USB-6009 data acquisition system.

While inserting the Millar catheters, the distance between the catheters was observed using ultrasound systems (BK500, BK Medical) to make sure they were within 3-4 cm apart. Datex Ohmeda Aestiva 5 (GE Healthcare) anesthesia machine was used to maintain mechanical ventilation greater than the respiratory rate of 0.2 Hz (12 bpm). At lower anesthetic levels a deviation from 0.2 Hz is expected. On the other hand, at higher anesthetic levels the pigs do not show self-breathing tendency and the respiratory rate shows very little deviation.

It should be noted that this chapter models the voltage signal recorded at the USB-6009 data acquisition system. The voltage (Volt) and the pressure (mmHg) are proportional to each other and are related in a linear relationship. Thus, the signal used in this study can have negative values.

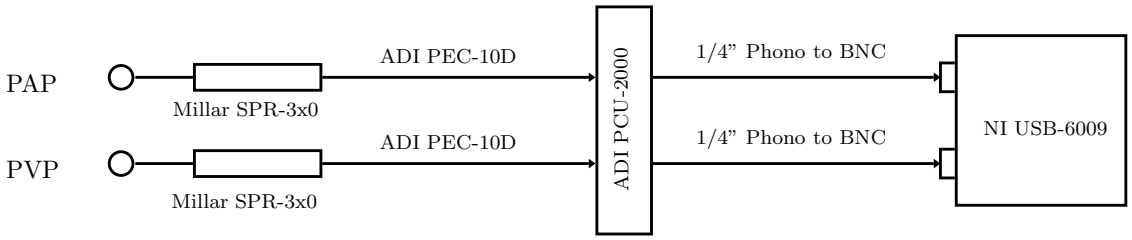

Figure 4.1: Schematic diagram of peripheral arterial pressure (PAP) and peripheral venous pressure (PVP) signals being collected using Millar SPR-3x0 (solid-state) catheters with National Instrument USB-6009 data acquisition system.

4.3.3 Data Collection Workflow

After inserting and attaching the transducers, the pigs were anesthetized using Isoflurane ($C_3H_2ClF_5O$). Isoflurane, a widely used general anesthetic, and vasodilator is given via inhalation. Vasodilation and vasoconstriction are the widening and narrowing of blood vessels respectively. The amount of Isoflurane was changed to 1.8%, 2.5%, 2.8%, 2.0%, and 1.5% over time. In terms of minimum alveolar concentration (MAC), Isoflurane 1.8%, 2.5%, and 2.8% are considered MAC-1, MAC-2, and MAC-3, respectively.

Propofol ($C_{12}H_{18}O$), another widely used anesthetic and vasodilator, was injected into the veins following Isoflurane delivery. The amount of Isoflurane was changed to 0.10 mg/kg/min, 0.15 mg/kg/min, 0.20 mg/kg/min, and 0.05 mg/kg/min. At this stage, all pigs went through bleeding from the femoral artery of the right leg. The amount of blood loss for each pig is listed in Table 4.1. After bleeding, Propofol 0.05 mg/kg/min is followed by Isoflurane 1.5% and 2.00%.

Propofol 0.10 mg/kg/min, 0.15 mg/kg/min, and 0.20 mg/kg/min are denoted as PRO-1, PRO-2, and PRO-3, respectively in this chapter. Also note that PRO-1 stage does not immediately start after MAC-3. The data collection workflow is summarized in Table 4.2.

In the original setup of this experiment, multiple other signals were collected. Based on the signal quality during the data collection procedure, this chapter only deals with PAP and PVP signals. PAP and PVP signal acquisition process do not show a blood clotting

Table 4.2: Data collection workflow. Asterisks (*) denote the data used in this chapter.

Action	Amount	Abbreviation
Isoflurane (%)	*1.80	MAC-1
	*2.50	MAC-2
	*2.80	MAC-3
	2.00	
	1.50	
Propofol (mg/kg/min)	*0.10	PRO-1
	*0.15	PRO-2
	*0.20	PRO-3
	0.05	
Bleeding		
Propofol	0.05	
Isoflurane	1.50	
	2.00	

tendency in the Millar catheter. It should be made clear that this chapter treats PAP and PVP signals as peripheral blood pressure signals as both are collected from the brachial artery and brachial vein respectively.

During dehydration or blood loss, humans and animals have the tendency to constrict blood vessels, i.e., narrowing of blood vessels that is also known as vasoconstriction. Vasodialators such as Isoflurane and Propofol can on the other hand widen the blood vessels. So, during the MAC-3 stage the blood vessels are wider compared to MAC-1. Thus compared to MAC-3, MAC-1 has a similar effects as the dehydrated state. The same reasoning holds for the PRO stages. This motivates the study of the various peripheral blood pressure signals under different stages of anesthesia in this experiment.

4.4 Parametric Signal Modeling

The two main components embedded in peripheral blood pressure signals are heart rate (with harmonics) and respiratory rate signals. These two signals are modulated with each other.

However, the heart rate does not remain constant over time. There is always a time dependent variation in heart rate. The variation in heart rate is called heart rate variability

(HRV). HRV is an important indicator of the autonomic nervous system (ANS). ANS is a network of nerves that handles unconscious tasks like heartbeat and breathing. This is why the HRV is an important physiological and neurological indicator of sympathetic (quick "fight-or-flight" response) and parasympathetic (slow response) activities of ANS. High HRV indicates a healthy nervous system that is responsive to both sympathetic and parasympathetic inputs. This is why HRV is still a highly active field of research.

HRV can be contributed by multiple sources. The main cause of HRV is respiration. Heart rate increases and decreases with inhalation and exhalation, respectively. Such respiratory-induced HRV is called respiratory sinus arrhythmia (RSA). There are also other long-term sources of HRV such as, ANS imbalance, heart diseases (like Arrhythmia), stress, poor sleep, unhealthy diet, and lack of exercise, etc. The signal model proposed in the chapter isolates the impact of respiration from heart rate signal.

The IPFM-based parametric model is developed based on the hypothesis that the occurrences of pulse onsets are initiated by a continuous-time modulating signal $m(t)$, which has certain physiological interpretations [109, 114]. A pulse onset trigger impulse is generated when the integration of $m(t)$ reaches a certain threshold. It is assumed that $|m(t)| \ll 1$ and its Fourier transform $\mathscr{F}\{m(t)\} = M(\omega)$ is bandlimited with negligible power spectral density beyond a frequency typically around 0.4-0.5 Hz [109, 114]. It is also assumed that $m(t)$ is a zero-mean signal, i.e., $M(0) = 0$.

Denote t_k as the time instant of the onset or occurrence of the k-th heartbeat. Without loss of generality, it is assumed that the first pulse onset occurs at time $t_0 = 0$ and $m(t)$ is causal [115], that is, $m(t) = 0$ for $t < 0$. Based on the principle of the IPFM model [109], the following equation relates t_k with $m(t)$ [109, Equation (1)]

$$T = \int_{t_k}^{t_{k+1}} [1 + m(t)]dt, \tag{4.1}$$

or alternatively

$$k = \int_0^{t_k} \frac{1 + m(t)}{T} dt, \tag{4.2}$$

where T is the mean heart rate interval in seconds, and $\frac{1}{T}$ is the mean heart rate in Hertz.

In the IPFM model, the continuous-time signal $m(t)$ models the dynamic variability of the heart rate. When $m(t) = 0$, there will be no HRV with $t_{k+1} - t_k = T$ for all $k \in \mathbb{Z}^+ = \{0, 1, 2, \cdots\}$.

Another related and important signal is the heart timing signal $\theta(t_k)$, which is defined as the deviation of heartbeat occurrence time t_k from the mean occurrence time kT [109]. From (4.2), we have

$$\theta(t_k) = kT - t_k = \int_0^{t_k} m(\tau) d\tau. \tag{4.3}$$

Equation (4.3) presents a non-uniformly sampled version of $\theta(t)$. The continuous-time heart timing signal $\theta(t)$ can be approximated by performing spline interpolation over the discrete time samples, $\theta(t_k)$, for $k \in \mathbb{Z}^+$. In this chapter, cubic spline interpolation is used to estimate $\theta(t)$ from $\theta(t_k)$. The knowledge of $\theta(t)$ can be used to estimate the signal $m(t)$ as

$$m(t) = \frac{d\theta(t)}{dt}. \tag{4.4}$$

The conventional IPFM framework can only be used to model the onset of pulses, and it cannot be readily applied to model continuous-time waveforms such as the PAP or PVP signals. In this chapter, we extend the IPFM model to model the continuous-time waveform of the signal, instead of just modeling the onset pulses as in the literature. Based on this proposed model, the respiratory signal, $r(t)$, is modulated with the heart rate signal, $x(t)$. The signal $x(t)$ is not just the heart rate and its higher-order harmonics. However, we refer to it as the heart rate signal for simplicity. Both $r(t)$ and $x(t)$ are constituent components of

PAP or PVP signals. Based on this idea, the continuous-time waveform $x(t)$ can be modeled by performing convolution between the IPFM pulses with the single heartbeat pulse of the heartbeat signal.

A single heartbeat pulse is defined as $p(t)$ with support T. Under the proposed model, the heart rate signal in the time domain can be represented as

$$x(t) = \sum_k p\left(\frac{T}{t_{k+1} - t_k} t\right) \otimes \delta(t - t_k), \tag{4.5}$$

where $\otimes$ denotes continuous time convolution. Given that $|m(t)| \ll 1$, we employ the approximation that $t_{i+1} - t_i \approx T$, and (4.5) can be simplified to

$$x(t) = p(t) \otimes \sum_k \delta(t - t_k). \tag{4.6}$$

This approximation is only used to facilitate the modeling process to demonstrate how different components of the signal are interacting with each other, and the final modeled signal still follows (4.5).

Once the pulse onset time instants $\{t_k\}_{k \in \mathbb{Z}^+}$ are known, a non-uniformly spaced pulse train is defined as

$$s(t) \triangleq \sum_k \delta(t - t_k), \tag{4.7}$$

$$= \frac{1 + m(t)}{T} \left[1 + 2 \sum_{n=1}^{\infty} \cos\left(\frac{2n\pi}{T}\big(t + \theta(t)\big)\right)\right]. \tag{4.8}$$

Detailed derivations of (4.8) is in Appendix 4.8. The second equality follows from the fact that $|1 + m(t)| = 1 + m(t)$ as $|m(t)| \ll 1$. As a sanity check, when $\theta(t)$ is constant, i.e., $m(t) = 0$, then $s(t)$ corresponds to an impulse train.

Using the definitions above, the observed PAP or PVP signal $y(t)$ can then be modeled

as

$$y(t) = \tau(t) + \big[\alpha + r(t)\big]\big[\beta + x(t)\big] \tag{4.9}$$

where $\tau(t)$ is a slow-changing bias or drift introduced by the data acquisition system, $r(t)$ is the respiratory signal, $x(t) = p(t) \otimes s(t)$ is the heart rate signal, and α and β are the DC offsets of $r(t)$ and $x(t)$, respectively. The signal $\tau(t)$ has zero mean.

Equation (4.9) incorporates the fact that $x(t)$ is amplitude modulated by $r(t)$. From (4.8), it is clear that the signal $s(t)$ and thus heart rate signal $x(t)$ are dependent of $m(t)$. The signal $x(t)$ does not contain any component of $r(t)$. This property separates the impact of the respiratory signal $r(t)$ from the modulating signal $m(t)$ and the heart rate signal $x(t)$. Based on the physiological conditions, the signals $m(t)$ and $r(t)$ can be correlated. In that case, the strength of the correlation between $m(t)$ and $r(t)$ can indicate the strength of RSA on HRV. The parametric model presented in (4.9) includes the effects of heartbeat, respiration, HRV, and their mutual interactions. The parameters of the model can be estimated or learned by fitting the model to the experimental data.

4.5 Model Fitting

In this section, both numerical and computational approaches are employed to model the experimental data collected from the pig subjects by using the theoretical foundations of the previous section.

4.5.1 Model Reformulation

Given that the measured signals contain impacts from multiple physiological factors at different frequency ranges, we reformulate the signal model to reflect the diverse signal compositions. The reformulation will facilitate the model fitting process.

From (4.8), define the heart rate and its harmonics as

$$h(t) \triangleq 2\sum_{n=1}^{\infty} \cos\left(\frac{2n\pi}{T}\big(t+\theta(t)\big)\right). \tag{4.10}$$

The pulse train $s(t)$ defined in (4.8) can then be alternatively represented as

$$\begin{aligned} s(t) &= \left[\frac{1+m(t)}{T}\right][1+h(t)], \\ &= \frac{1}{T} + \frac{1}{T}m(t) + \frac{1}{T}h(t) + \frac{1}{T}m(t)h(t). \end{aligned} \tag{4.11}$$

Combining the definition of $s(t)$ from (4.7) with (4.6) gives

$$x(t) = p(t) \otimes s(t) = \gamma + \frac{1}{T}p(t) \otimes h(t) + q(t), \tag{4.12}$$

where

$$\gamma \triangleq \frac{1}{T}\int_0^T p(t)dt, \tag{4.13}$$

$$q(t) \triangleq \frac{1}{T}p(t) \otimes [m(t)(1+h(t))]. \tag{4.14}$$

Combining (4.9) and (4.12) yields,

$$\begin{aligned} y(t) &= \tau(t) + \alpha\beta + \beta r(t) + \alpha x(t) + x(t)r(t), \\ &= \tau(t) + y_{\text{LF}}(t) + y_{\text{HF}}(t), \end{aligned} \tag{4.15}$$

where

$$
\begin{aligned}
y_{\text{LF}}(t) &\triangleq \alpha(\beta+\gamma) + (\beta+\gamma)r(t), \\
&= \big[\beta+\gamma\big]\big[\alpha + r(t)\big], \qquad (4.16) \\
y_{\text{HF}}(t) &\triangleq [\alpha + r(t)] \left[\frac{1}{T}p(t) \otimes h(t) + q(t)\right], \\
&= [\alpha + r(t)]\tilde{x}(t). \qquad (4.17)
\end{aligned}
$$

Here, $\tilde{x}(t) \triangleq x(t) - \gamma = \frac{1}{T}p(t) \otimes h(t) + q(t)$.

In the reformulated model given in (4.15), the pressure signal $y(t)$ is decomposed into three components, the very slow changing bias $\tau(t)$, the low-frequency component due to respiratory activities $y_{\text{LF}}(t)$, and the high-frequency component $y_{\text{HF}}(t)$ dominated by heart rate. Also, it should be noted that γ is the DC value of the signal $x(t)$ and thus $\tilde{x}(t)$ is a zero-mean signal.

4.5.2 Model Fitting

With the reformulated model presented in (4.15), we can numerically estimate the reformulated model parameters from the experimental data $y_r(t)$. Details of the model fitting process are given in Algorithm 4.

The model fitting process is developed by separating the bias $\tau(t)$, the low-frequency component $y_{\text{LF}}(t)$, and the high-frequency component $y_{\text{HF}}(t)$, from the experimental data.

At first, the recorded experimental data $y_r(t)$ is bandlimited by passing through a low pass filter with cutoff frequency $f_c = 15$ Hz to remove noises from the data collection process. The stopband attenuation of the low pass filter is 60 dB. The cutoff frequency is chosen based on the fact that the fundamental frequency of heart rate is usually at 2 Hz or lower, thus its harmonics are usually well below 15 Hz. Denote the output of the low pass filter as $y(t)$.

The bias $\tau(t)$ is a small drift introduced by the data acquisition system, and it changes

much slower than the respiratory or heart rate signals. The bias signal is estimated using a simple moving average (low pass) filter of a window length of 30 seconds. Fig. 4.2 shows an example of the experimental data, $y_r(t)$, the corresponding bandlimited signal, $y(t)$, and the slow changing bias, $\tau(t)$, from the PAP signal collected from Pig-2 during MAC-2.

The bias-corrected signal, $\tilde{y}(t) \triangleq y(t) - \tau(t)$, passes through a low pass filter with cutoff frequency f_0 to separate the low and high-frequency components, $y_{\text{LF}}(t)$ and $y_{\text{HF}}(t)$, respectively. The cutoff frequency f_0 is chosen below the heart rate but above the respiratory harmonics. In this chapter, we use $f_0 = 0.5$ Hz, and the stopband attenuation of this low pass filter is 60 dB. Examples of $\tilde{y}(t)$ with the corresponding $y_{\text{LF}}(t)$ and $y_{\text{HF}}(t)$ from Pig-2 during MAC-2 are shown in Fig. 4.3.

The model parameters, α, $\beta + \gamma$, $r(t)$, and $\tilde{x}(t)$ can then be estimated by using the filtered signals, $y_{\text{LF}}(t)$ and $y_{\text{HF}}(t)$. From (4.17), the high-frequency component $y_{\text{HF}}(t)$ can be equivalently considered as $\tilde{x}(t)$ amplitude modulated by the low-frequency signal $\alpha + r(t)$. Thus the low-frequency modulating signal, $\alpha + r(t)$, is embedded in the amplitude of $y_{\text{HF}}(t)$. Consequently, $\alpha + r(t)$ can be estimated by performing envelope detection over $y_{\text{HF}}(t)$ as shown in Fig. 4.4. Denote the upper and lower envelopes of $y_{\text{HF}}(t)$ as $e_u(t)$ and $e_l(t)$, respectively. Under the assumption of ideal envelope detection, we have $e_u(t) = -e_l(t) = \alpha + r(t)$. Then the DC offset α can be estimated as

$$\alpha = \frac{\bar{e}_u - \bar{e}_l}{2}, \tag{4.18}$$

where $\bar{e}$ denotes the time average of the signal $e(t)$ and $r(t)$ is assumed to be a zero-mean signal. Envelopes are estimated using spline interpolation over local maxima using the "envelope(_, _, 'peak')" function of Matlab. As can be seen from (4.16), the mean value of the y_{LF} signal is $\bar{y}_{\text{LF}} = \alpha(\beta + \gamma)$. Similarly, the parameter $\beta + \gamma$ can be estimated as

$$\beta + \gamma = \frac{\bar{y}_{\text{LF}}}{\alpha}. \tag{4.19}$$

Given the parameters α and $\beta + \gamma$, the signals $r(t)$ and $\tilde{x}(t)$ can then be estimated by using (4.16) and (4.17) as

$$r(t) = \frac{y_{\text{LF}}(t)}{\beta + \gamma} - \alpha, \tag{4.20}$$

$$\tilde{x}(t) = \frac{y_{\text{HF}}(t)}{\alpha + r(t)}. \tag{4.21}$$

The signal $\tilde{x}(t)$ is the heart rate signal without the DC offset γ. A peak detection algorithm is applied to $\tilde{x}(t)$ to identify the time of systolic cycle onset or the beginning of a pulse as shown in Fig. 4.4. Details of the onset detection algorithm are given in Section 4.5.3. It should be noted that the onset detection algorithm is not applied on $y_{\text{HF}}(t)$ because, based on (4.17), it is modulated by $\alpha + r(t)$. To ignore the effect of $r(t)$, the onset detection is performed over $\tilde{x}(t)$ instead of $y_{\text{HF}}(t)$. This can be also observed in Fig. 4.4.

The onset detection algorithm is used to estimate the systolic cycle onset or the pulse onset time $\{t_k\}_{k\in\mathbb{Z}^+}$, which is then used to construct an estimate of the impulse train $s(t)$ defined in (4.7). After estimating $\{t_k\}_{k\in\mathbb{Z}^+}$, the mean pulse shape $\tilde{p}(t)$ is estimated by taking average from aggregating all the pulses. Ideally, the initial value $p(0)$ and the final value $p(T)$ of the pulse $p(t)$ should be equal. To make this happen, a linear normalization is performed on the mean pulse shape $\tilde{p}(t)$ as

$$p(t) = \tilde{p}(t) - \frac{\tilde{p}(T) - \tilde{p}(0)}{T} t.$$

Fig. 4.8 is an example of the estimated $p(t)$. Based on (4.13), the parameter γ can be easily estimated by taking the time average of the pulse $p(t)$. With the estimated parameters $\{t_k\}_{k\in\mathbb{Z}^+}$, T and $p(t)$, we can construct a synthesized model signal $\hat{x}(t)$ by using (4.5) (Step 7 in Algorithm 4). Once $\hat{x}(t)$ is constructed, we can construct a synthesized version of the overall signal by following Step 8 in Algorithm 4.

It should be noted that at Step 5 of Algorithm 4, the signal $\alpha + r(t)$ must be non-zero

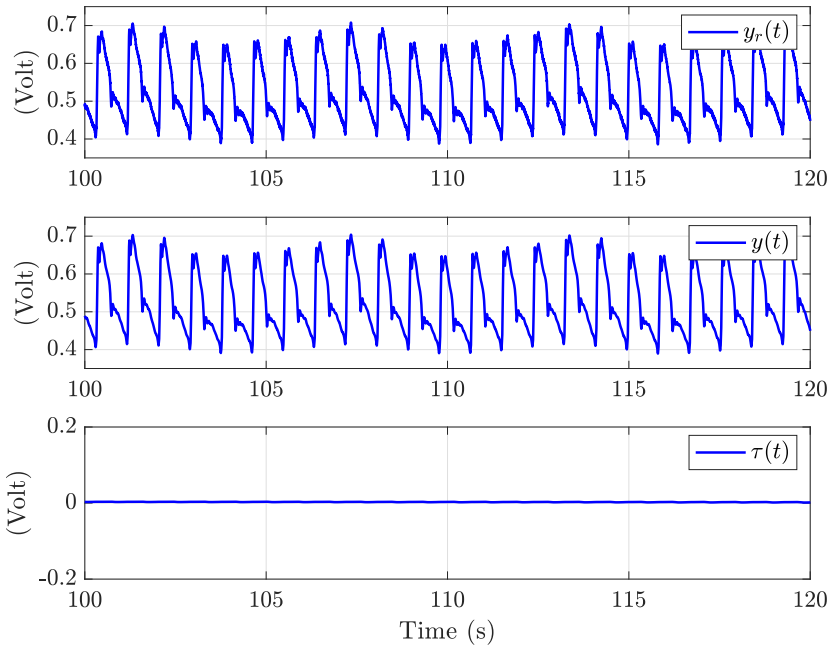

Figure 4.2: Example of separating high-frequency noise and trend from recorded signal $y_r(t)$ (Pig-2 MAC-2 PAP). Here, $y(t)$ obtained after filtering $y_r(t)$ and $\tau(t)$ is zero-mean trend of the signal $y(t)$. Regarding $\tau(t)$, the whole signal recording has a mean zero.

at all times to avoid dividing by zero. To make sure that this signal does not cross zero, a DC offset is added to the signal $y(t)$, when needed, without loss of generality. In the case of the PVP signal, an offset of 1 has been added to $y(t)$.

4.5.3 Pulse Onset Detection

Pulse onset detection is a crucial part of the signal modeling process. Based on the nature of the signal, two different kinds of onset detection algorithms are used.

For the PVP signals, the onset detection can be performed by using peak detection algorithms that are similar to the ones used for ECG [116], PPG [117] and other related signals [118]. In this chapter, the onset is detected using "findpeaks()" function of Matlab on $-\tilde{x}(t)$. After detecting a peak, a certain time (in the range between 500-800 ms) is waited to find the next peak. The wait time depends on the fundamental heart rate frequency of the signal.

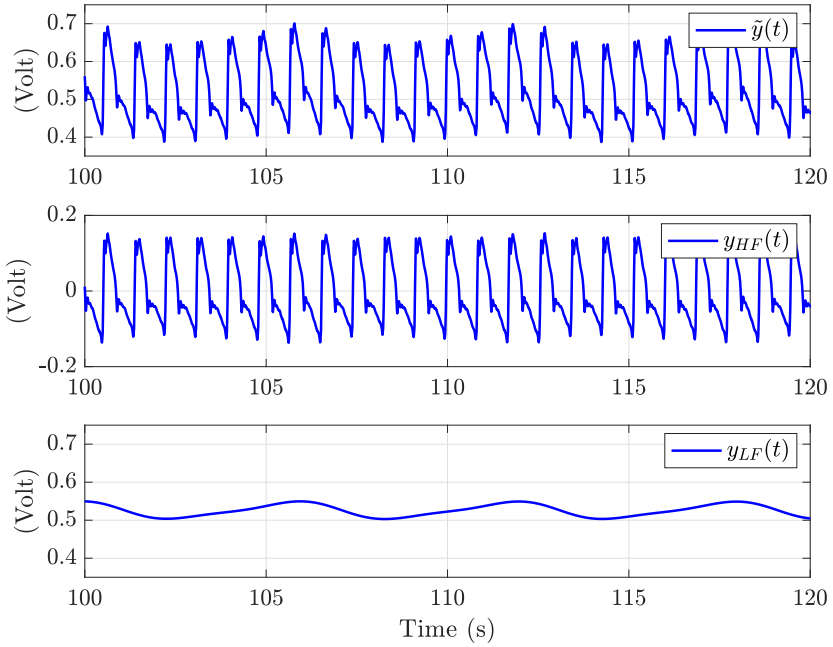

Figure 4.3: Separating high-frequency component $y_{\text{HF}}(t)$ and low-frequency component $y_{\text{LF}}(t)$ from $\tilde{y}(t) = y(t) - \tau(t)$ signal (Pig-2 MAC-2 PAP). $y_{\text{LF}}(t)$ is a linear transformation of respiratory signal $r(t)$.

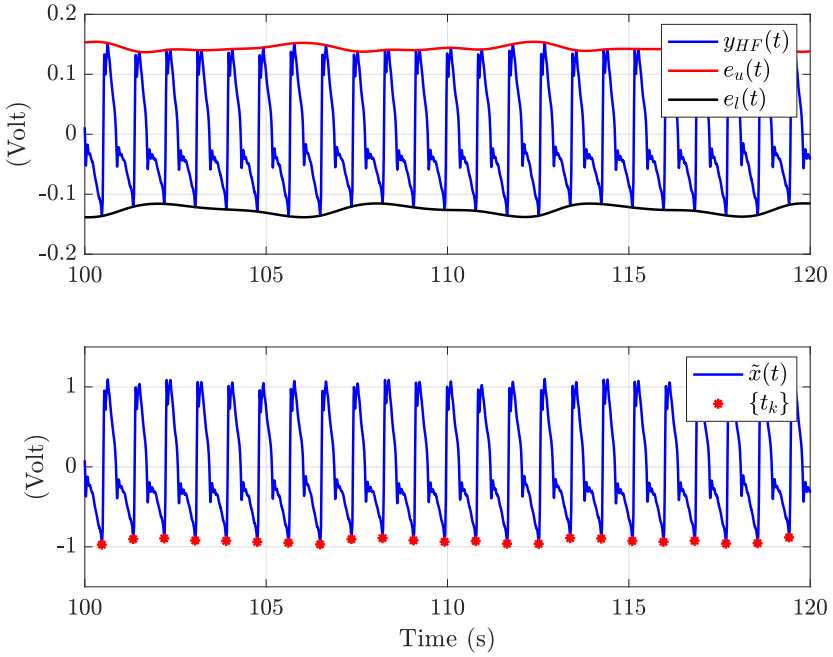

Figure 4.4: Estimating envelope $e_u(t)$ and $e_l(t)$ from $y_{\text{HF}}(t)$ signal (Pig-2 MAC-2 PAP). The model aims to flatten the envelope for $\hat{x}(t)$ signal.

Algorithm 4 IPFM model fitting of blood pressure signal

Require: Experimental signal $y_r(t)$

Ensure: Synthesized signal, $\hat{y}(t)$

1: Band-limiting the input signal $y_r(t)$ to discard frequency components beyond 15 Hz and this filtered signal is denoted as $y(t)$.

2: Apply a simple moving average filter with a window length of 30 s to extract the zero-mean bias $\tau(t)$ of the signal.

3: Apply a low-pass filter with cutoff frequency $f_0 = 0.5$ Hz to the bias-corrected signal, $\tilde{y}(t) = y(t) - \tau(t)$, to separate low and high frequency component $y_{\text{LF}}(t)$ and $y_{\text{HF}}(t)$ respectively as

$$\tilde{y}(t) = y(t) - \hat{\tau}(t) = y_{\text{LF}}(t) + y_{\text{HF}}(t).$$

4: Find the upper envelope $e_u(t)$ and lower envelope $e_l(t)$ of the signal $y_{\text{HF}}(t)$.

5: Estimate α, $\beta + \gamma$, $r(t)$ and $\tilde{x}(t)$ using the following equations

$$\begin{aligned}
\alpha &= \frac{\bar{e}_u - \bar{e}_l}{2}, \\
\beta + \gamma &= \frac{\bar{y}_{\text{LF}}}{\alpha}, \\
r(t) &= \frac{y_{\text{LF}}(t)}{\beta + \gamma} - \alpha, \\
\tilde{x}(t) &= \frac{y_{\text{HF}}(t)}{\alpha + r(t)}.
\end{aligned}$$

6: The signal $\tilde{x}(t)$ is used to estimate $\{t_k\}_{k \in \mathbb{Z}^+}$ or $s(t)$, $p(t)$ and T using onset detection algorithm in 4.5.3. Pulse $p(t)$ is used to estimate γ as

$$\gamma = \int_0^T p(t)dt.$$

7: To incorporate pulse width modulation, using (4.5)

$$\hat{x}(t) = \sum_k p\left(\frac{T}{t_{k+1} - t_k}t\right) \otimes \delta(t - t_k).$$

8: Synthesize $\hat{y}(t)$ as follows

$$\begin{aligned}
y_{\text{LF}}(t) &= \big[\beta + \gamma\big]\big[\alpha + r(t)\big] \\
y_{\text{HF}}(t) &= \big[\alpha + r(t)\big]\big[\hat{x}(t) - \gamma\big] \\
\hat{y}(t) &= \tau(t) + y_{\text{LF}}(t) + y_{\text{HF}}(t)
\end{aligned}$$

For the PAP signal, the onset detection is performed by using a surrogate signal defined as the Slope sum function (SSF) [119]. This SSF is basically the accumulation of non-zero slopes looking back in time over a certain duration (120 ms in this chapter). The SSF signal has a sharp rise after the systolic onset. A threshold is selected (non-adaptive in this case as it is applied on $\tilde{x}(t)$ signal) by taking the average of the first 8 seconds of the SSF signal. If a point exceeds the threshold, a search is performed on the neighborhood (60 ms before and after the crossing) to detect the onset. Once an onset is found, the process of searching for a new onset is paused for the next 500 ms. This method works well for PAP signals because it has a sharp change at systolic onset and the location and shape of the dicrotic notch do not cause trouble in the pulse onset detection procedure.

It should be also made clear that these algorithms are not always 100% correct in detecting onset time. There is false detection of onset time and the synthesized signal quality is compromised at those time instances.

4.6 Performance Evaluation

In this section, the proposed model is applied to experimental PAP and PVP signals. The performance of the new model is evaluated by comparing the model-synthesized signals with their respective experimental counterparts in terms of both signal fidelity and their capabilities in classifying the different anesthetic stages of the animal subjects.

4.6.1 Dataset Description

The duration of PAP and PVP signals under different MAC and PRO levels used in this chapter are shown in Table 4.3. At each stage, the PAP and PVP signals are of the same duration and are synchronized. The sampling rate of all the signals is 1,000 samples/s.

4.6.2 Model Fitting Performance

To evaluate the similarity between the experimental and model-synthesized signals, Pearson's correlation coefficient has been used. It has been used in similar chapters [120, 121]

Table 4.3: PAP and PVP Signal duration (unit: minutes).

Subject	MAC-1	MAC-2	MAC-3	PRO-1	PRO-2	PRO-3
Pig-1	21.08	13.79	2.42	20.58	16.94	9.89
Pig-2	18.74	6.45	22.30	20.01	12.50	13.31
Pig-3	20.84	20.37	19.89	33.43	19.43	30.38
Pig-4	20.12	19.97	10.00	20.39	21.34	20.39

and defined as

$$\rho = \frac{\sum_i (y[i] - \bar{y})(\hat{y}[i] - \bar{\hat{y}})}{\sqrt{\sum_i (y[i] - \bar{y})^2} \sqrt{\sum_i (\hat{y}[i] - \bar{\hat{y}})^2}}. \tag{4.22}$$

where $y[n]$ and $\hat{y}[n]$ are discrete time samples of the signal $y(t)$ and $\hat{y}(t)$, respectively. Also, $\bar{y}$ and $\bar{\hat{y}}$ are the mean of the samples $y[n]$ and $\hat{y}[n]$. A value closer to $\rho = 1$ corresponds to a perfect fit.

The ρ values obtained by fitting the proposed IPFM model to PAP and PVP signals are tabulated in Tables 4.4, and 4.5 respectively. For the PAP signals, the ρ between the synthesized and experimental data is, in general, greater than 0.991 except in one case: Pig-4 at PRO-1 and ρ is still greater than 0.980, which indicates an excellent matching between the synthesized signals and their experimental counterparts. For PVP signals, the ρ values are greater than 0.90 except for five stages (three MAC and two PRO). PVP signals with short duration in time have low correlation values.

Figs. 4.5, and 4.6 show examples of the experimental and synthesized signals for PAP and PVP signals, respectively, both collected from Pig-2 during MAC-2. The synthesized signals match very well with their experimental counterparts, and this provides a visual verification of the ρ or the correlation results in Tables 4.4, and 4.5.

We tried to explore some similar studies to compare the performance with our proposed model. As there does not exist any peripheral blood pressure signal models to compare with, we tried to compare our results with similar studies on other signals. In [120, 121], a model was proposed to reconstruct ECG signals from PPG signals. The reconstruction

Table 4.4: Correlation coefficient ρ for the IPFM model in PAP signal.

Subject	MAC-1	MAC-2	MAC-3	PRO-1	PRO-2	PRO-3
Pig-1	0.994	0.992	0.992	0.995	0.995	0.995
Pig-2	0.996	0.997	0.997	0.992	0.995	0.994
Pig-3	0.999	0.999	0.998	0.995	0.992	0.996
Pig-4	0.998	0.997	0.993	0.983	0.996	0.997

Table 4.5: Correlation coefficient ρ for the PVP signal.

Subject	MAC-1	MAC-2	MAC-3	PRO-1	PRO-2	PRO-3
Pig-1	0.930	0.927	0.784	0.922	0.912	0.905
Pig-2	0.921	0.904	0.913	0.952	0.954	0.941
Pig-3	0.952	0.959	0.952	0.947	0.957	0.940
Pig-4	0.902	0.884	0.828	0.905	0.871	0.879

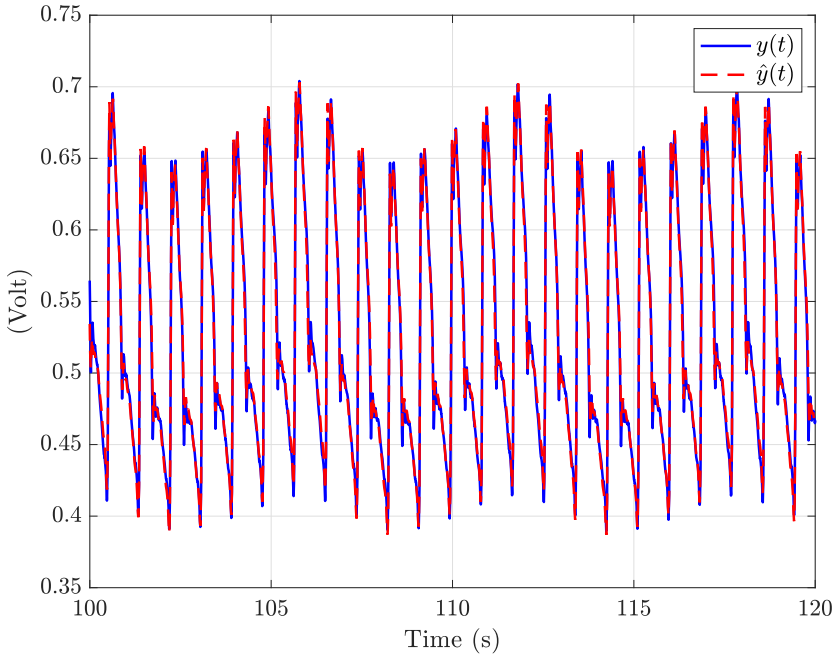

Figure 4.5: The PAP signal $y(t)$ and corresponding synthesized signal $\hat{y}(t)$ from Pig-2 MAC-2 with $\rho = 0.997$.

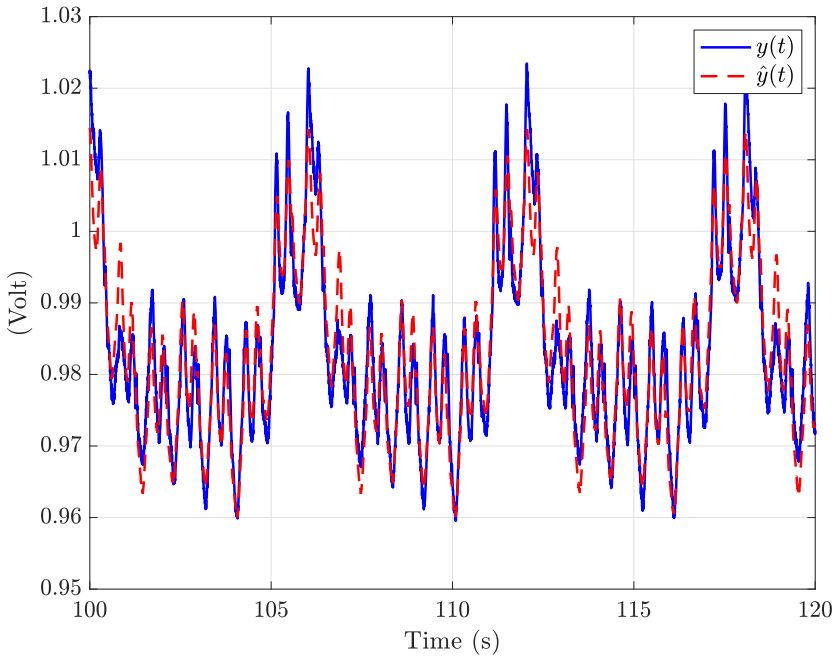

Figure 4.6: The PVP signal $y(t)$ and corresponding synthesized signal $\hat{y}(t)$ for Pig-2 MAC-2 with $\rho = 0.904$.

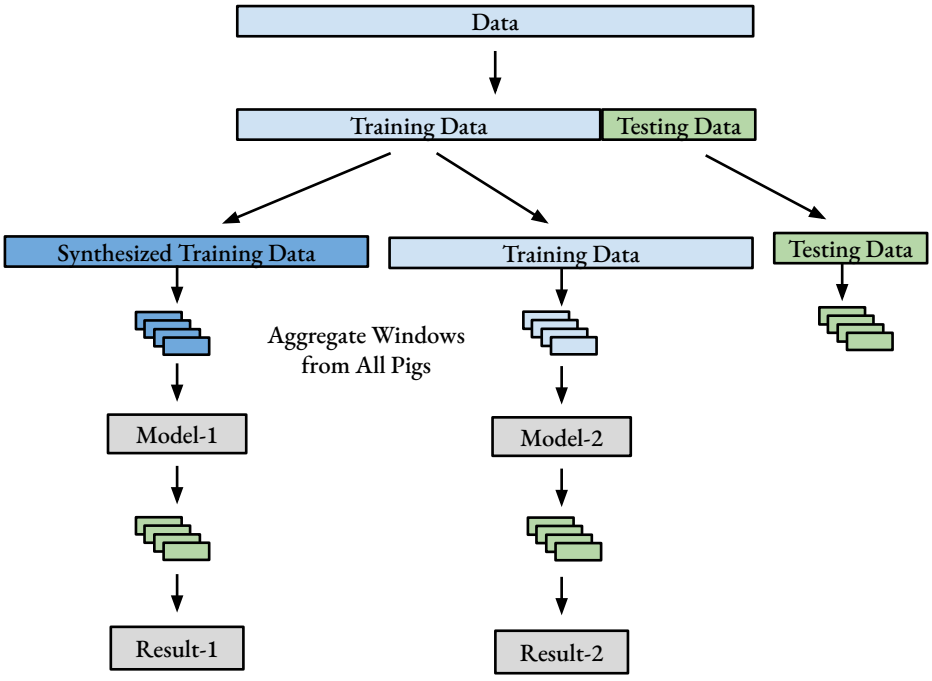

Figure 4.7: This diagram explains the training and testing procedures used in Section 4.6.3. The testing results are reported in Table 4.6.

results are evaluated by using relative root mean-squared error (rRMS) and Pearson's correlation coefficient (ρ). The value of ρ is 0.823 and 0.940 on MIMIC-III dataset under two different segmentation methods used in [121]. In comparison, our average ρ on PAP and PVP signals is 0.995 and 0.914, respectively. We did not use rRMS as a performance metric because it is a biased estimator of performance. Given the contextual nature of model fitting, it is not possible to have a head to head comparison with other models. Given the correlation results, our model works well enough when compared to similar models for other signals.

4.6.3 MAC and PRO Classification Using Logistic Regression

In this subsection, the experimental and synthesized pressure signals are used to classify whether a given pig was injected with Isoflurane (different MAC stages), or both Isoflurane and Propofol (different PRO stages). Both are vasodilators but Isoflurane is inhaled and Propofol is injected into the veins. Distinguishing the impacts of these two different vasodilators on the peripheral signals is important. All MAC stages under Isoflurane are combined as a single class. Similarly, all PRO stages under propofol are combined as a second class. The classification is performed by using logistic regression in the frequency domain. For a given pig and for each stage, the total data is split into 70%-30% for training and testing, respectively. These training and testing time domain signals are divided into overlapping 10 s segments or windows. To make the training and testing procedure unbiased, each stage of each pig is divided into equal numbers (70 windows from the training segment and 30 windows from the testing segment) of overlapping time domain windows. Each time domain segment of 10 s is converted to the frequency domain by using the fast Fourier transform (FFT). Frequency domain magnitudes between 0 and 10 Hz are used as the feature vector for logistic regression. Since the window size is 10 s, the frequency domain resolution is 0.1 Hz, thus there are 100 frequency domain samples between 0 and 10 Hz. Finally, the windows are aggregated based on the class. The total number of windows used for MAC and PRO classes during training are $4\times3\times70 = 840$ each. Similarly, $4\times3\times30 = 360$ windows

are used for each class during testing. Two different classifiers are trained; one uses the experimental data and another one is based on model-synthesized data. The testing is done on experimental testing data for both cases. The whole procedure of classification has been explained in Fig. 4.7. The classification results are shown in Table 4.6. In the tables, the results obtained by using the experimental data are shown in parenthesis, and those obtained by using the model-synthesized data are shown outside of the parenthesis.

The classifier based on model-synthesized PAP signals has a better overall accuracy (92.50%) compared to that trained with the experimental data (90.56%), partly due to the fact the modeling process removed some of the noise and artifacts in the experimental data. This proves the utility of the model-synthesized signal to retain the valuable information embedded in the signal vital for class discrimination. On the other hand, synthesized PVP signal results have less accuracy (85.42%) compared to the experimental PVP signal results (94.44%). This can be understood from the ρ results presented in Table 4.5. Here, five stages (three MAC and two PRO) have comparatively bad fitting ($\rho < 0.89$). PVP signals with short duration in time have low correlation values. This relatively bad fitting is probably the reason behind the reduced accuracy in the PVP signal with the model-synthesized data.

Table 4.6: MAC and PRO classification results. (The results obtained by using the raw experimental data are shown in parenthesis, and those obtained by using the model-based synthesized data are shown outside of the parenthesis.)

(a) MAC and PRO classification using PAP data.

-	MAC Estimated	PRO Estimated
MAC	93.33% (90.56%)	6.67% (9.44%)
PRO	8.33% (9.44%)	91.67% (90.56%)
Accuracy	92.50% (90.56%)	

(b) MAC and PRO classification using PVP data.

-	MAC Estimated	PRO Estimated
MAC	85.56% (98.33%)	14.44% (1.67%)
PRO	14.72% (9.44%)	85.28% (90.56%)
Accuracy	85.42% (94.44%)	

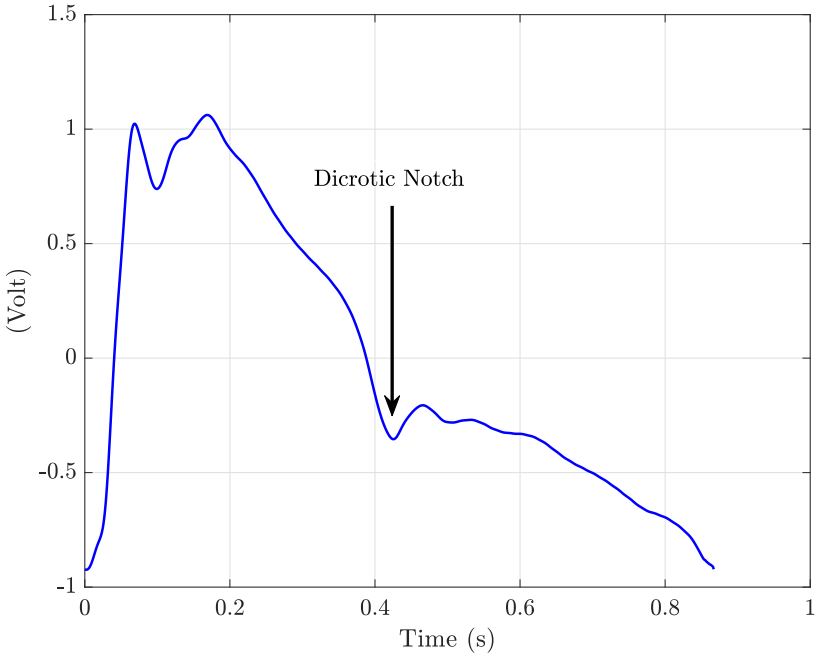

Figure 4.8: Dicrotic notch in a PAP pulse $p(t)$ in Pig-2 MAC-2. This notch marks the end of systole and the beginning of diastole.

4.6.4 Proposed Signal Model on CHARIS Dataset

The proposed IPFM model is also applied to an open-source blood pressure dataset, the CHARIS dataset [122]. This dataset is part of the larger collection of biomedical signals in the PhysioNet database [123]. CHARIS contains arterial blood pressure recordings from radial arteries of 13 traumatic brain injury human subjects. The signals were sampled at 50 samples/s. Each recording has 1 million samples (5.56 hours of recording). The data contain collection anomalies such as abrupt signal level changes and signal discontinuities. In this chapter, the IPFM model is fitted to 15-minute continuous snippets from each of the recordings, and the correlation coefficient (ρ) is presented in Table 4.7. The ρ values are consistently above 0.95 for all 13 human subjects. Thus the proposed model is effective in modeling both pig and human ABP signals. Only 15-minute signals are used because the proposed IPFM-based model assumes the pulse shape to be stationary. For onset detection, "findpeaks()" function of Matlab has been used as described in 4.5.3.

Table 4.7: Correlation coefficient ρ for the IPFM model in CHARIS data with human subjects.

Subject	ρ	Subject	ρ	Subject	ρ
1	0.959	6	0.989	11	0.978
2	0.960	7	0.992	12	0.974
3	0.988	8	0.990	13	0.989
4	0.964	9	0.988	-	-
5	0.983	10	0.990	-	-

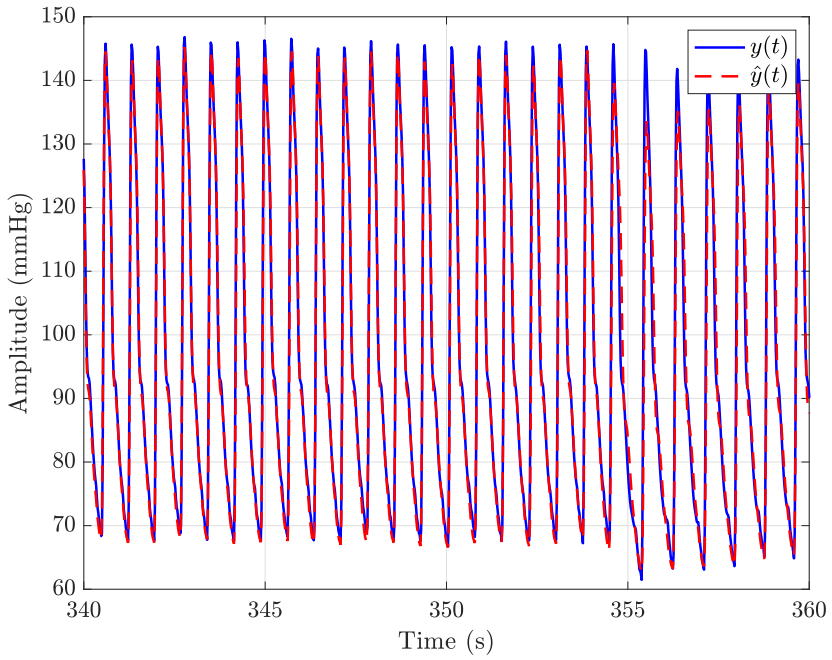

Figure 4.9: Comparison of arterial pressure signal $y(t)$ and corresponding model fit $\hat{y}(t)$ from subject 10 of the CHARIS dataset. This 15-minute data fit has $\rho = 0.990$.

4.7 Significance of Model Parameters

This section discusses the significance of different model parameters under different physiological conditions both qualitatively and quantitatively.

4.7.1 Unit Pulse Shapes

For each pulse, the notch after the systolic peak is called the dicrotic notch (Fig. 4.8). This marks the ending of the systole where the aortic valve closes and the beginning of diastole and ventricular filling. With increasing MAC and PRO levels, the dicrotic notch in $p(t)$ becomes less prominent in the PAP signals (Fig. 4.10). A similar phenomenon is observed in peripheral arterial and peripheral venous Doppler waveforms [124] and in arterial signals [125]. The dicrotic notch is visible or prominent when the resistance in the arteries and veins is dominant. At a lower resistance, the dicrotic notch is absent and the visual demarcation of systole and diastole is lost. Both Propofol and Isoflurane are vasodilators. Increase in the anesthetic level results in widening of blood vessel and lower vessel impedance. The variations in pulse shape at different anesthesia levels provide the distinguishing features that can be exploited by the classifier. The notch is less visually prominent in PRO signals. It is interesting to note that in PVP signals we do not see any consistent notch and there is no consistent visual distinction in pulses (Fig. 4.11). This seems quite obvious because the effect of aortic valve closing does not directly impact the PVP signal. There are other factors that contribute to the classification of these signals but pulse shape seems to be a dominant characteristic in PAP signals. It can also be said that under dehydration, the PAP signal shows strong docrotic notch. On the other hand, the dicrotic notch loses its visual prominence with hydration.

4.7.2 Variation in Pulse Duration

The inverse of the pulse duration (T) is the heart rate ($\frac{1}{T}$). The heart rates of various signals at different anesthesia levels are shown in Fig. 4.12. The heart rate decreases with

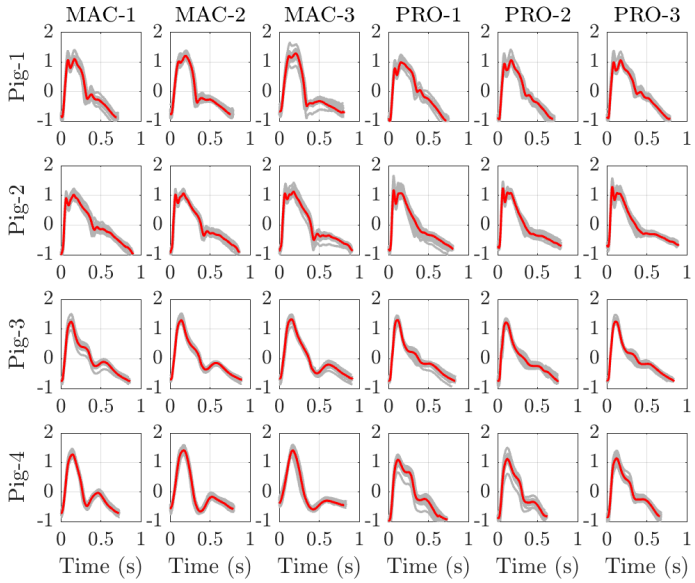

Figure 4.10: Pulse shapes in PAP signals at the different anesthetic stages. The red pulses are the aggregate averages of gray pulses and correspond to $\tilde{p}(t)$ of the proposed model, not $p(t)$.

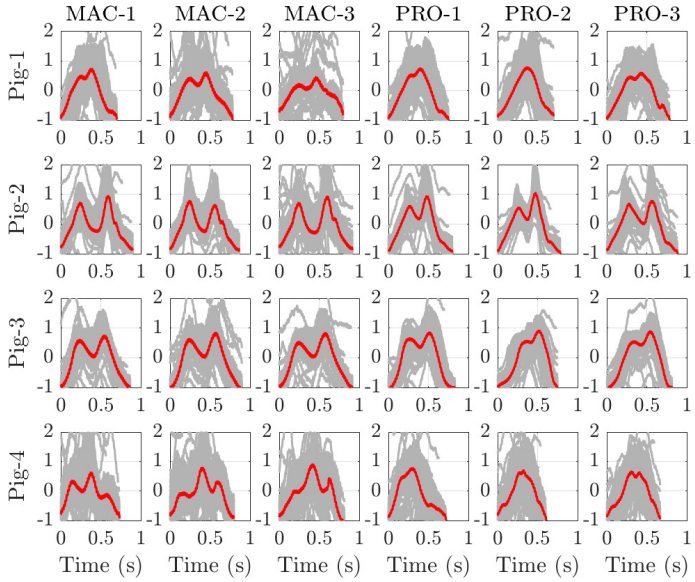

Figure 4.11: Pulse shapes in PVP signals at the different anesthetic stages. The red pulses are the aggregate averages of gray pulses and correspond to $\tilde{p}(t)$ of the proposed model, not $p(t)$.

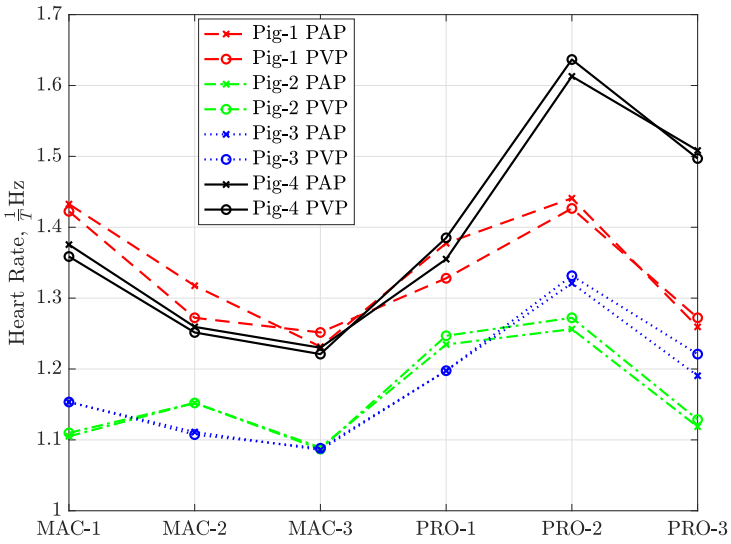

Figure 4.12: Change in pulse duration (T) of $\hat{y}(t)$ under different levels of Isoflurane and Propofol based on PAP, and PVP signals.

the increase of the MAC levels. This is also similar to what happens during dehydration. Under dehydration, the heart beats faster to maintain the cardiac output. With the increase of MAC levels, the heart rate slows down, which is similar to increased hydration. After introducing Propofol, the heart rate initially increases at PRO-1 and PRO-2 but drops at PRO-3. This trend is consistent in all pigs. A similar explanation is not valid for PRO-1 stage, because the trend of heart rate is different here partly due to the fact that before PRO-1 there were other two stages of Isoflurane (2.00% and 1.50%) and the residual effect of these stages (see Table 4.2). We presume that this is due to the dominant effects of the large doses of vasodilators at this stage. It increases at PRO-2 and decreases at PRO-3, this follows the pattern of heart rate decrease at MAC stages. Initial heart rate increase (at MAC-1 and PRO-1) should be compensatory to vasodilation to maintain cardiac output. The sharp final heart rate drop at PRO-3 is most likely due to the fact that the heart loses its ability to beat spontaneously for higher doses of anesthesia [126].

4.7.3 Correlation Between Respiratory and Modulating Signals

To understand the relationship between the respiratory signal $r(t)$ and the modulating signal $m(t)$, a simple Pearson correlation coefficient has been calculated between $m(t)$ and $r(t)$ for each state that is presented in Fig. 4.13. These correlation measures can work as indicators of respiratory-induced heart rate variability or respiratory sinus arrhythmia (RSA). A higher correlation between $m(t)$ and $r(t)$ signals indicates that the HRV caused by $m(t)$ is passively influenced by $r(t)$. Based on the results, it can be inferred qualitatively that the negative correlation in the range -0.35 to -0.75 between $m(t)$ and $r(t)$ signal is strong in PAP MAC signals and becomes weaker in PAP PRO signals. This means that the HRV in PAP signal under MAC doses is influenced by the respiratory rate. This is not the case during the PRO stages because by this time the pigs lose their ability to breathe independently and are supported by mechanical ventilation. On the other hand, the correlation is almost nonexistent in PVP signals. Thus it is better to use PAP signals to measure the strength and sources of HRV compared to PVP signals. Based on the data from the four subjects, a stronger correlation between $r(t)$ and $m(t)$ components of the PAP signal can possibly be a good marker or indicator to distinguish MAC and PRO stages. Both being vasodilators, MAC is showing a strong presence of RSA whereas PRO does not.

4.8 *Derivation of Equation (4.8)*

From (4.3), $\theta(t_k) = kT - t_k$ and $\theta'(t_k) = m(t_k)$. Using the properties of delta function [127], if a function $g(t)$ has only one root at $t = t_k$, then

$$\delta[g(t)] = \frac{1}{|g'(t_k)|}\delta(t - t_k) = \frac{1}{|g'(t)|}\delta(t - t_k),$$

or

$$\delta[g(t)] \cdot |g'(t_k)| = \delta[g(t)] \cdot |g'(t)| = \delta(t - t_k).$$

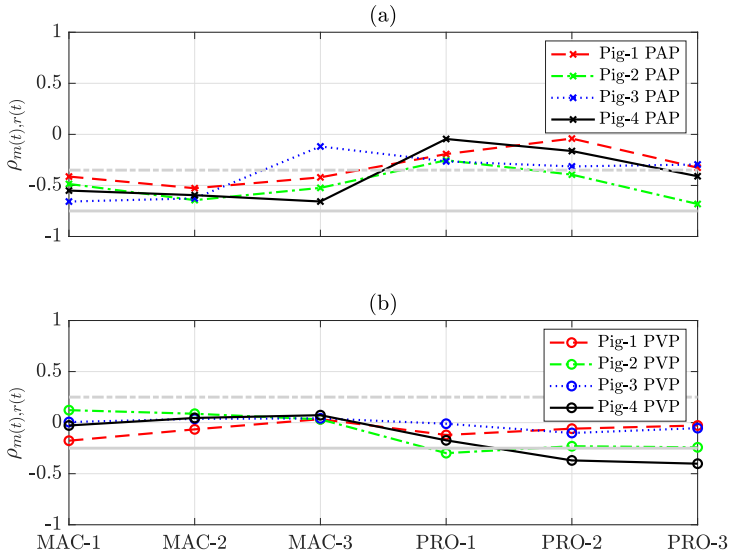

Figure 4.13: Correlation coefficients of different states between $m(t)$ and $r(t)$ signals. Panels (a) and (b) are based on PAP and PVP signals respectively. The gray horizontal lines in (a) represent $\rho = -0.35$ and $\rho = -0.75$ respectively. Similarly, in (b) the gray lines represent $\rho = 0.25$ and $\rho = -0.25$ respectively.

Assuming $g(t) = t - kT + \theta(t)$ yields $g'(t) = 1 + m(t)$. Replacing $g(t)$ and $g'(t)$ in (4.7) gives

$$\begin{aligned} s(t) &= \sum_k \delta(t - t_k), \\ &= |1 + m(t)| \sum_k \delta\big[t - kT + \theta(t)\big]. \end{aligned}$$

Using Fourier series

$$\sum_{k=-\infty}^{\infty} \delta(t - kT) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{jn\frac{2\pi}{T}t},$$

$$\sum_{k=-\infty}^{\infty} \delta\big[t - kT + \theta(t)\big] = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{jn\frac{2\pi}{T}\left[t+\theta(t)\right]}$$

Using the relation above,

$$
\begin{aligned}
s(t) &= \frac{|1+m(t)|}{T} \sum_{n=-\infty}^{\infty} e^{jn\frac{2\pi}{T}\left[t+\theta(t)\right]} \\
&= \frac{|1+m(t)|}{T} \left\{ 1 + 2\sum_{n=1}^{\infty} \cos\left[n\frac{2\pi}{T}\big(t+\theta(t)\big)\right] \right\} \\
&= \frac{1+m(t)}{T} \left\{ 1 + 2\sum_{n=1}^{\infty} \cos\left[n\frac{2\pi}{T}\big(t+\theta(t)\big)\right] \right\}
\end{aligned}
$$

where the last equality follows from the fact that $|1+m(t)| = 1+m(t)$ as $|m(t)| \ll 1$. This proof assumes $k \in \mathbb{Z}$. Whereas in the rest of the chapter assumes $k \in \mathbb{Z}^+$. The only reason to assume $k \in \mathbb{Z}^+$ is for practical numerical convenience without loss of generality.

4.9 Conclusion

In this chapter, we have proposed a new integral pulse frequency modulation-based model for peripheral arterial and venous pressure signals. The model was developed by using arterial and venous pressure data collected from in vivo porcine animal experiments with four female pigs under different levels of anesthesia.

Under proposed model assumption, respiratory signal $r(t)$ is responsible for RSA and modulating signal $m(t)$ is responsible for HRV. The high frequency components of the PVP or PAP signal are both functions of $r(t)$ and $m(t)$. The proposed model isolates these two signals. If the RSA is the sole reason behind HRV then the correlation between $m(t)$ and $r(t)$ should reflect that. To the best of our knowledge, this is the first mathematical model that tries to isolate these two components in the peripheral blood pressure signals.

During dehydration, the blood vessels constrict. On the other hand, vasodialators like Isoflurane and Propofol used in this study broaden the blood vessels. So, increasing doses of vasodilators have similar effect of going from dehydration to hydration. Thus the results obtained in this study are applicable to dehydration-like scenario too. This is why vasodilators were used in this study.

The model captures the impacts of multiple physiological factors, such as the heartbeat pulse shape variation, heart rate variation, and correlation between $m(t)$ and $r(t)$. With model parameters estimated from the experimental data, it has been demonstrated that the model-synthesized data match very well with their experimental counterparts, especially for PAP signals. The correlation coefficients of PAP signals are consistently above 0.99 for almost all signals. The correlation coefficient values drop to around 0.90 for PVP signals. Similar results (correlation coefficient > 0.95) have been obtained using the CHARIS dataset of human arterial blood pressure signals. We strongly believe this model can be extended to other similar biomedical signals like PPG, cerebral blood flow velocity, and Doppler waveforms related to cardiac physiology.

The model-synthesized data and experimental data were also used to train a classifier to distinguish between two different types of anesthesia. The proposed model was able to retain meaningful features and achieved comparable classification results with experimental data. The results indicate that both PAP signal and PVP signals can be used to classify different anesthesia levels (and similarly can be used for dehydration). This chapter also suggests a possible way to distinguish respiratory-induced heart rate variability from other causes.

For future work, the classification performance can be improved by using information from multiple signals collected from peripheral regions in a minimally invasive way, instead of just using one signal. In addition, we would like to further improve the model by addressing some of its current limitations. Some limitations of the current model include: (1) The pulse shape is assumed stationary across the entire signal duration. A more effective approach would be to divide the signal into segments and consider stationary pulse shapes in each smaller time segment. (2) The assumption of $e_u(t) = -e_l(t)$ was used in the model development. This assumption is not always valid. A more general approach would be to assume $e_u(t) = \phi e_l(t)$, where $\phi \in \mathbb{R}$. A more efficient upper and lower envelope detection algorithm can help to further improve the estimation accuracy of the modulating signal $m(t)$. (3) A better pulse onset detection method can be used for PVP signals. (4) Although time was

allotted for inhalation concentration to fall to sub-therapeutic levels when transitioning from Isoflurane to Propofol, the experimental setup lacked the equipment to confirm the end-tidal concentration of Isoflurane was at or near zero. Residual inhalation anesthesia (Isoflurane) could be contributing to initial Propofol measurements. Ideally resourced, future work would examine intravenous agents independently from the moment of induction.

All supplementary materials and codes of this chapter can be found at this Github repository: `https://github.com/i2pt/MPAVPSwIPFM`.

5 Conclusion

This chapter summarizes the main contributions of this dissertation on peripheral blood pressure signal processing. Some possible directions of future research has also been discussed that might help the new researchers in this field.

5.1 Contribution

The prime objective of this research is to identify or find relationship between PVP signals with dehydration. This research is based on two different dataset. One is collected from human and another one from pigs.

For human dataset, the data is collected using regular PIV cathteres, this runs into the problem of blood clotting. Due to the nature of the PVP signal and blood clotting issue, we have to detect anomalies in the signal before performing any analysis. In this research, we have proposed an anomaly detection algorithm using hidden Markov model and Kalman filter to discard the part of the signal that is corrupted. This method is objective and free from human bias. Using the PVP signal from human subjects, we have been able to classify hydrated and dehydrated subjects using regularized logistic regression with high accuracy.

Using the other dataset collected from pigs, we have developed a model of the PAP and PVP signals based on integral pulse frequency modulation (IPFM). This model can successfully separate and find out relationship between HRV and respiration. It should be made clear that the signal was collected at brachial artery and brachial vein using Millar catheter which is different from regular PIV catheters. This is why the nature of the PVP signal in this dataset is different from first dataset collected from human subjects. It does not get corrupted by blood clotting. During the data collection process, two different anesthetics were used. These anesthetics work as vasodilator that broadens the blood vessels. This is just the opposite of dehydration where the blood vessels constrict. With increasing anesthesia doses, the blood vessel is going to be wider. This is similar to what happens during a subject becomes hydrated from dehydration.

We proposed an IPFM-based parametric model for PAP and PVP Signals. Based on

this model, the model parameters are estimated by using experimental data. We demonstrate our model performance using Pearson's correlation coefficient and demonstrated that our model-synthesized data is highly correlated with the experimental data. We also compare the MAC and PRO signal classification with logistic regression. Two logistic regression models were trained by using model-synthesized and experimental data, respectively. The classification results showed that we can achieve similar classification results by using either the model-synthesized data or experimental data. This affirms that the model-synthesized signal retains the class-discriminatory features present in the experimental data. We discussed the utility of the model on CHARIS dataset collected from human subjects. We qualitatively discuss the model estimated pulse shape at different stages of anesthesia and the pattern of model estimated heart rate changes under different anesthetic stages. This discussion is also going to be valuable to subjects changing from dehydrated state to hydrated state because of the similarity of the effect of vasodilators. We also performed quantitative evaluation on the effects of HRV and respiratory rate. All these results are basically dependent on the the newly proposed IPFM-model.

5.2 Future Works

The ultimate goal of any long-term future work should include creation of a massive dataset from hydrated and dehydrated subjects under different circumstances (under controlled environment or uncontrolled environment). Using modern feature extraction methods like self-supervised approaches of deep learning, we can extract meaningful features that can be used for various downstream classification tasks. This data driven approach is going to be expensive and time-intensive. In ideal scenario, the dataset should be open and available. Signals with open dataset get a lot more attention from scientific community compared to the signals which do not have any open dataset.

Apart from collecting PVP signals, a careful design of experiment should also include other signals like ECG, PPG, breathing rate and other related signals. This is going to help us understand the interplay between the different signals and to develop a mutivariate

model to understand the physiological changes. Also, this might facilitate the utility of one signal over another under different circumstances. The peripheral blood pressure signals being invasive, it is not going to be a primary choice unless it outperforms other non-invasive signals like PPG. So, there should be a study that compares side-by-side the utility of PVP signal and PPG in predicting dehydration or other similar open-ended problems. This is why collection of other signals along with PVP is also important.

Last but not the least, PVP signals should be used in some interesting applications as done with other signal. For example, there are some signal processing algorithms that can reconstruct PPG signal from EEG signals. A similar algorithm can be developed to prove the utility of PVP signal reconstruction through ECG signals. Other interesting application of ECG signals include predicting the cardiac age from ECG. A similar study using PVP signal, without doubt, is going to be interesting.

Bibliography

[1] E. Pierce. (2006) Diagram of the human heart. [Online]. Available: https://commons.wikimedia.org/wiki/File:Diagram_of_the_human_heart_(cropped).svg

[2] M. Niescierenko and R. Bachur, "Advances in pediatric dehydration therapy," *Current opinion in pediatrics*, vol. 25, no. 3, pp. 304–309, 2013.

[3] T. Hastie and R. Tibshirani, "Friedman," *The Elements of Statistical Learning," Springer*, p. 52, 2001.

[4] J. E. Wathen, T. MacKenzie, and J. P. Bothner, "Usefulness of the serum electrolyte panel in the management of pediatric dehydration treated with intravenously administered fluids," *Pediatrics*, vol. 114, no. 5, pp. 1227–1234, 2004.

[5] M. J. Steiner, D. A. DeWalt, and J. S. Byerley, "Is this child dehydrated?" *Jama*, vol. 291, no. 22, pp. 2746–2754, 2004.

[6] R. Vega and J. R. Avner, "A prospective study of the usefulness of clinical and laboratory parameters for predicting percentage of dehydration in," *Pediatric Emergency (CARE)*, 1997.

[7] R. Vega, "Dehydration, pediatric," *StatPearls. Treasure Island, FL: StatPearls Publishing*, 2018.

[8] H. W. Yang, W. Jeon, Y. G. Min, and J. S. Lee, "Usefulness of end-tidal carbon dioxide as an indicator of dehydration in pediatric emergency departments: A retrospective observational study," *Medicine*, vol. 96, no. 35, 2017.

[9] J. Van den Berg and M. Y. Berger, "Guidelines on acute gastroenteritis in children: a critical appraisal of their quality and applicability in primary care," *BMC family practice*, vol. 12, no. 1, pp. 1–13, 2011.

[10] S. J. Hanson, R. J. Berens, P. L. Havens, M. K. Kim, and G. M. Hoffman, "Effect of volume resuscitation on regional perfusion in dehydrated pediatric patients as measured by two-site near-infrared spectroscopy," *Pediatric emergency care*, vol. 25, no. 3, pp. 150–153, 2009.

[11] R. Wardhan and K. Shelley, "Peripheral venous pressure waveform," *Current Opinion in Anesthesiology*, vol. 22, no. 6, pp. 814–821, 2009.

[12] E. A. Tansey, L. E. Montgomery, J. G. Quinn, S. M. Roe, and C. D. Johnson, "Understanding basic vein physiology and venous blood pressure through simple physical assessments," *Advances in physiology education*, 2019.

[13] D. Chang, P. J. Leisy, J. H. Sobey, S. K. Reddy, C. Brophy, B. D. Alvis, K. Hocking, and M. Polcz, "Physiology and clinical utility of the peripheral venous waveform," *JRSM Cardiovascular Disease*, vol. 9, p. 2048004020970038, 2020.

[14] K. H. Shelley, M. Dickstein, and S. M. Shulman, "The detection of peripheral venous pulsation using the pulse oximeter as a plethysmograph," *Journal of clinical monitoring*, vol. 9, no. 4, pp. 283–287, 1993.

[15] A. A. Alian and K. H. Shelley, "Respiratory physiology and the impact of different modes of ventilation on the photoplethysmographic waveform," *Sensors*, vol. 12, no. 2, pp. 2236–2254, 2012.

[16] I. Liang, S. Nie, L. Hingula, A. J. Shelley, D. G. Silverman, K. H. Shelley, and A. A. Alian, "Impact of different breathing patterns on peripheral venous pressure (pvp)," in *Anesthesia & Analgesia*, vol. 119, 2014.

[17] M. C. Khoo and P. Chalacheva, "Respiratory modulation of peripheral vasoconstriction: a modeling perspective," *Journal of Applied Physiology*, vol. 127, no. 5, pp. 1177–1186, 2019.

[18] D. De Backer and J.-L. Vincent, "Should we measure the central venous pressure to guide fluid management? ten answers to 10 questions," *Critical care*, vol. 22, no. 1, p. 43, 2018.

[19] J. R. Munis, S. Bhatia, and L. J. Lozada, "Peripheral venous pressure as a hemodynamic variable in neurosurgical patients," *Anesthesia & Analgesia*, vol. 92, no. 1, pp. 172–179, 2001.

[20] D. Amar, J. A. Melendez, H. Zhang, C. Dobres, D. H. Leung, and R. E. Padilla, "Correlation of peripheral venous pressure and central venous pressure in surgical patients," *Journal of cardiothoracic and vascular anesthesia*, vol. 15, no. 1, pp. 40–43, 2001.

[21] M. Tugrul, E. Camci, K. Pembeci, A. Al-Darsani, and L. Telci, "Relationship between peripheral and central venous pressures in different patient positions, catheter sizes, and insertion sites," *Journal of cardiothoracic and vascular anesthesia*, vol. 18, no. 4, pp. 446–450, 2004.

[22] R. Desjardins, A. Y. Denault, S. Bélisle, M. Carrier, D. Babin, S. Lévesque, and R. Martineau, "Can peripheral venous pressure be interchangeable with central venous pressure in patients undergoing cardiac surgery?" *Intensive care medicine*, vol. 30, no. 4, pp. 627–632, 2004.

[23] A. Anter and R. Bondok, "Peripheral venous pressure is an alternative to central venous pressure in paediatric surgery patients," *Acta anaesthesiologica scandinavica*, vol. 48, no. 9, pp. 1101–1104, 2004.

[24] N. Hadimioglu, Z. Ertug, A. Yegin, S. Sanli, A. Gurkan, and A. Demirbas, "Correlation of peripheral venous pressure and central venous pressure in kidney recipients," in *Transplantation proceedings*, vol. 38, no. 2. Elsevier, 2006, pp. 440–442.

[25] N. Hoftman, M. Braunfeld, G. Hoftman, and A. Mahajan, "Peripheral venous pressure as a predictor of central venous pressure during orthotopic liver transplantation," *Journal of clinical anesthesia*, vol. 18, no. 4, pp. 251–255, 2006.

[26] S. J. Choi, M. S. Gwak, J. S. Ko, G. S. Kim, T. H. Kim, H. Ahn, J. A. Kim, M. Yang, S. Lee, and M. Kim, "Can peripheral venous pressure be an alternative to central venous pressure during right hepatectomy in living donors?" *Liver transplantation*, vol. 13, no. 10, pp. 1414–1421, 2007.

[27] G. Cave and M. Harvey, "The difference between peripheral venous pressure and central venous pressure (cvp) decreases with increasing cvp," *European journal of anaesthesiology*, vol. 25, no. 12, pp. 1037–1040, 2008.

[28] L. Baty, P. Russo, and J. D. Tobias, "Measurement of central venous pressure from a peripheral intravenous catheter following cardiopulmonary bypass in infants and children with congenital heart disease," *Journal of Intensive Care Medicine*, vol. 23, no. 2, pp. 136–142, 2008.

[29] H. Amoozgar, G. H. Ajami, M. Borzuoee, A. Amirghofran, and P. Ebrahimi, "Peripheral venous pressure as a predictor of central venous pressure in continuous monitoring in children," *Iranian Red Crescent Medical Journal*, vol. 13, no. 5, p. 342, 2011.

[30] P. P. Vlismas, E. Wiesenfeld, K. T. Oh, S. Murthy, S. Vukelic, O. Saeed, S. Patel, J. J. Shin, U. P. Jorde, and D. B. Sims, "Relation of peripheral venous pressure to central venous pressure in patients with heart failure, heart transplant, and left ventricular assist device," *The American Journal of Cardiology*, vol. 138, pp. 80–84, 2021.

[31] S. Kim, S. Park, J. Cui, J. Lee, S. Cho, W. Chae, H. Jin, and K. Hwang, "Peripheral venous pressure as an alternative to central venous pressure in patients undergoing laparoscopic colorectal surgery," *British journal of anaesthesia*, vol. 106, no. 3, pp. 305–311, 2011.

[32] S. Masutani, C. Kurishima, A. Yana, S. Kuwata, Y. Iwamoto, H. Saiki, H. Ishido, and H. Senzaki, "Assessment of central venous physiology of fontan circulation using peripheral venous pressure," *The Journal of thoracic and cardiovascular surgery*, vol. 153, no. 4, pp. 912–920, 2017.

[33] W. Tan, A. Small, R. Gallotti, J. Moore, and J. Aboulhosn, "Peripheral venous pressure accurately predicts central venous pressure in the adult fontan circulation," *International Journal of Cardiology*, vol. 326, pp. 77–80, 2021.

[34] B. W. Sperry, J. Campbell, M. Yanavitski, S. Kapadia, W. W. Tang, and M. Hanna, "Peripheral venous pressure measurements in patients with acute decompensated heart failure (pvp-hf)," *Circulation: Heart Failure*, vol. 10, no. 7, p. e004130, 2017.

[35] A. A. Alian, N. J. Galante, N. S. Stachenfeld, D. G. Silverman, and K. H. Shelley, "Impact of lower body negative pressure induced hypovolemia on peripheral venous pressure waveform parameters in healthy volunteers," *Physiological Measurement*, vol. 35, no. 7, p. 1509, 2014.

[36] A. Alian, K. Shelley, and H.-T. Wu, "Amplitude and phase measurements from harmonic analysis may lead to new physiologic insights: lower body negative pressure photoplethysmographic waveforms as an example," *Journal of Clinical Monitoring and Computing*, pp. 1–11, 2022.

[37] H.-T. Wu, A. Alian, and K. Shelley, "A new approach to complicated and noisy physiological waveforms analysis: peripheral venous pressure waveform as an example," *Journal of Clinical Monitoring and Computing*, vol. 35, no. 3, pp. 637–653, 2021.

[38] K. M. Hocking, B. Sileshi, F. J. Baudenbacher, R. B. Boyer, K. L. Kohorst, C. M. Brophy, and S. S. Eagle, "Peripheral venous waveform analysis for detecting hemorrhage and iatrogenic volume overload in a porcine model," *Shock*, vol. 46, no. 4, pp. 447–452, 2016.

[39] K. Hocking, B. Alvis, F. Baudenbacher, R. Boyer, C. Brophy, I. Beer, and S. Eagle, "Peripheral iv analysis (piva) of venous waveforms for volume assessment in patients undergoing haemodialysis," *BJA: British Journal of Anaesthesia*, vol. 119, no. 6, pp. 1135–1140, 2017.

[40] M. Miles, B. D. Alvis, K. Hocking, F. Baudenbacher, C. Guth, J. Lindenfeld, C. Brophy, and S. Eagle, "Peripheral intravenous volume analysis (piva) for quantitating volume overload in patients hospitalized with acute decompensated heart failure—a pilot study," *Journal of Cardiac Failure*, vol. 24, no. 8, pp. 525–532, 2018.

[41] P. C. Bonasso, M. S. Dassinger, M. O. Jensen, S. D. Smith, J. M. Burford, and K. W. Sexton, "Optimizing peripheral venous pressure waveforms in an awake pediatric patient by decreasing signal interference," *Journal of clinical monitoring and computing*, vol. 32, no. 6, pp. 1149–1153, 2018.

[42] S. Raju, W. Crim, and W. Buck, "Factors influencing peripheral venous pressure in an experimental model," *Journal of Vascular Surgery: Venous and Lymphatic Disorders*, vol. 5, no. 6, pp. 864–874, 2017.

[43] "Volumetrix - the next wave in vital monitoring," https://www.volumetrix.com/, accessed: 2022-08-20.

[44] D. Hernando, R. McCallister, J. Lázaro, K. Hocking, E. Gil, B. Alvis, P. Laguna, C. Brophy, and R. Bailón, "Validity of venous waveform signal for heart rate variability monitoring," in *2018 Computing in Cardiology Conference (CinC)*, vol. 45. IEEE, 2018, pp. 1–4.

[45] A. Alalawi, *Using Peripheral Venous Pressure Waveforms to Predict Key Hemodynamic Parameters*. University of Arkansas, 2019.

[46] A. Z. Al-Alawi, K. R. Henry, L. D. Crimmins, P. C. Bonasso, M. A. Hayat, M. S. Dassinger, J. M. Burford, H. K. Jensen, J. Sanford, J. Wu *et al.*, "Anesthetics af-

fect peripheral venous pressure waveforms and the cross-talk with arterial pressure," *Journal of clinical monitoring and computing*, vol. 36, no. 1, pp. 147–159, 2022.

[47] L. D. Crimmins-Pierce, G. P. Bonvillain, K. R. Henry, M. A. Hayat, A. A. Villafranca, S. E. Stephens, H. K. Jensen, J. A. Sanford, J. Wu, K. W. Sexton *et al.*, "Critical information from high fidelity arterial and venous pressure waveforms during anesthesia and hemorrhage," *Cardiovascular Engineering and Technology*, pp. 1–13, 2022.

[48] J. R. Mitchell and J.-J. Wang, "Expanding application of the wiggers diagram to teach cardiovascular physiology," *Advances in physiology education*, vol. 38, no. 2, pp. 170–175, 2014.

[49] J. N. Friedman, R. D. Goldman, R. Srivastava, and P. C. Parkin, "Development of a clinical dehydration scale for use in children between 1 and 36 months of age," *The Journal of pediatrics*, vol. 145, no. 2, pp. 201–207, 2004.

[50] R. D. Goldman, J. N. Friedman, and P. C. Parkin, "Validation of the clinical dehydration scale for children with acute gastroenteritis," *Pediatrics*, vol. 122, no. 3, pp. 545–549, 2008.

[51] A. P. Yoganathan, R. Gupta, and W. H. Corcoran, "Fast fourier transform in the analysis of biomedical data," *Medical and biological engineering*, vol. 14, no. 2, pp. 239–245, 1976.

[52] D. J. Wilkinson, R. A. Chapman, A. Owen, S. Olpin, and S. S. Marven, "Hypertrophic pyloric stenosis: predicting the resolution of biochemical abnormalities," *Pediatric surgery international*, vol. 27, no. 7, pp. 695–698, 2011.

[53] N. D. Taylor, D. T. Cass, and A. J. Holland, "Infantile hypertrophic pyloric stenosis: has anything changed?" *Journal of Paediatrics and Child Health*, vol. 49, no. 1, pp. 33–37, 2013.

[54] G. J. Tutay, G. Capraro, B. Spirko, J. Garb, and H. Smithline, "Electrolyte profile of pediatric patients with hypertrophic pyloric stenosis," *Pediatric Emergency Care*, vol. 29, no. 4, pp. 465–468, 2013.

[55] S. Pandya and K. Heiss, "Pyloric stenosis in pediatric surgery: an evidence-based review," *Surgical Clinics*, vol. 92, no. 3, pp. 527–539, 2012.

[56] B. G. Dalton, K. W. Gonzalez, S. R. Boda, P. G. Thomas, A. K. Sherman, and S. D. S. Peter, "Optimizing fluid resuscitation in hypertrophic pyloric stenosis," *Journal of pediatric surgery*, vol. 51, no. 8, pp. 1279–1282, 2016.

[57] C. Howe and L. Le Quesne, "Pyloric stenosis: The metabolic effects," *Journal of British Surgery*, vol. 51, no. 12, pp. 923–932, 1964.

[58] G. Aspelund and J. C. Langer, "Current management of hypertrophic pyloric stenosis," in *Seminars in pediatric surgery*, vol. 16, no. 1. Elsevier, 2007, pp. 27–33.

[59] C. Vlachopoulos, M. O'Rourke, and W. W. Nichols, *McDonald's blood flow in arteries: theoretical, experimental and clinical principles*. CRC press, 2011.

[60] J. Gareth, W. Daniela, H. Trevor, and T. Robert, *An introduction to statistical learning: with applications in R*. Spinger, 2013.

[61] T. Hastie, R. Tibshirani, and M. Wainwright, "Statistical learning with sparsity," *Monographs on statistics and applied probability*, vol. 143, p. 143, 2015.

[62] F. J. Massey Jr, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[63] A. C. Levine, S. P. Shah, I. Umulisa, R. B. Mark Munyaneza, J. M. Dushimiyimana, K. Stegmann, J. Musavuli, P. Ngabitsinze, S. Stulac, H. M. Epino *et al.*, "Ultrasound assessment of severe dehydration in children with diarrhea and vomiting," *Academic Emergency Medicine*, vol. 17, no. 10, pp. 1035–1041, 2010.

[64] D. L. Wyrick, S. D. Smith, J. M. Burford, C. J. Swearingen, and M. S. Dassinger, "Surgeon-performed bedside ultrasound to assess volume status: a feasibility study," *Pediatric surgery international*, vol. 31, no. 12, pp. 1165–1169, 2015.

[65] L. Chen, A. Hsiao, M. Langhan, A. Riera, and K. A. Santucci, "Use of bedside ultrasound to assess degree of dehydration in children with gastroenteritis," *Academic Emergency Medicine*, vol. 17, no. 10, pp. 1042–1047, 2010.

[66] D. El Amrousy, R. Gamal, S. Elrifaey, and S. Hassan, "Non-invasive assessment of significant dehydration in infants using the inferior vena cava to aortic ratio: is it useful?" *Journal of pediatric gastroenterology and nutrition*, vol. 66, no. 6, pp. 882–886, 2018.

[67] J. Jauregui, D. Nelson, E. Choo, B. Stearns, A. C. Levine, O. Liebmann, and S. P. Shah, "The buddy (bedside ultrasound to detect dehydration in youth) study," *Critical ultrasound journal*, vol. 6, no. 1, pp. 1–8, 2014.

[68] A. G. Kovách, P. Sándor, and M. Kollai, *Cardiovascular Physiology Neural Control Mechanisms: Proceedings of the 28th International Congress of Physiological Sciences, Budapest, 1980*. Elsevier, 2014, vol. 9.

[69] J. P. Mynard and J. J. Smolich, "One-dimensional haemodynamic modeling and wave dynamics in the entire adult circulation," *Annals of biomedical engineering*, vol. 43, no. 6, pp. 1443–1460, 2015.

[70] B. Sileshi, K. M. Hocking, R. B. Boyer, F. J. Baudenbacher, K. L. Kohurst, C. M. Brophy, and S. Eagle, "Peripheral venous waveform analysis for detecting early hemorrhage: a pilot study," *Intensive care medicine*, vol. 41, no. 6, pp. 1147–1148, 2015.

[71] D. Farcy, A. Jain, M. Dalley, and T. Scalea, "Pitfalls in using central venous pressure as a marker of fluid responsiveness," *Emerg Med*, vol. 48, no. 1, pp. 18–28, 2016.

[72] P. E. Marik and R. Cavallazzi, "Does the central venous pressure predict fluid responsiveness? an updated meta-analysis and a plea for some common sense," *Critical care medicine*, vol. 41, no. 7, pp. 1774–1781, 2013.

[73] L. Paladino, R. Sinert, D. Wallace, T. Anderson, K. Yadav, and S. Zehtabchi, "The utility of base deficit and arterial lactate in differentiating major from minor injury in trauma patients with normal vital signs," *Resuscitation*, vol. 77, no. 3, pp. 363–368, 2008.

[74] V. A. Convertino, "Blood pressure measurement for accurate assessment of patient status in emergency medical settings," *Aviation, space, and environmental medicine*, vol. 83, no. 6, pp. 614–619, 2012.

[75] P. C. Bonasso, M. S. Dassinger, B. McLaughlin, J. M. Burford, and K. W. Sexton, "Fast fourier transformation of peripheral venous pressure changes more than vital signs with hemorrhage," *Military medicine*, vol. 184, no. Supplement_1, pp. 318–321, 2019.

[76] P. C. Bonasso, K. W. Sexton, M. A. Hayat, J. Wu, H. K. Jensen, M. O. Jensen, J. M. Burford, and M. S. Dassinger, "Venous physiology predicts dehydration in the pediatric population," *Journal of Surgical Research*, vol. 238, pp. 232–239, 2019.

[77] P. C. Bonasso, K. W. Sexton, S. C. Mehl, M. S. Golinko, M. A. Hayat, J. Wu, M. O. Jensen, S. D. Smith, J. M. Burford, and M. S. Dassinger, "Lessons learned measuring peripheral venous pressure waveforms in an anesthetized pediatric population," *Biomedical Physics & Engineering Express*, vol. 5, no. 3, p. 035020, 2019.

[78] P. C. Bonasso, K. W. Sexton, A. Hayat, A. Al-Alawi, J. Wu, H. K. Jensen, M. O. Jensen, S. D. Smith, J. M. Burford, and M. S. Dassinger, "Venous physiology predicts anesthetic induced hypotension in infants," *Journal of the American College of Surgeons*, vol. 227, no. 4, p. e116, 2018.

[79] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[80] J. Lee, D. D. McManus, S. Merchant, and K. H. Chon, "Automatic motion and noise artifact detection in holter ecg data using empirical mode decomposition and statistical approaches," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 6, pp. 1499–1506, 2011.

[81] M. R. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in ppg signals based on as-lms adaptive filter," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 5, pp. 1445–1457, 2011.

[82] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "Long short term memory networks for anomaly detection in time series," in *Proceedings*. Presses universitaires de Louvain, 2015, p. 89.

[83] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "Lstm fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, 2017.

[84] G. Petris, S. Petrone, and P. Campagnoli, "Dynamic linear models," in *Dynamic Linear Models with R*. Springer, 2009, pp. 31–84.

[85] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[86] A. Krogh, B. Larsson, G. Von Heijne, and E. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden markov model: application to complete genomes," *Journal of molecular biology*, vol. 305, no. 3, pp. 567–580, 2001.

[87] V. Narasimhan, P. Danecek, A. Scally, Y. Xue, C. Tyler-Smith, and R. Durbin, "Bcftools/roh: a hidden markov model approach for detecting autozygosity from next-generation sequencing data," *Bioinformatics*, vol. 32, no. 11, pp. 1749–1751, 2016.

[88] M. Gales, S. Young *et al.*, "The application of hidden markov models in speech recognition," *Foundations and Trends® in Signal Processing*, vol. 1, no. 3, pp. 195–304, 2008.

[89] C.-M. Chen, D.-J. Guan, Y.-Z. Huang, and Y.-H. Ou, "Anomaly network intrusion detection using hidden markov model," *Int. J. Innov. Comput. Inform. Control*, vol. 12, pp. 569–580, 2016.

[90] C. Y. Goh, J. Dauwels, N. Mitrovic, M. T. Asif, A. Oran, and P. Jaillet, "Online map-matching based on hidden markov model for real-time traffic sensing applications," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*. IEEE, 2012, pp. 776–781.

[91] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *The annals of mathematical statistics*, vol. 41, no. 1, pp. 164–171, 1970.

[92] G. Petris and R. An, "An r package for dynamic linear models," *Journal of Statistical Software*, vol. 36, no. 12, pp. 1–16, 2010.

[93] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[94] W. Zucchini, I. L. MacDonald, and R. Langrock, *Hidden Markov models for time series: an introduction using R*. Chapman and Hall/CRC, 2016.

[95] M. A. Hayat, J. Wu, P. C. Bonasso, K. W. Sexton, H. K. Jensen, M. S. Dassinger, and M. O. Jensen, "Unsupervised anomaly detection in peripheral venous pressure signals with hidden markov models," *Biomedical Signal Processing and Control*, vol. 62, p. 102126, 2020.

[96] A. Z. Al-Alawi, K. R. Henry, L. D. Crimmins, P. C. Bonasso, M. A. Hayat, M. S. Dassinger, J. M. Burford, H. K. Jensen, J. Sanford, J. Wu *et al.*, "Anesthetics affect peripheral venous pressure waveforms and the cross-talk with arterial pressure," *Journal of Clinical Monitoring and Computing*, pp. 1–13, 2021.

[97] L. Nilsson, A. Johansson, and S. Kalman, "Respiratory variations in the reflection mode photoplethysmographic signal. relationships to peripheral venous pressure," *Medical and Biological Engineering and Computing*, vol. 41, no. 3, pp. 249–254, 2003.

[98] D. Martin-Martinez, P. Casaseca-de-la Higuera, M. Martin-Fernandez, and C. Alberola-López, "Stochastic modeling of the ppg signal: a synthesis-by-analysis approach with applications," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 9, pp. 2432–2441, 2013.

[99] Q. Tang, Z. Chen, J. Allen, A. Alian, C. Menon, R. Ward, and M. Elgendi, "Ppgsynth: An innovative toolbox for synthesizing regular and irregular photoplethysmography waveforms," *Frontiers in Medicine*, p. 735, 2020.

[100] Q. Tang, Z. Chen, R. Ward, and M. Elgendi, "Synthetic photoplethysmogram generation using two gaussian functions," *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.

[101] M. Chen, Q. Zhu, M. Wu, and Q. Wang, "Modulation model of the photoplethysmography signal for vital sign extraction," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 969–977, 2020.

[102] N. Jafarnia-Dabanloo, D. McLernon, H. Zhang, A. Ayatollahi, and V. Johari-Majd, "A modified zeeman model for producing hrv signals and its application to ecg signal generation," *Journal of theoretical biology*, vol. 244, no. 2, pp. 180–189, 2007.

[103] M. Boulakia, S. Cazeau, M. A. Fernández, J.-F. Gerbeau, and N. Zemzemi, "Mathematical modeling of electrocardiograms: a numerical study," *Annals of biomedical engineering*, vol. 38, no. 3, pp. 1071–1097, 2010.

[104] M. Quiroz-Juárez, O. Jiménez-Ramírez, R. Vázquez-Medina, V. Breña-Medina, J. Aragón, and R. Barrio, "Generation of ecg signals from a reaction-diffusion model spatially discretized," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[105] M. A. Awal, S. S. Mostafa, M. Ahmad, M. A. Alahe, M. A. Rashid, A. Z. Kouzani, and M. P. Mahmud, "Design and optimization of ecg modeling for generating different cardiac dysrhythmias," *Sensors*, vol. 21, no. 5, p. 1638, 2021.

[106] E. J. Bayly, "Spectral analysis of pulse frequency modulation in the nervous systems," *IEEE Transactions on Biomedical Engineering*, no. 4, pp. 257–265, 1968.

[107] O. Rompelman, J. B. Snijders, and C. J. Van Spronsen, "The measurement of heart rate variability spectra with the help of a personal computer," *IEEE Transactions on Biomedical Engineering*, no. 7, pp. 503–510, 1982.

[108] J. Mateo and P. Laguna, "New heart rate variability time-domain signal construction from the beat occurrence time and the ipfm model," in *Computers in Cardiology 1996*. IEEE, 1996, pp. 185–188.

[109] ——, "Improved heart rate variability signal analysis from the beat occurrence times according to the ipfm model," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 8, pp. 985–996, 2000.

[110] ——, "Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal," *IEEE Transactions on biomedical engineering*, vol. 50, no. 3, pp. 334–343, 2003.

[111] R. Bailón, P. Laguna, L. Mainardi, and L. Sornmo, "Analysis of heart rate variability using time-varying frequency bands based on respiratory frequency," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2007, pp. 6674–6677.

[112] J. Lázaro, E. Gil, J. M. Vergara, and P. Laguna, "Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children," *IEEE journal of biomedical and health informatics*, vol. 18, no. 1, pp. 240–246, 2013.

[113] ADInstruments, "Pressure catheters," Accessed: 2023-03-20. [Online]. Available: https://www.adinstruments.com/products/pressure-catheters

[114] P. Laguna and L. Sörnmo, "Modelling heart rate variability," in *Proceedings of the sixteenth international symposium on mathematical theory of networks and systems. Loeven*, 2004, pp. 1–6.

[115] M. Brennan, M. Palaniswami, and P. Kamen, "Do existing measures of poincare plot geometry reflect nonlinear features of heart rate variability?" *IEEE transactions on biomedical engineering*, vol. 48, no. 11, pp. 1342–1347, 2001.

[116] M. S. Manikandan and K. Soman, "A novel method for detecting r-peaks in electrocardiogram (ecg) signal," *Biomedical Signal Processing and Control*, vol. 7, no. 2, pp. 118–128, 2012.

[117] H. S. Shin, C. Lee, and M. Lee, "Adaptive threshold method for the peak detection of photoplethysmographic waveform," *Computers in biology and medicine*, vol. 39, no. 12, pp. 1145–1152, 2009.

[118] S. Asgari, N. Canac, R. Hamilton, and F. Scalzo, "Identification of pulse onset on cerebral blood flow velocity waveforms: A comparative study," *BioMed research international*, vol. 2019, 2019.

[119] W. Zong, T. Heldt, G. Moody, and R. Mark, "An open-source algorithm to detect onset of arterial blood pressure pulses," in *Computers in Cardiology, 2003.* IEEE, 2003, pp. 259–262.

[120] Q. Zhu, X. Tian, C.-W. Wong, and M. Wu, "Ecg reconstruction via ppg: A pilot study," in *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE, 2019, pp. 1–4.

[121] ——, "Learning your heart actions from pulse: Ecg waveform reconstruction from ppg," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16 734–16 748, 2021.

[122] N. Kim, A. Krasner, C. Kosinski, M. Wininger, M. Qadri, Z. Kappus, S. Danish, and W. Craelius, "Trending autoregulatory indices during treatment for traumatic brain injury," *Journal of clinical monitoring and computing*, vol. 30, no. 6, pp. 821–831, 2016.

[123] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[124] E. S. Kim, A. M. Sharma, R. Scissons, D. Dawson, R. T. Eberhardt, M. Gerhard-Herman, J. P. Hughes, S. Knight, A. Marie Kupinski, G. Mahe *et al.*, "Interpretation of peripheral arterial and venous doppler waveforms: A consensus statement from the society for vascular medicine and society for vascular ultrasound," *Vascular Medicine*, vol. 25, no. 5, pp. 484–506, 2020.

[125] M. Nirmalan and P. M. Dark, "Broader applications of arterial pressure wave form analysis," *Continuing Education in Anaesthesia, Critical Care & Pain*, vol. 14, no. 6, pp. 285–290, 2014.

[126] S. Eis and J. Kramer, "Anesthesia inhalation agents cardiovascular effects," in *StatPearls [Internet]*. StatPearls Publishing, 2021, pp. 1–8.

[127] E. W. Weisstein, "Delta function," https://mathworld.wolfram.com/DeltaFunction.html, Accessed: 2023-03-20.

A All Publications Published, Submitted, and Planned

A.1 Journals

1. **M. A. Hayat**, J. Wu, S. Stephens, H. K. Jensen, A. A. Villafranca, J. A. Sanford, K. W. Sexton and M. O. Jensen, "Modeling peripheral arterial and venous pressure signals with integral pulse frequency modulation," *Biomedical Signal Processing and Control*, vol. 86, p. 105240, 2023. https://doi.org/10.1016/j.bspc.2023.105240

2. L. D. Crimmins-Pierce, G. P. Bonvillain, K. R. Henry, **M. A. Hayat** et al., "Critical information from high fidelity arterial and venous pressure waveforms during anesthesia and hemorrhage," *Cardiovascular Engineering and Technology*, vol. 13, no. 6, pp. 886–898, 2022. https://doi.org/10.1007/s13239-022-00624-4

3. A. Z. Al-Alawi, K. R. Henry, L. D. Crimmins, P. C. Bonasso, **M. A. Hayat**, et al., "Anesthetics affect peripheral venous pressure waveforms and the cross-talk with arterial pressure," *Journal of Clinical Monitoring and Computing*, pp. 1–13, 2021. https://doi.org/10.1007/s10877-020-00632-6

4. **M. A. Hayat**, G. Stein, P. Harrington, Z. Lukić, and M. Mustafa, "Self-supervised representation learning for astronomical images," *The Astrophysical Journal Letters*, vol. 911, no. 2, p. L33, 2021. https://doi.org/10.3847/2041-8213/abf2c7

5. **M. A. Hayat**, J. Wu, P. C. Bonasso et al., "Unsupervised anomaly detection in peripheral venous pressure signals with hidden Markov models," *Biomedical Signal Processing and Control*, vol. 62, p. 102126, 2020. https://doi.org/10.1016/j.bspc.2020.102126

6. **M. A. Hayat**, J. Wu, and Y. Cao, "Unsupervised bayesian learning for rice panicle segmentation with UAV images," *Plant Methods*, vol. 16, no. 1, pp. 1–13, 2020. https://doi.org/10.1186/s13007-020-00567-8

7. P. C. Bonasso, K. W. Sexton, **M. A. Hayat**, J. Wu, H. K. Jensen, M. O. Jensen, J. M. Burford, and M. S. Dassinger, "Venous physiology predicts dehydration in the

pediatric population," *Journal of Surgical Research*, vol. 238, pp. 232–239, 2019. https://doi.org/10.1016/j.jss.2019.01.036

8. P. C. Bonasso, K. W. Sexton, S. C. Mehl, M. S. Golinko, **M. A. Hayat**, J. Wu, M. O. Jensen, S. D. Smith, J. M. Burford, and M. S. Dassinger, "Lessons learned measuring peripheral venous pressure waveforms in an anesthetized pediatric population," *Biomedical Physics & Engineering Express*, vol. 5, no. 3, p. 035020, 2019. https://doi.org/10.1088/2057-1976/ab0ea8

A.2 Poster Presentations

1. **M. A. Hayat**, P. Harrington, G. Stein, Z. Lukić, and M. Mustafa, "Estimating galactic distances from images using self-supervised representation learning," Machine Learning and the Physical Sciences Workshop, *36th conference on Neural Information Processing Systems (NeurIPS)*, 2020.

2. **M. A. Hayat**, J. Wu, and Y. Cao, "Rice Panicle Segmentation from UAV Images Using Multivariate Gaussian Mixture Model," *Spring Lecture Series, University of Arkansas*, 2019.

3. K. R. Henry, A. Z. Al-Alawi, **M. A. Hayat** et al., "Novel Method of Detecting the Effect from Inhaled Anesthetics on Peripheral Venous Pressure Waveforms," *Summer Biomechanics, Bioengineering, and Biotransport Conference*, 2019. http://dx.doi.org/10.13140/RG.2.2.36345.31849

4. A. Z. Al-Alawi, K. R. Henry, **M. A. Hayat** et al.,"Propofol Affects Peripheral Venous Tone in Anesthetized Patients," *Midsouth Computational Biology & Bioinfomatics Society Conference*, 2019. http://dx.doi.org/10.13140/RG.2.2.25425.38241

5. P. C. Bonasso, K. W. Sexton, **M. A. Hayat** et al., "Venous physiology predicts anesthetic induced hypotension in infants," *Journal of the American College of Surgeons*, vol. 227, no. 4, p. e116, 2018. http://dx.doi.org/10.1016/j.jamcollsurg.2018.08.313

6. **M. A. Hayat**, J. Wu, H. K. Jensen, M. O. Jensen, M. S. Dassinger, P. C. Bonasso, K. W. Sexton, and J. M. Burford, "Predicting dehydration in pediatric patients with peripheral venous waveforms," *Midsouth Computational Biology & Bioinfomatics Society Conference*, 2018. http://dx.doi.org/10.13140/RG.2.2.35817.54887

Vita

Md Abul Hayat was born and raised in Rajshahi, Bangladesh. He received the following degrees: B.S. in Electrical & Electronic Engineering from Bangladesh University of Engineering & Technology (2015), an M.S. in Statistics & Analytics (2021), and a Ph.D. in Electrical Engineering (2023) from the University of Arkansas. He has worked at Grameenphone - Telenor (2015-17), Nokia Bell Labs (2019), Lawrence Berkeley National Laboratory (2020), and Amazon Web Services (2021).