Cuffless Blood Pressure Estimation

Mémoire

Guillaume Weber-Boisvert

Maîtrise en génie électrique - avec mémoire
Maître ès sciences (M. Sc.)

Québec, Canada

Cuffless Blood Pressure Estimation

Mémoire

Guillaume Weber-Boisvert

Sous la direction de :

Benoit Gosselin

Résumé

L'hypertension est une maladie qui affecte plus d'un milliard de personnes dans le monde. Il s'agit d'une des principales causes de décès; le suivi et la gestion de cette maladie sont donc cruciaux. La technologie de mesure de la pression artérielle la plus répandue, utilisant le brassard pressurisé, ne permet cependant pas un suivi en continu de la pression, ce qui limite l'étendue de son utilisation. Ces obstacles pourraient être surmontés par la mesure indirecte de la pression par l'entremise de l'électrocardiographie ou de la photo-pléthysmographie, qui se prêtent à la création d'appareils portables, confortables et peu coûteux.

Ce travail de recherche, réalisé en collaboration avec le département d'ingénierie biomédicale de l'université de Lund, en Suède, porte principalement sur la base de données publique *Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) Waveform Dataset* de PhysioNet, largement utilisée dans la littérature portant sur le développement et la validation d'algorithmes d'estimation de la pression artérielle sans brassard pressurisé. Puisque ces données proviennent d'unités de soins intensifs et ont été recueillies dans des conditions non contrôlées, plusieurs chercheurs ont avancé que les modèles d'estimation de la pression artérielle se basant sur ces données ne sont pas valides pour la population générale. Pour la première fois dans la littérature, cette hypothèse est ici mise à l'épreuve en comparant les données de MIMIC à un ensemble de données de référence plus représentatif de la population générale et recueilli selon une procédure expérimentale bien définie. Des tests statistiques révèlent une différence significative entre les ensembles de données, ainsi qu'une réponse différente aux changements de pression artérielle, et ce, pour la majorité des caractéristiques extraites du photopléthysmogramme. De plus, les répercussions de ces différences sont démontrées à l'aide d'un test pratique d'estimation de la pression artérielle par apprentissage machine. En effet, un modèle entraîné sur l'un des ensembles de données perd en grande partie sa capacité prédictive lorsque validé sur l'autre ensemble, par rapport à sa performance en validation croisée sur l'ensemble d'entraînement. Ces résultats constituent les contributions principales de ce travail et ont été soumis sous forme d'article à la revue *Physiological Measurement*.

Un volet additionnel de la recherche portant sur l'analyse du pouls par décomposition (*pulse decomposition analysis* ou *PDA*) est présenté dans un deuxième temps. La PDA est une technique permettant de séparer l'onde du pouls en une composante excitative et ses

réflexions, utilisée pour extraire des caractéristiques du signal dans le contexte de l'estimation de la pression artérielle. Les résultats obtenus démontrent que l'estimation de la position temporelle des réflexions à partir de points de référence de la dérivée seconde du signal donne d'aussi bons résultats que leur détermination par la méthode traditionnelle d'approximation successive, tout en étant beaucoup plus rapide. Une méthode récursive rapide de PDA est également étudiée, mais démontrée comme inadéquate dans un contexte de comparaison intersujet.

Abstract

Hypertension affects more than one billion people worldwide. As one of the leading causes of death, tracking and management of the condition is critical, but is impeded by the current cuff-based blood pressure monitoring technology. Continuous and more ubiquitous blood pressure monitoring may be achieved through simpler, cheaper and less invasive cuff-less devices, performing an indirect measure through electrocardiography or photo-plethysmography.

Produced in collaboration with the department of biomedical engineering of Lund University in Sweden, this work focuses on public data that has been widely used in the literature to develop and validate cuffless blood pressure estimation algorithms: The Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) Waveform Dataset from PhysioNet. Because it is sourced from intensive care units and collected in absence of controlled conditions, it has many times been hypothesized that blood pressure estimation models based on its data may not generalize to the normal population. This work tests that hypothesis for the first time by comparing the MIMIC dataset to another reference dataset more representative of the general population and obtained under controlled experimental conditions. Through statistical testing, a majority of photoplethysmogram based features extracted from MIMIC are shown to differ significantly from the reference dataset and to respond differently to blood pressure changes. In addition, the practical impact of those differences is tested through the training and cross validating of machine learning models on both datasets, demonstrating an acute loss of predictive powers of models facing data from outside the dataset used in the training phase. As the main contribution of this work, these findings have been submitted as a journal paper to Physiological Measurement.

Additional original research is also presented in relation to pulse decomposition analysis (PDA), a technique used to separate the pulse wave from its reflections, in the context of blood pressure estimation. The results obtained through this work show that when using the timing of reflections as part of blood pressure predictors, estimating those timings from fiducial points in the second derivative works as well as using the traditional and computationally costly successive approximation PDA method, while being many times faster. An alternative fast recursive PDA algorithm is also presented and shown to perform inadequately in an inter-subject comparison context.

Table of content

List of figures

List of tables

List of abbreviations

- Calib: Calibration
- Eq: Equation
- Opt: Optional

List of acronyms

- AC: Alternating current
- AAMI: Association for the Advancement of Medical Instrumentation
- ABP: Arterial blood pressure
- ABPM: Ambulatory blood pressure monitoring
- AI: Augmentation Index
- ANSI: American National Standards Institute
- ANN: Artificial neural network
- APG: Acceleration photoplethysmography
- AX : Aging index
- BHS: British Hypertension Society
- BP: Blood pressure
- BPM : Beats per minute
- CDC: Center for Disease Control
- CNN: Convolutional neural network
- CVD : Cardiovascular diseases
- DBP: Diastolic blood pressure
- DC : Direct current
- DNN: Deep neural network
- ECG: Electrocardiogram, Electrocardiography
- EHS: European Hypertension Society
- HBPM: Home blood pressure monitoring
- HR: Heart rate
- HRV: Heart rate variability
- ICU : Intensive care unit
- IEEE: Institute of Electrical and Electronics Engineers
- IPA: Inflection point area ratio
- ISO: International Standards Organisation
- KNN: K-nearest neighbors
- KS : Kolmogorov–Smirnov
- LASI: Large artery stiffness index
- LSTM: Long short-term memory (neural network)
- MAD: Mean absolute difference
- MAE: Mean absolute error
- MAP: Mean arterial pressure
- mmHg : Millimeters of mercury (unit)
- MIMIC : Multiparameter Intelligent Monitoring in Intensive Care (dataset)
- mNPV: Modified normalised pulse volume
- NPV: Normalised pulse volume
- PAT: Pulse arrival time
- PCA: Principal component analysis
- PDA: Pulse decomposition analysis
- PEP: Pre-ejection period
- PPG: Photoplethysmogram, Photoplethysmography
- PTT: Pulse transit time

- PWA: Pulse wave analysis
- PWV: Pulse wave velocity
- RBF : Radial basis function
- RI: Reflection index
- SBP: Systolic blood pressure
- SI: Stiffness index
- STD : Standard deviation
- SVD: Singular value decomposition
- SVM: Support vector machine
- SVR : Support vector regression
- UCI : University of California, Irvine (dataset)

Acknowledgments

I would like to thank my research supervisors Benoit Gosselin and Frida Sandberg who supported me with tips, advice, helped me develop my ideas and drove me to always reach for the highest scientific rigor. This project would not have been possible without them. I would also like to thank the SMAART program, Sentinel North and the Canada Research Chair in Smart Biomedical Microsystems for their support.

Foreword

This work was produced in collaboration with Frida Sandberg from the department of biomedical engineering of Lund University, Sweden. She acted in an advisory role, lending her strong background in cardiovascular signal processing to the project.

Chapter two of this work consists of the pre-print version of a journal paper submitted to Physiological Measurement, a part of IOP Science, on 2022/08/02. The formatting was adapted to the present thesis and the acknowledgement section was omitted, but no other changes were made compared to the submitted version. Writing the code and algorithms, performing the experiments, and writing the paper was done by myself in its entirety. B. Gosselin and F. Sandberg's involvement were in providing guidance, ideas, validating the methodology and reviewing the paper throughout the project. In addition, B. Gosselin was involved in managing and ensuring the smooth running of the project.

Reference:

> Guillaume Weber-Boisvert, Benoit Gosselin, Frida Sandberg, "Intensive care photoplethysmogram datasets and machine-learning for blood pressure estimation: generalization not guarantied*", Physiological Measurement, 2022* (Pending review)

Introduction

Hypertension is one of the greatest threats to health of our time. Statistics from the Center for Disease Control (CDC) in the United States provide a glimpse at the extent of the issue. According to the CDC, 45% of the adult population suffer from hypertension, and only 24% of those with hypertension have their condition under control [1]. Worldwide, it is estimated that close to 1 billion people have hypertension. High blood pressure (BP) is known to be an independent risk factor for cardiovascular diseases such as heart attack, stroke, and kidney disease; all leading causes of death in the United States [2], [3], where hypertension also ranks second amongst the preventable causes of death, second only to cigarette smoking [4].

Blood pressure is measured using the millimeter of mercury (mmHG), and the pressure varies through the cardiac cycle. Systolic blood pressure (SBP) is the maximum pressure, which happens when the heart is contracting. Diastolic blood pressure (DBP) is the lowest pressure between two heart beats, when the heart is at rest [5]. Individuals can be classified into different categories, as shown in Table 1, which are defined by different thresholds in blood pressure [6, p. 21]. A meta-analysis comprising data from more than 1 million adults has shown that the risk of cardiovascular disease can more than double for an increase of 20 mmHG of systolic pressure or 10 mmHG of diastolic pressure above normal, depending on age [7].

Table 1: Categories of blood pressure ranges according to the 2017 Clinical Practice Guideline of the American College of Cardiology and the American Heart Association

Category	DBP (mmHg)	SBP (mmHg)
Normal	< 80	< 120
Elevated	< 80	120-129
Stage 1 Hypertension	80-89	130-139
Stage 2 Hypertension	≥ 90	≥ 140

Accurate blood pressure measurement and monitoring is therefore critical to the well-being of the population. For reasons that will be covered in Chapter 1, there is currently a shift in practice towards fully automated and cuffless ambulatory blood pressure monitoring. Those devices would rely on indirect BP measurement through other biological signals such as through electrocardiography (ECG) or photoplethysmography (PPG), with the latter being

seen as the most promising way to make cheap and practical BP monitoring devices. Over the last decade, many methods and algorithms have been proposed to estimate BP from those other signals and reported various degrees of success. However, the use of different datasets collected with different acquisition platforms and comprising of different sub-groups of the population prevents comparing the performance of published solutions [8, p. 159].

The Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) Waveform Dataset from PhysioNet is a large public database that has been widely used to develop and validate BP estimation algorithms. However, because it is sourced from intensive care units with patients under the effect of drugs and in critical conditions, it is suspected that models created from the data would not generalize to the normal population.

While this project started with the objective of improving BP estimation methods, the challenges of accessing quality human data as a basis for the development and validation of BP estimation models made the use of public data necessary. The potential issue posed by the MIMIC dataset then became clear and pressing, thus the focus of the project pivoted to investigating that issue, which had only been a hypothesis until now.

Research objectives

The purpose of this research project is therefore twofold. First, it aims to investigate the concerns voiced in the literature about using MIMIC, and by extension ICU data, for developing PPG based BP estimation models:

- Determine if significant differences exist between MIMIC and reference data more representative of the general population for:
    - BP data and morphological PPG features of interest.
    - The relationship between morphological features and BP data.
    - The relationship between different morphological features.
- Determine if the use of MIMIC to create data-driven models may impair their generalization to the general population.

Second, it aims to improve upon existing methods used in BP estimation:

- Evaluate different algorithmic or signal processing methods used in the context of BP estimation and implement changes improving processing speed or accuracy.

Structure of the thesis and contributions

Chapter 1 presents a literature review of the wide field of cuffless BP estimation, with a strong focus on the algorithmic component.

The main scientific contribution of this work, which has been submitted as a journal paper to Physiological Measurement, is presented in Chapter 2. It consists of the investigation into the clinical implications of using the MIMIC dataset in the development and validation of BP estimation algorithms using PPG. Chapter 2 demonstrates that the MIMIC PPG data differs significantly from a reference dataset obtained under controlled conditions, and that features extracted from MIMIC responds differently to blood pressure changes compared to the reference. In addition, the practical impact of those differences is tested through the training and cross validating machine learning models on both datasets, demonstrating an acute loss of predictive powers of models when facing data from outside the dataset used to generate the model.

Chapter 3 presents additional contributions related to pulse decomposition analysis, a technique used to identify the timing and contribution to the PPG pulse of the reflections in the arterial system. In that chapter, two published methods are compared in the context of blood pressure estimation. Results show that using timings estimated from the second derivative, as proposed in Chapter 2, results in similar or better estimation performance than when using timings extracted using more complex alternatives.

Following is a short discussion positioning the results against remaining challenges and discussing possible directions for future work. The thesis then concludes with a summary of the results and their significance.

1 Literature review

1.1 Historical perspective

The circulatory system has been studied since antiquity, but it is not until the 18th century that the concept of, and interest in, blood pressure surfaced [9]. In 1733 Reverend Stephen Hales published Haemastaticks describing the first recorded blood pressure measurements, wherein the elevation of blood was observed in a glass tube connected to the left crural artery of a live horse [10]. In the early 19th century, Hales' experiments inspired Jean Léonard Marie Poiseuille, a medical student, to improve upon his method and measure blood pressure using a mercury manometer, which he called an hæmodynameter [11]. In his doctoral thesis, Poiseuille proposed that blood pressure is related to the force of the heart and the radial area of the artery, and also introduced the millimeter of mercury (mmHg) as a unit of blood pressure [9], [12, pp. 26, 43].

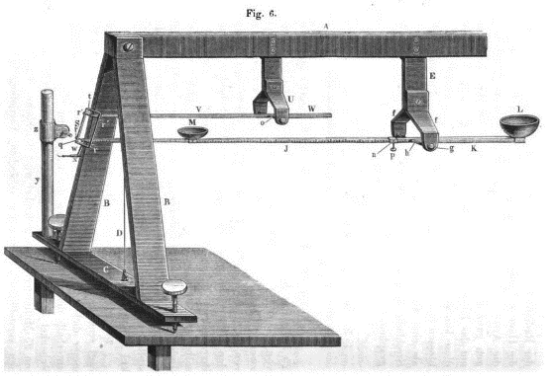

Figure 1: Karl Von Vierordt's sphygmograph (K. Vierordt,1855)

The second half of the 19th century brings a revolution with the development of non-invasive blood pressure measurement devices. In 1854, Karl von Vierordt builds the first device capable of measuring blood pressure non-invasively, the sphygmograph. This device, depicted in Figure 1, was imprecise and cumbersome with long protruding parts. However, it successfully demonstrated that blood pressure can be estimated from the external pressure needed to stop transmission of the pulse in the artery [9], [13]. When the weights were adjusted properly, the movement of the artery would be transmitted to a lever, which could be used to trace the pulse contour on paper [11]. In 1860, Etienne-Jules Marey created a new sphygmograph, shown in Figure 2, small enough to be attached to the wrist [14]. By replacing the weights with

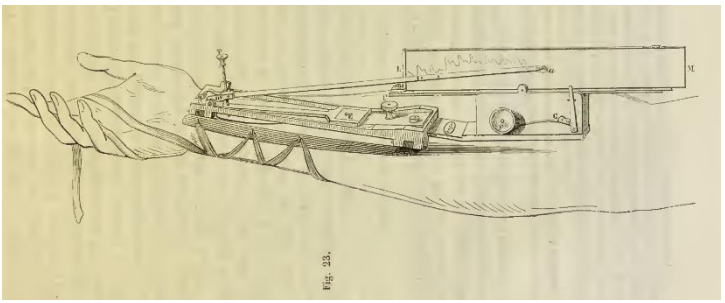

Figure 2: Marey's sphygmograph (E.-J. Marey, 1863)

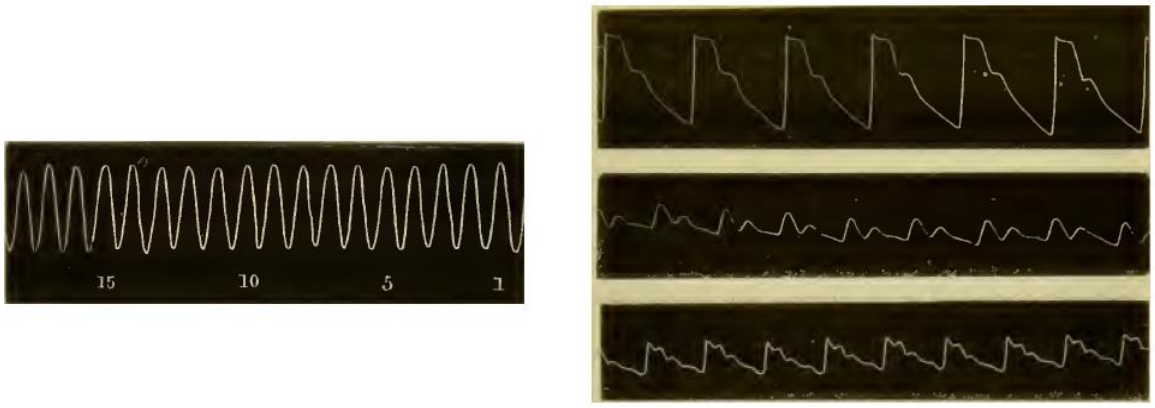

Figure 3: A typical record from Vierordt's device (left) compared to records from Marey's device (right). (E.-J. Marey, 1863)

springs and thus eliminating the inertia, his device could show details of the pulse contour, as shown in Figure 3, which were not clearly visible on Vierordt's recordings [15, pp. 182–184]. However, physicians found the device hard to use and several improved versions were produced by a variety of people over the next decade. Frederick Akbar Mahomed's design, which could not only trace the pulse contour but also quantify the pressure in ounces, became popular for clinical use [11, pp. 18–22], [16]. Diederik Korteweg and Adriaan Isebree Moens published their work on the propagation of waves in elastic tubes in 1878, establishing the well-known Moens-Kortweg equation, which set the foundations for modeling the pulse wave in relation to physical parameters [17]. In 1879, with the use of his improved device, Mahomed was the first to identify a form of hypertension without any secondary cause, today called essential hypertension, and to describe many of its characteristics [16], [18]. Two years later, in 1881, Samuel Siegfried Karl von Basch created the sphygmomanometer. The device was simple: in essence a water filled bag connected to a mercury manometer. However, the arm of the patient needed to be clamped on the device and the water filled bag incrementally pressed against the radial artery using an adjustment screw, until the physician felt the pulse disappear at a distal point by palpation [19]. Even though it was more precise than any previous device, use of the sphygmomanometer did not become widespread until it was simplified and improved on by Scipione Riva-Rocci in 1896. Riva-Rocci realised that taking the measure from the brachial artery would be more representative of the aortic pressure and built a device similar to those we use today: an inflatable cuff, which was wrapped around the patient's arm, attached to a manometer. Because it applied pressure evenly around the arm, the measurements were more reliable than before, and its small size allowed physicians to provide bedside blood pressure measurements easily [19], [20]. Since it was based on the concept of finding the

pressure point at which the pulse wave is blocked, the sphygmomanometer was at first only successful at determining systolic pressure [9]. That changed in 1905, when Nikolai Korotkoff discovered that while deflating the pressurized cuff of the sphygmomanometer, the arteries emit an increasing and then decreasing murmur that can be heard with a stethoscope. He correctly observed that these sounds start when the manometer is at the systolic pressure and end at the diastolic pressure [19], [21]. Thus, it was from now on not only possible to find the diastolic pressure with the sphygmomanometer, but it was also no longer necessary to use palpation to find the obliteration of the pulse in order to find the systolic pressure. Those sounds, now known as Korotkoff sounds, are still used today, even by some automated BP measurement devices [9], [21].

Most of the advances in the later part of the century did not result in any major changes in the way blood pressure was measured. There are however a few turning points that both improved understanding of hemodynamics and paved the way for modern research in the field of indirect measurement. In 1923, John Crighton Bramwell and Archibald Vivian Hill published a series of papers demonstrating the relationship between blood pressure, arterial extensibility, and the velocity of the pulse wave, based on post-mortem experiments on humans [22]. Around 1940, William Ferguson Hamilton published several articles developing the idea that the pulse contour is shaped by wave reflection in arteries. Through different experiments on dogs, he demonstrated and explained the appearance of reflections in relation to wall elasticity and changes in resistance in the arterial tree [23], [24]. Hamilton's conclusions were not immediately accepted as relevant for humans, and several experiments published in the 1950s sowed doubt about their significance [25, p. 197], [26]. By the 1960s, thanks to evidence resulting from frequency domain analysis, the concept of pulse wave reflections gained acceptance. There was, however, still no consensus about the mechanisms causing the reflection, nor of the reflection sites [27, pp. 156–159], [28]. In 1967, Michael F. O'Rourke identified two discrete reflection sites affecting the pulse wave in dogs, and finally in humans in 1980, effectively ending the debate [29], [30]. Meanwhile, 1976 brings the introduction of the first automated oscillometric blood pressure meter called Dinamap, an acronym for "Device for Indirect Non-invasive Mean Arterial Pressure". In 1978 a new version is released, capable of measuring systolic and diastolic blood pressure, as well as the heart rate [31]. The oscillometric method used by those devices is based on the observation of oscillations during the decompression of the sphygmomanometer's cuff, first

demonstrated by Marey in 1876 [32]. It has now become the clinical standard for non-invasive measurement [33].

1.2 State of the art

1.2.1 Current measurement methods

This section summarizes the blood pressure measurement methods currently used in clinical or research environments.

Catheterization is the modern version of the direct arterial manometric measurement method. It is still invasive, but the tools used have changed significantly since Poiseuille's experiments in the 19th century. Using a strain gauge placed at the tip of a catheter, which is inserted into one of the patient's arteries, the method allows measuring instantaneous (also called beat-to-beat) blood pressure [34]. When paired with a Millar semiconductor catheter-tip micromanometer, this method is considered the gold standard for high-fidelity blood pressure wave recording [25, p. 141].

Auscultation, the method devised by Korotkoff in 1905, is still considered the gold standard for non-invasive measurement. Its usage is however in decline due to the ban of mercury sphygmomanometers, and the inadequate performance of their replacement [32]. The method remains unchanged and consists of deflating the sphygmomanometer cuff slowly while listening for the Korotkoff sounds with a stethoscope. The pressure on the manometer at their onset is the systolic pressure, while the diastolic pressure is indicated when they disappear [34]. The method has been used in automated devices with the use of a microphone to detect the Korotkoff sounds [25, p. 143].

Ultrasounds can be used in combination with a sphygmomanometer cuff and rely on the Doppler effect to determine when the systolic and diastolic pressure readings should be taken. An emitter-receiver is applied over the brachial artery under the cuff, which is then deflated as in the auscultation method. The reflected ultrasounds display a phase shift at the systolic pressure caused by induced motion in the arterial walls. The diastolic pressure is recorded when a reduction of arterial motion is detected. This can be helpful in patients with weak Korotkoff sounds where auscultation would be difficult [32].

The oscillometric method has become very popular since the use of mercury in sphygmomanometer has progressively been banned in many countries. Because of its

simplicity of use, low cost and low maintenance required by the equipment, the oscillometric method has become the clinical standard for both home and medical office settings [33]. This method also uses a pressurized cuff placed around a limb, which is slowly deflated. The particularity of the method is that during deflation, the system monitors small oscillations in the internal pressure of the cuff. These are caused by vibrations of the arterial walls, induced by the pulsatile blood flow in the partially occluded artery. The oscillations increase as the cuff is deflated and reach a maximum at the mean arterial pressure. There is however no specific oscillometric pattern directly related to SBP and DBP. The algorithms for estimating those metrics are proprietary and differ from one device to another [25, p. 144], [33]. Nonetheless, studies have shown that commercial oscillometric devices approved for medical use provide sufficient accuracy, as their measurements fall within 5 mmHg of error of an intra-arterial measurement [25, p. 144], [32]. Some medical conditions, such as diabetes and atrial fibrillation, may negatively affect accuracy of the device using this method, and care should be taken to use a device appropriate for the patient [25, p. 144], [33].

Vascular unloading, also called volume clamping or method of Peňáz after its inventor Jan Peňáz, is a non-invasive method that has the advantage of being able to measure beat-to-beat blood pressure. In this system, a photoplethysmogram measures the blood volume under a pressurized finger cuff. The PPG signal is then used as feedback to regulate the pressure in the cuff in order to keep a constant blood volume throughout the cardiac cycle. Thus, the pressure applied to the cuff is directly correlated to the arterial blood pressure and can be measured easily [25, pp. 145–146], [32]. This method can track beat-to-beat variability in arterial blood pressure, but the measured pressure level differs from that measured in the arteries. With proper filtering and calibration, the instantaneous arterial pressure wave can however be reconstructed from the finger reading with sufficient accuracy [35]–[37]. While devices using this method once lacked the precision needed for clinical use [38], the more recent devices generally compare favorably to invasive intra-arterial measurements, and can meet the performance requirements for medical grade BP measurement devices [37], [39]–[42]. The method is however not a universal replacement for intra-arterial pressure measurement, as results have been shown to differ significantly in certain situations, such as during intense exercise [41], in elderly comorbid patients [43], and in critically ill patients [44].

Applanation Tonometry is another non-invasive method, capable of recording the beat-to-beat pressure wave. It consists of applying a micromanometer over a superficial artery to sense its expansion and retraction during the pulse cycle. A difficulty lies in flattening the artery so that its wall tension is perpendicular to the sensor. It has been used mostly in research for observation and analysis of pressure waveforms [25], [34, p. 146]. This method is otherwise impractical for actual non-invasive quantification of blood pressure, as it has been shown that it can only provide acceptable accuracy when calibrated against an invasive BP measurement device [45], [46].

1.2.2 Towards cuffless continuous ambulatory monitoring

Since 1940, it has been known that some otherwise normotensive patients, that is patients with normal blood pressure, show elevated pressure when having it measured at the doctor's office [47]. It is however not until 1988 that Tomas Pickering et al. quantified the prevalence of the phenomenon and coined the term "white coat hypertension" [48]. Pickering et al. then proceeded to attract attention to the inverse phenomenon, where hypertensive patients show normal pressure when measured at the doctor's office, and coined the term "masked hypertension" [49]. It is estimated that white coat hypertension affects between 15 to 30 percent of patients who show a high blood pressure at the doctor's office [50], while masked hypertension is estimated to occur in 9 to 23 percent of patients with normal pressure readings at the doctor's office [51].

Since the discovery of those phenomenon, there has been mounting evidence from prospective studies that home blood pressure monitoring (HBPM) and ambulatory blood pressure monitoring (ABPM) are much better predictors of risk than office blood pressure measurements [32]. As a result, the national blood pressure measurement guidelines now recommend HBPM or ABPM in many cases, either to rule out white coat hypertension and masked hypertension, or to help the patients monitor their condition [6, pp. e26–e28], [52, pp. 599–602].

To this day, ABPM and HBPM is done through cuff-based automated devices using the oscillometric or auscultatory method, taking readings periodically through the day. While this has improved ability to diagnose and monitor the patient's condition, the current methods still have major limitations. Since the readings are discontinuous, they do not provide information about the entire dynamic range and patterns of BP fluctuations of the patient. Moreover, the pressurized cuff may be uncomfortable, which is particularly problematic for

nocturnal BP measurement, as it can disturb sleep and may interfere with measuring pressure during the night [8, pp. 9–11]. This last point is of particular importance as nocturnal high blood pressure has been shown by a large-scale systematic review to be a significant risk factor for mortality, not only in hypertensive, but also in otherwise normotensive individuals. That review also suggested that, for treated hypertensive patients, nocturnal blood pressure is the main risk determinant, while the risk associated with daytime BP becomes insignificant once adjusted for nighttime BP [53]. Another risk factor now associated with increase in mortality is lack of the normal pressure drop during the night, which also affects otherwise normotensive individuals [53], [54]. Finally, current limitations of night time BP measurement also hinder research on important issues, such as on the link between sleep apnea and hypertension, and on why strokes and heart attacks happen mostly during sleep [8, p. 11].

For the reasons explained above, there is now interest in developing cuffless BP measuring devices, and technological advancements have brought us to a turning point. The ideal cuffless BP monitor would offer high accuracy, be comfortable for long term use, be easy to use without medical supervision, allow continuous monitoring, and automated data storage [25, p. 11], [55]. While a variety of transducers have been used in the development of cuffless blood pressure monitors [8, p. 40], most efforts have turned towards PPG and ECG as they can be used without medical supervision and be implemented in a comfortable, low cost device [55].

1.3 Blood pressure modeling for cuffless devices

The principal BP estimation methods used in conjunction with ECG and PPG signals will be presented in this section. It should be noted that some of the techniques that will be presented can also be used with other types of signals, though adjustments may be needed, and performance may vary.

1.3.1 The ECG and PPG signals

To understand how blood pressure is modeled around the ECG and PPG signals, a brief description of those signals is first necessary.

The ECG signal is a measure of the electrical activity of the heart. It is the sum of three discrete waveforms: the P wave associated with the depolarization of the atria, the QRS complex associated with depolarization of the ventricles, and the T wave associated with

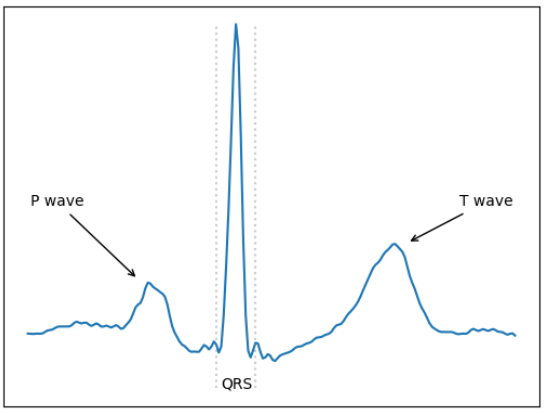

Figure 4: Parts of the ECG pulse wave

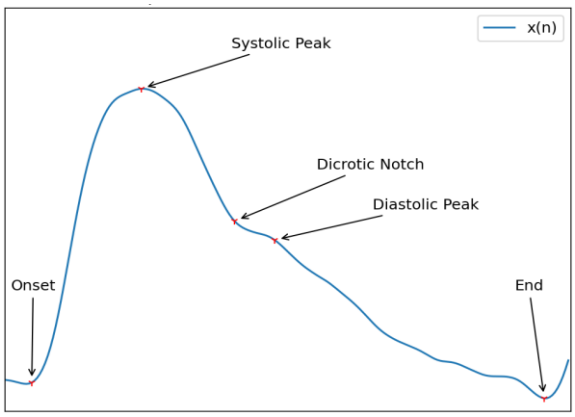

Figure 5: Main fiducial points of the PPG pulse wave

their repolarization. The morphology changes depending on the recording location, but the sequence is always in the P-QRS-T order, as illustrated in Figure 4 [56]. The QRS complex has a duration of about 100 ms and is itself composed of distinct waves, some of which are not always present. The Q wave is any negative deflection before the central positive R wave, and the S wave is any negative deflection after it [57, p. 26,27]. The most common ECG systems use ten electrodes, six placed on the chest and one on each of the arms and legs. From these electrodes, a set of 12 signals, also called leads, are usually recorded by looking at the potential difference between two or more electrodes [57, pp. 15–19].

The PPG signal is a measure of changes in blood volume in the tissues. It is acquired non-invasively by placing a light source against the skin and a receptor to capture the optical signal modulated by the vascular system. The receptor can either be placed nearby the emitter to capture the reflection, or on the opposite side of the body to capture the light that passes through it. Reflection PPG is more often used on the thicker parts of the body such as the forehead, chest, or wrist, where light cannot go through; while transmission PPG can be used on thinner parts, such as the finger or ear lobe for example. Infrared (940 nm) and red (660 nm) are often used for transparency application due to their greater penetration depth and stability at different oxygen blood saturation levels. Green (530 nm) is often used for reflectance, because its lower penetration depth provides a stronger response to the changes in blood flow near the skin surface. The PPG signal has an important quasi-DC component linked to respiration, thermoregulation, skin, fat, tissues and non-pulsatile blood. The part of interest is however a smaller pulsatile AC component linked to the changes in blood volume in the vasculature [8, p. 45,46], [58]. A typical PPG wave, as shown in Figure 5, is usually separated in a rising anacrotic phase associated with the systole, and a catacrotic

phase associated with the diastole and wave reflections from the periphery. In healthy individuals, it is usually possible to see a small deflection, or hollow, called the dicrotic notch between the two phases. The systolic peak is defined as the signal's maximum point in the cardiac cycle. The diastolic peak can be defined as the point corresponding to the second negative zero crossing of the first derivative of the signal. If no clear second peak is present in the signal, the diastolic peak may be defined as an inflection point in the signal, which can be inferred from the second derivative [59].

1.3.2 Pulse Wave Velocity based models

The pulse wave velocity (PWV) is the speed at which the pulse wave travels through the arteries. The relationship between PWV and blood pressure has been well studied and demonstrated empirically. It is based on two models: the arterial wall model associating BP with arterial elasticity, and the arterial wave propagation model associating arterial elasticity to PWV. Those relationships can be seen through the Moens-Korteweg (1) and Frank/Bramwell-Hill (2) equations [34], [60]. The empirical Hughes equation (3) is also often used in conjunction with the Moens-Korteweg equation (1) to establish a relation with blood pressure [8, p. 65], [60].

$$PWV = \sqrt{\frac{Eh}{2\rho r}} \tag{1}$$

$$PWV = \sqrt{\frac{A}{\rho}\frac{dP}{dA}} \tag{2}$$

$$E = E_0 e^{yP} \tag{3}$$

Where E is the elastic modulus of the artery, h the wall thickness, ρ the blood density and r the arterial radius, P the arterial blood pressure, A the arterial cross-sectional area, E_0 the elastic modulus at zero pressure, and y a material coefficient of the artery.

This direct relationship implies that, in theory, an accurate measurement of PWV would allow a more precise and robust blood pressure estimation than the oscillometric or auscultatory methods, which rely on less direct relationships [8, p. 4].

For estimating blood pressure, the average PWV over a certain distance is used. It can be derived by measuring the pulse transit time (PTT), the time necessary for the pulse wave to travel from the heart to a distal point at a distance *L*, as follow [8, p. 66]:

$$PWV = \frac{L}{PTT} \tag{4}$$

Isolating *P* from the combination of (1) and (3) results in a logarithmic relationship. The equation can then be rephrased by combining all the patient specific physical parameters into the α_1 and β_1 coefficients [8, p. 66]:

$$P = \left(-\frac{2}{y}\right)\ln(PTT) + \left(\frac{1}{y}\right)\ln\left(\frac{2\rho r L^2}{hE_0}\right) = \alpha_1 \cdot \ln(PTT) + \beta_1 \tag{5}$$

Several linear and nonlinear variations on the model have been proposed. Those models make various assumptions about the biological characteristics, and the range of variations, or approximate the model by series expansion. Of all the variations, a very simple linear approximation model (6) has been widely used due to its low computational cost [8, p. 67], [55]. It has also been observed that addition of the Heart Rate and other metrics relating to arterial stiffness as additional terms in the equation could improve the correlation to blood pressure [61].

$$P = \alpha_2 \cdot PTT + \beta_2 \tag{6}$$

More recently, it has been shown that the relationship between BP and PWV in humans only follows the Moens-Korteweg equation at very low blood pressures. It has been proposed that the relationship between PWV and BP is better represented as follow [60]:

$$P = \alpha_3 \cdot PWV^2 + \beta_3 \tag{7}$$

The physical parameters of the subject, represented by the alpha and beta coefficients in the model, have several implications. First, the accuracy of the BP estimation is highly dependent on determination of those unknown parameters by calibration of the model [55]. Since most of those parameters are patient specific, calibration needs to be performed on a per patient basis for good results. Moreover, several of those parameters change with age and time, which means that the models need periodic recalibration to maintain accuracy [8, p. 64], [55].

1.3.2.1 Limitations and challenges of PVW based models

The PTT cannot be easily calculated from ECG and PPG signals, as the R-peak used as reference for the start of the pulse on the ECG does not correspond to the start of the mechanical pulse creating the pressure wave. This unwanted delay is called the pre-ejection period (PEP). As a result, the Pulse Arrival Time (PAT), defined as the sum of PTT and PEP, is widely used as a surrogate of PTT [34]. The PAT is easily extracted by measuring the delay between the R-peak of the ECG signal with a specific point in the PPG signal, usually the foot of the main wave, the peak, or the maximum point of the slope [55].

However, PAT does not correlate as strongly with blood pressure as PTT, because of the variability of PEP, which can be between 10 to 35% of PTT [34]. Despite those difficulties, there are a few medical-grade wearable systems based on PAT [8, pp. 56, 57]. They are not fully ambulatory as they rely on cables to connect a chest ECG sensor and a finger PPG sensor to a wrist worn unit, which, as a whole, would be in the way of daily life activities [62], [63]. Although some completely wrist worn ECG/PPG based systems, relying on PVW for BP measurement exist, none have been validated and approved as medical devices [8, pp. 56, 57]. The wrist may not be a suitable location for a PPG sensor in the context of a PVW based system, as the small surface vessels from where PPG is taken are prone to vasodilation and vasoconstriction. This can not only cause morphology changes but can also affect timing, which is critical in the context of PVW based models [8, p. 46].

1.3.3 Pulse Wave Analysis

Pulse wave analysis (PWA) is an analysis of the pressure pulse contour to extract features correlated to physical properties of the arterial system. It has been used with tonometry since the 1960's but has recently been gaining traction in conjunction with PPG for the purpose of estimating blood pressure [8, p. 107]. The pulse volume waveform of PPG has similarities to the blood pressure waveform, both in its morphology and in the changes it undergoes in diseased patients, and has good potential for extracting cardiovascular diagnostic information [58, p. 8], [64]. The foray in use of the combination of PPG and PWA for BP estimation was traced to a small study by Teng & Zeng in 2003 [65]. They showed that a correlation between BP and the PPG signal could be established, and thereby paved the way for subsequent studies.

To estimate BP from the features extracted from the PPG signal, multilinear or nonlinear regression models have been used. Those allow mapping the non-linearities of the features extracted from peripheral vessels to central blood pressure. A user specific calibration step is usually necessary to initialize the translation model, but efforts are being made towards generalizing the calibration step so that a single model can be used for a large part of the population [8, p. 119].

1.3.3.1 Common PWA features

This section will briefly describe the most common PWA features encountered in the literature and their reported link to physical properties. Some of those features were first defined in relation to the pulse pressure wave. To better fit the current purpose, they will here be described only in relation to the PPG signal.

The Augmentation Index, abbreviated AI, or AIx, was first described by Kelly et al. in 1989. It is a measure of the contribution of the first pulse wave reflection to the maximum pressure point in the pressure waveform. In the PPG signal, the measurement points used in the pressure wave correspond approximately to the systolic peak and diastolic peaks [66]. It is an indicator of the elastic compliance of the artery, as reduced compliance causes early reflections and a higher contribution to the systolic peak [59]. It has also been shown to be correlated to age [66]. To follow Kelly's definition, the AI in PPG can be defined as equation (8). There is however confusion in the literature and the augmentation index is often confused with reflection index, which is closely related [59].

The Reflection Index or RI is a variation on AI. It has been shown to be linearly correlated to AI [67].

$$AI = \frac{x\text{-}y}{y} \tag{8}$$

$$RI = \frac{y}{x} \tag{9}$$

where x is the amplitude of the systolic peak and y is the amplitude of the diastolic peak.

The Stiffness Index or SI, also called Large Artery Stiffness Index or LASI, is a measure of arterial stiffness proposed by Millaseau et al. in 2002. Based on the concept of PWV and the idea that it is related to arterial stiffness, it aims to measure the velocity between the

pulse and its reflection. To do so, one simply divides the height of the patient *h* by *Δt,* the time delay between the systolic and diastolic peak. Millaseau also showed that SI is correlated to PWV and age, and weakly correlated to blood pressure. Moreover, after administration of Nitroglycerin, the change in SI was more correlated to the change in PWV and mean arterial pressure (MAP) than the change observed in RI, which could make it a more robust predictor than AI and RI in relation to changes in the tone of small arteries [68]. In many publications, the subject's height is omitted, and the stiffness index is simply defined as being equal to the time delay, to which it is inversely proportional according to the original definition. The inflection point between the dicrotic notch and the diastolic peak is also sometimes used instead of the diastolic peak as a reference point to calculate the delay, such as in [69]–[71]. No information could however be found as to how omitting the height affects the useful information carried by the metric, or its direct correlation to BP.

$$SI = \frac{h}{\Delta t} \tag{10}$$

The Inflection Point Area Ratio or IPA seems to have been first designed by Wang et al. in 2009 as a measure of total peripheral resistance for the purpose of calculating the cardiac output. It was defined as the ratio between the area under the curve of the first and second peak of the PPG signal, separated at the dicrotic notch [72]. The idea has been reused for blood pressure measurement, sometimes with ratios between additional zones defined by other inflection points in the signal, such as by Kachuee et al. [71].

$$IPA = \frac{A_1}{A_2} \tag{11}$$

where A_1 and A_2 are the areas under the curve before and after the dicrotic notch, respectively.

The Heart Rate or HR, Peak to Peak Interval, and Pulse Interval are all metrics that carry information about the duration of a cardiac cycle. The peak to peak interval is defined as the delay between two systolic peaks, the pulse interval is the delay between the beginning and the end of the pulse wave, and the heart rate is defined as the number of cardiac pulsations per second or minute [55], [59]. The heart rate has been shown to correlate with BP and to contribute to better results when added to some BP estimation algorithms [8, p. 67], [55].

The Pulse Width, usually defined by the width of the pulse wave at 50% of its height, has been shown to be correlated to systemic vascular resistance and to blood pressure [73]–[75]. The width of the wave from the foot to the systolic peak, also called crest time or systolic time, has been known to be longer in hypertensive patients since the 1940's [76]. Some small scale studies have suggested a better correlation with BP in the diastolic time, which is the width of the second half of the wave after the systolic peak, respectively [65], [77].

The Normalized Pulse Volume (NPV) has been proposed as an indicator of vascular tone. It is defined as the ratio of the amplitude of the pulsatile component of the PPG signal ΔI_{ac} to its slowly drifting DC component I_{dc}, as shown in equation (12). Based on the Beer-Lambert law of light attenuation, it allows measuring the pulse volume as an absolute value rather than as the relative measure given by the PPG signal [78]. NPV has also been derived from the modified Beer-Lambert law to take into account light-scattering, and named modified NPV or mNPV, but the resulting equation is exactly the same as NPV [79]. It has recently been proposed that BP can be estimated from the heart rate and NPV through their relationship to cardiac output and total peripheral resistance, using equation (13), without any individual calibration [80].

$$NPV = \Delta I_{\mathrm{ac}} / I_{\mathrm{dc}} \tag{12}$$

$$P = \exp(a \cdot \ln(HR) + b \cdot \ln(NPV) + c) \tag{13}$$

Finally, a set of features known as second derivative PPG, acceleration PPG or APG has been used extensively in the literature [8, p. 118]. In 1998, Takazawa et al. identified a group of fiducial points in the second derivative of the PPG signal, shown on Figure 6, whose amplitude ratios were shown to correlate to vascular aging and arterial stiffness. Through a "second derivative aging index" computed from those points, Takazawa et al. also reported an association to hypertension [81]. The ratios are usually formed in respect to the a-point (*b*/*a*, *c*/*a*, *d*/*a* and *e*/*a*). It should also be

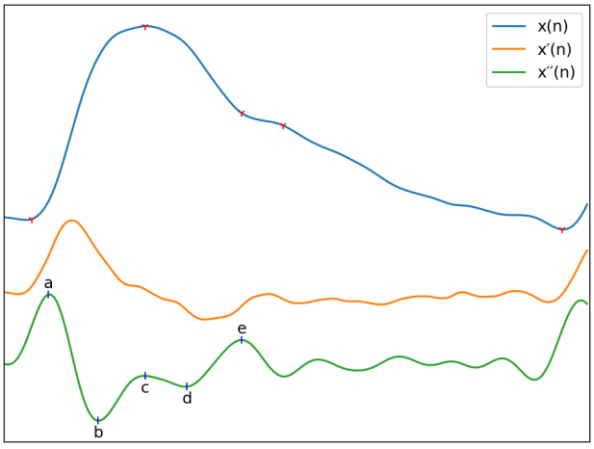

Figure 6: The PPG pulse aligned with its first two derivatives (scaled) with the second derivative fiducial points indicated

noted that the e-point is aligned with the dicrotic notch, which is sometimes used to locate it in the PPG signal [8, p. 118].

1.3.4 Pulse Decomposition Analysis

A weakness of PWA models is that it can be hard to extract the desired parameters from the signal if it differs too much from the ideal PPG waveform [71]. The morphology of real-world PPG signals is rarely ideal. For example, it has been said that 80% of the PPG signals from the popular PhysioNet MIMIC II database from ICU patients are non-ideal and inappropriate for direct parameter extraction [82]. To palliate this issue, pulse decomposition analysis (PDA) can be used to identify the position of important points in non-ideal waveforms [83], or to extract information about hemodynamic parameters directly [8, p. 76].

The circulatory system is still sometimes modeled as a resistor-capacitor system, referred to as the Windkessel model, which does not allow modeling the wave reflections observed in reality. Thus, more complex models of wave propagation have now largely replaced the old Windkessel model [25, pp. 278–282]. Based on this concept, PDA is the decomposition of the observed PPG waveform into its different components: the main pulse wave generated by the heart and its reflections from different locations in the arterial tree. Two main reflection sites are recognized, one at the junction between the aorta to the renal arteries, and one at the branching to the iliac arteries. Other minor reflection sites also exist but are not well defined. The PPG signal is thus usually modeled as the sum of the main pulse wave and up to five reflections, represented as either Gaussian, Rayleigh or Logarithmic Normal functions [83]. Fitting the model to the signal is usually done by expressing the system as an optimization problem, and the optimal parameters found through various means, such as successive approximation [83]–[85]. While that method can yield optimal fitting, they may not be suitable for real-time applications. A lighter method has been proposed that consists of fitting one component at a time. That is however based on the assumption that the components do not overlap for more than fifty percent of their width, which may increase the fitting error if the assumption is not true [86]. That method has been implemented as a recursive algorithm [87], [88], which may be suitable for real-time processing. Those two methods will be detailed in Chapter 3.

The development of PDA has largely happened around a finger cuff based blood pressure monitor, the CareTaker [8, pp. 102–104]. The basis of the CareTaker PDA model is the Moens-Korteweg equation (1), which relates blood pressure to the velocity and amplitude of

the reflections, based on elasticity of the arteries. The main publication supporting the model made two interesting demonstrations. First, the amplitude ratio of the renal reflection wave to the main wave is closely linked to systolic pressure. Second, the time delay between the third reflection wave and the main wave is correlated to the pulse pressure. While methodological details are missing, the general relations upon which the CareTaker PDA model is built and a validation study have been published [89], [90].

1.3.5 Data driven and non-parametric models

Several blood pressure estimation models that do not use physical based parameters have been proposed and are often called “non-parametric”. They rely on more generic time-frequency features and signal processing techniques, as well as machine learning. Many of the proposed methods claim to offer better accuracy than parameter based models, as well as calibration-free operation, but they can be computationally complex [55]. With its ability to learn complex models and incorporate concealed factors not included in the analytical model, data-driven machine learning holds great promise in the field of BP estimation [8, p. 143]. More recently, recurrent neural network (RNN) in a long-short-term-memory (LSTM) architecture has been shown to provide better performance than traditional machine learning [8, p. 148]. It has been suggested that such an architecture is better suited to provide long term recalibration-free accuracy [91]. While most BP estimation studies with machine learning have used handcrafted features and ECG [8, p. 154], some have also used the raw signal, also called the “whole-based” method, and sometimes supplemented it with ECG signal information [71], [82]. In addition to its algorithmic complexity, there are other objections to using machine learning for BP estimation. A well-trained machine learning model performs well within the scope of the problem presented to it, but is hard to generalize, as quality data representing all the sub-groups of the population would need to be included in the training set. In addition, the resulting model is opaque, uninterpretable, and thus harder to review and trust for regulators, scientists and patients [8, p. 146,147,158,159].

Beyond machine learning, Millasseau et al. have proposed using a transfer function to estimate instantaneous blood pressure from the finger PPG signal. By using Fourier analysis and looking at the first ten harmonics of the signal, they established a generalized transfer function matching the frequency content of the PPG signal to that of the pressure wave. An estimation of the pressure waveform was then produced using a transfer function on the spectrum of the PPG signal and recovering the time domain signal via an inverse Fourier transform. Good results were reported on normotensive and hypertensive patients, as well

as for normotensive patients after administration of nitroglycerin [92]. The method seems to have been largely eclipsed by machine learning as only a few subsequent publications were found on the subject [93], [94].

1.4 Regulations and standards

Since this review is not focused on the hardware aspect of blood pressure monitoring, the standards and regulations governing the hardware aspect of medical devices will not be touched here. Instead, this section will focus on the aspects of regulations touching the performance requirements for blood pressure monitors, and thus the underlying algorithms. There is currently no standard covering all aspects of cuffless blood pressure monitors, especially not in relation to differences compared to traditional monitors, such as continuous measurements during daily activities, calibration, and temporal stability. A new standard, ISO/CD 81060-3, is being developed by the ANSI/AAMI/ISO working group for that purpose [8, pp. 203–209]. In the meantime, researchers have relied mostly on existing standards designed for traditional cuff-based monitors.

The British Hypertension Society (BHS) protocol for the evaluation of blood pressure measuring devices [95] is often used in the literature for evaluating performance of BP estimation algorithms. It uses a three-tier grading system illustrated in Table 2 below.

Table 2: BHS performance requirements for BP measuring devices

Grade	Minimum percentage of samples with error smaller than…		
	5 mmHg	10 mmHg	15 mmHg
A	60%	85%	95%
B	50%	75%	90%
C	40%	65%	85%

To receive a grade, the cumulative error percentage of all measurement made with the device during the trial must meet the three requirements for that grade. For example, obtaining an A grade means that at least 60 % of the measurement had an error of 5 mmHg or less, that 85 % had an error of 10 mmHg or less and that 95 % had an error of 15 mmHg or less. To be recommended for clinical use, the device must obtain an A or B grade for both SBP and DBP measurements.

For the main validation phase, the BHS protocol requires a minimum of 85 subjects, wherein sex and age distribution is left to chance. The protocol also defines five pressure groups, that must include a minimum of 8 to 20 subjects depending on the group. Three measurements must be taken from each subject for a total of 255. An additional validation phase with 30 subjects can also be done, to include subgroups of the population such as pregnant women or the elderly, if the device is intended for their use. Finally, if the device is intended for use in special circumstances, such as during exercise, a final validation test should be done under those circumstances.

The ANSI/AAMI SP10 standard, or its replacement the AINSI/AAMI/ISO 81060, are also frequently used in the literature. The full text could unfortunately not be obtained for this review, but the most important parts are discussed by Solà et al. [8, pp. 206–221]. Validation under the standard requires a minimum of three measurements on 85 subjects, with at least 30 % of either sex, for a total of 255 measurements. Under the AINSI standard, grading is dichotomous. To obtain a “pass” grade, the accuracy criterion only requires the mean error of the measurements to be below 5 mmHg with a standard deviation under 8 mmHg, while any values above those result in a “fail” grade. If the device is intended for use in special subgroups of the population, additional tests on 35 subjects of each sub-group should be performed.

The European Society of Hypertension (ESH) International Protocol for the validation of blood pressure measuring devices in adults [96] is not as widely used for validating algorithms in the literature as the BHS or ANSI standards. Validation under the ESH standard requires 33 subjects of at least 25 years of age, at least 10 of each sex. It defines three SBP and three DBP recruitment ranges, which should contain between 10 and 12 subjects. Three measurements should be taken for each subject for a total of 99 measurements. The performance criterion is similar to the BHS system: the error of a fixed number of measurements must be under 5, 10 and 15 mmHg. In addition, at least 24 subjects must have the error of 2/3 of their measurements under 5 mmHg, and at most 3 subjects can have the error of all of their measurements above 5 mmHg. Because of concerns regarding its statistical power, the ESH standard has been discontinued in favor of the international ISO standard currently in development [8, p. 214].

The IEEE 1708 Standard for Wearable, Cuffless Blood Pressure Measuring Devices [97] was published in 2014, in an attempt to provide some guidance for the development of

cuffless BP measurement devices. The latest amendment was released in 2019 [98], but the standard has not been widely used in the literature so far. Validation under the standard is a two-phase process, where the only difference between phases is the number of subjects. The first phase is composed of 20 subjects and the second phase of 65 subjects, who should be distributed evenly both between sexes and four pressure ranges. All subjects should be between 21 and 50 years of age. The IEEE standard mandates several types of tests, namely a static test, a test with BP change from the calibration point, and a third test after a time from the calibration. The standard does not define any minimum time period between re-calibrations, but states that the device should be re-tested after "a sufficient time" no longer than the recalibration time period indicated by the manufacturer. Three measurements are required per subject for each test. The performance criterion is simple and uses the mean absolute difference (MAD) between the device under validation and a reference. Grading is done on a scale from A to D, as presented in Table 3. Grades from A to C are acceptable under IEEE but must be obtained independently for all tests. Amendment 1 of the standard adds a new clause for ABPM, separating devices into four categories, and mandating additional tests for ambulatory and motion situations. It does not however provide any guidance regarding patient posture changes but do mention that this aspect will be covered in a later revision.

Table 3: Overall accuracy requirement for the various IEEE 1708 grades

Grade	A	B	C	D (fail)
MAD range (mmHG)	≤ 5	5-6	6-7	≥ 7

1.5 Comparison of recent studies

A comparison of recent studies is given in Table 4, where N is the number of subject and Calib. Is for calibration. While some studies presented algorithms without patient-specific calibration, few were truly generalized models. Those marked as *no** in the calibration field indicate that patient overlap between their training and testing sets was suspected or confirmed, effectively calibrating the model to the subjects it was trained on. Studies using the subject's biometrics in the algorithm were counted as requiring calibration, even though calibration against a reference device may not be necessary. When calibration was optional (Opt.), results without calibration are shown.

Table 4: Comparison of recent studies on blood pressure estimation with PPG and ECG.

Year	Author	Signals	Solver & features	Calib.	N	ME ± SD SBP, DBP (mmHg)	MAD ± SD SBP, DBP (mmHg)
2017	M. Kachuee et al. [71]	PPG + ECG	Adaboost with PAT, HR, LASI	Opt.	942		11.2 ± 10.1 5.4 ± 6.1
2017	Miao et al. [99]	PPG + ECG	Subject-specific linear regression model with adaptative feature selection from 14 morphological features	Yes	73	0.00 ± 3.4 0.00 ± 2.47	
2018	L. Wang et al. [100]	PPG	ANN with systolic time, diastolic time and 20 spectral features	No*	72		4.02 ± 2.79 2.27 ± 1.82
2018	Y. Wang et al. [101]	2x PPG	Peripheral PTT calculated between two PPG sensors	Yes	30		7.61 6.82
2018	Lin et al. [102]	PPG	Linear regression using 1 of 19 morphological features	Yes	22		4.06 ± 9.16 3.99 ± 5.24
2018	Simjanoska et al. [103]	ECG	Random forest with age and 5 generic signal processing features	Yes	15		8.64 ± 10.74 18.20 ± 8.45
2018	Tanveer et al. [70]	PPG + ECG	LSTM neural network with 13 PWA and PTT based features	No*	39	0.02 ± 1.26 0.00 ± 0.73	0.93 0.51
2018	Yoon et al. [104]	PPG + ECG	Linear regression with PAT, HR, and 11 PWA features	Yes	23		8.7 ± 3.2 4.4 ± 1.6
2019	Mousavi et al. [82]	PPG	Adaboost with 43 features from PCA of whole-based features	No*	441	0.19 ± 4.17 -0.05 ± 8.90	
2019	Yan et al. [105]	PPG	Subject-specific SVM using whole-based features	Yes	70	0.04 ± 5.00 0.01 ± 3.69	

2019	Xing et al. [106]	PPG	Random forest with 19 features including biometric information, second derivative PPG, SVD of signal and its second derivative	Yes	1249	2.1 ± 13.6 2.3 ± 9.5	
2019	Baek et al. [107]	PPG + ECG	CNN using raw signals in time and frequency domain	Opt.	942		9.3 ± 8.85 5.12 ± 5.52
2020	Hasanzadeh et al. [108]	PPG	Adaboost with HR, Pulse width, RI, LASI, IPA, mNPV, HRV	No	1000		8.22 ± 10.38 4.17 ± 4.22
2020	Khalid et al. [109]	PPG	Stacked KNN and regression tree with 16 morphologic features	No*	272	0.07 ± 7.10 -0.08 ± 6.00	
2020	Song et al. [110]	PPG + ECG	DNN with PAT, and PPG morphological features	Yes	110		4.8 ± 4.7 4.8 ± 3.9
2020	Landry et al. [111]	PPG + ECG	Subject specific ANN with raw signal	Yes	15		<5 ± <8 <5 ± <8
2021	Chakraborty et al. [112]	PPG	Learning-based regression with PVW extracted from two waves PDA model	?	150		2.50 ± 2.07 2.12 ± 1.79

1.6 Conclusion

In July 2019, a PPG based wrist-worn cuffless blood pressure monitor was approved for the first time by the Food and Drugs Administration in the United States [113]. The device, from BioBeat, show performance within the acceptable margin of error of a cuff based sphygmomanometer [114]. It has however not been fully approved in relation to European standards for medical devices [115]. Another PPG based device from Aktiia is currently undergoing validation, and early published results show a good agreement between readings from the device and intra-arterial measurement [116].

These recent developments are certainly proof of concept for ambulatory cuffless blood pressure measurement devices. However, the proprietary algorithms and their underlying models have not been published, and a lot of uncertainty remains around the clinical potential of algorithms published in the literature. Most approaches to measuring blood

pressure through ECG and PPG are based on physical properties with inter-individual variability in their correlation to blood pressure [8, pp. 131, 164]. Thus, some form of calibration seems inevitable. The forms of calibration proposed in the literature vary greatly, ranging from training a machine learning model specifically for a subject such as in [105], to simply using the patient's biometrics such as in [106]. Temporal stability of the calibration, an important aspect of an algorithm's performance [8, p. 207], is rarely reported. Other critical aspects of the algorithms' clinical performance, such as their stability in reaction to large or rapid pressure changes [8, pp. 207–208] and posture [117], are usually not investigated. In addition, the large variation in sample sizes and composition of data used to test algorithms, different ways of reporting performance, and lack of a complete and coherent regulatory backdrop make published algorithms for cuffless ambulatory devices difficult to evaluate.

In conclusion, a relationship between BP and various features of the ECG and PPG signals has been demonstrated. How to reliably map those features to a subject's blood pressure is, however, still not well defined. Many proposed methods show promise, but that they can be applied to a large and varied population, in a reliable and time-stable way, remains to be proven.

2 Intensive care photoplethysmogram datasets and machine-learning for blood pressure estimation: generalization not guarantied

Reference: Guillaume Weber-Boisvert, Benoit Gosselin, Frida Sandberg, “Intensive care photoplethysmogram datasets and machine-learning for blood pressure estimation: generalization not guarantied*", Physiological Measurement, 2022* (Pending review)

2.1 Résumé

Objectif : Le vaste ensemble de données *MIMIC waveform dataset* a été largement utilisé pour le développement d’algorithmes d’estimation de la pression artérielle à partir de la photopléthysmographie (PPG). Cependant, puisque les données proviennent de patients ayant de graves conditions médicales et étant souvent médicamentés, il est fréquemment mentionné que la relation entre le signal PPG et la pression artérielle n’est peut-être pas normale. Cet argument sera ici examiné.

Approche : Un échantillon de 12 000 enregistrements provenant de MIMIC sont comparés aux 219 enregistrements de l’ensemble de données PPG-BP, un ensemble de données obtenu dans le cadre de conditions expérimentales contrôlées. Les distributions de la pression systolique et diastolique, ainsi que de 31 caractéristiques morphologiques du pouls sont d’abord comparées entre les ensembles de données. Ensuite, la corrélation entre les caractéristiques et la pression artérielle, ainsi que la corrélation entre les caractéristiques, est analysée. Enfin, des modèles de régression sont entraînés et validés pour chaque ensemble de données, puis validés de nouveau en utilisant l’autre ensemble n’ayant pas servi à l’entrainement.

Résultats principaux : L’analyse statistique démontre des différences significatives ($p < 0{,}001$) entre les ensembles de données pour ce qui est de la pression diastolique et de 20 des 31 caractéristiques, après ajustement pour la fréquence cardiaque. Les huit caractéristiques dont les coefficients de corrélation de rang sont les plus élevés ($|\rho| > 0{,}40$) relativement à la pression systolique pour PPG-BP ont une corrélation très faible ($|\rho| < 0{,}10$) pour MIMIC. Les tests de régression démontrent une capacité de prédiction de base deux fois plus élevée avec PPG-BP qu’avec MIMIC. Quant aux tests de régression entre les bases de données, ils révèlent une perte pratiquement complète de la capacité prédictive dans chacun des cas.

Signification : Les différences entre les ensembles de données MIMIC et PPG-BP démontrées dans cette étude indiquent que les modèles d'estimation de la pression artérielle se basant sur les données de MIMIC pourraient voir leur capacité prédictive réduite lorsqu'appliqués à la population générale.

2.2 Abstract

Objective: The large MIMIC waveform dataset, sourced from intensive care units, has been used extensively for the development of Photoplethysmography (PPG) based blood pressure (BP) estimation algorithms. Yet, because the data comes from patients in severe conditions – often under the effect of drugs – it is regularly noted that the relationship between BP and PPG signal characteristics may be anomalous, a claim that we investigate here.

Approach: A sample of 12000 records from the MIMIC waveform dataset was stacked up against the 219 records of the PPG-BP dataset, an alternative public dataset obtained under controlled experimental conditions. The distribution of systolic and diastolic BP data and 31 PPG pulse morphological features was first compared between datasets. Then, the correlation between features and BP, as well as between the features themselves, was analysed. Finally, regression models were trained for each dataset and validated against the other.

Main Results: Statistical analysis showed significant ($p < 0.001$) differences between the datasets in diastolic BP and in 20 out of 31 features when adjusting for heart rate differences. The eight features showing the highest rank correlation ($|\rho| > 0.40$) to SBP in PPG-BP all displayed muted correlation levels ($|\rho| < 0.10$) in MIMIC. Regression tests showed twice higher baseline predictive power with PPG-BP than with MIMIC. Cross-dataset regression displayed a practically complete loss of predictive power for all models.

Significance: The differences between the MIMIC and PPG-BP dataset exposed in this study suggest that BP estimation models based on the MIMIC dataset have reduced predictive power on the general population.

2.3 Introduction

Hypertension is one of the greatest challenges to public health of our time. According to the Centre for Disease Control, 45% of the adult population in the United States suffer from

hypertension, and only 24% of those with hypertension have their condition under control [1]. Hypertension is an independent risk factor for cardiovascular diseases such as heart attack, stroke, and kidney disease, and ranks second amongst the preventable causes of death in the U.S., trailing cigarette smoking only [2]–[4]. It is now widely accepted that home blood pressure (BP) monitoring and ambulatory BP monitoring are much better at predicting risks associated with hypertension than in-clinic BP measurements [32], with night time BP increasingly seen as an important risk determinant [53], [118]. Devices presently used for home BP monitoring utilize an inflatable cuff, which only provides intermittent readings instead of presenting the entire dynamic range and patterns of BP fluctuations. Moreover, the discomfort caused by cuff inflation is particularly problematic for nocturnal BP measurement, as it can disturb sleep and thereby interfere with measurements [8].

Photoplethysmography (PPG) based BP estimation shows promises to be a low-cost and convenient technique that enables wearable designs and has the potential to replace cuff-based devices [119]. However, the lack of open access, standardized PPG datasets for training and testing BP estimation algorithms is an obstacle to researchers in the field. Most studies are based on private databases where composition of the data and methods of acquisition vary considerably, making a direct comparison between the published BP estimation algorithms impossible [8].

At the time of writing, several public datasets that include BP and PPG signal are available. There are two large datasets sourced from intensive care and surgical units: the Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC) Waveform Dataset [120] from the Massachusetts Institute of Technology, released on PhysioNet [121] in 2011, and the VitalDB from the Seoul National University Hospital [122] released in 2017. Several smaller datasets also exist, often with a focus on a specific condition. A few examples are: the University of Queensland Vital Sign Dataset [123], a 32 patient dataset focusing on anaesthesia acquired at the Royal Adelaide Hospital in Adelaide, Australia, released in 2012; the Bed-Based Ballistocardiography Dataset [124], a 40 patient dataset from the Kansas State University, released at the end of 2020; and the PPG-BP dataset [125], a 219 patients dataset from the Guilin People's Hospital, released in 2018, with a focus on the screening of cardiovascular diseases (CVD) from PPG.

The PPG-BP dataset can be considered a middle ground among the available datasets. It contains 657 short PPG segments three for each of the 219 patients and recorded at rest

under controlled experimental conditions. Each patient is associated with a single BP measurement, as well as patient biometric data and diagnosed CVD, if any. In contrast, MIMIC contains more than 25000 records of variable length and varying measurement types, at times including PPG and arterial blood pressure (ABP). The data was acquired from bedside monitoring devices at intensive care units (ICU), including surgery and cardiac care units, at the Beth Israel Deaconess Medical Center in Boston, USA. Among all the public datasets, MIMIC has been available the longest and has been used the most extensively in the field of BP estimation. The other datasets have seen little use in comparison, and some are not well suited for developing and validating BP estimation algorithms due to the limited number of subjects, the special conditions of data collection and the sporadicity of BP measurements.

MIMIC has been used in many BP estimation studies. Kachuee et al. used a sample of 3663 records from 942 subjects to estimate systolic blood pressure (SBP) using 10 PPG and ECG morphological features. Their best results were a mean absolute error (MAE) of 11.17 mmHG without calibration and 8.21 mmHG with calibration, using AdaBoost for regression [71]. In 2020 Hasanzadeh et al. used a sample of about 1000 subjects to estimate SBP from one spectral and 18 morphological features using PPG only. Their best results were obtained with AdaBoost regression, giving a MAE of 8.22 mmHg [108]. In 2021, a subset of 200 subjects has been used by Esmaelpoor et al. to compare of 56 machine-learned features generated by convolutional neural network (CNN) against a set of 27 frequently used morphological features from PPG and ECG. Eight regression methods were tested and the best results were obtained with squared exponential Gaussian regression or Gaussian process regression depending on the test parameters, providing SBP with a MAE under 6 mmHg using morphological features, and under 3.5 mmHg using machine-learned features [126]. As in this last example, the dataset has been used many times with pulse transit time and pulse arrival time algorithms, despite that variability in the ECG sampling time makes it unsuitable for transit and arrival time calculation [119]. The breadth and variable quality of the dataset also resulted in uneven sampling by researchers, and as such hardly makes performance comparison easier, even between two studies using it. A more serious concern is the frequently mentioned hypothesis that because the data is sourced from ICU, with patients having received medication and being in varied critical conditions, the MIMIC population may exhibit abnormalities or a different relation between PPG and BP

than would be seen in a more controlled setting [71], [108], [127], casting doubt on the validity of results beyond the dataset itself.

The aim of this study is to evaluate if the relationship between PPG pulse characteristics and BP in MIMIC is truly different from that in data acquired under controlled conditions. To achieve this goal a subset of MIMIC was compared to PPG-BP in a two-step approach. First, a statistical comparison of the datasets was performed. It comprised comparing the distribution of features characterizing PPG pulse morphology, as well as comparing the correlation between features on each dataset. Second, the correlation between BP and features was compared between datasets to see if similar morphological variations could be observed on both datasets in relation to BP changes. To illustrate the implication of the differences between the datasets, Support Vector Regression models were trained on each dataset and their cross-validated performance on the training set were compared to their performance on the other dataset, in order to assert whether predictive powers were retained.

This research project was approved by the ethics committee of Université Laval, Canada (approval number: 2022-174).

2.4 Material and methods

2.4.1 Datasets

A subset of the MIMIC database, prepared especially for BP estimation by Kachue et al. [128], is used in this study. Because it is hosted by the University of California, Irvine, the subset is sometimes called the “UCI” dataset, which will be used hereafter. This subset, which excludes segments with missing signals and abnormal values from MIMIC, contains 12000 records of lengths varying between 8 seconds and 10 minutes. Each record is sampled at 125 Hz and contain fingertip PPG, electrocardiogram, and instantaneous ABP. No additional information about the subjects is provided.

The PPG-BP dataset contains 657 fingertip PPG segments from 219 subjects of 21 to 86 years of age. Each segment has a duration of 2.1 seconds and a sampling rate of 1 kHz. A single SBP and DBP measurement is provided for each subject, as well as the sex, age, height, weight, heart rate, and disease records. While also sourced from hospital patients, the PPG-BP data does not come from ICU and was acquired under controlled conditions following an experimental protocol. Data acquisition was conducted in private, following a

relaxation and adaptation period of 10 minutes, with the patients sitting in an office chair and their arms resting on a desk. The same acquisition devices were used for all subjects. Furthermore, a screening process excluded patients diagnosed with diseases other than cardiovascular diseases and diabetes. The data was also screened for abnormal and missing values, while a consistent signal quality was ensured by computing a signal quality index and excluding subjects with low values. [125]

2.4.2 Pre-processing

All signal processing was done in Python and references to functions are, otherwise noted, part of the standard library or of the *SciPy* scientific library [129].

For UCI, five evenly spaced segments of a duration of five seconds were first extracted from each of the records in the dataset. Records shorter than 25 seconds were rejected. SBP and diastolic blood pressure (DBP) were extracted from the continuous ABP signal by averaging all the peak values in the sequence, using function *find_peaks*. Records with less than three ABP peaks, due to non-pulsatile ABP segments, were rejected. Even though the UCI dataset had already been pre-processed to eliminate invalid or excessively noisy signals found in MIMIC, signal segments with movement artefacts, as well as sequences with large variations in pulsatile amplitude remained. To eliminate those issues and ensure coherence between the datasets, the following pre-processing steps were applied to both UCI and PPG-BP. First, all segments had their mean removed and were then filtered using a 0.7-12 Hz zero-phase fourth order Butterworth bandpass filter. The resulting PPG signals of each segment $x(n)$ and the corresponding first derivative $x'(n)$ were then screened with the following three rejection criteria:

$$\max(|x'(n)|) > \mu\left(x'(n)\right) + 5\sigma\left(x'(n)\right) \tag{14}$$

where σ is the standard deviations (STD) and μ the mean. This excluded signal segments with very rapid changes associated with signal artefacts such as those caused by body movements or device disconnection.

$$\begin{aligned} &\left(\max(x_i) - \min(x_i)\right) > 1.5\left(\max\left(x_j\right) - \min\left(x_j\right)\right) \\ &\text{for } i,j \in \{1,2,3\} \text{ and } i \neq j \end{aligned} \tag{15}$$

where x_1, x_2 and x_3 are three equally sized subdivisions of $x(n)$. This ensures similar pulsatile amplitude throughout the segment. It was not applied to PPG-BP because of the shorter segment duration.

$$PR < 40 \ \vee PR > 220 \tag{16}$$

where PR is the pulse rate in beat per minute (BPM) estimated as the average first derivative $x'(n)$ peak to peak interval, considering only peaks with a prominence greater than 60% of the maximum prominence. The prominence of a peak quantifies the difference between the peak amplitude and its bases, computed by function *peak_prominences*. This removed segments with extreme heart rate or with characteristics interfering with peak detection.

Finally, to be able to compare the various time-based features, both datasets were resampled to a matching frequency of 250 Hz.

2.4.3 Fiducial points extraction

The fiducial points used for feature extraction are shown in Figure 7. The second derivative of the signal was first computed and low-pass filtered with a 12 Hz zero-phase sixth order Butterwort filter to obtain $x''(n)$, after which the third derivative $x'''(n)$ was computed. The PPG pulses peak positions n_p, and their maximum upslope positions n_u were then established by finding the peak positions of $x(n)$ and $x'(n)$ with the *find_peaks* function, considering only peaks with a prominence greater than 60% of the maximum prominence. Boundaries for each pulse were established by finding the pulse onset, n_0, associated with

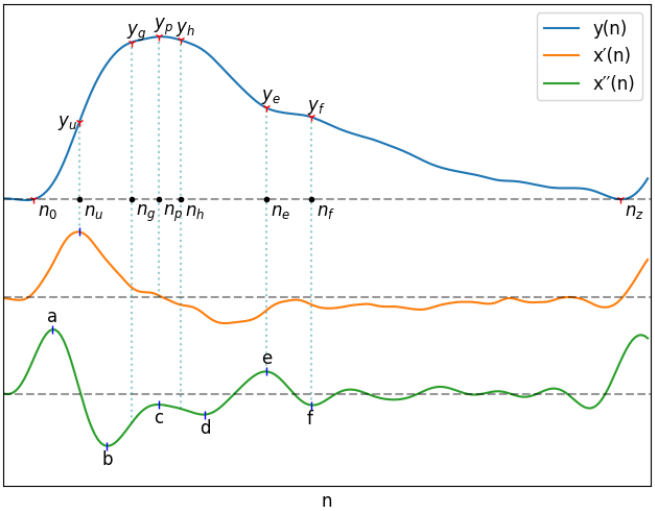

Figure 7: A typical PPG signal as well as its first and second derivatives with their most important fiducial points

each n_u. The position of n_0 was chosen as the first positive peak of $x''''(n)$ left of n_u, subject to $x'''(n) > 0.4 \max(x'''(n))$ to ignore minor peaks. If a positive zero crossing of $x'(n)$ could be found between that point and n_u, it was used instead. This strategy allowed a robust detection of onset even for pulses preceded with a slow rise before the onset. The end of the pulse, n_z, was defined as the next pulse onset. Pulses with a marked difference between the amplitude at onset and end point, satisfying $|x(n_0)-x(n_z)| > 0.12\,(x(n_p)-x(n_0))$, were discarded.

With pulse boundaries and peaks established, the remaining fiducial points were extracted from $x''(n)$. Five of those points are the *a, b, c, d* and *e* points described by Takazawa [81]. Since the *e* point also marks the position of the dicrotic notch, the same nomenclature was kept for the additionnal *f*, *g* and *h* points designating the second derivative estimates of the diastolic peak, early systolic peak and late systolic peak positions. To ensure robust extraction of the fiducial points, the following five step process was used:

1. Set the position of *a*, n_a, to the point where $x''(n)$ is at its maximum and the position of *b*, n_b, where it is at its minimum, subject to $n < n_p$.

2. Set the position of the dicrotic notch *e*, n_e as the earliest $x''(n)$ peak with n constrained by: $n_P < n < \frac{2}{3}(n_z-n_0) \wedge x(n) < 0.7\,x(n_p) \wedge x''(n) > 0.05\,x''(n_a)$.

3. Set the position of the diastolic peak *f*, n_f, as the earliest downward peak satisfying the condition $n_e < n < \frac{2}{3}(n_z-n_0) \wedge x''(n) < 0$.

4. Set the position of *c* and *d*, n_c and n_d, as the $x''(n)$ upward and downward peaks with the greatest difference between them, constrained by $n_b < n < n_e$. For pulses where those peaks did not exist, the positions were estimated as the position of the maximum inflection points of $x''(n)$, that is the maximum downward and upward peaks of the fourth derivative constrained by $n_b < n < n_e$.

5. Estimate the position of the early and late systolic peak by setting $n_g = n_b + \frac{n_c-n_b}{2}$ and $n_h = n_c + \frac{n_d-n_c}{2}$.

All peaks of $x''(n)$ and $x''''(n)$ were extracted by detection of the zero-crossings of the next higher order derivative.

2.4.4 Features extraction

All features were extracted on a pulse-by-pulse basis. The trend of the signal of each pulse was first removed by subtracting the linear slope connecting the start point of each pulse to its end point, as described in [130]. Thus, all pulses in the resulting detrended signal, $y(n)$, have value of zero at their starting and ending point. The amplitudes of the detrended signal at various fiducial point are hereafter designated by the form y_i where $y_i = y(n_i)$.

2.4.4.1 Amplitude ratios

The reflection index (RI) along with the augmentation index (AI) measure the contribution of the peripheral wave reflections to the overall pulse [59]. As a measure of reflected waves, AI can also be computed in regards to the early and late systolic peaks as in eq. (19) and eq. (20) while Y_{gh} defined in eq. (21) is an estimate of amplitude ratio of the late to early systolic peak, which has been correlated with changes in systolic pressure [89].

$$\mathrm{RI} = \frac{y_f}{y_p} \tag{17}$$

$$\mathrm{AI} = \frac{y_p - y_f}{y_p} = 1 - \mathrm{RI} \tag{18}$$

$$\mathrm{AI}_{gh} = \frac{y_g - y_h}{y_g} \tag{19}$$

$$\mathrm{AI}_{gf} = \frac{y_g - y_f}{y_g} \tag{20}$$

$$Y_{\mathrm{gh}} = \frac{y_h}{y_g} \tag{21}$$

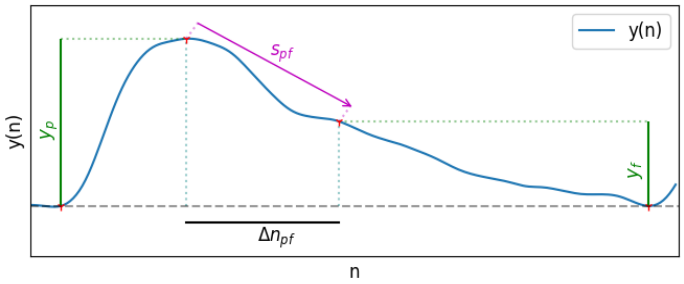

Figure 8: Different types of measurements used in feature extraction

2.4.4.2 Area ratios

The Inflection Point Area ratio (IPA) is the ratio of area under the curve up to the dicrotic notch and after it, and is an indicator of total peripheral resistance [59].

$$\mathrm{IPA}=\frac{\sum_{n=n_0}^{n_e} y(n)}{\sum_{n=n_e}^{n_z} y(n)} \tag{22}$$

2.4.4.3 Time spans

Time spans all take the same general form, given in eq. (23), and can be visualised in Figure 8. The duration of the systolic phase, Δn_{0p}, has been associated with hypertension [59], [76] while the duration of the diastolic phase, Δn_{pz}, has been associated with DBP [77]. The time spans Δn_{0g}, Δn_{0h}, Δn_{gh}, and Δn_{gf}, are spans between reflected waves components, of which Δn_{gf} has been associated with pulse pressure (PP) [89]. Δn_{pf} is the time between the peak and the diastolic peak. Δn_{up}, Δn_{ue}, Δn_{uf} are time spans in relation to the maximum upslope point, of which the last has been shown to have a strong correlation with SBP and DBP [131].

$$\Delta n_{ij} = n_j - n_i \tag{23}$$

The HR estimation used as a feature is also, in essence, a time span, and was calculated based on the pulse duration as shown in eq. (24), where f_s is the sampling rate.

$$\mathrm{HR} = \frac{60 f_s}{n_z - n_0} \tag{24}$$

2.4.4.4 Time ratios

Three different time ratios have been included in this study, each representing the pulse duration ratio before and after a fiducial point, taking the form shown in eq. (25). Included are N_p, in relation to the peak, N_e, in relation to the *e* point, and N_f, in relation to the *f* point. The time ratio of systole to diastole, N_f, was shown to be correlated to SBP [132].

$$N_i = \frac{\Delta n_{oi}}{\Delta n_{iz}} \tag{25}$$

2.4.4.5 Acceleration PPG

Acceleration PPG, or second derivative PPG, is a group of features extracted from the fiducial points in the second derivative of the signal. They have been associated with arterial stiffness and vascular aging [81]. The features *b/a*, *c/a*, *d/a* and *e/a* are amplitude ratios of the second derivative at those fiducial points, while the aging index (AX) is shown in eq. (26).

$$\mathrm{AX} = \frac{b\text{-}c\text{-}d\text{-}e}{a} \tag{26}$$

2.4.4.6 Slopes

The slopes from the peak to the dicrotic notch, S_{pe}, and to the diastolic peak, S_{pf}, have been investigated as BP predictors. S_{pe} was shown, although with low certainty, to have a weak correlation to DBP [131], and has also been associated with peripheral resistance [133]. Slopes used in this study are normalized, as in [131], in relation to the pulse peak value.

$$S_{pe} = \frac{y_e - y_p}{y_p \Delta n_{pe}} \tag{27}$$

$$S_{pf} = \frac{y_f - y_p}{y_p \Delta n_{pf}} \tag{28}$$

2.4.4.7 Widths

Widths are conceptually the same as time spans, but they are not calculated from specific fiducial points in the pulse. Rather, the span is the width of the pulse at a certain percentage of its amplitude. It has been used as a BP predictor [75] and is associated with systemic vascular resistance [73]. In this study, the pulse width is measured at 30%, 50%, 70% and 90% of y_p for W_{30}, W_{50}, W_{70} and W_{90}, respectively.

2.4.4.8 Outlier exclusion and feature vector construction

Outlier exclusion was performed on a per-pulse basis. Morphologically abnormal values for IPA were identified first and any values below 0.5, usually caused by an abnormal shape of the diastolic part of the pulse, were rejected. The feature vectors of both datasets were then temporarily joined to compute the global mean and the global STD, σ_{global}, of each feature. Pulses where any feature diverged more than 4 STD from the mean were considered outliers and rejected.

The remaining feature vectors for pulses in the same segment were then averaged and saved. Since only a single BP measurement is provided per subject in the PPG-BP dataset,

features extracted from different segments but from the same subject were also averaged together.

2.4.5 Statistical comparison of the datasets

To characterise the differences between PPG-BP and UCI, the distribution of features and BP data compiled in Section 2.4.4.8 were first examined.

For each feature, as well as SBP and DBP a two sample Kolmogorov–Smirnov (KS) test was performed with α=0.001 to determine if differences between distributions were significant.

For each dataset, the mean and STD of each feature was calculated. For each feature, the difference between the mean value of the two datasets, was determined as per eq. (29). The same was also done for the STD value as in eq. (30). The results were computed as a percentage of σ_{global} to bring them on a comparable scale. This analysis was also done on SBP and DBP.

$$\mu_{\%} = \frac{\mu_{uci} - \mu_{ppg\text{-}bp}}{|\sigma_{global}|} \cdot 100 \tag{29}$$

$$\sigma_{\%} = \frac{\sigma_{uci} - \sigma_{ppg\text{-}bp}}{|\sigma_{global}|} \cdot 100 \tag{30}$$

Since many features are affected by the pulse duration, those tests were then repeated with HR compensation. That is to say that all time spans (Section 2.4.4.3) and widths (Section 2.4.4.7) were multiplied by HR while slopes (Section 2.4.4.6) were divided by HR before recomputing σ_{global}, eq. (29) and eq. (30), yielding $\mu_{adj\%}$ and $\sigma_{adj\%}$.

Finally, the feature correlation matrix was computed: for each feature, the Pearson correlation coefficient (r) was calculated against every other feature. The difference between the correlation matrix of each dataset was then produced to highlight their discrepancies.

2.4.6 Response to BP variations and shared predictive power

2.4.6.1 BP correlation test

The Spearman rank correlation coefficient (ρ) was computed to assess correlation between each feature and SBP as well as DBP, respectively. Spearman correlation was selected here instead of Pearson for its ability to identify both linear and non-linear relationships. The

difference between the datasets was then computed to reveal any divergence in BP-features relationship.

2.4.6.2 BP estimation test

For this section, the *Scikit-Learn* machine learning library was used [134]. Using the *svm.svr* module, a support vector regression (SVR) model with a radial basis function (RBF) kernel was trained for SBP estimation on the PPG-BP dataset and another on UCI, keeping one random sample per subject. Therefore, when splitting a dataset into a training and testing set, data from one subject was never included into both the training and testing set.

To account for non uniformity of the sample distribution, sample weights were passed to the model for training and also in subsequent evaluation of performance. Samples were first split into 12 equally spaced bins based on their BP value. The weight *g* of each sample was $g_i = k_{max}/k_i$ where k_{max} is the number of samples in the bin with the most samples and k_i is the number of samples in the current sample's bin.

The features were centered to zero mean and scaled to unit variance before being handed to the model.

The model regularization parameter *C*, controling penalization of estimation errors during training, and the kernel function scale parameter γ, were optimized first through a coarse then a fine parameter grid search, as described in [135]. A leave-one-out cross-validation strategy was used to maximize the ammount of useable data for training.

Backward feature elimination was used to find the optimal feature set for each dataset, following this method:

1. Using 10-fold cross validation, sequentially train and test the SVR using all features but one, until all features have been left out once.

2. Compare the results and save the reduced feature set with the best cross-validated performance.

3. Restart from step one using the reduced feature set until only 4 features remain.

4. Select the optimal feature set, that is the one that had the best performance throughout the entire process.

At every step, performance was evaluated using the weighted coefficient of determination R^2, as defined in eq. (31), where i is the sample index, g_i is the sample weight, u_i is a sample's true BP, $\hat{u}_i$ is a sample's estimated BP, $\bar{u}$ is the weighted mean of the true BP of all samples defined in eq. (32), and k is the number of samples.

$$R^2 = 1 - \frac{\sum_{i=1}^{k} g_i (u_i - \hat{u}_i)^2}{\sum_{i=1}^{k} g_i (u_i - \bar{u})^2} \tag{31}$$

$$\bar{u} = \frac{\sum_{i=1}^{k} g_i u_i}{\sum_{i=1}^{k} g_i} \tag{32}$$

The Pearson correlation coefficient between the estimated BP and true BP, as well as the MAE of the estimated BP were used as secondary metrics. Final performance evalution with the optimised model parameters was carried out through leave-one-out cross-validation on the training dataset. The models were then retrained separately on their entire respective training dataset without leaving out any samples, but keeping the same set of features and the same C and γ values. Those retrained models were then validated against the other dataset to see if the predictive power would be retained.

2.5 Results

2.5.1 Pre-processing and feature extraction

For PPG-BP, 16 of the dataset's 657 segments were rejected by criterion (14) before feature extraction. No segments were rejected due to criterion (15) while criterion (16) did not apply to PPG-BP because of the short segment duration. From the remaining segments, 742 pulses were identified, of which 22 (3%) were rejected as outliers based on extracted feature values. Among the outliers, 10 had out-of-bounds IPA values, while other feature values were out-of-bounds in at most three pulses each. Averaging the remaining pulse features per segment yeilded 533 valid segments with complete feature vectors, for an overall segment rejection rate of 19%. After averaging per subject, the dataset had 211 feature vectors.

For UCI, 2376 records were too short to generate the segments and were ignored. The remaining records yielded 48120 segments, of which 1791 were rejected due to non-pulsatile ABP signals, 1228 because criterion (14), 2663 because of criterion (15) and 78 because of criterion (16). From the remaining segments, 83903 pulses were identified, of which 7104 (8%) were rejected as outliers based on extracted feature values. Again, IPA values were out-of-bound in 4069 or 5% of the total pulses, while other features had out-of-bound values in less than 1% of the pulses each.

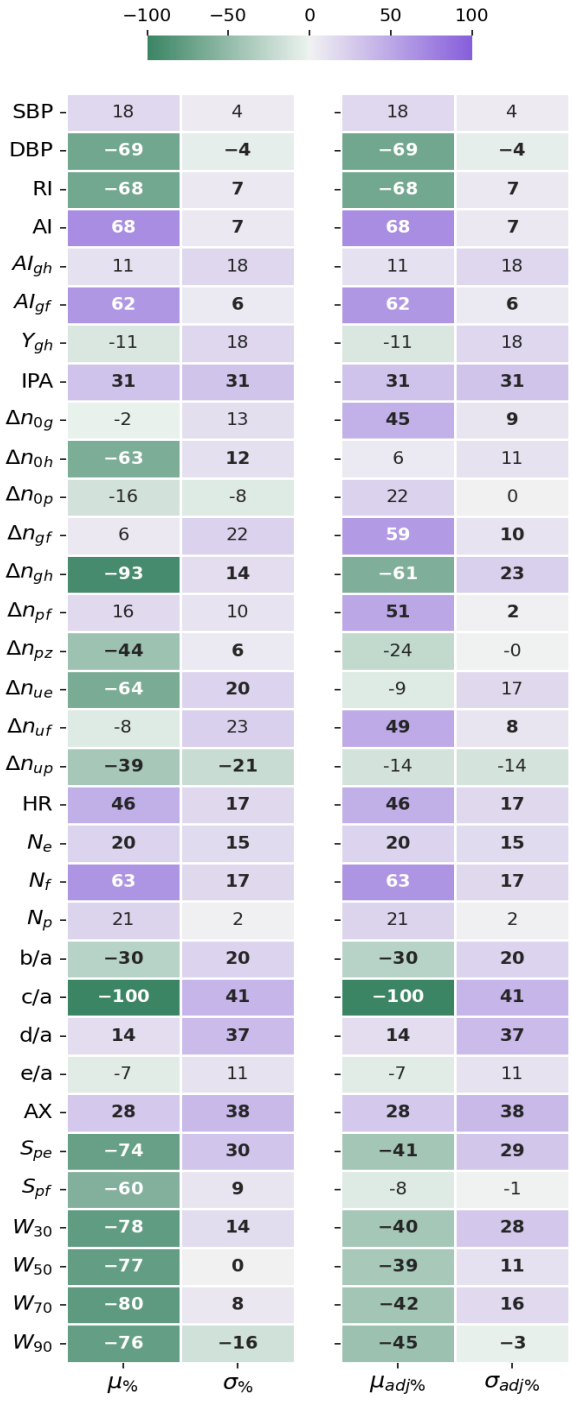

Figure 9: The difference in mean ($\mu_{\%}$) and standard deviation ($\sigma_{\%}$) between the datasets, given as a percentage standard deviation of the joined datasets. The HR adjusted forms ($\mu_{adj\%}$) and ($\sigma_{adj\%}$), compensate for the different HR distributions affecting time sensitive features. Negative values indicate that the mean or STD values for UCI are lower than for PPG-BP. Values in bold indicate significantly different distributions ($p < 0.001$) according to the Kolmogorov-Smirnov test.

Averaging the valid pulses per segment yeilded 21698 valid segments with complete feature vectors, for an overall segment rejection rate of 55%.

2.5.2 Statistical comparison

Results of the statistical comparison of the datasets are aggregated in Figure 9.

According to the KS-test, the differences between feature distributions were significant ($p < 0.001$) for 22 out of 31 features for the original features and 21 out of 31 features for the HR adjusted features.

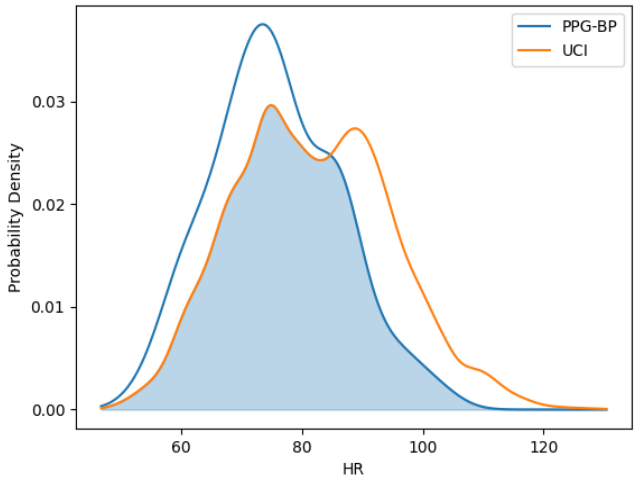

Figure 10: Comparison of the HR distribution for the PPG-BP and UCI datasets

Looking at $\mu_{\%}$ the difference in mean original feature values between the datasets, *c/a* stood out among all features, registering a difference of -100% of the standard deviation on UCI compared to PPG-BP. Several other features displayed a large difference, the second highest being Δn_{gh} (-93%), followed by the width features all showing at least -75% difference, S_{pe} (-74%), AI (68%) and RI (-68%). The difference between HR distributions (46%) is worth noting because of its direct physiological implication and its effect on other features. As shown in Figure 10, the UCI HR distribution is bimodal with a first peak positioned around 75 BPM, similar to PPG-BP, and a second peak close to 90 BPM. The average HR was 6.2 BPM higher in UCI and 28% of segments had a HR above 90 compared to 8% in PPG-BP.

Because HR directly affects the value of many features, looking at the HR corrected difference in mean $\mu_{adj\%}$ reveals what part of $\mu_{\%}$ is not explained by the difference in HR

distribution, and highlights fundamental differences in the pulse shapes. Values significantly higher on UCI were AI (68%), N_f (68%), AI_{gf} (62%), Δn_{gf} (59%), Δn_{pf} (51%), Δn_{uf} (49%), HR (46%), Δn_{0g} (45%), IPA (31%), AX (28%) and *d/a* (14%). Values significantly lower on UCI were *c/a* (-100%), RI (-68%), Δn_{gh} (-61%), W_{90} (-45%), W_{70} (-42%), S_{pe} (-41%), W_{30} (-40%), W_{50} (-39%) and *b/a* (-30%).

The five features with the highest STD difference were *c/*a (41%), AX (38%), *d/*a (37%), IPA (31%) and S_{pe} (30% or 29% adjusted for HR), all higher on UCI. In fact, STD was higher

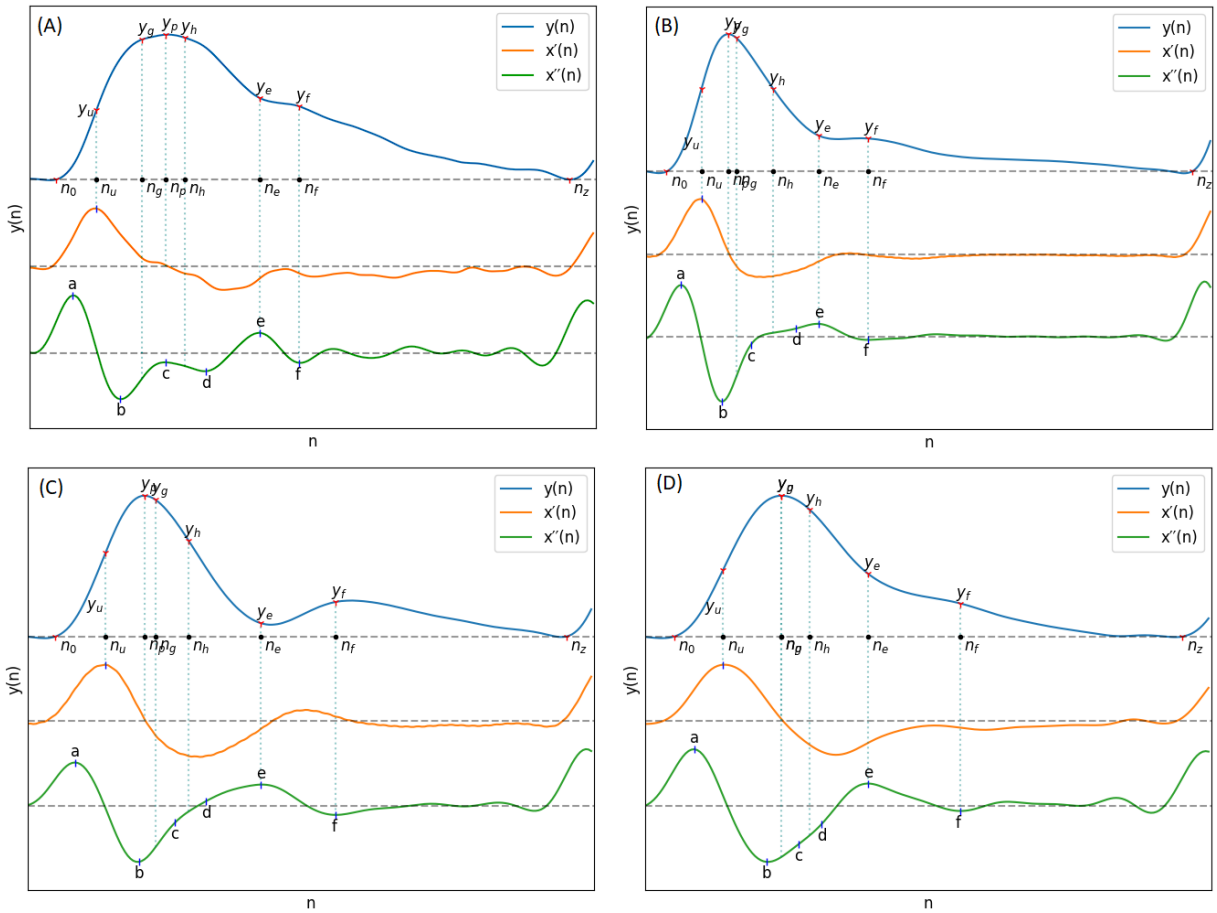

Figure 11: (A) Pulse from PPG-BP with characteristics representative of the dataset. (B), (C) and (D) Pulses from UCI illustrating some of the differences observed with PPG-BP. In general, the pulse shape was more pointed and narrower, dropping sharply after the peak. The amplitude of the PPG signal was usually lower at the e and f points, and the f point was often encountered later in the pulse. The second derivative showed a lot of variability, but compared to PPG-BP, the b point had usually a lower amplitude and the c and d points were often not well-defined peaks in the second derivative and were thus estimated from the inflection points. This resulted in highly variable but general lower amplitude values for the c point especially, compared to PPG-BP where it more consistently appears as a peak with a value closer to zero. Note that the pulse duration is normalised in all four pulses of this figure.

in UCI for 87% (or 80% adjusted for HR) of features, indicating a greater variability in pulse morphology within the dataset.

The relation of those differences to differences in pulse morphology between UCI and PPG-BP are illustrated in Figure 11. For example, the PPG-BP pulse with typical values (A) had a well defined second derivative peak for the *c* point with *c/a* = -0.15 while the depression at the *c* position for the bottom two pulses, (C) and (D), gave lower values of *c/a* = - 0.40 and *c/a* = -0.68. The *g* and *f* points were also positioned later in the pulse for (C) and (D), resulting in larger time spans. Pulse (C) had S_{pe} = -0.022, Δn_{0g} = 36, Δn_{uf} = 83, Δn_{pf} = 69 and N_f = 1.22 which can be directly compared to the values of (A), S_{pe} = -0.012, Δn_{0g} = 32, Δn_{uf} = 76, Δn_{pf} = 50 and N_f = 0.90, since both had similar heart rates. Pulse (C) also had a very narrow peak section with W_{90} = 14 while the PPG-BP pulse (A) had a wider one with W_{90} = 30. The heart rate of the UCI pulse (B) was 20 BPM lower than the PPG-BP pulse (A) but still only had W_{90} = 15. Pulse (B) also had AI = 0.76 because of the larger amplitude difference between *p* and *f* as well as a lower *b/a* = -1.26 caused by its more pronounced *b* peak in the second derivative. In comparison the PPG-BP pulse (A) had AI = 0.50 and *b/a* = -0.79. The variability of *c*/a in UCI is also illustrated in Figure 11, where the amplitude of *c* can be seen fluctuating between zero and the amplitude of *b* in the three UCI pulses.

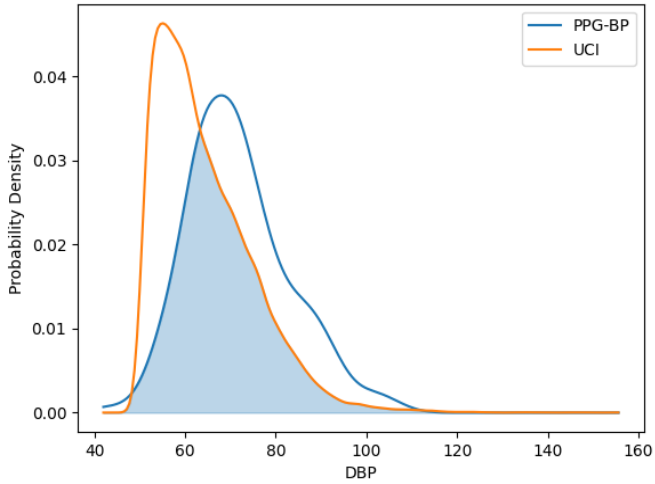

Figure 12: Comparison of the DBP distribution of the PPG-BP and UCI datasets

In regards to BP, the SBP distribution was similar on both datasets and close to normality. However, the DBP distributions had significant differences. The average DBP value for UCI was lower at 64.3 mmHG, compared to 71.8 mmHG for PPG-BP, or a difference equivalent

to -64% of the global sandard deviation. The DBP distribution of UCI was also found to deviate significantly from normality, as shown in Figure 12, with a slightly leptokurtic shape and a significant skew towards lower values.

Overall, the datasets had a similar degree of internal correlation, with $mean(|r|) = 0.36$ compared to $mean(|r|) = 0.35$ for PPG-BP. UCI had 61% of feature pairs with $|r|>0.25$ and 28% of feature pairs with $|r|>0.50$, as compared to 57% and 25%, respectively, for PPG-BP.

As for the correlation between features, the largest differences between datasets were observed with S_{pe}, a feature that also displayed a significant mean and STD differences between the datasets. Compared to PPG-BP, the correlation level $|r|$ of S_{pe} increased on average by 0.40 with seven other features in UCI: *b/a*, *d/a*, AX, Δn_{up}, Δn_{0p}, Y_{gh} and AI_{gh}. Another important difference was *e/a*, which had a correlation of $r = -0.42$ with *d/a* for PPG-BP, while that correlation fell to $r = 0.02$ for UCI.

The differences observed between the datasets were in large part associated with the presence of particularily pointed pulses in UCI and rare in PPG-BP. Those pulses hold a different relationship between features compared not only to most pulses in PPG-BP, but also to other types of pulses in UCI, increasing variability. Their caracteristics can be seen in the UCI Pulses of Figure 11. In general their *c* and *d* points were not well defined peaks in the second derivative, but inflection points in a curve between *b* and *e*. The amplitude of *d* tended to be higher as *e* also got higher and the S_{pe} slope became more pronounced. AX, which is calculated from the amplitude of the second derivative fiducial points, was in turn affected. Those pulses were also associated with a quick pulse onset with shorter Δn_{up}, Δn_{0p} and a proportionnaly narrower pulse wave. Finally the *g* point also tended to be situated around the peak while the *h* point came later at a much decreased PPG amplitude, wherea in PPG-BP the amplitude at *g* and *h* was not related to S_{pe} due to their more varied positions around a generally flatter pulse peak.

2.5.3 Response to BP variations and shared predictive power

2.5.3.1 Correlation to BP

The Spearman rank correlation coefficient (ρ) of each dataset's features against SBP and DBP is presented in Figure 13.

For SBP, significant correlation could be established for 15 features in PPG-BP. The three most correlating features were AI_{gh}(ρ = -0.50), Y_{gh}(ρ = 0.50), S_{pf}(ρ = -0.48), and a total of 14 features had a correlation of $|\rho| > 0.25$. For UCI, the three most correlating features were *c/a* (ρ = 0.24), W_{90} (ρ = 0.22) and Δn_{gh} (ρ = 0.20). It should be noted that those features all had major mean and STD differences with PPG-BP (see Section 2.5.2). In total, significant correlations with SBP could be established for 22 features in UCI, although coefficients were lower at $|\rho| \leq 0.25$ for all features.

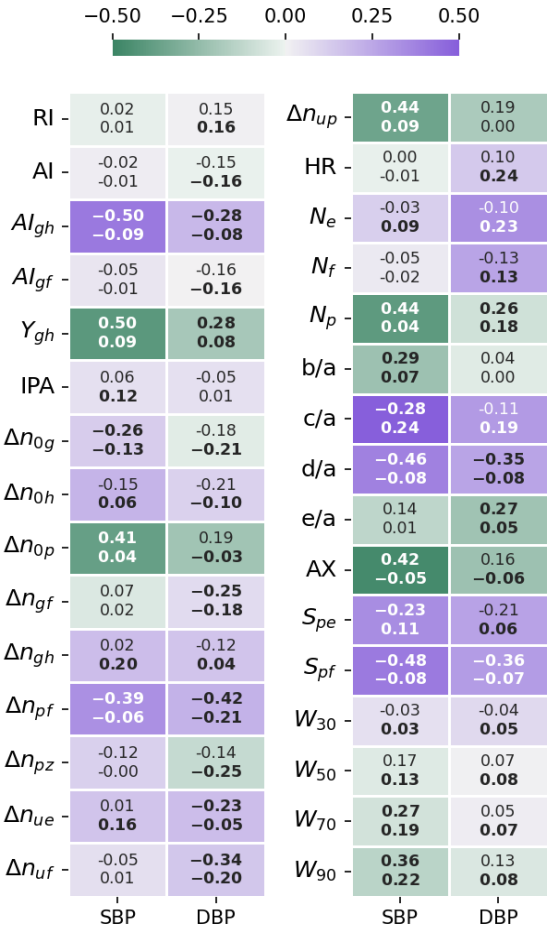

Figure 13: Feature-BP Spearman correlation test results. The top value is the correlation coefficient for PPG-BP, the bottom value the correlation coefficient for UCI, while the colour and intensity show by how much UCI differs from PPG-BP. Values in bold indicate that the correlation is significant ($p < 0.001$)

Stronger correlation with SBP in one dataset was not associated to a stronger correlation with SBP in the other dataset. For example, the three most correlating features of PPG-BP, AI_{gh}, Y_{gh} and S_{pf} ($|\rho| \geq 0.48$) were not among the higest in UCI, where their SBP correlation reached at most $|\rho| = 0.09$. As for the most correlating features in UCI, W_{90} obtained ρ = 0.36 in PPG-BP, Δn_{gh} was not significant and *c/a* had a stronger but opposite correlation of ρ = -0.28. Two other features showed significant but reversed correlation, although to a lesser degree: S_{pe} with ρ = -0.23 for PPG-BP and ρ = 0.11 for UCI, and AX with ρ = 0.42 for PPG-BP and ρ = -0.05 in UCI.

A similar pattern was observed for DBP. Significant correlations were established for ten features for PPG-BP. Those with the highest correlation were Δn_{pf} (ρ = -0.42), S_{pf} (ρ = -0.36), and *d/a* (ρ = -0.35). For UCI, significant correlations were established for a total of 28 features. Those with the highest correlation were Δn_{pz} (ρ = -0.25), HR (ρ = 0.24), and N_e(ρ = 0.23). In addition, relatively strong correlation (for UCI) was shared with one of

the most correlating features of PPG-BP: Δn_{pf} (ρ = -0.21). Again for DBP, correlation levels of $|\rho|>0.25$ were only reached on PPG-BP, and that for nine of the ten features where significance was attained.

2.5.3.2 BP estimation

Sampling one feature vector per subject in UCI for the BP estimation test yielded a total of 7087 vectors. Parameter selection for the PPG-BP trained model resulted in C = 75, γ = 0.1 while selected parameters for UCI were C = 0.25, γ = 0.03. For PPG-BP, eight features were retained during feature selection: N_e, S_{pf}, W_{90}, Δn_{gf}, Δn_{gh}, Δn_{pf}, AX and HR. For UCI, sixteen features were retained: AI_{gf}, N_f, S_{pf}, W_{30}, W_{50}, W_{90}, Δn_{0g}, Δn_{0h}, Δn_{uf}, Δn_{gh}, HR, AX, RI, b/a, c/a and d/a.

SBP estimation results for the PPG-BP trained model are presented in Figure 14 for cross-validated tests on PPG-BP. During cross-validated tests, the model tended to overestimate samples with low BP values and underestimated samples with high BP. Nonetheless, it showed significant predictive power over the entire range of BP values, as shown by the R^2 score of 0.63. Secondary metrics were r = 0.63 and MAE = 13.96 mmHg with an STD of 10.50 mmHg. When applied to predict SBP for the UCI dataset, a model with the same parameters trained with the entire PPG-BP did not retain any predictive power, as shown in Figure 15, giving worse results than a mean predictor, as shown by the R^2 score of -0.07. Secondary metrics were r = 0.09 and MAE = 21.03 mmHg with an STD of 16.95 mmHg.

Cross-validated results for UCI, shown in Figure 16 were considerably worse than for PPG-BP, achieving only R^2 = 0.31, with secondary metrics r = 0.42 and MAE = 16.56 mmHg with an STD of 12.82 mmHg. Applying the UCI trained model to PPG-BP resulted again in a loss of predictive power, as shown in Figure 17, although not as dramatic as for the PPG-BP trained model applied to UCI. It resulted in an R^2 score of 0.12, barely better than a mean predictor. Secondary metrics were r = 0.45 and MAE = 16.61 mmHg with an STD of 11.91 mmHg. The model can in fact be said to almost act as a mean predictor, as the produced BP values always remain close to the mean BP, with an STD of 5.74 mmHg.

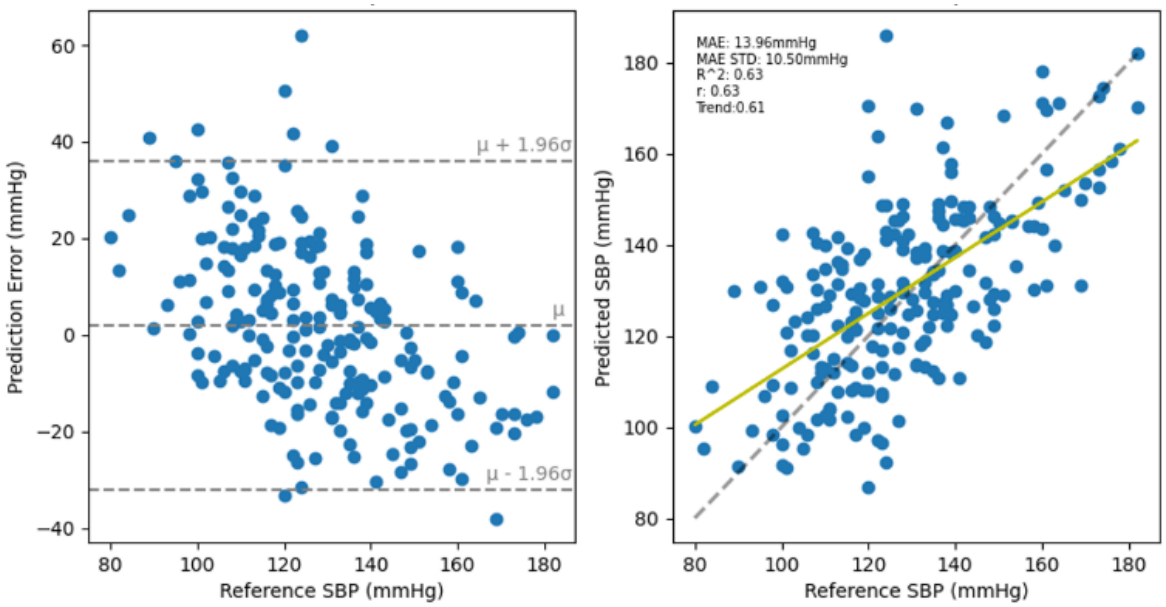

Figure 14: Cross validated results for the PPG-BP trained model

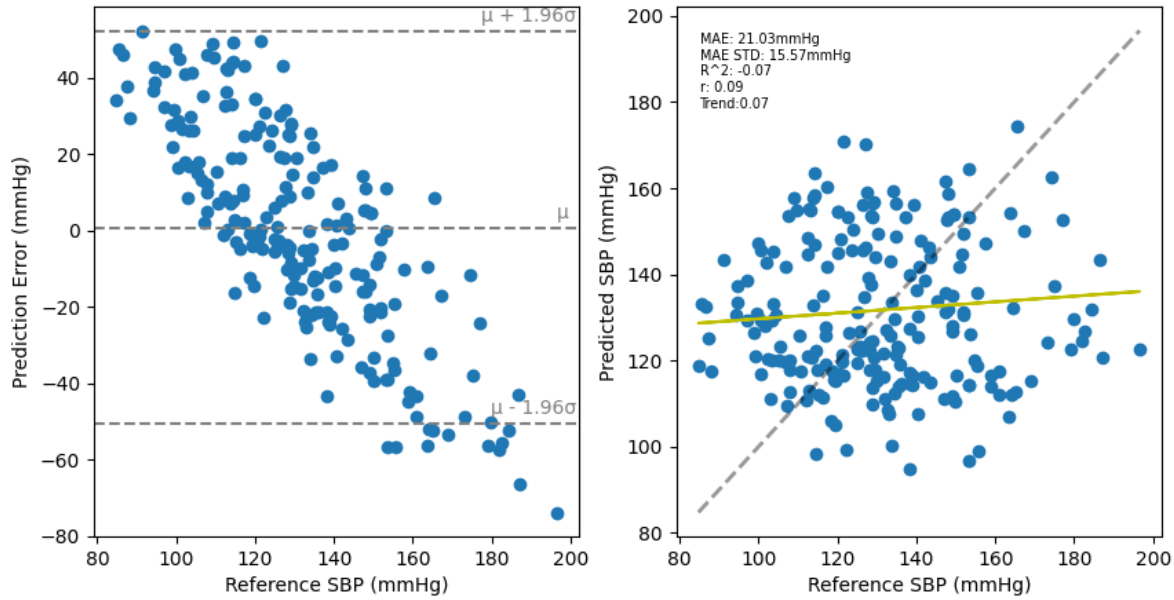

Figure 15: Results of the PPG-BP trained model tested with the UCI data. A random sample of the BP estimations are shown but the metrics are for the entire dataset.

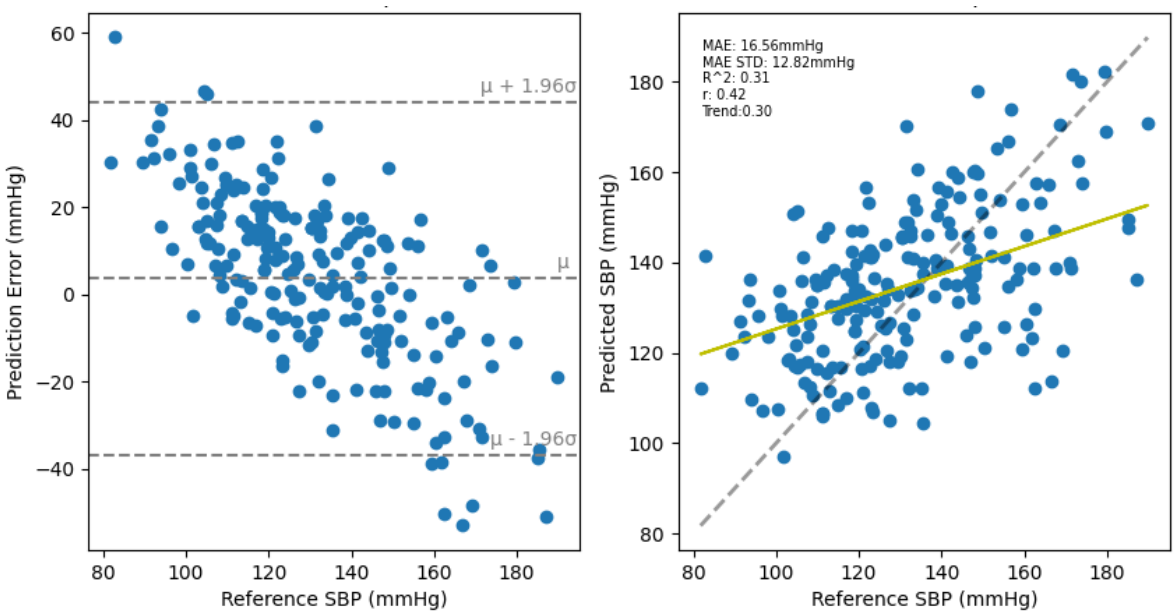

Figure 16: Cross validated results for the UCI trained model. A random sample of the BP estimations are shown but the metrics are for the entire dataset.

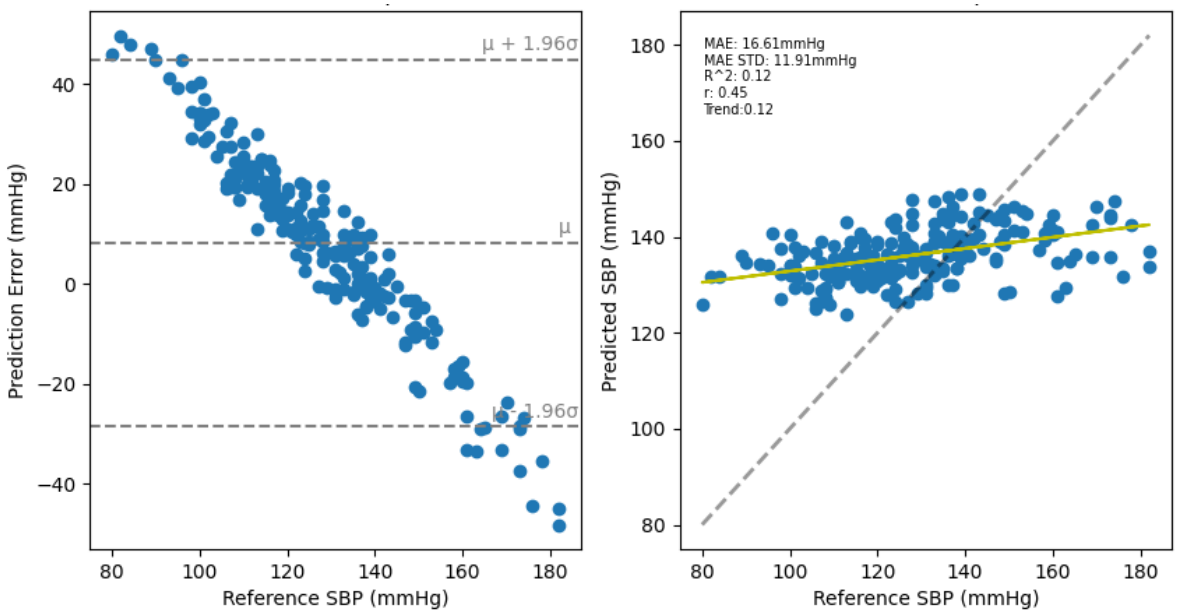

Figure 17: Results of the UCI trained model tested with the PPG-BP data.

2.6 Discussion

Analysis of the BP and features distributions showed fundamental differences between datasets. For several features, the difference can be partly attributed to the higher mean HR in UCI which results in narrower pulses. That characteristic of UCI and more specifically the large portion of data associated with a HR above 90, could be linked with stress or poor health. It supports the idea that differences in the conditions in which the data was obtained, or in the condition of the subject, has an influence on the data.

However, even when scaled by the heart rate, the difference in mean values between the datasets remained significant. That remainder was linked to morphological variations between the datasets, notably to the UCI pulse types illustrated in Figure 11. Those have a more pointed peak, more of their energy concentrated early in their period, and often lack well defined *c* and *d* peaks in the second derivative.

The loss of correlation between *d/a* and *e/a* on UCI may give some insight into the physiological origin behind the differences. The *e/a* ratio is associated with an increased inflection at the dicrotic notch while the *d/a* ratio is associated with inflection at the late systolic peak. A lower *d/a* ratio often equates to a flatter PPG peak with a sharp drop, as compared to a pointier PPG peak with a more progressive decline for higher ratios. In the wave-reflection based PPG model, this can be seen as the effect of timings and amplitude between the main systolic peak and the renal and iliac reflections [89]. The correlation between those features in the PPG-BP indicates that relatively uniform parameters in the circulatory system of the subjects define both reflections, while the loss of correlation on UCI indicates less uniformity, since the renal and iliac reflections appear modulated by different parameters. The *d/a* and *e/a* ratios have been shown to change independently with the administration of vasodilators or vasoconstrictors [81], which hints at possible differences in subject or environmental conditions between UCI and PPG-BP.

The relatively high degree of linear correlation between features was expected, as many features are similar in nature and are influenced in the same way by the pulse characteristics. For example widths and timings are all expected, to a certain degree, to vary together with the pulse duration.

The BP correlation test showed a different relationship between features and BP for each dataset. The higher correlation coefficients generaly found on PPG-BP indicated a more

uniform response between the subjects, which is coherent with the controlled data collection and subject selection methodology of PPG-BP, whereas the data from UCI lacks any control over environmental and subject conditions. For UCI, two of the features correlating the most with SBP were *c/a* and Δn_{gh}, features that also had the most difference between the datasets. Since those features had an opposite or null correlation on PPG-BP, the difference points to possible clusters of patients or conditions in UCI where consistant BP changes accompany those morphological changes. In fact, 35% of the pulses had a *c* amplitude lower than *d* in UCI, while it is the case for only 5% of pulses in PPG-BP. In UCI, those pulses were associated with an average SBP lower by 10.7 mmHg and average DBP lower by 3.8 mmHg compared to pulses with well defined second derivative peaks where $c > d$.

Of the features retained by the SBP estimation model for PPG-BP, four out of eight (S_{pf}, W_{90}, Δn_{pf} and AX) had significant correlation to SBP. Some of the feature that showed among the strongest correlation were not retained, which may be attributed to information redundance due to the generally strong correlation between features. For UCI, the large sample size allowed establishing significance at lower correlation levels, despite the increased variability of the data. The three features with the highest correlation to SBP (W_{90}, Δn_{gh} and *c/a*) were retained for estimation, the latter two also being the two features with the largest difference in mean value between datasets. Despite significant SBP correlation being present for *b/a*, *c/a* and *d/a*, no second derivative ratios were retained for PPG-BP. The fact that those three ratios were retained for UCI, and especially *c/a* with its opposite correlation profile compared to PPG-BP, maybe related to the aforementioned presence of clusters of patients with significant differences in the second derivative. It is also interesting to note that HR was retained for both datasets despite the absence of direct correlation to SBP, which suggests that scaling of some features in relation to the pulse duration played an important part in the estimation process.

Performance of BP estimation algorithms are extremely difficult to compare. The absence of a standard test dataset and the tradition of reporting the results in mmHg mean error or MAE make the results very sensitive to sample selection and BP range. Sample size, preprocessing and sampling methods vary widely, and are not always clearly described in published studies. Comparison with a few other calibration free studies can be made but should not be seen as decisive. Kachuee et al. obtained an MAE of 11.17 ±10.09 mmHG and $r = 0.59$ on UCI using Adaboost, but also making use of ECG [71]. Slapnicar et al.

obtained an MAE of 15.41 mmHg on 510 MIMIC subjects with a deep neural network on the raw PPG signal and the two first derivatives, while 18.34 mmHG was obtained when using a random forest algorithm with hand crafted features [136]. As a last example, Maqsood et al. tested the same algorithms on both PPG-BP and MIMIC (although without cross-dataset tests) and reported an MAE of 5.32 ± 4.26 mmHG for PPG-BP and 5.59 ± 5.92 mmHg for MIMIC with a bidirectional long short-term memory neural network (Bi-LSTM) and time domain features, while they obtained 15.48 ± 3.52 mmHG for PPG-BP and 12.14 ± 6.67 mmHG with an SVR [137].

While the use of more complex models such as Bi-LSTM may potentially bring uncalibrated BP estimation closer to medical device requirements, optimal performance was not the goal of this study and a simpler model was prefered to illustrate the impact of observed differences. The present results are thus more in line with those of other simpler models. More importantly, the present results clearly show that what was learned on one dataset does not apply, or applies only weakly, to the other. It is interesting to note that with fewer features, the cross-validated model of PPG-BP obtained an R^2 twice as that of UCI. The fact that less features and thus less information is necessary to get those results in PPG-BP indicates a more uniform response in the subjects, which may be due to the more controlled data collection conditions. This would also explain why the PPG-BP trained model retains no predictive power at all with UCI, since it would not cover the wider variety of patients and recording conditions present in UCI, while the UCI trained model, having knowledge of a wider variety of conditions, may be able to retain some power, even though very weak, when applied to PPG-BP.

Absolute values of the PPG signal can vary greatly depending on the recording conditions and equipment calibration. To ensure a consistent comparison between the different records, and especially between datasets possible, no raw amplitudes were used as features, neither was the DC component of the signal. Thus, a part of the signal information, which could potentially improve performance, was not used. The added benefit of this information in the case of UCI is however doubtful, as amplitudes were uneven between segments, with the pulsatile amplitude actually following a strict bimodal distribution with a wide separating gap.

Although not presented here, two pulse decomposition algorithms were evaluated to extract the *g* and *h* points: The recursive algorithm described by Kontaxis et al. [88] and the

Gaussian fitting algorithm described by Couceiro et al. [83]. The first one gave very inconsistant results for pulses with different shapes, such as more pointed or wider top pulses and may not be appropriate to compare between subjects with such differences. R^2 estimation results were also lower by as much as 0.18 with that method compared to the estimation method based on the second derivative described in Section 2.4.3. For the Gaussian fitting method, R^2 estimation results were similar while computation time for feature extraction was several times longer. The observation that some points in the second derivative were highly correlated with the position of the fitted Gaussian components resulted in using those points directly, as described in Section 2.4.3.

To conclude, the various private datasets used in the indirect BP measurement literature make comparing the published algorithms difficult, and researchers have called for the creation of a standardised dataset suitable to compare and validate BP estimation algorithms [8]. MIMIC, and by extension UCI, are publicly available and contain a large quantity of data, which may seem like a good basis for comparison. However, results presented in this paper reinforces suspicions of many researchers: that data sourced in intensive care units, under unknown conditions, may have a skewed response to BP and impair the generalisation of BP estimation algorithms. While the results do not postulate that a general model is impossible, they highlight significant differences in the relationship between BP and pulse features between the UCI data in contrast to data obtained under more controlled conditions.

3 Pulse decomposition analysis in the context of blood pressure estimation

3.1 Introduction

Pulse Decomposition Analysis (PDA) was introduced in Section 1.3.4 as the concept of decomposing the PPG pulse wave into an initial forward pulse wave and its reflection in the circulatory system. Notably, the first reflection, corresponding to the late systolic component of the pulse, is caused by the renal reflection site; while the second main reflection, appearing after the dicrotic notch, is caused by the iliac reflection site [83]. Two PDA methods were also briefly discussed: the Gaussian fitting method, a successive approximation method; as well as a lighter recursive method better suited for real-time applications. In Chapter 2, the early and late systolic components are also estimated through another method: a second derivative estimate of the position of those components. This chapter expands on this subject by presenting experiments comparing the three methods in the context of blood pressure estimation as well as supporting the use of the second derivative estimate in Chapter 2. At the time of writing, no comparison of these methods in the context of blood pressure estimation could be found in the literature.

3.2 Methodology

3.1.1 Gaussian fitting

The Gaussian fitting PDA method models the detrended PPG signal $y(n)$ as the sum of Z Gaussian functions, resulting in its approximation $\breve{y}(n)$. This summation is expressed by eq. (33), where i is the index of the Gaussian; e is Euler's number; and a_i, b_i, and c_i are the Gaussians' amplitude, location, and width, respectively. The optimal values of a_i, b_i, and c_i can then be found by minimizing the error function defined as eq. (34), where N is the length of the signal.

$$\breve{y}(n, \beta) = \sum_{i=1}^{Z} a_i e^{-\frac{(n+b_i)^2}{2c_i^2}} \quad \text{for } \beta=\{a_i, b_i, c_i\} \tag{33}$$

$$f_{error}(\beta) = \frac{1}{N}\sum_{n=1}^{N}[y(n)-\check{y}(n,\beta)]^2 \quad (34)$$

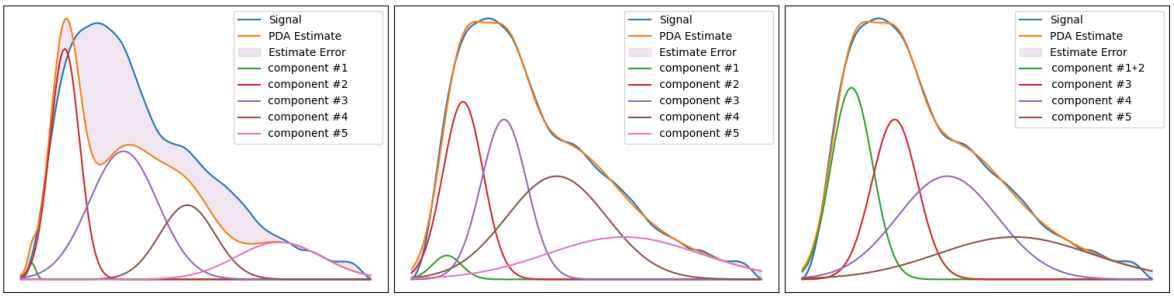

Figure 18: Position and size of the PDA components in relation to the PPG pulse with pre-optimization values (left), optimised values (center), and optimised values with the two first components combined (right)

The implementation for this project used Z=5, which has been shown to adequately approximate PPG [85]. The optimization was performed using Python function *optimize.minimize* with the sequential least square solver and a limit of 500 iterations. Initial values were set in relation to the pulses' fiducial points, as shown in Table 5; while boundaries and constrains are shown in Table 6 and Table 7. As suggested in [83], the first component was designed as an early and low amplitude component to improve fitting of the systolic rise, which is otherwise not approximated well by a Gaussian function. Thus, the first two components were merged after fitting to form the early systolic component to be used for feature extraction. The components and their sum, the PDA estimate, is visualized in Figure 18 in relation to the PPG signal prior to optimization, after optimization, and after merging the components.

Table 5: Parameter values used to initialize the Gaussian PDA model

Parameter	Initial values for components				
	#1	#2	#3	#4	#5
a_i	$0.7y(n_a)$	$0.9y(n_p)$	$0.5y(n_p)$	$0.4y(n_f)$	$0.2y(n_f)$
b_i	n_a	n_b	n_d	n_f	$n_f + (n_f - n_z) / 2$
c_i	$n_a / 3$	$n_b / 3$	$n_d / 3$	$n_f / 6$	$(n_f + (n_f - n_z) / 2) / 6$

Table 6: Boundaries applied to the parameters during optimization of the Gaussian PDA model

Parameter	Boundaries for components				
	#1	#2	#3	#4	#5
min(a_i)	0.1y(n_a)	0.3y(n_p)	0.2y(n_p)	0.1y(n_f)	0
max(a_i)	y(n_a)	y(n_p)	0.8y(n_p)	y(n_f)	y(n_f)
min(b_i)	n_a	n_a	n_c	n_e	n_f + (n_f - n_z) / 4
max(b_i)	n_b	n_c	n_d + (n_e - n_d) / 2	n_z	n_z
min(c_i)	0.01	n_a / 3	n_b / 3	0.01	0.01
max(c_i)	n_b / 3	n_d / 3	n_d + (n_e - n_d) / 6	n_f / 2	n_f / 2

Table 7: Inequality constrains applied to the parameters during optimization of the Gaussian PDA model

Parameter	Constrains for components				
	#1	#2	#3	#4	#5
a_i	< a_3	-	< 0.9a_2	< a_3	< 0.3a_4
b_i	< b_2	< b_3	< b_4	< b_5	-
c_i	-	-	-	-	-

3.1.2 Recursive pulse decomposition

The recursive method is a low computational complexity variation of PDA. The approach detailed in this section was first described in [88]. With this method, the detrended PPG signal $y(n)$ is approximated as per eq. (35) as the sum of Z components $y_i(n)$, where i is the index of the component, resulting in the approximation of the signal $\breve{y}(n)$. Components are determined sequentially by running the computation recursively on a running residual of the signal $\tilde{y}_i(n)$ defined as eq. (36). A component is defined in two equal-length parts: the first is equal to the running residual $\tilde{y}_i(n)$ from its slope onset n_{0_i} to its local maximum n_{m_i}, while the second part consists of a symmetrical reflection of the first part centered on n_{m_i}, as expressed by eq. (37). To avoid false detection of n_{m_i} in ripples often present early in $\tilde{y}_i(n)$, valid n_{m_i} positions were constrained to positions beyond the maximum upslope point of $\tilde{y}_i(n)$.

$$\check{y}(n) = \sum_{i=1}^{z} y_i(n) \tag{35}$$

$$\tilde{y}_i(n) = \begin{cases} y(n) & i = 1 \\ \tilde{y}_{i-1}(n) - y_{i-1}(n) & i > 1 \end{cases} \tag{36}$$

$$y_i(n) = \begin{cases} \tilde{y}_i(n), & n \in [n_{0_i}, n_{m_i}] \\ \tilde{y}_i(-n + 2n_{m_i}), & n \in [n_{m_i}, 2n_{m_i} - n_{0_i}] \\ 0, & \text{otherwise} \end{cases} \tag{37}$$

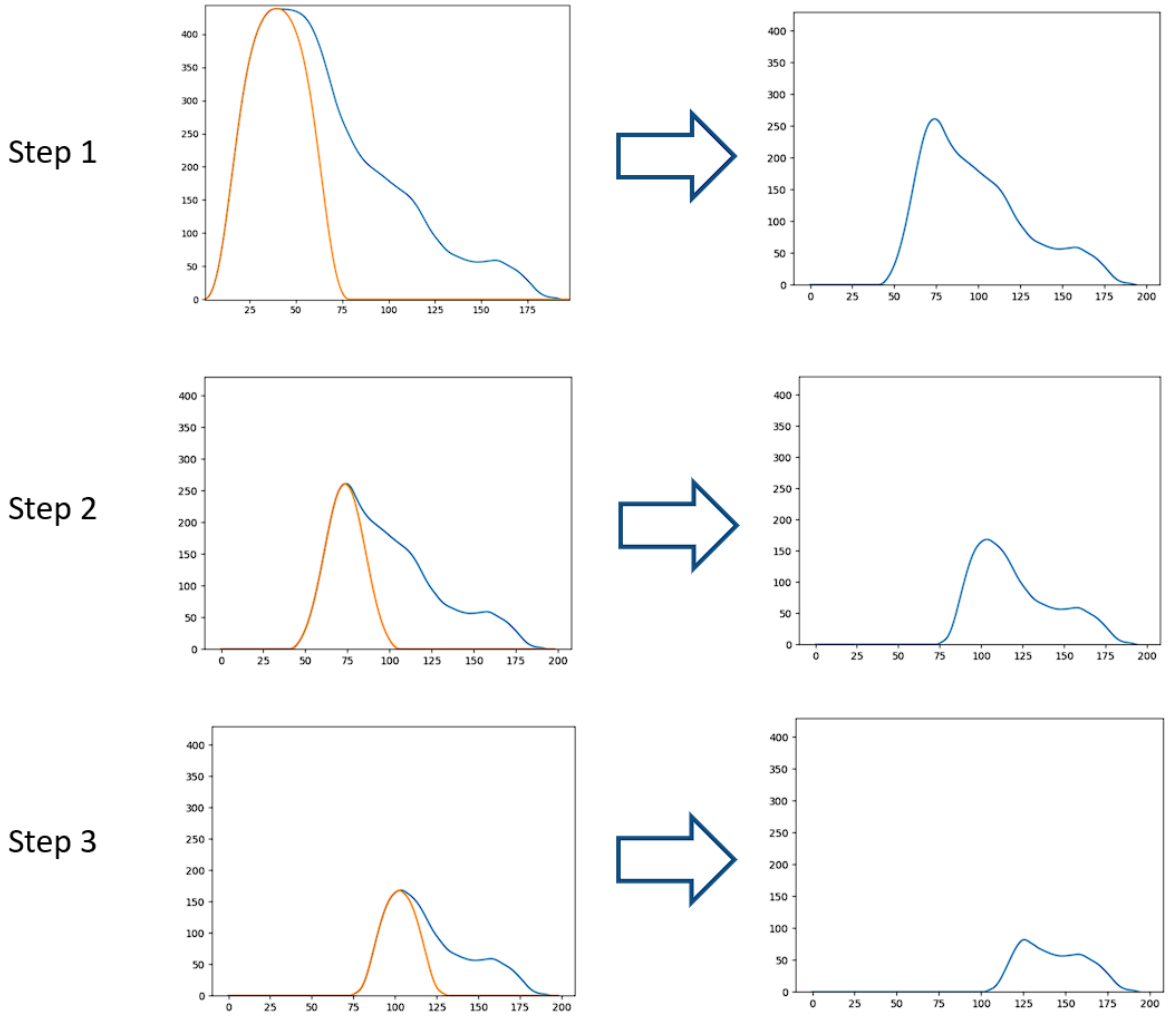

Figure 19: Illustration of the recursive PDA process. One component is identified from the running residual (left) and subtracted to create a new running residual (right) passed on to the next step

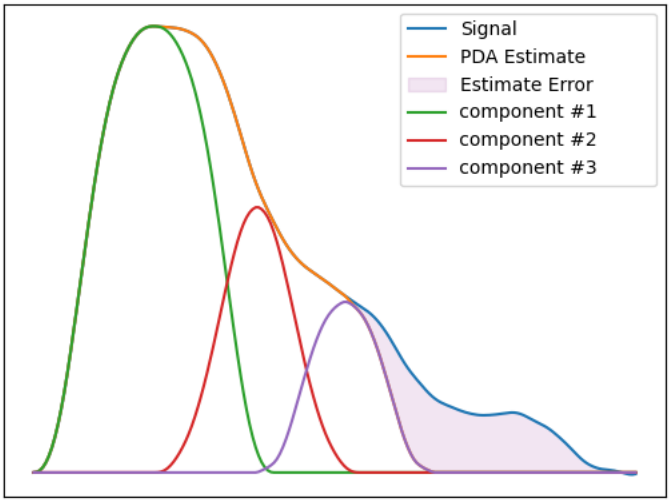

Figure 20: PDA for a PPG pulse produced with the recursive PDA algorithm with three components

Since determination of each component is not influenced by subsequent components, and the features analyzed called only for the first three components, Z=3 was chosen. The pulse decomposition process can be visualized in Figure 19, while an example of the results compared to the PPG pulse are shown in Figure 20. This technique results in an estimation error that is practically null over the zone of interest, while the residuals leave a rather large, but in the present case inconsequential, error at the end of the pulse. If needed, that error could be eliminated by adding more components.

3.1.3 Comparative tests

Using the PPG-BP dataset, the methodology previously described in Chapter 2 was used to re-extract features, with the difference that the *g*, *h,* and *f* fiducial points positions were set to the center point of the first, second, and third PDA component respectively. For each of the *g*, *h,* and *f* points, and for the two PDA methods, the Pearson correlation coefficient, as well as the mean absolute difference, was computed between the positions obtained from PDA and those obtained from the second derivative estimate. Spearman correlation for the PDA related features (AI_{gf}, AI_{gh}, Y_{gh}, Δn_{0g}, Δn_{0h}, Δn_{gf}, $\Delta n_{gh,}$ and Δn_{pf}) was then computed against SBP and DBP. The results were then compared to those obtained in Chapter 2, from features computed with the *g*, *h,* and *f* points obtained from the second derivative. Again, using the methodology described in Chapter 2, a SVM was optimized and cross validated for each of the PDA methods, and compared to the results previously reported. Finally, visual

analysis was performed on different pulse types to explain the statistical differences observed.

3.2 Results

3.2.1 Correlation of component positions against the second derivative estimate

As shown in Table 8, the position of components generated through Gaussian fitting showed almost perfect correlation to the second derivative estimates, the weakest coefficient being $r = 0.998$ for the third component. The recursive method also displayed lower, albeit very strong, correlation with $r \geq 0.960$ for all components. The mean absolute time difference between the positions were overall lower with Gaussian fitting, especially for the second component, which had an MAD of 17 ms compared to 92 ms for the recursive method. The third component showed approximately the same difference for both methods, with Gaussian fitting resulting in a MAD of 63 ms compared to 62 ms for the recursive method. Considering the average pulse duration of 791 ms, this represents up to 11.6% of the pulse duration, and up to 7.9% of the pulse duration for the largest mean error with the recursive and Gaussian fitting methods respectively. The close to perfect correlation of the Gaussian fitting component positions with those of the second derivative estimate implies that despite the rather large time difference it produces, the second derivative estimate can be used to track changes in the position of Gaussian fitted components with high accuracy from one pulse to the other.

Table 8: Correlation and mean absolute time difference between the position of PDA components and their second derivative estimates

Method	Pearson correlation coefficient for components ($p < 0.001$)			Mean absolute time difference for components (ms)		
	#1	#2	#3	#1	#2	#3
Gaussian fitting	0.999	0.999	0.998	30	17	63
Recursive	0.960	0.960	0.974	40	92	62

3.2.2 Correlation of features with BP

Correlation between the PDA related features and SBP is presented in Table 9, and their correlation to DBP is presented in Table 10. Overall, correlation to both SBP and DBP was stronger with Gaussian fitting. Similarities exist between the Gaussian fitting and second derivative method for both SBP and DBP. All features computed from the *g* and *h* points have similar correlation levels, except for $\Delta n_{0g,}$ which has significant correlation with SBP with the second derivative method but not with Gaussian fitting. Significantly correlated features also all share the same sign, indicating that features vary in the same direction in relation to BP changes. Differences were however present in features using the *f* point, where the second derivative method generally showed lower correlation levels compared to Gaussian fitting. The *f* point corresponds to the third component, which also displayed lower correlation and higher MAD between the two methods in Section 3.2.1.

The recursive method exhibited an important difference compared to the other methods for both SBP and DBP: the sign of the correlation coefficient of all shared significant features was opposed. Thus, if a variation in BP causes features produced through the second derivative or Gaussian fitting method to increase, those same features would decrease if computed through the recursive method. Another difference is that the recursive method produced significant correlation levels for more features with SBP than the two other methods. However, those correlation levels were overall lower, with a maximum for SBP of $|r| = 0.39$, while the other methods had maximums of $|r| \geq 0.50$. For DBP, the recursive method had only two significant correlating features, with a maximum $|r| = 0.23$, while the other methods had at least four significant features with maximums of $|r| \geq 0.40$.

Table 9: Spearman correlation of features against SBP for different PDA methods. Results in bold are significant ($p < 0.001$)

Method	Spearman correlation coefficient against SBP							
	AI_{gf}	AI_{gh}	Y_{gh}	Δn_{0g}	Δn_{0h}	Δn_{gf}	Δn_{gh}	Δn_{pf}
Second derivative	-0.05	**-0.50**	**0.50**	**-0.26**	-0.15	0.07	0.02	**-0.39**
Gaussian fitting	**-0.29**	**-0.50**	**0.50**	0.15	0.10	**-0.45**	0.04	**-0.53**
Recursive	**0.31**	**0.35**	**-0.35**	**0.39**	**0.39**	0.20	**0.38**	0.17

Table 10: Spearman correlation of features against DBP for different PDA methods. Results in bold are significant ($p < 0.001$)

Method	Spearman correlation coefficient against DBP							
	AI_{gf}	AI_{gh}	Y_{gh}	Δn_{0g}	Δn_{0h}	Δn_{gf}	Δn_{gh}	Δn_{pf}
Second derivative	-0.16	**-0.28**	**0.28**	-0.18	-0.21	**-0.25**	-0.12	**-0.42**
Gaussian fitting	**-0.26**	**-0.31**	**0.31**	-0.01	-0.09	**-0.40**	-0.13	**-0.39**
Recursive	0.12	**0.23**	**-0.23**	0.17	0.17	0.06	0.17	0.06

3.2.3 Effect of the PDA method on BP estimation

For each of the two PDA methods, the SVR model parameter selection process produced the same optimal parameter values as those produced for the second derivative method: $C = 75$, $\gamma = 0.1$. The features retained during backward feature elimination however differed based on the PDA method, and are given in Table 11. Among the three methods, only S_{pf}, AX, and HR are constant. All of the methods retained several features computed from the *g, h,* and *f* points, indicating that some usable information was carried by each of the PDA methods. The Gaussian fitting method obtained optimal results with ten features, similar to the second derivative method, which had eight. The recursive method however required 16 features to reach optimal results.

Table 11: Features retained during backward feature elimination for each PDA method

Method	Feature count	Retained features
Second derivative	8	N_e, S_{pf}, W_{90}, Δn_{gf}, Δn_{gh}, Δn_{pf}, AX and HR
Gaussian fitting	10	AI_{gh}, N_p, S_{pf}, W_{70}, W_{90}, Y_{gh}, Δn_{0g}, Δn_{pf}, AX and HR
Recursive	16	b/a, AI_{gh}, N_p, S_{pf}, W_{30}, W_{70}, Y_{gh}, Δn_{0h}, Δn_{gf}, Δn_{gh}, Δn_{pf}, Δn_{pz}, Δn_{uf}, Δn_{up}, AX and HR

As shown in Table 12, similar results were obtained by the Gaussian fitting and the second derivative method in the BP estimation test. For the Gaussian fitting method, the similar *r* value, lower MAE, slightly lower weighed R^2, and higher STD compared to the second derivative method indicates that estimation were more precise towards the mean BP value; while more variability was present at the extremes of the BP range. Thus, the second derivative provided a slightly better fit over the entire BP range.

The recursive method produced worse BP estimation results compared to the two other methods by all metrics. R^2 was lower by up to 0.18 points; *r* was lower by up to 0.12 points, while the MAE was up to 3.15 mmHg higher with an STD slightly higher than all other methods.

Table 12: BP estimation results for each PDA method

Method	R^2	*r*	MAE (mmHg)	STD (mmHg)
Second derivative	0.63	0.63	13.96	10.50
Gaussian fitting	0.61	0.64	12.95	11.08
Recursive	0.45	0.52	16.10	11.14

3.2.4 Variability of recursive method results in relation to pulse shape

Because the Gaussian fitting method is constrained by the second derivative fiducial points, positions of the components cannot diverge widely from expected positions. The recursive method is however unconstrained, and the central position of a component is only the result of the position of the next local maximum n_{m_i} in the running residual $\tilde{y}_i(n)$, while the width of a component is the result of the distance between the slope onset n_{0_i} and n_{m_i}. As a result, the components often diverged widely from expected positions, based on subtle variations in the pulse shape, such as the skewness of the systolic pulse wave. As shown in Figure 21, a peak skewed to the left would tend to compress the components early in the pulse, with the third component often present before the dicrotic notch. As previously mentioned, this third component, representing the second major reflection and coming from the iliac arteries, is known to contribute to the diastolic peak past the dicrotic notch. Figure 21 also illustrate another limitation of the recursive method: even though the positions of the two first components are similar to the Gaussian fitted positions, the amplitudes and widths imposed

on the components by the running residual prevent correct positioning of the third component.

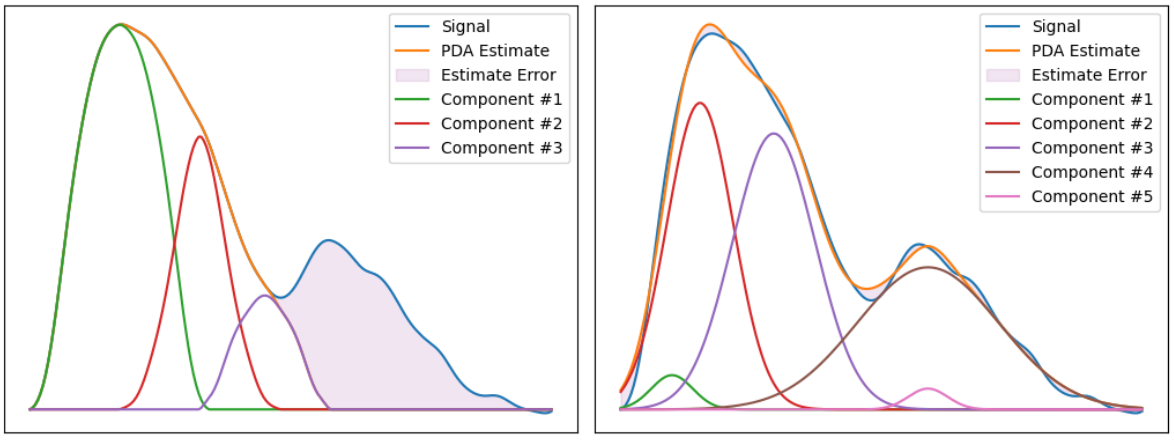

Figure 21: Comparison of PDA methods for a pulse with a left skewed systolic pulse. Results for the recursive method are shown on the left while the results for Gaussian fitting are shown on the right

Pulses skewed to the right display the opposite effect, where components are pushed to the right, and the second component is sometimes found after the dicrotic notch. This was also observed in systolic pulses with little or no skew, as shown in Figure 22: if the downstroke was relatively symmetrical to the upstroke, the next slope onset n_{0_i} in the running residual could be delayed, pushing other components to a later point in the pulse.

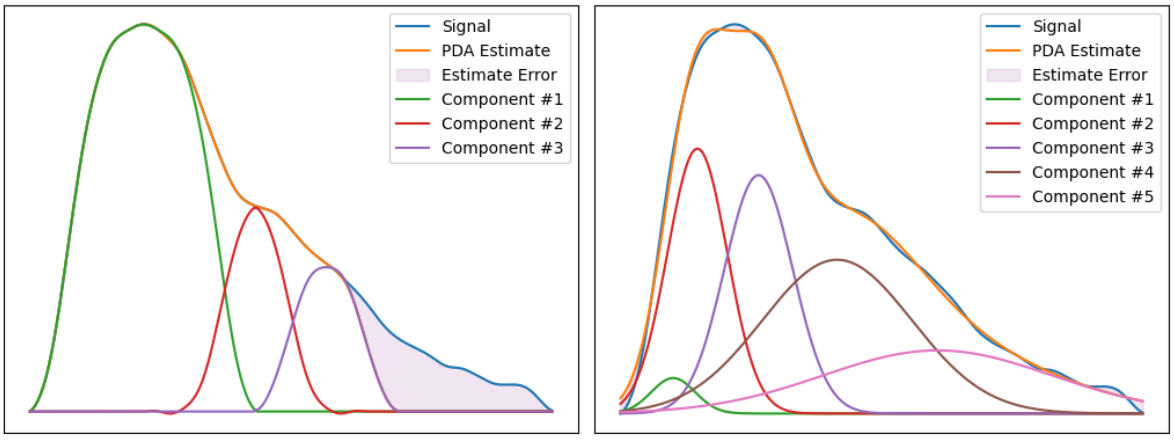

Figure 22: Comparison of PDA methods for a pulse with a certain symmetry in the upstroke and downstroke part of the systolic pulse wave. Results for the recursive method are shown on the left while the results for Gaussian fitting are shown on the right

3.3 Discussion

The general good concordance between the second derivative method and the Gaussian fitting method at all levels led to its use in the work described in Chapter 2. Surprisingly, the BP estimation results for the second derivative method are even better according to R^2, the main performance metric. The higher weighted R^2 indicates a better tracking of BP changes far from the population mean. This indicates that the second derivative points may carry additional useful information not provided by the Gaussian fitting method. Those better results with the second derivative estimate were also obtained, despite the better correlation of the *f* point features to SBP produced by Gaussian fitting, and the use of Δn_{pf} in the BP estimation model for both methods.

Although not used in the features presented here, amplitude of individual or combined Gaussian components (as opposed to the PPG amplitude at their center point) has been used to compute features correlated to physiological information, such as in [83]. Using those amplitudes might be beneficial for BP estimation. Further experimentation would be needed to determine if it provides an advantage over using the PPG signal amplitude, and if so, whether those amplitudes can somehow be estimated from the second derivative. The use of an estimation method instead of progressive approximation is desirable not only for real time applications, but also for large databases as the fitting time can take several seconds per pulse, rendering the extraction time impractical. Additionally, a better position estimation of the *f* point from the second derivative might also bring performance improvements, but no robust way of providing a better estimate was found in the context of this work.

The recursive method was found to be an inadequate alternative in the present context. Relatively small differences in PPG pulse shape types could position components outside of their expected range, differing greatly from the two other methods. This resulted in lower correlation levels of features with BP, and significantly reduced BP estimation performance. The recursive method is limited by the fact that the components are never allowed to overlap by more than half their width, limiting the range of possible positions. In contrast, a larger overlap is often observed when using the Gaussian fitting method. A secondary effect of that limitation is that the amplitude of individual components is always equal to the PPG signal at their center point, and thus does not provide the additional potentially useful information that is provided by Gaussian fitting. The recursive method does deliver its

promise of being light weight, with extraction times similar to the second derivative method. While the method was unhelpful for comparing pulse component positions between different individuals, where the type of pulse shape varies greatly, it may be valid and useful in detecting changes in a single individual with a known baseline level.

Discussion

The results presented in Chapter 2 provided, for the first time, evidence that the use of ICU data for the development of BP estimation algorithms may prevent them from generalizing to the normal population. The results also added weight to the calls from the research community for the creation of a standardized public database for comparing and validating BP estimation algorithms. Unfortunately, the currently limited nature of publicly available data imposed some limitations on the present work. Notably, because of the single BP record per subject available in the PPG-BP dataset, the comparison could only be done in the context of population-wide BP estimation. Calibration is however often seen as a necessary step to achieve acceptable accuracy for indirect BP measurement methods, a step that may also make models adapt better to different subgroups in the population. Therefore, future work could further investigate this topic by comparing public ICU data to another reference datasets with multiple records per subject, thus allowing calibration. One thing is however clear from the present results: The clinical potential of algorithms should not be judged based on their results obtained on ICU data alone.

Chapter 3 investigated the viability of two PDA methods and justification for instead estimate component timing through the second derivative. The second derivative method proposed in this work presents several advantages compared to the other methods, such as simplicity, speed and BP estimation performance. While the method works well for tracking component timing changes between pulses, it is poor at providing accurate absolute timings, especially for the *f* point, which may be an issue for some applications. The second derivative method also doesn't provide the amplitude of the individual PDA components and generating an estimate in a similar way has yet to be attempted. That additional information, characterizing the subject's arterial system, could potentially bring improvement in BP estimation performance and would be worth investigating in future work.

Other research opportunities also abound in the field. The temporal stability of calibration is an important topic that has been touched on by only a few published studies. Calibration methods without a reference device, such as through biometrics or additional sensors, have also yet to be proven effective, but would make future devices much easier to initialize.

At this time, the first PPG-based BP monitors are seeking medical certification and will be heading to market. Little is known of their internal working principles, and the current

regulatory void surrounding those new ambulatory devices may not give enough confidence to the medical community to replace the trusted cuff-based devices. But with time, additional research, transparency, and thorough testing, they may one day become the new standard.

Conclusion

The first objective of this research project was to investigate if the concerns were warranted regarding the use of ICU-sourced data, such as the MIMIC dataset, for the development of BP estimation methods. That topic was investigated in Chapter 2, which compared the MIMIC dataset to the PPG-BP dataset representing the general population. The analysis revealed significant differences in the distribution of diastolic BP and 20 out of 31 PPG morphological features, as well as significant differences in the response of features to BP changes. Furthermore, it was shown that data-driven models created following the same methodology and using the same features, with either MIMIC or the PPG-BP dataset, lost most of their predictive power when applied to predicting BP on the other dataset. While this does not ascertain that a model created with ICU data will not perform well on the general population, it shows that the peculiarity of ICU data may prevent generalization and that results obtained from ICU data should not be used as a benchmark of performance for the wider population.

As the main scientific contribution of this work, the content of Chapter 2 was submitted as a journal article to Physiological Measurement and should provide some much-needed clarity on the question of the use of ICU data in the field.

> Guillaume Weber-Boisvert, Benoit Gosselin, Frida Sandberg, “Intensive care photoplethysmogram datasets and machine-learning for blood pressure estimation: generalization not guarantied*", Physiological Measurement, 2022* (Pending review)

The second objective of this work was to evaluate and improve upon existing methods used in the field of BP estimation. That aspect was covered Chapter 3, which compared the benchmark Gaussian fitting method and the recursive method to a new method estimating the PDA timings from the second derivative fiducial points. The results demonstrated that the recursive method provides inconsistent results when comparing pulses with marked morphology differences and performs poorly in regression tests. In contrast, the second derivative method was shown to track component timing changes with high accuracy, while performing as well as the much slower Gaussian fitting method in the BP regression test. Therefore, the new method opens the door to accessing component timing information in situations where the complexity of the Gaussian fitting method is prohibitive. That could and will hopefully be beneficial for enabling more accurate BP estimation on wearable devices.

Bibliography

[1] "Facts About Hypertension," *Center for Disease Control and Prevention*. https://www.cdc.gov/bloodpressure/facts.htm

[2] "The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure," US Department of Health and Human Services, Dec. 2003.

[3] "Deaths: Final Data for 2017," US Department of Health and Human Services, Volume 68, Number 9, Jun. 2019.

[4] G. Danaei *et al.*, "The Preventable Causes of Death in the United States: Comparative Risk Assessment of Dietary, Lifestyle, and Metabolic Risk Factors," *PLoS Med.*, vol. 6, no. 4, p. e1000058, Apr. 2009, doi: 10.1371/journal.pmed.1000058.

[5] B. A. Bushman, "Blood Pressure Basics and Beyond," *ACSMs Health Fit. J.*, vol. 20, no. 3, pp. 5–9, Jun. 2016, doi: 10.1249/FIT.0000000000000198.

[6] "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines," p. 103.

[7] S. Lewington, R. Clarke, and N. Qizilbash, "Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies," *The Lancet*, vol. 360, pp. 1903–1913, Dec. 2002.

[8] J. Solà and R. Delgado-Gonzalo, Eds., *The Handbook of Cuffless Blood Pressure Monitoring: A Practical Guide for Clinicians, Researchers, and Engineers*. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-24701-0.

[9] A. Roguin, "Scipione Riva-Rocci and the men behind the mercury sphygmomanometer: The History Behind the Mercury Sphygmomanometer," *Int. J. Clin. Pract.*, vol. 60, no. 1, pp. 73–79, Sep. 2005, doi: 10.1111/j.1742-1241.2005.00548.x.

[10] S. B. George, "Hales' 1733 Haemastaticks," *Anesthesiology*, vol. 112, no. 1, Jan. 2010.

[11] R. E. Dudgeon, *The sphygmograph: its history and use as an aid to diagnosis in ordinary practice*, Baillière, Tindall, And Cox. London, 1882.

[12] J. L. M. Poiseuille, “Recherches sur la force du coeur aortique,” Paris, 1828.

[13] K. Vierordt, *Die Lehre vom Arterienpuls in gesunden und kranken Zuständen*. Braunschweig, 1855.

[14] C. Lawrence, “Physiological apparatus in the Wellcome Museum. 1. The Marey sphygmograph,” *Med. Hist.*, vol. 22, no. 2, pp. 196–200, Apr. 1978, doi: 10.1017/S0025727300032324.

[15] E.-J. Marey, *Physiologie médicale de la circulation du sang*. Paris: Adrien Delahaye, 1863.

[16] M. F. O’Rourke, “Frederick Akbar Mahomed,” *Hypertension*, vol. 19, no. 2, pp. 212–217, Feb. 1992.

[17] A. S. Tijsseling and A. Anderson, “A. Isebree Moens and D.J. Korteweg: on the speed of propagation of waves in elastic tubes,” p. 19.

[18] F. A. Mahomed, “On Chronic Bright’s Disease, and its Essential Symptoms,” *The Lancet*, Mar. 1879.

[19] J. Booth, “A Short History of Blood Pressure Measurement,” *Proc. R. Soc. Med.*, vol. 70, pp. 793–799, Nov. 1970.

[20] G. Mancia, “Scipione Riva-Rocci,” *Clin. Cardiol.*, vol. 20, no. 5, pp. 503–504, May 1997, doi: 10.1002/clc.4960200520.

[21] D. Paskalev, A. Kircheva, and S. Krivoshiev, “A Centenary of Auscultatory Blood Pressure Measurement: A Tribute to Nikolai Korotkoff,” *Kidney Blood Press. Res.*, vol. 28, no. 4, pp. 259–263, 2005, doi: 10.1159/000090084.

[22] C. Bramwell, “Blood-Pressure and its estimation,” *The Lancet*, pp. 138–140, Jan. 1940.

[23] W. F. Hamilton and P. Dow, “An experimental study of the standing waves in the pulse propagated through the aorta,” 1938.

[24] W. F. Hamilton, “The patterns of the arterial pressure pulse,” *Am. J. Physiol.*, vol. 141, no. 2, pp. 235–241, 1944, doi: 10.1152/ajplegacy.1944.141.2.235.

[25] W. W. Nichols, W. W. Nichols, and D. A. McDonald, Eds., *McDonald’s blood flow in arteries: theoretic, experimental, and clinical principles*, 6th ed. London: Hodder Arnold, 2011.

[26] F. O’Rourke, “The arterial pulse in health and disease,” *Fundam. Clin. Cardiol.*, vol. 82, no. 5, 1971.

[27] D. A. McDonald and M. G. Taylor, “The Hydrodynamics of the Arterial Circulation,” *Prog. Biophys. Biophys. Chem.*, vol. 9, pp. 105–173, 1959, doi: 10.1016/S0096-4174(18)30167-7.

[28] M. F. O’Rourke, “Pressure and flow waves in systemic arteries and the anatomical design of the arterial system,” *J. Appl. Physiol.*, vol. 23, no. 2, pp. 139–149, Aug. 1967, doi: 10.1152/jappl.1967.23.2.139.

[29] M. F. O’Rourke and M. G. Taylor, “Input Impedance of the Systemic Circulation,” *Circ. Res.*, vol. 20, no. 4, pp. 365–380, Apr. 1967, doi: 10.1161/01.RES.20.4.365.

[30] M. F. O’Rourke and A. P. Avolio, “Pulsatile flow and pressure in human systemic arteries. Studies in man and in a multibranched model of the human systemic arterial tree.,” *Circ. Res.*, vol. 46, no. 3, pp. 363–372, Mar. 1980, doi: 10.1161/01.RES.46.3.363.

[31] R. N. Westhorpe and C. Ball, “Blood Pressure Monitoring – Automated Non-invasive Blood Pressure Monitors,” *Anaesth. Intensive Care*, vol. 37, no. 3, pp. 343–343, May 2009, doi: 10.1177/0310057X0903700321.

[32] G. Ogedegbe and T. Pickering, “Principles and techniques of blood pressure measurement,” *Cardiol. Clin.*, Nov. 2010.

[33] B. S. Alpert, D. Quinn, and D. Gallick, “Oscillometric blood pressure: a review for clinicians,” *J. Am. Soc. Hypertens.*, vol. 8, no. 12, pp. 930–938, Dec. 2014, doi: 10.1016/j.jash.2014.08.014.

[34] R. Mukkamala *et al.*, “Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice,” *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1879–1901, Aug. 2015, doi: 10.1109/TBME.2015.2441951.

[35] I. Guelen *et al.*, “Validation of brachial artery pressure reconstruction from finger arterial pressure,” *J. Hypertens.*, vol. 26, no. 7, pp. 1321–1327, Jul. 2008, doi: 10.1097/HJH.0b013e3282fe1d28.

[36] W. J. W. Bos, J. van Goudoever, G. A. van Montfrans, A. H. van den Meiracker, and K. H. Wesseling, “Reconstruction of Brachial Artery Pressure From Noninvasive Finger Pressure Measurements,” *Circulation*, vol. 94, no. 8, pp. 1870–1875, Oct. 1996, doi: 10.1161/01.CIR.94.8.1870.

[37] B. Imholz, “Fifteen years experience with finger arterial pressure monitoring: assessment of the technology,” *Cardiovasc. Res.*, vol. 38, no. 3, pp. 605–616, Jun. 1998, doi: 10.1016/S0008-6363(98)00067-4.

[38] D. N. Stokes, T. Clutton-Brock, C. Patil, J. M. Thompson, and P. Hutton, “Comparison of invasive and non-invasive measurement of continuous arterial pressure using the Finapres,” *Br. J. Anaesth.*, vol. 67, no. 1, pp. 26–35, Jul. 1991, doi: 10.1093/bja/67.1.26.

[39] M. O. Fischer *et al.*, “Non-invasive continuous arterial pressure and cardiac index monitoring with Nexfin after cardiac surgery,” *Br. J. Anaesth.*, vol. 109, no. 4, pp. 514–521, Oct. 2012, doi: 10.1093/bja/aes215.

[40] J. J. Vos *et al.*, “Comparison of continuous non-invasive finger arterial pressure monitoring with conventional intermittent automated arm arterial pressure measurement in patients under general anaesthesia,” *Br. J. Anaesth.*, vol. 113, no. 1, pp. 67–74, Jul. 2014, doi: 10.1093/bja/aeu091.

[41] S. Eckert and D. Horstkotte, “Comparison of Portapres non-invasive blood pressure measurement in the finger with intra-aortic pressure measurement during incremental bicycle exercise,” *Blood Press. Monit.*, vol. 7, no. 3, pp. 179–183, Jun. 2002, doi: 10.1097/00126097-200206000-00006.

[42] J. R. Martina *et al.*, “Noninvasive Continuous Arterial Blood Pressure Monitoring with Nexfin®,” *Anesthesiology*, vol. 116, no. 5, pp. 1092–1103, May 2012, doi: 10.1097/ALN.0b013e31824f94ed.

[43] P. Klose *et al.*, “Continuous noninvasive monitoring of arterial pressure using the vascular unloading technique in comparison to the invasive gold standard in elderly comorbid patients: A prospective observational study,” *Health Sci. Rep.*, vol. 3, no. 4, Dec. 2020, doi: 10.1002/hsr2.204.

[44] J. F. Stover *et al.*, “Noninvasive cardiac output and blood pressure monitoring cannot replace an invasive monitoring system in critically ill patients,” *BMC Anesthesiol.*, vol. 9, no. 1, p. 6, Dec. 2009, doi: 10.1186/1471-2253-9-6.

[45] S. Hansen and M. Staber, “Oscillometric blood pressure measurement used for calibration of the arterial tonometry method contributes significantly to error,” *Eur. J. Anaesthesiol.*, vol. 23, no. 9, pp. 781–787, Sep. 2006, doi: 10.1017/S0265021506000688.

[46] H.-M. Cheng, D. Lang, C. Tufanaru, and A. Pearson, “Measurement accuracy of non-invasively obtained central blood pressure by applanation tonometry: A systematic review and meta-analysis,” *Int. J. Cardiol.*, vol. 167, no. 5, pp. 1867–1876, Sep. 2013, doi: 10.1016/j.ijcard.2012.04.155.

[47] D. M. D. Ayman and A. D. M. D. Goldshine, “BLOOD PRESSURE DETERMINATIONS BY PATIENTS WITH ESSENTIAL HYPERTENSION: I. The Difference Between Clinic and Home Readings Before Treatment,” *J. Med. Sci.*, vol. 200, no. 4, pp. 465–474, Oct. 1940.

[48] T. G. Pickering, G. D. James, C. Boddie, G. A. Harshfield, S. Blank, and J. H. Laragh, “How Common Is White Coat Hypertension?,” *JAMA*, vol. 259, no. 2, pp. 225–228, Jan. 1988, doi: 10.1001/jama.1988.03720020027031.

[49] T. G. Pickering, K. Davidson, W. Gerin, and J. E. Schwartz, “Masked Hypertension,” *Hypertension*, vol. 40, no. 6, pp. 795–796, Dec. 2002, doi: 10.1161/01.HYP.0000038733.08436.98.

[50] Franklin Stanley S., Thijs Lutgarde, Hansen Tine W., O'Brien Eoin, and Staessen Jan A., "White-Coat Hypertension - New insights from recent studies," *Hypertension*, vol. 62, no. 6, pp. 982–987, Dec. 2013, doi: 10.1161/HYPERTENSIONAHA.113.01275.

[51] T. G. Pickering, K. Eguchi, and K. Kario, "Masked Hypertension: A Review," *Hypertens. Res.*, vol. 30, no. 6, pp. 479–488, 2007, doi: 10.1291/hypres.30.479.

[52] D. M. Rabi *et al.*, "Hypertension Canada's 2020 Comprehensive Guidelines for the Prevention, Diagnosis, Risk Assessment, and Treatment of Hypertension in Adults and Children," *Can. J. Cardiol.*, vol. 36, no. 5, pp. 596–624, May 2020, doi: 10.1016/j.cjca.2020.02.086.

[53] T. W. Hansen, Y. Li, J. Boggia, L. Thijs, T. Richart, and J. A. Staessen, "Predictive Role of the Nighttime Blood Pressure," *Hypertension*, vol. 57, no. 1, pp. 3–10, Jan. 2011, doi: 10.1161/HYPERTENSIONAHA.109.133900.

[54] Y. Yano and K. Kario, "Nocturnal blood pressure and cardiovascular disease: a review of recent advances," *Hypertens. Res.*, vol. 35, no. 7, pp. 695–701, Jul. 2012, doi: 10.1038/hr.2012.26.

[55] G. Wang, M. Atef, and Y. Lian, "Towards a Continuous Non-Invasive Cuffless Blood Pressure Monitoring System Using PPG: Systems and Circuits Review," *IEEE Circuits Syst. Mag.*, vol. 18, no. 3, pp. 6–26, 2018, doi: 10.1109/MCAS.2018.2849261.

[56] A. Bayés de Luna, *Basic electrocardiography: normal and abnormal ECG patterns*. Malden, MA: Blackwell Futura, 2007.

[57] F. Kusumoto, *ECG Interpretation: From Pathophysiology to Clinical Application*. Cham: Springer International Publishing, 2020. doi: 10.1007/978-3-030-40341-6.

[58] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, Mar. 2007, doi: 10.1088/0967-3334/28/3/R01.

[59] M. Elgendi, "On the Analysis of Fingertip Photoplethysmogram Signals," *Curr. Cardiol. Rev.*, vol. 8, no. 1, pp. 14–25, Jun. 2012, doi: 10.2174/157340312801215782.

[60] Y. Ma *et al.*, “Relation between blood pressure and pulse wave velocity for human arteries,” *Proc. Natl. Acad. Sci.*, vol. 115, no. 44, pp. 11144–11149, Oct. 2018, doi: 10.1073/pnas.1814392115.

[61] H. J. Baek, K. K. Kim, J. S. Kim, B. Lee, and K. S. Park, “Enhancing the estimation of blood pressure using pulse arrival time and two confounding factors,” *Physiol. Meas.*, vol. 31, no. 2, pp. 145–157, Feb. 2010, doi: 10.1088/0967-3334/31/2/002.

[62] “SomnoTouch NIBP brochure.” Somnomedics. Accessed: Jan. 05, 2021. [Online]. Available: https://somnomedics.de/wp-content/uploads/2020/02/Brosch%C3%BCre_SOT-NIBP_EN_Rev_4_2020_01_final.pdf

[63] “ViSi Mobile Brochure.” Sotera Wireless. Accessed: Jan. 05, 2021. [Online]. Available: https://www.soterawireless.com/wp-content/uploads/2020/02/9-004020_B-ViSi-Mobile-Brochure.pdf

[64] W. B. Murray and P. A. Foster, “The peripheral pulse wave: Information overlooked,” *J. CIinical Monit.*, vol. 12, no. 5, Sep. 1995.

[65] X. F. Teng and Y. T. Zhang, “Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach,” in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, Cancun, Mexico, 2003, pp. 3153–3156. doi: 10.1109/IEMBS.2003.1280811.

[66] R. Kelly, C. Hayward, A. Avolio, and M. O’Rourke, “Noninvasive determination of age-related changes in the human arterial pulse.,” *Circulation*, vol. 80, no. 6, pp. 1652–1659, Dec. 1989, doi: 10.1161/01.CIR.80.6.1652.

[67] U. Rubins, A. Grabovskis, J. Grube, and I. Kukulis, “Photoplethysmography Analysis of Artery Properties in Patients with Cardiovascular Diseases,” in *14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics*, vol. 20, A. Katashev, Y. Dekhtyar, and J. Spigulis, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 319–322. doi: 10.1007/978-3-540-69367-3_85.

[68] S. C. Millasseau, R. P. Kelly, J. M. Ritter, and P. J. Chowienczyk, “Determination of age-related increases in large artery stiffness by digital pulse contour analysis,” *Clin. Sci.*, vol. 103, no. 4, pp. 371–377, Oct. 2002, doi: 10.1042/cs1030371.

[69] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, “Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time,” in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, Portugal, May 2015, pp. 1006–1009. doi: 10.1109/ISCAS.2015.7168806.

[70] M. S. Tanveer and M. K. Hasan, “Cuffless Blood Pressure Estimation from Electrocardiogram and Photoplethysmogram Using Waveform Based ANN-LSTM Network,” *ArXiv181102214 Eess*, Nov. 2018, Accessed: Jan. 06, 2021. [Online]. Available: http://arxiv.org/abs/1811.02214

[71] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, “Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 4, pp. 859–869, Apr. 2017, doi: 10.1109/TBME.2016.2580904.

[72] L. Wang, E. Pickwell-MacPherson, Y. P. Liang, and Y. T. Zhang, “Noninvasive cardiac output estimation using a novel photoplethysmogram index,” in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, Sep. 2009, pp. 1746–1749. doi: 10.1109/IEMBS.2009.5333091.

[73] A. A. Awad *et al.*, “The relationship between the photoplethysmographic waveform and systemic vascular resistance,” *J. Clin. Monit. Comput.*, vol. 21, no. 6, pp. 365–372, Nov. 2007, doi: 10.1007/s10877-007-9097-5.

[74] A. A. Awad, M. A. M. Ghobashy, R. G. Stout, D. G. Silverman, and K. H. Shelley, “How Does the Plethysmogram Derived from the Pulse Oximeter Relate to Arterial Blood Pressure in Coronary Artery Bypass Graft Patients?,” *Anesth. Analg.*, vol. 93, no. 6, pp. 1466–1471, Dec. 2001, doi: 10.1097/00000539-200112000-00022.

[75] X. Ding, B. P. Yan, Y.-T. Zhang, J. Liu, P. Su, and N. Zhao, “Feature Exploration for Knowledge-guided and Data-driven Approach Based Cuffless Blood Pressure Measurement,” p. 4, 2019.

[76] J. B. Dillon and A. B. Hertzman, “The form of the volume pulse in the finger pad in health, arteriosclerosis, and hypertension,” *Am. Heart J.*, 1941.

[77] Youngzoon Yoon, J. H. Cho, and Gilwon Yoon, “Non-constrained Blood Pressure Monitoring Using ECG and PPG for Personal Healthcare,” *J. Med. Syst.*, vol. 33, no. 4, pp. 261–266, Aug. 2009, doi: 10.1007/s10916-008-9186-0.

[78] Y. Sawada, G. Tanaka, and K. Yamakoshi, “Normalized pulse volume (NPV) derived photo-plethysmographically as a more valid measure of the finger vascular tone,” *Int. J. Psychophysiol.*, vol. 41, no. 1, 2001.

[79] J. Lee *et al.*, “Validation of normalized pulse volume in the outer ear as a simple measure of sympathetic activity using warm and cold pressor tests: towards applications in ambulatory monitoring,” *Physiol. Meas.*, vol. 34, no. 3, pp. 359–375, Mar. 2013, doi: 10.1088/0967-3334/34/3/359.

[80] K. Matsumura, P. Rolfe, S. Toda, and T. Yamakoshi, “Cuffless blood pressure estimation using only a smartphone,” *Sci. Rep.*, vol. 8, no. 1, p. 7298, Dec. 2018, doi: 10.1038/s41598-018-25681-5.

[81] K. Takazawa *et al.*, “Assessment of Vasoactive Agents and Vascular Aging by the Second Derivative of Photoplethysmogram Waveform,” *Hypertension*, vol. 32, no. 2, pp. 365–370, Aug. 1998, doi: 10.1161/01.HYP.32.2.365.

[82] S. S. Mousavi, M. Firouzmand, M. Charmi, M. Hemmati, M. Moghadam, and Y. Ghorbani, “Blood pressure estimation from appropriate and inappropriate PPG signals using A whole-based method,” *Biomed. Signal Process. Control*, vol. 47, pp. 196–206, Jan. 2019, doi: 10.1016/j.bspc.2018.08.022.

[83] R. Couceiro *et al.*, “Assessment of cardiovascular function from multi-Gaussian fitting of a finger photoplethysmogram,” *Physiol. Meas.*, vol. 36, no. 9, pp. 1801–1825, Sep. 2015, doi: 10.1088/0967-3334/36/9/1801.

[84] C. Liu, D. Zheng, A. Murray, and C. Liu, “Modeling carotid and radial artery pulse pressure waveforms by curve fitting with Gaussian functions,” *Biomed. Signal Process. Control*, vol. 8, no. 5, pp. 449–454, Sep. 2013, doi: 10.1016/j.bspc.2013.01.003.

[85] L. Wang, L. Xu, S. Feng, M. Q.-H. Meng, and K. Wang, “Multi-Gaussian fitting for pulse waveform using Weighted Least Squares and multi-criteria decision making method,” *Comput. Biol. Med.*, vol. 43, no. 11, pp. 1661–1672, Nov. 2013, doi: 10.1016/j.compbiomed.2013.08.004.

[86] S. Banerjee, R. Bailon, J. Lazaro, V. Marozas, P. Laguna, and E. Gil, “A Two Step Gaussian Modelling to Assess PPG Morphological Variability Induced by Psychological Stress,” Sep. 2017. doi: 10.22489/CinC.2017.270-035.

[87] J. Lázaro, E. Gil, M. Orini, P. Laguna, and R. Bailón, “Baroreflex Sensitivity Measured by Pulse Photoplethysmography,” *Front. Neurosci.*, vol. 13, p. 339, Apr. 2019, doi: 10.3389/fnins.2019.00339.

[88] S. Kontaxis *et al.*, “Photoplethysmographic Waveform Analysis for Autonomic Reactivity Assessment in Depression,” *IEEE Trans. Biomed. Eng.*, pp. 1–1, 2020, doi: 10.1109/TBME.2020.3025908.

[89] M. C. Baruch, D. E. Warburton, S. S. Bredin, A. Cote, D. W. Gerdt, and C. M. Adkins, “Pulse Decomposition Analysis of the digital arterial pulse during hemorrhage simulation,” *Nonlinear Biomed. Phys.*, vol. 5, no. 1, p. 1, Dec. 2011, doi: 10.1186/1753-4631-5-1.

[90] M. C. Baruch, K. Kalantari, D. W. Gerdt, and C. M. Adkins, “Validation of the pulse decomposition analysis algorithm using central arterial blood pressure,” *Biomed. Eng. OnLine*, vol. 13, no. 1, p. 96, 2014, doi: 10.1186/1475-925X-13-96.

[91] P. Su, X.-R. Ding, Y.-T. Zhang, J. Liu, F. Miao, and N. Zhao, “Long-term blood pressure prediction with deep recurrent neural networks,” in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Las Vegas, NV, USA, Mar. 2018, pp. 323–328. doi: 10.1109/BHI.2018.8333434.

[92] S. C. Millasseau *et al.*, “Noninvasive Assessment of the Digital Volume Pulse: Comparison With the Peripheral Pressure Pulse,” *Hypertension*, vol. 36, no. 6, pp. 952–956, Dec. 2000, doi: 10.1161/01.HYP.36.6.952.

[93] E. Zahedi, V. Sohani, M. Ali, K. Chellappan, and G. Beng, “Experimental Feasibility Study of Estimation of the Normalized Central Blood Pressure Waveform from Radial

Photoplethysmogram," *J. Healthc. Eng.*, vol. 6, no. 1, pp. 121–144, Mar. 2015, doi: 10.1260/2040-2295.6.1.121.

[94] L. Wang, E. Pickwell-MacPherson, and Y. T. Zhang, "Blood pressure contour analysis after exercise by the photoplethysmogram using a Transfer Function method," in *2008 5th International Summer School and Symposium on Medical Devices and Biosensors*, Hong Kong, China, 2008, pp. 82–85. doi: 10.1109/ISSMDBS.2008.4575022.

[95] E. O'brien, N. Atkins, F. Mee, and K. O'malley, "The British Hypertension Society protocol for the evaluation of blood pressure measuring devices," *Clin. Exp. Hypertens.*, vol. 15, no. 6, pp. 1087–1097, Jan. 1993, doi: 10.3109/10641969309037096.

[96] E. O'Brien *et al.*, "European Society of Hypertension International Protocol revision 2010 for the validation of blood pressure measuring devices in adults," *Blood Press. Monit.*, vol. 15, no. 1, pp. 23–38, Feb. 2010, doi: 10.1097/MBP.0b013e3283360e98.

[97] "IEEE Standard for Wearable Cuffless Blood Pressure Measuring Devices," IEEE, 2014. doi: 10.1109/IEEESTD.2014.6882122.

[98] "IEEE Standard for Wearable, Cuffless Blood Pressure Measuring Devices - Amendment 1," IEEE, 2019. doi: 10.1109/IEEESTD.2019.8859685.

[99] F. Miao *et al.*, "A Novel Continuous Blood Pressure Estimation Approach Based on Data Mining Techniques," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 6, pp. 1730–1740, Nov. 2017, doi: 10.1109/JBHI.2017.2691715.

[100] L. Wang, W. Zhou, Y. Xing, and X. Zhou, "A Novel Neural Network Model for Blood Pressure Estimation Using Photoplethesmography without Electrocardiogram," *J. Healthc. Eng.*, vol. 2018, pp. 1–9, 2018, doi: 10.1155/2018/7804243.

[101] Y. Wang, Z. Liu, and S. Ma, "Cuff-less blood pressure measurement from dual-channel photoplethysmographic signals via peripheral pulse transit time with singular spectrum analysis," *Physiol. Meas.*, vol. 39, no. 2, p. 025010, Feb. 2018, doi: 10.1088/1361-6579/aa996d.

[102] W.-H. Lin, H. Wang, O. W. Samuel, G. Liu, Z. Huang, and G. Li, "New photoplethysmogram indicators for improving cuffless and continuous blood pressure

estimation accuracy," *Physiol. Meas.*, vol. 39, no. 2, p. 025005, Feb. 2018, doi: 10.1088/1361-6579/aaa454.

[103] M. Simjanoska, M. Gjoreski, M. Gams, and A. Madevska Bogdanova, "Non-Invasive Blood Pressure Estimation from ECG Using Machine Learning Techniques," *Sensors*, vol. 18, no. 4, p. 1160, Apr. 2018, doi: 10.3390/s18041160.

[104] Y.-Z. Yoon *et al.*, "Cuff-Less Blood Pressure Estimation Using Pulse Waveform Analysis and Pulse Arrival Time," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 4, pp. 1068–1074, Jul. 2018, doi: 10.1109/JBHI.2017.2714674.

[105] W.-R. Yan, R.-C. Peng, Y.-T. Zhang, and D. Ho, "Cuffless Continuous Blood Pressure Estimation From Pulse Morphology of Photoplethysmograms," *IEEE Access*, vol. 7, pp. 141970–141977, 2019, doi: 10.1109/ACCESS.2019.2942936.

[106] X. Xing, Z. Ma, M. Zhang, Y. Zhou, W. Dong, and M. Song, "An Unobtrusive and Calibration-free Blood Pressure Estimation Method using Photoplethysmography and Biometrics," *Sci. Rep.*, vol. 9, no. 1, p. 8611, Dec. 2019, doi: 10.1038/s41598-019-45175-2.

[107] S. Baek, J. Jang, and S. Yoon, "End-to-End Blood Pressure Prediction via Fully Convolutional Networks," *IEEE Access*, vol. 7, pp. 185458–185468, 2019, doi: 10.1109/ACCESS.2019.2960844.

[108] N. Hasanzadeh, M. M. Ahmadi, and H. Mohammadzade, "Blood Pressure Estimation Using Photoplethysmogram Signal and Its Morphological Features," *IEEE Sens. J.*, vol. 20, no. 8, pp. 4300–4310, Apr. 2020, doi: 10.1109/JSEN.2019.2961411.

[109] S. G. Khalid, H. Liu, T. Zia, J. Zhang, F. Chen, and D. Zheng, "Cuffless Blood Pressure Estimation Using Single Channel Photoplethysmography: A Two-Step Method," *IEEE Access*, vol. 8, pp. 58146–58154, 2020, doi: 10.1109/ACCESS.2020.2981903.

[110] K. Song, K. Chung, and J.-H. Chang, "Cuffless Deep Learning-Based Blood Pressure Estimation for Smart Wristwatches," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4292–4302, Jul. 2020, doi: 10.1109/TIM.2019.2947103.

[111] C. Landry, S. D. Peterson, and A. Arami, “Nonlinear Dynamic Modeling of Blood Pressure Waveform: Towards an Accurate Cuffless Monitoring System,” *IEEE Sens. J.*, vol. 20, no. 10, pp. 5368–5378, May 2020, doi: 10.1109/JSEN.2020.2967759.

[112] A. Chakraborty, D. Goswami, J. Mukhopadhyay, and S. Chakrabarti, “Measurement of Arterial Blood Pressure Through Single-Site Acquisition of Photoplethysmograph Signal,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021, doi: 10.1109/TIM.2020.3011304.

[113] “BioBeat BB-613 WP 510k FDA Clearance.” FDA, Aug. 22, 2019.

[114] D. Nachman *et al.*, “Comparing blood pressure measurements between a photoplethysmography-based and a standard cuff-based manometry device,” *Sci. Rep.*, vol. 10, no. 1, p. 16116, Dec. 2020, doi: 10.1038/s41598-020-73172-3.

[115] “Biobeat BB-613WP (Wrist-monitor) Blood Pressure Monitor,” *Medaval*. https://medaval.ie/device/biobeat-bb-613wp-wrist-monitor/ (accessed Jan. 05, 2021).

[116] C. Pellaton *et al.*, “Accuracy testing of a new optical device for noninvasive estimation of systolic and diastolic blood pressure compared to intra-arterial measurements,” *Blood Press. Monit.*, vol. 25, no. 2, pp. 105–109, Apr. 2020, doi: 10.1097/MBP.0000000000000421.

[117] J. Sola *et al.*, “Are cuffless devices challenged enough? Design of a validation protocol for ambulatory blood pressure monitors at the wrist: the case of the Aktiia Bracelet,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, Jul. 2020, pp. 4437–4440. doi: 10.1109/EMBC44109.2020.9176286.

[118] J. Gehring, H. Gesche, G. Drewniok, G. Küchler, and A. Patzak, “Nocturnal blood pressure fluctuations measured by using pulse transit time in patients with severe obstructive sleep apnea syndrome,” *Sleep Breath.*, vol. 22, no. 2, pp. 337–343, May 2018, doi: 10.1007/s11325-017-1555-9.

[119] M. Elgendi *et al.*, “The use of photoplethysmography for assessing hypertension,” *Npj Digit. Med.*, vol. 2, no. 1, p. 60, Dec. 2019, doi: 10.1038/s41746-019-0136-7.

[120] M. Saeed *et al.*, “Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): A public-access intensive care unit database,” *Crit. Care Med.*, vol. 39, no. 5, pp. 952–960, May 2011, doi: 10.1097/CCM.0b013e31820a92c6.

[121] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet,” *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000, doi: 10.1161/01.CIR.101.23.e215.

[122] H.-C. Lee, Y. Park, S. B. Yoon, S. M. Yang, D. Park, and C.-W. Jung, “VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients,” *Sci. Data*, vol. 9, no. 1, p. 279, Dec. 2022, doi: 10.1038/s41597-022-01411-5.

[123] D. Liu, M. Görges, and S. A. Jenkins, “University of Queensland Vital Signs Dataset: Development of an Accessible Repository of Anesthesia Patient Monitoring Data for Research,” *Anesth. Analg.*, vol. 114, no. 3, pp. 584–589, Mar. 2012, doi: 10.1213/ANE.0b013e318241f7c0.

[124] C. Carlson, V.-R. Turpin, A. Suliman, C. Ade, S. Warren, and D. E. Thompson, “Bed-Based Ballistocardiography: Dataset and Ability to Track Cardiovascular Parameters,” *Sensors*, vol. 21, no. 1, p. 156, Dec. 2020, doi: 10.3390/s21010156.

[125] Y. Liang, Z. Chen, G. Liu, and M. Elgendi, “A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China,” *Sci. Data*, vol. 5, no. 1, p. 180020, Dec. 2018, doi: 10.1038/sdata.2018.20.

[126] J. Esmaelpoor, M. H. Moradi, and A. Kadkhodamohammadi, “Cuffless blood pressure estimation methods: physiological model parameters versus machine-learned features,” *Physiol. Meas.*, vol. 42, no. 3, p. 035006, Mar. 2021, doi: 10.1088/1361-6579/abeae8.

[127] P. C.-P. Chao, C.-C. Wu, D. H. Nguyen, B.-S. Nguyen, P.-C. Huang, and V.-H. Le, “The Machine Learnings Leading the Cuffless PPG Blood Pressure Sensors Into the Next Stage,” *IEEE Sens. J.*, vol. 21, no. 11, pp. 12498–12510, Jun. 2021, doi: 10.1109/JSEN.2021.3073850.

[128] M. Kachuee, “UCI Machine Learning Repository: Cuff-Less Blood Pressure Estimation Data Set,” Jul. 2015. https://archive.ics.uci.edu/ml/datasets/Cuff-Less+Blood+Pressure+Estimation# (accessed Apr. 19, 2021).

[129] P. Virtanen *et al.*, “SciPy 1.0: fundamental algorithms for scientific computing in Python,” *Nat. Methods*, vol. 17, no. 3, Art. no. 3, Mar. 2020, doi: 10.1038/s41592-019-0686-2.

[130] X. Xing *et al.*, “Robust blood pressure estimation from finger photoplethysmography using age-dependent linear models,” *Physiol. Meas.*, vol. 41, no. 2, p. 025007, Mar. 2020, doi: 10.1088/1361-6579/ab755d.

[131] J. S. Kim, K. K. Kim, H. J. Baek, and K. S. Park, “Effect of confounding factors on blood pressure estimation using pulse arrival time,” *Physiol. Meas.*, vol. 29, no. 5, pp. 615–624, May 2008, doi: 10.1088/0967-3334/29/5/007.

[132] Y. Li, Z. Wang, L. Zhang, X. Yang, and J. Song, “Characters available in photoplethysmogram for blood pressure estimation: beyond the pulse transit time,” *Australas. Phys. Eng. Sci. Med.*, vol. 37, no. 2, pp. 367–376, Jun. 2014, doi: 10.1007/s13246-014-0269-6.

[133] W.-H. Lin, X. Li, Y. Li, G. Li, and F. Chen, “Investigating the physiological mechanisms of the photoplethysmogram features for blood pressure estimation,” *Physiol. Meas.*, vol. 41, no. 4, p. 044003, May 2020, doi: 10.1088/1361-6579/ab7d78.

[134] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, 2011.

[135] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, “A Practical Guide to Support Vector Classification,” National Taiwan University, 2016. [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[136] G. Slapničar, N. Mlakar, and M. Luštrek, “Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network,” *Sensors*, vol. 19, no. 15, p. 3420, Aug. 2019, doi: 10.3390/s19153420.

[137] S. Maqsood, S. Xu, M. Springer, and R. Mohawesh, “A Benchmark Study of Machine Learning for Analysis of Signal Feature Extraction Techniques for Blood Pressure Estimation Using Photoplethysmography (PPG),” *IEEE Access*, vol. 9, pp. 138817–138833, 2021, doi: 10.1109/ACCESS.2021.3117969.