Circuits and Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
https://cas.tudelft.nl/

CAS-2023-00

M.Sc. Thesis

Machine learning algorithm to estimate cardiac output based on non-invasive arterial blood pressure measurements

Alan Hamo

Abstract

Cardiac output (CO), a vital hemodynamic parameter that reflects the blood volume pumped by the heart per minute, is crucial for determining tissue oxygen delivery and the heart's ability to meet the body's demands. Researchers have developed various methods to measure cardiac output, including thermodilution using pulmonary artery catheters (PAC), also called Swan-Ganz catheters, the gold standard for cardiac output measurements. Such an approach involves an invasive procedure associated with complications, and it requires specialized equipment and expertise, limiting its use to critically ill patients undergoing operations in intensive care units (ICUs). An alternative, less invasive way to estimate CO is by analyzing arterial blood pressure (ABP) waveforms. However, the relationship between cardiac output and blood pressure is unknown. This study uses machine learning and feature engineering techniques to discover the relationship between CO and ABP. We applied the sparse identification non-linear dynamics (SINDy) algorithm to discover features that significantly contribute to the relationship between CO and ABP. Additionally, we investigated the optimum number of cardiac cycles required for feature extraction to achieve the best performance providing insights into the temporal dynamics of CO estimation. The proposed approach achieved clinically acceptable performance regarding radial limits of agreement and bias. Further, the proposed approach was validated on an external dataset and achieved comparable performance. Finally, the learned model was interpreted as a differential equation describing the blood flow where CO acts as an external force to the system. All materials used in this study, including code, model, raw data, processed data, and extracted features, are available on GitHub to facilitate further development.

Machine learning algorithm to estimate cardiac output based on non-invasive arterial blood pressure measurements

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Alan Hamo
born in Aleppo, Syria

Dated: July 23, 2023

This work was performed in:

Circuits and Systems Group
Department of Microelectronics
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Delft University of Technology

Delft University of Technology
Department of
Microelectronics

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **"Machine learning algorithm to estimate cardiac output based on non-invasive arterial blood pressure measurements"** by **Alan Hamo** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: July 23, 2023

Chairman: _____
Dr.ir.J. Dauwels

Advisor: _____
Dr.ir.J. Dauwels

Committee Members: _____
Dr. Jie Yang

_____

Abstract

Cardiac output (CO), a vital hemodynamic parameter that reflects the blood volume pumped by the heart per minute, is crucial for determining tissue oxygen delivery and the heart's ability to meet the body's demands. Researchers have developed various methods to measure cardiac output, including thermodilution using pulmonary artery catheters (PAC), also called Swan-Ganz catheters, the gold standard for cardiac output measurements. Such an approach involves an invasive procedure associated with complications, and it requires specialized equipment and expertise, limiting its use to critically ill patients undergoing operations in intensive care units (ICUs). An alternative, less invasive way to estimate CO is by analyzing arterial blood pressure (ABP) waveforms. However, the relationship between cardiac output and blood pressure is unknown. This study uses machine learning and feature engineering techniques to discover the relationship between CO and ABP. We applied the sparse identification non-linear dynamics (SINDy) algorithm to discover features that significantly contribute to the relationship between CO and ABP. Additionally, we investigated the optimum number of cardiac cycles required for feature extraction to achieve the best performance providing insights into the temporal dynamics of CO estimation. The proposed approach achieved clinically acceptable performance regarding radial limits of agreement and bias. Further, the proposed approach was validated on an external dataset and achieved comparable performance. Finally, the learned model was interpreted as a differential equation describing the blood flow where CO acts as an external force to the system. All materials used in this study, including code, model, raw data, processed data, and extracted features, are available on GitHub to facilitate further development.

Acknowledgments

I want to express my heartfelt gratitude to the individuals who have supported and guided me throughout this challenging journey in completing my thesis.

First, I sincerely thank my supervisor, Dr. Justin Dauwels, your expertise and knowledge have been instrumental in shaping this thesis and expanding my understanding of the research field. Under your guidance, I experienced continuous learning throughout the 35 weeks of this project. It is impressive that you ensured I keep acquiring new knowledge and skills until my last work day. You would say, "This is my job," but I am genuinely grateful for your mentorship; During this journey, I learned a lot of new things from you that shaped me as an engineer.

I also appreciate the collaboration with the clinicians from Erasmus MC, Dr. Niki Ottenhof and Dr. Jan-Wiebe Korstanje. Their extensive knowledge and dedication to their field have significantly enriched the quality of this research.

Next to my academic mentors, I owe a special thanks to my mother, Falek Bejou; your encouragement and endless sacrifices have been a constant source of motivation. Thank you for instilling in me a passion for knowledge and perseverance.

I am also grateful to my wife, Amina Joulak; your understanding and patience have constantly motivated me. It was a challenging year for us; thanks for believing in me and being supportive during this journey.

I am equally grateful to my brother Jan, sister, Lilan, and colleagues from the Hogeschool Rotterdam, who supported and encouraged me throughout this thesis journey.

Lastly, I would like to dedicate this work to my father's soul, whom we lost in the war in Syria.

To all those mentioned and many others who have been part of my life's journey, thank you from the depths of my heart. Your presence, love, and encouragement have played a vital role in shaping this thesis and making me who I am today.

Alan Hamo
Delft, The Netherlands
July 23, 2023

Contents

List of Figures

List of Tables

Acronyms

ABP Arterial blood pressure. 1–3, 8

APCO arterial pressure cardiac output. 1, 2, 10, 11

CNN Convolutional Neural Network. 10

CO cardiac output. 1–8, 16, 19, 23, 25, 33

CR Concordance rate. 10, 30, 31, 35

DBP Diastolic blood pressure. 6, 7

EDA Electrodermal activity. 29

EDV end-diastolic volume. 6

ESV End-systolic volume. 6

HR heart rate. 1, 6–8

LOA Limits of agreement. 23, 31, 35

MAE Mean absolute error. 22, 31, 35, 41

MAP Mean arterial pressure. 7, 8

MIMIC Medical Information Mart for Intensive Care. v, 8, 11, 39–42

PAC pulmonary artery catheters. 1, 11

PE Percentage error. 23, 31, 35

PP Pulse pressure. 6

RBias Radial biase. 25, 31, 35, 41

RLOA Radial limits of agreement. 25, 31, 35, 41

RMS Root mean square error. 11, 22, 31, 35, 41

SAI Signal abnormality index. vii, 28, 29

SBP Systolic blood pressure. 6, 7

SD Standard deviation. 23, 25

SINDy Sparse identification of non-linear dynamics. 13, 18, 19, 29, 30, 38, 39

SV stroke volume. 1, 6

1 Introduction

Hemodynamic optimization is critical in reducing postoperative complications that belong to the leading causes of death worldwide [37]. The key to such optimization is early interventions, where research shows that early interventions for hemodynamic control decrease mortality rates by over 20% in high-risk patients [23]. Such an improvement can be achieved by maintaining cardiac output (CO) or stroke volume (SV) within specific ranges [11]. CO, a vital hemodynamic parameter, reflects the volume of blood pumped by the heart per minute from a single ventricle. On the other hand, stroke volume (SV) represents the volume of blood pumped by the heart per heartbeat. The relationship between CO and SV can be expressed mathematically:

$$\mathrm{CO[L/min]} = \mathrm{SV[L/beat]} \times \mathrm{HR[beats/min]}. \tag{1.1}$$

Here HR is the heart rate. Cardiac output is crucial for determining tissue oxygen delivery and indicates the heart's ability to meet the body's demands. Furthermore, it plays a fundamental role in understanding the causes of high blood pressure. Consequently, measuring CO has been a topic of interest to researchers and clinicians [47]. Researchers have developed various methods to measure cardiac output, including thermodilution using pulmonary artery catheters (PAC), also called Swan-Ganz catheters. This technique, illustrated in Figure 1.1, involves injecting a cold saline solution into the proximal port of the catheter and measuring the resulting temperature change using a thermistor [9, 36]. Therefore, such an approach involves an invasive procedure associated with complications, and it requires specialized equipment and expertise, limiting its use to critically ill patients undergoing operations in intensive care units (ICUs). Although thermodilution is considered the gold standard for CO measurements, it only provides an approximation since measuring the actual CO is extremely hard in clinical practices.

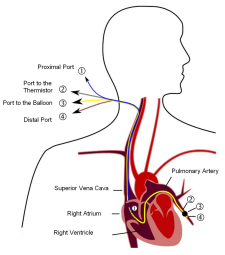


Figure 1.1: Invasive measurements Swan-Ganz catheter.[54]

Researchers have explored less invasive alternatives for estimating cardiac output throughout the past century, focusing on utilizing arterial blood pressure ABP waveforms. Various techniques for estimating CO from arterial blood pressure were developed. These techniques are called arterial pressure cardiac output (APCO) and utilize the continuous arterial pressure waveform obtained from an arterial line (shown in Figure 1.2) to estimate CO. Pressure-based techniques APCO can estimate CO and uncover hidden information

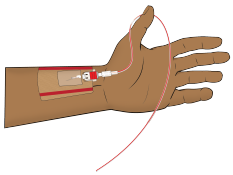


Figure 1.2: Alternative method for measuring ABP using arterial line.[53]

about CO by analyzing the arterial waveform during one cardiac cycle (beat-to-beat analysis) and using hemodynamic features (pressure information). Many models were developed over the years based on arterial blood pressure, including models based on solid physical principles called classical models throughout this thesis and models based on data-driven approaches [21, 46].

The classical models are simplified versions of the actual model and rely on the systematic circulation in blood flow where the heart models the generator (for example, a current source). Furthermore, capacitors and resistors model the arterial system. These models could provide a physical interpretation of the estimated value. However, such models must be more accurate and are associated with many assumptions that simplify the model. Some commercially available devices, like FloTrac/Vigileo, employ an APCO technique, offering a more straightforward and less invasive approach. However, these devices have limitations, as their algorithms are only partially released. Hence no improvement is possible [55].

Data-driven techniques such as machine learning and deep learning have been increasingly applied in medical research and clinical practice to improve patient outcomes and have demonstrated promising results. In the last decade, there has been growing interest in leveraging machine learning algorithms to estimate cardiac output from arterial pressure waveform, and this was shown in many studies to provide improvements. Such an approach estimates the CO by analyzing 10 to 20 seconds of ABP waveform. Such an analysis includes utilizing hemodynamic, waveform, and demographic features and expressing complex relationships between these features. However, no algorithm results in a clinically acceptable algorithm regarding accuracy and precision. Additionally, one drawback of the advanced data-driven techniques is their black-box nature, which makes them not readily accepted in clinical practice as they need more explanation of the results.

1.1 Problem statement

This study, conducted in collaboration with the Erasmus Medical Center in Rotterdam, the Netherlands, aims to design a system for the less invasive estimation of cardiac output. This estimation will be based on continuous arterial blood pressure measurements obtained through an arterial line, specifically for anesthesia patients. The proposed algorithm should have performance comparable to the gold standard method, namely the Swan-Ganz catheter. This study will utilize real-world ABP waveforms obtained from publically available datasets, and the Erasmus MC will provide their expertise in the field.

Figure 1.3 summarizes the main ideas behind this thesis. The gold standard is the Swan-Gans catheter, and the assignment is to develop a new algorithm that uses less invasive measurements to estimate cardiac output.

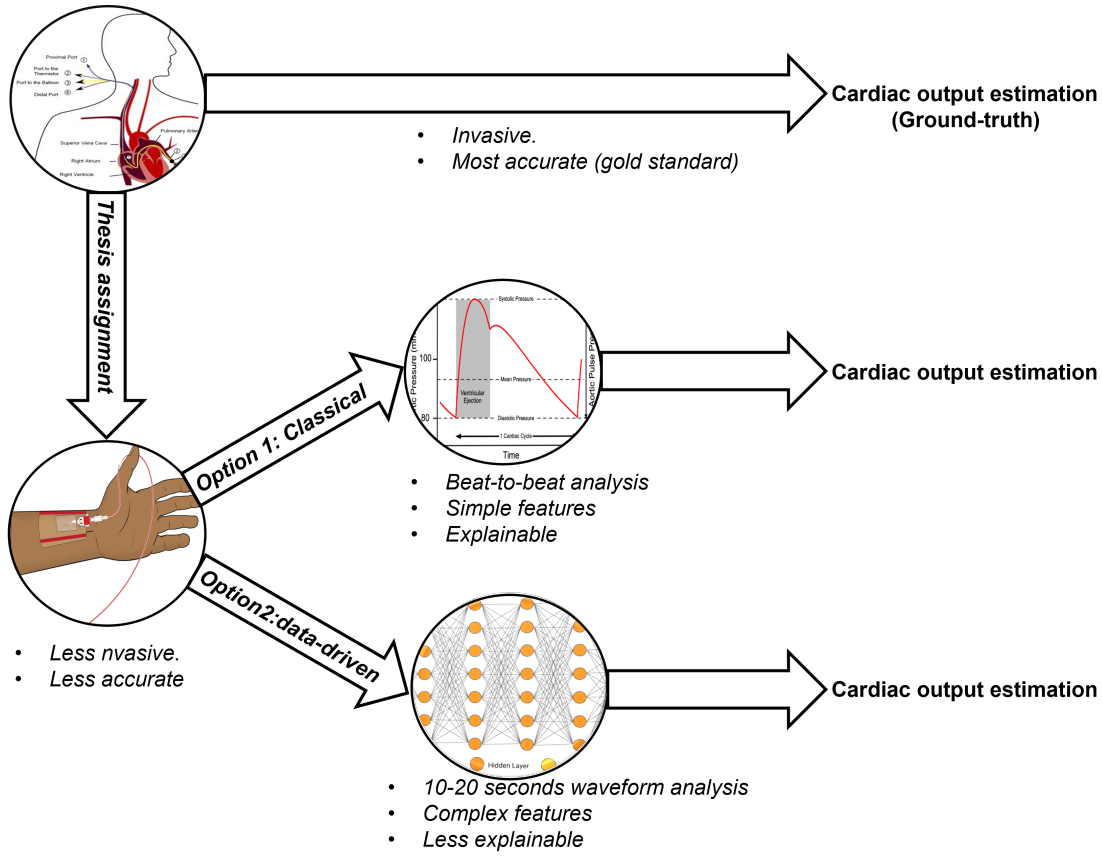


Figure 1.3: Summary of the idea behind the project

Developing accurate and precise predictive models for estimating CO from ABP can significantly enhance cardiovascular healthcare, improve patient outcomes, and empower clinicians to make well-informed decisions. One key advantage of this approach is its less invasive nature, enabling cardiac output estimation for a broader range of patients.

In the classical approach, the features (pressure information) are simplistic and must fully capture the complex model that links CO with ABP. Nonetheless, the underlying concept was to represent the dynamics of the arterial system as an electrical circuit using capacitors and resistors. It was believed that there is a covering differential equation that describes the system's behavior. However, during that period, the limited computational power of computing machines hindered the ability to test various settings, resulting in a reliance on assumptions.

In contrast, the data-driven approach incorporates more features and builds complex models to capture better the relationship between CO and ABP. Advanced learning techniques can capture non-linear relationships empowered by modern computing capabilities. However, despite achieving superior performance compared to classical es-

timators, data-driven approaches often need more interpretability inherent in classical models.

To address these challenges, this study proposes using machine learning techniques and feature engineering to develop an explainable algorithm for estimating cardiac output from arterial blood pressure measurements. We combine the core idea of the classical approach, which suggests the presence of a differential equation governing the system's behavior, with the power of advanced techniques to discover these equations. To this end, we aim to explore a novel set of features that have not been previously investigated, offering a fresh perspective on the problem. This approach enables us to uncover new insights and provide an understanding of the problem from a different perspective.

Furthermore, this study will examine the optimal duration of the algorithm's input. Classical approaches typically analyze data beat-to-beat, considering only one cardiac cycle. On the other hand, data-driven approaches often utilize input data spanning 10 to 20 seconds of the arterial waveform, equivalent to approximately 8 to 20 cardiac cycles. However, the rationale behind these specific input lengths remains to be determined, and their selection varies across different studies. To address this knowledge gap, we will investigate the ideal number of cardiac cycles for the input, aiming to reconcile the disparities between studies and provide valuable insights into determining the appropriate input duration for the algorithm.

By leveraging the capabilities of machine learning and feature engineering capabilities, this study will contribute to the progress of less invasive CO estimation by connecting the missing dots between classical and data-driven models. Specifically, this study aims to:

Design a machine learning algorithm to estimate cardiac output from continuous arterial blood pressure waveform features for patients undergoing anesthesia such that the learned model is as explainable as possible and achieves a clinically accepted level regarding accuracy and precision compared to the gold standard, the Swan-Ganz catheter.

In this study, the following research question will be addressed:

RQ 1: *Which arterial blood pressure features should be used as input for the algorithm?*

RQ 2: *What is the optimal number of cardiac cycles required for extracting features that yield the best performance?*

RQ 3: *How does the developed model compare to the traditional invasive method (Swan-Ganz catheter) regarding accuracy and precision?*

1.2 Thesis contribution

This thesis makes a novel contribution to less invasive cardiac output estimation based on arterial blood pressure waveform. It investigates new features that have not been utilized before, such as derivatives, time-related, and non-linear features. Considering these features, the learned model can be expressed as a differential equation, providing an explainable equation that describes the blood flow and its relation to cardiac output.

Furthermore, the study will investigate the optimal number of cardiac cycles required to achieve the best performance. Experimenting with different cardiac cycles will provide valuable insights into the input length and determine how far back in time we need to go to estimate CO accurately, bridging the gap between classical and data-driven approaches. Figure 1.4 visualizes the contribution of this study, highlighting its aim to connect the dots between classical models and data-driven approaches.

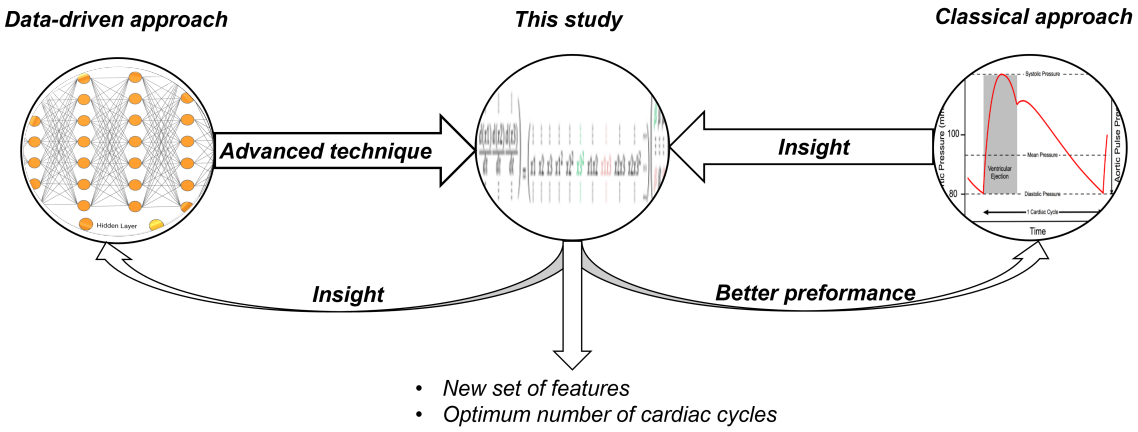


Figure 1.4: Visualization of the thesis contribution

1.3 Thesis outlines

This thesis is organized into several chapters. Chapter 2 will review the existing literature, followed by the methods in Chapter 3. The results will be presented and explained in Chapter 4. The obtained results will be discussed in chapter 5. Finally, the thesis will be concluded in Chapter 6, and future work will be elaborated.

2 Literature review

This chapter will review and summarize the literature on pressure-based cardiac output estimators. Part of the literature review was performed as a part of the extra project course ET4399. The literature review of the extra project can be found in the appendix B. This chapter first explains the cardiac cycle and its relation to blood pressure. A review of classical models will be followed this. Classical models can be categorized into lumped parameter models and pressure area models. Finally, the literature on data-driven approaches will be reviewed; such an approach utilizes machine learning and deep learning techniques.

2.1 Pressure cardiac cycle

Before delving into the review of both approaches, the classical and data-driven. Knowing what a cardiac cycle is and how it relates to blood pressure is essential. During diastole, the heart relaxes and fills with blood as venous returns bring blood back to the heart. The end-diastolic volume EDV represents the maximum amount of blood filling the heart at the end of relaxation. The diastole period follows the dicrotic notch in the arterial blood pressure waveform (Figure 2.1). The diastolic blood pressure DBP corresponds to the pressure at the end of the diastole period and serves as the starting pressure for the subsequent cardiac cycle. Subsequently, the heart contracts during systole, expelling blood from the heart into the arteries. This contraction is referred to as systole. At the end of systole, the heart has ejected the maximum amount of blood it can expel, leaving a residual volume known as the end-systolic volume ESV. The difference between the EDV and ESV represents the stroke volume SV, the volume of blood ejected per heartbeat, and is also related to the Frank-Starling mechanism [16], linking stroke volume to ESV. Cardiac output CO is calculated by multiplying the stroke volume SV by the heart rate HR, as shown in equation 1.1. In the arterial blood pressure waveform, the difference between diastolic blood pressure DBP and systolic blood pressure SBP is referred to as pulse pressure PP.

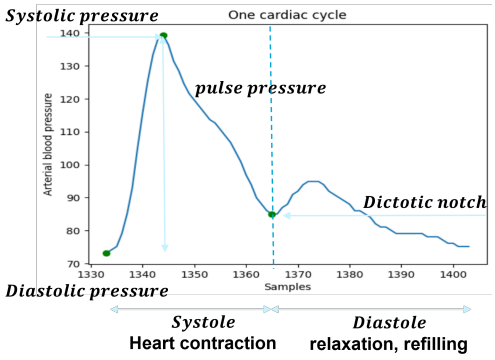


Figure 2.1: Example showing one cardiac cycle with most important features

2.2 Lumped parameter models

Lumped parameter models are simplified mathematical representations of the cardiovascular system, where the heart acts as the generator for blood flow, and the systematic circulation is modeled using lumped elements such as resistors and capacitors. Many models were proposed based on systematic circulations where the cardiac output is calculated beat-to-beat. Almost all classical estimators calculate the stroke volume first and then multiply it by the heart rate to get the cardiac output. All these estimators are proportional to the cardiac output and include a proportionality constant encapsulating non-linear terms.

One of the classical lumped parameter models is the Windkessel model [10]. This model assumes that the amount of blood entering a blood vessel equals the amount of blood leaving the vessel during a cardiac cycle. It also assumes a steady hemodynamic state of blood flow. The systematic circulation is modeled using resistors and capacitors, and the model suggests that cardiac output is proportional to the pulse pressure:

$$\mathrm{CO_{Wind}} = \mathrm{k} \cdot (\mathrm{SBP} - \mathrm{DBP}) \cdot \mathrm{HR}. \tag{2.1}$$

Here, SBP represents systolic blood pressure, DBP represents diastolic blood pressure, and k is the proportionality constant that encapsulates resistance and compliance. Another variant of the Windkessel model is the Windkessel with RC decay [3, 51], which calculates cardiac output differently by incorporating the decay constant of the capacitor:

$$\mathrm{CO_{WindRC}} = \mathrm{k} \cdot \frac{\mathrm{MAP}}{\mathrm{T}} \cdot \ln(\frac{\mathrm{SBP}}{\mathrm{DBP}}) \cdot \mathrm{HR}. \tag{2.2}$$

Another estimator called the Herd estimator [13] was developed for teaching purposes. Therefore, it was simple and less complicated. This model suggests that cardiac output is proportional to the difference between mean arterial blood pressure MAP and diastolic blood pressure DBP:

$$\mathrm{CO_{Herd}} = \mathrm{k} \cdot (\mathrm{MAP} - \mathrm{DBP}) \cdot \mathrm{HR}. \tag{2.3}$$

The MAP is calculated based on the systolic blood pressure SBP and diastolic blood pressure DBP using the following equation:

$$\mathrm{MAP} = \mathrm{DBP} + \frac{1}{3}(\mathrm{SBP} - \mathrm{DBP}). \tag{2.4}$$

In another study, MAP was considered in itself to be proportional to cardiac output[46].

$$\mathrm{CO_{MAP}} = \mathrm{k} \cdot \mathrm{MAP} \cdot \mathrm{HR}. \tag{2.5}$$

The last lumped parameter model is the Liljestrand-Zender[28]. This model takes into account the non-linear variation of arterial capacitance and models cardiac output as follows:

$$\mathrm{CO_{LZ}} = \mathrm{k} \cdot \frac{\mathrm{SBP} - \mathrm{DBP}}{\mathrm{SBP} + \mathrm{DBP}} \cdot \mathrm{HR}. \tag{2.6}$$

2.3 Pressure area models

Pressure area models are another approach that utilizes the area under the arterial pressure curve to estimate cardiac output. This technique involves integrating the arterial blood pressure waveform over time, appropriate calibration, or using proportional constants to convert the area into volumetric units. One pressure-area estimator is the systolic area [50] that states that the cardiac output is proportional to the area under the systole area in an arterial blood pressure waveform:

$$\mathrm{CO_{SA}} = \mathrm{k} \cdot \int_{\mathrm{T_{sys}}} \mathrm{ABP(t).dt} \cdot \mathrm{HR}. \tag{2.7}$$

Later, a correction factor was proposed for the previous estimator (systolic area), and the following estimator was proposed [24]:

$$\mathrm{CO_{SAC}} = \mathrm{k} \cdot \left(1 + \frac{\mathrm{T_{sys}}}{\mathrm{T_{dis}}}\right) \cdot \int_{\mathrm{T_{sys}}} \mathrm{ABP(t).dt} \cdot \mathrm{HR}. \tag{2.8}$$

Here $\mathrm{T_{dis}}$ and $\mathrm{T_{sys}}$ are the diastolic and systolic duration, respectively. Another correction to the pressure-area estimators was introduced in [52] where a correction factor to the impedance was added and hence the name "systolic area" with corrected impedance:

$$\mathrm{CO_{SACI}} = \mathrm{k} \cdot (163 + \mathrm{HR} - 0.48 \cdot \mathrm{MAP}) \cdot \int_{\mathrm{T_{sys}}} \mathrm{ABP(t).dt} \cdot \mathrm{HR}. \tag{2.9}$$

Another estimator was proposed in [19] where the belief was that cardiac output is proportional to the root-mean-square of each cardiac cycle, therefore, modeling cardiac output as the AC power of the cardiac cycle in an arterial blood pressure waveform:

$$\mathrm{CO_{rms}} = \mathrm{k} \cdot \sqrt{\int_{\mathrm{T}_{sys}} (\mathrm{ABP(t)} - \mathrm{MAP})^2 \mathrm{dt}} \cdot \mathrm{HR}. \tag{2.10}$$

All estimators mentioned above have several limitations, as was discussed in [48]. These models are a simplified version of the actual model, even the most complex model. In addition to that, these estimators are patient-dependent, meaning all methods requires calibration to obtain a CO estimation. Finally, these mods are derived based on solid physical principles and less physiological models. In [46], all these models were evaluated using the MIMIC dataset [40], as a result, the Liljestrand-Zender model performs the best, achieving an error of 0.8 L/min at one standard deviation. However, it should be noted that the proportionality constant was estimated using the ground truth cardiac output measurements. Table 2.1 summarizes the classical models.

Table 2.1: Summary of classical models used to estimate cardiac output on beat-to-beat basis

Model	**Equation**
MAP	$\mathrm{CO_{MAP}} = \mathrm{k} \cdot \mathrm{MAP} \cdot \mathrm{HR}$
Windkessel 1904 [10]	$\mathrm{CO_{Wind}} = \mathrm{k} \cdot (\mathrm{SBP} - \mathrm{DBP}) \cdot \mathrm{HR}$
Windkessel with RC 1974,1976 [3],[51]	$\mathrm{CO_{WindRC}} = \mathrm{k} \cdot \frac{\mathrm{MAP}}{\mathrm{T}} \cdot \ln(\frac{\mathrm{SBP}}{\mathrm{DBP}}) \cdot \mathrm{HR}$
Herd 1966 [13]	$\mathrm{CO_{Herd}} = \mathrm{k} \cdot (\mathrm{MAP} - \mathrm{DBP}) \cdot \mathrm{HR}$
Liljestrand-Zender 1928 [28]	$\mathrm{CO_{LZ}} = \mathrm{k} \cdot \frac{\mathrm{SBP}-\mathrm{DBP}}{\mathrm{SBP}+\mathrm{DBP}} \cdot \mathrm{HR}$
systolic area 1953 [50]	$\mathrm{CO_{SA}} = \mathrm{k} \cdot \int_{\mathrm{T_{sys}}} \mathrm{ABP(t).dt} \cdot \mathrm{HR}$
systolic area with correction 1970 [24]	$\mathrm{CO_{SAC}} = \mathrm{k} \cdot (1 + \frac{\mathrm{T_{sys}}}{\mathrm{T_{dis}}}) \cdot \int_{\mathrm{T_{sys}}} \mathrm{ABP(t).dt} \cdot \mathrm{HR}$
systolic area with corrected impedance 1983 [52]	$\mathrm{CO_{SACI}} = \mathrm{k} \cdot (163 + \mathrm{HR} - 0.48 \cdot \mathrm{MAP}) \cdot \int_{\mathrm{T_{sys}}} \mathrm{ABP(t).dt} \cdot \mathrm{HR}$
root-mean-square 2002 [19]	$\mathrm{CO_{rms}} = \mathrm{k} \cdot \sqrt{\int_{\mathrm{T_{sys}}} (\mathrm{ABP(t)} - \mathrm{MAP})^2 \mathrm{dt}} \cdot \mathrm{HR}$

2.4 Data-driven models

The reliability of classical estimators for estimating cardiac output from arterial blood pressure is limited due to their simplifying linear assumptions, which may only hold across a narrow range of hemodynamic conditions. Researchers have explored data-driven techniques as an alternative approach to overcome this limitation. One of the earliest attempts in this direction was proposed in 1994, where pattern recognition and image processing techniques were employed to derive cardiac output from arterial blood pressure waveforms [32]. The study aimed to utilize the Typical Shape Function (TSF) algorithm for waveform classification and recognition. Two sets of features were utilized: physiological and waveform. The physiological features included mean values, minimum and maximum values, initial values, final values, and maximum slopes of the arterial blood pressure waveform.

On the other hand, the waveform features encompass Fourier coefficients, statistical moments, and beat-spline values derived from the waveform. The study utilized animal models to evaluate the performance of their approach and compared the results with those obtained from three classical estimators. It used an ultrasonic probe as a reference, achieving 2.8 as the mean error and 9.8 standard deviations. The findings of this study demonstrated the feasibility and potential of using data-driven techniques for estimating cardiac output from arterial blood pressure, providing a proof-of-concept for this approach.

In a subsequent study conducted 17 years later, in 2011 [8], the performance of two neural network models, namely multi-layer perception (MLP) and radial basis function (RBF), was evaluated for estimating cardiac output. The performance of these neural network models was compared with three classical estimators: Liljestrand-Zender, Herd, and Wesseling. Diastolic, systolic, pulse, and pressure area during systole features were extracted to calculate the classical estimators. To assess the performance of the models, data from the Multi Parameter Intelligent Patient Monitoring for Intensive Care (MIMIC) database [5] were utilized. From 121 patients, 27 were selected, and

data were split randomly for training and testing. The findings of this study revealed a reduction in the mean absolute error between the cardiac output derived from the Swan-Ganz catheter and the cardiac output estimated by the neural network models compared to classical models.

Four years later, an application of data-driven techniques emerged to address the challenge of estimating cardiac output from arterial blood pressure [29] proposed using machine learning techniques to develop a rule for determining when the calibration of classical estimators is required. In their study, two classical estimators, specifically the mean arterial pressure and Windkessel, were combined through a fusion approach. The aim was to leverage the strengths of each estimator and improve overall performance. This approach was evaluated on a dataset comprising ten sheep records. The study's findings demonstrated that incorporating machine learning techniques enhanced performance compared to the standalone classical estimators. The fusion of the mean arterial pressure and Windkessel estimators, guided by the developed rule, yielded improved accuracy in estimating cardiac output from arterial blood pressure waveforms. The reference method in this study was the pulmonary arterial blood flow which is not the gold standard for cardiac output measurement. The agreement was -0.13%, and the concordance rate CR was 92% with a 12% exclusion zone.

In recent years, several studies have utilized deep learning techniques to address the problem of estimating cardiac output from arterial blood pressure. In 2019, a study was conducted where a Convolutional Neural Network CNN was employed to predict stroke volume [34]. The input to the model consisted of 10.24-second raw arterial blood pressure waveforms. The data used in this study was collected in their institution. In this study, patients were divided into training and testing groups. The CNN model demonstrated superior performance compared to commercially available devices, achieving a high concordance rate CR of 77.74% and a correlation coefficient R of 0.840.

Another study, published in 2021, utilized a 1-D Convolutional Neural Network CNN to estimate stroke volume [25]. The input data for this model consisted of three channels: the first channel represented the 10-second normalized raw arterial blood pressure waveform, the second channel contained frequency data obtained from the Fourier transform, and the third channel represented the slope of the arterial blood pressure waveform. The data used in this study was collected using the vital Recorder in their institution. Patients were divided into training and testing groups. The model demonstrated excellent performance, with a concordance rate of 96.26%, a Pearson correlation R of 0.95, a bias of -0.85, and limits of agreement of -2.88 and 0.71. However, it is essential to note that these results were compared to a commercially available arterial pressure cardiac output APCO device rather than the gold standard Swan-Ganz catheter for cardiac output measurement.

Another development in 2021 expanded the deep learning approach by incorporating additional steps in the system [55]. Firstly, they trained the system using arterial

pressure cardiac output APCO data, where the input consisted of 20-second raw arterial blood pressure waveforms to extract features. These extracted features were then fused with demographic information and used as input to a regressor for predicting stroke volume. After training the model with APCO data, transfer learning was applied using Pulmonary Artery Catheter PAC or (Swan-Ganz catheter) data, considered the gold standard for cardiac output measurement. Data used in this study were obtained from the publicly available dataset called Vitaldb [26]. The developed model exhibited superior performance compared to commercially available APCO devices. The model achieved 0.64 Spearman roh value and 53% concordance rate. It should be noted that this study estimated stroke volume, and hence the comparison is tough since stroke volume values should be converted to cardiac output before comparison.

In line with the present master's thesis, a recent study conducted in 2023 also employed machine learning techniques to predict cardiac output from arterial blood pressure [21]. The study evaluated 19 machine learning models, including linear regression, tree, ensemble, and Bayesian models. The data used in this study was obtained from the MIMIC database. Data was split into train, test, and validation sets. To validate the model, 10-fold cross-validation was used. However, this was performed in segmentation and not in patients. Therefore, it could be the case that a test was performed on seen data during training, especially that in this study, gender and age had the highest F-scores.

The study fused hemodynamic, waveform, and demographic features to enhance the predictive performance, collectively serving as input to the machine learning models. Notably, the performance of the proposed approach surpassed that of other models analyzing arterial pressure waveforms. The study achieved notable results, including a mean squared error RMS of 1.421, a bias of -0.01, limits of agreement of -2.35 and 2.32, and a percentage error of 39.44%. Specifically, the XGBoost model stood out by achieving a clinically acceptable level of radial limits of agreement with a value of +28.89. These results highlight the effectiveness of machine learning techniques in accurately predicting cardiac output from arterial blood pressure. However, it should be mentioned that in this study, it could be the case that the patient was used for training and testing since the splitting was based on the segmentation of samples and not patients.

Table 2.2: Summary of data-driven approaches used to estimate cardiac output on beat-to-beat basis

Paper	Technique	Input	Pos	Cons
[32] 1994	pattern recognition	physiological and waveform features	proved the concept of using data-driven techniques	Evaluated on Animal models and results were compared to ultrasonic probe
[8] 2011	Neural networks	Hemodynamic features	Rea-world data set and achieved better performance than classical models	Did not report the clinically accepted requirements and results were compared to classical models
[29] 2015	Machine learning	Classical model features	Provided an automated way when a method needs calibration	Evaluated on ten sheep records and compared to a different reference.
[34] 2019	Convolutional neural network	10.24s raw arterial blood pressure waveform	high concordance rate and correlation	lack of explainable AI, did not report all the clinically accepted requirements, and did not test on all patients
[25] 2021	1-D convolutional neural network	three channels: raw pressure, Fourier frequency, and slope information	High accuracy and concordance rate	Was not compared to the standard gold method and lack of explainable AI and did not report all the clinically accepted requirements
[55] 2021	deep learning+ transfer learning	20s raw data	High accuracy, found a smart way to deal with a small dataset and performed better than commercial devices	lack of explainable AI, did not report all the clinically accepted requirements, and did not test on all patients
[21] 2023	Machine learning	Hemodynamic, waveform, and demographic features	High accuracy and fused waveform, hemodynamic, and demographic features in a smart way	lack of explainable AI and did not split the data based on patients

2.5 Conclusion

This chapter reviewed the literature on cardiac output estimation based on arterial blood pressure waveform analysis. The aim was to explore the existing methods for cardiac output estimation, focusing on the features used and the input length such that the proposed research questions in section 1.1 can be answered. Table 2.1 and 2.2 summarize the existing methods.

RQ 1: *Which arterial blood pressure features should be used as input for the algorithm?*

In this literature review, two major categories were found, namely, the classical and

data-driven approaches. Classical estimators aimed to establish a relationship using features derived from arterial blood pressure. These features were based on well-established theories and provided physiological interpretation. In contrast, data-driven approaches leveraged the capabilities of machine learning and deep learning techniques but sacrificed the intuitive understanding of the features and their interpretability.

Classical estimators relied on the computational resources available at the time, which limited their ability to explore and incorporate additional sets of features. The emphasis was placed on developing models with a clear physiological basis, enabling researchers to interpret the relationship between the selected features and cardiac output. However, these constrained computational resources may have overlooked more complex relationships and interactions within the data.

This study aims to explore and identify a novel set of features that effectively capture the relationship between cardiac output and arterial blood pressure. To achieve our goal, we propose incorporating the temporal changes in the features within cardiac cycles and introducing augmented features that capture different dynamics between cardiac output and arterial blood pressure. The augmentation of the features will be based on the classical models. The literature review on classical methods found that logarithm, quadratic, non-linear (ratio) features, and derivative and time-related features were used. In this study, we propose augmenting features similar to the classical approach alongside derivatives and time-related features. We propose using sparse promoting techniques such as minimizing the $l1$norm to select the relevant features. The mentioned approach is called the Sparse Identification of Non-linear Dynamics or the SINDy algorithm [4]. SINDy has demonstrated its ability to discover sparse non-linear dynamics representations from measured data. By leveraging the sparsity-promoting nature of SINDy, we can identify the subset of features that significantly contribute to the relationship between cardiac output and arterial blood pressure. The discovery of these informative features will enhance our understanding of the underlying dynamics and enable a more accurate estimation of cardiac output based on arterial blood pressure measurements.

RQ 2: *What is the optimal number of cardiac cycles required for extracting features that yield the best performance?*

Cardiac output estimation traditionally relies on beat-to-beat estimators, while the data-driven approach involves analyzing waveforms over a specific time duration in seconds. This study also aims to investigate the optimal time window or the number of cardiac cycles required to accurately estimate cardiac output, explicitly addressing how far back in time we need to go to estimate cardiac output accurately.

By exploring the relationship between the length of the time window and the accuracy of cardiac output estimation, we aim to determine the appropriate duration for capturing relevant information from the arterial waveforms. This investigation will provide insights into the temporal dynamics of cardiac output estimation and guide the selection of an optimal time window for accurate and reliable estimation.

RQ 3: *How does the developed model compare to the traditional invasive method (Swan-Ganz catheter) regarding accuracy and precision?*

Measuring the actual cardiac output is extremely difficult in clinical practice, and the reference method can only provide an approximation. The precision level of the reference method is +-20% [39]. Therefore, measuring the correlation could be misleading, and the agreement between the two methods should be evaluated.

A new algorithm should be accurate and precise compared to the reference method. To assess these two concepts, two types of analysis with requirements were suggested in the literature [33, 39], namely, Bland-Altman analysis and Trending ability analysis. Reporting these two types of analysis will help clinicians assess the algorithm's acceptability in clinical practices.

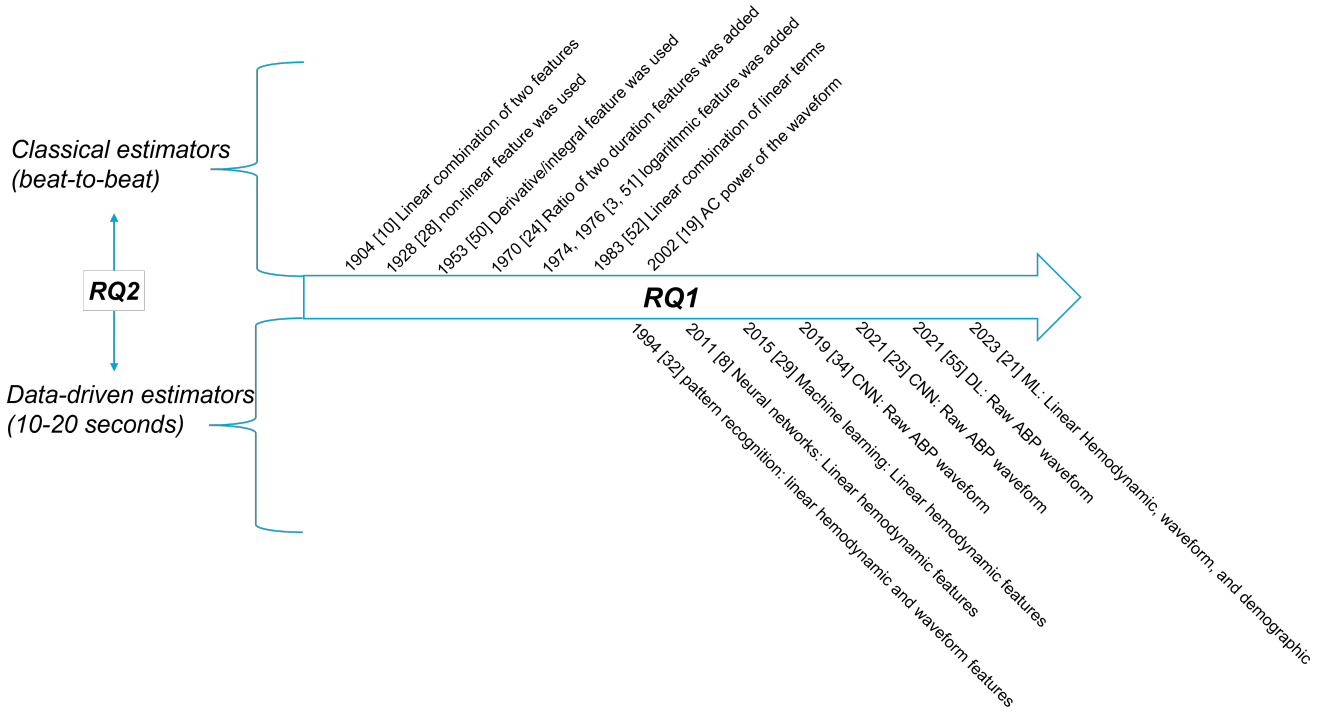


Figure 2.2: Summary of the literature review showing the difference in used features between classical and data-driven approaches.

3 Methods

In this chapter, the methodologies used throughout this project will be discussed. The methodology employed in this research encompasses several vital steps. Initially, the arterial blood pressure signal processing will be discussed. This includes downsampling the signal and filtering the estimated values. Following this, a signal abnormality index is presented. This is used to assess the quality of the signal, ensuring that only reliable data is used in subsequent analyses. After that, features extraction and two experiments will be explained. The first experiment will involve feature discovery, and the second will focus on determining the optimum number of cardiac cycles required for accurate prediction. Finally, the requirements for a clinically accepted algorithm will be addressed. Figure 3.1 shows the methodologies flow used in this research.

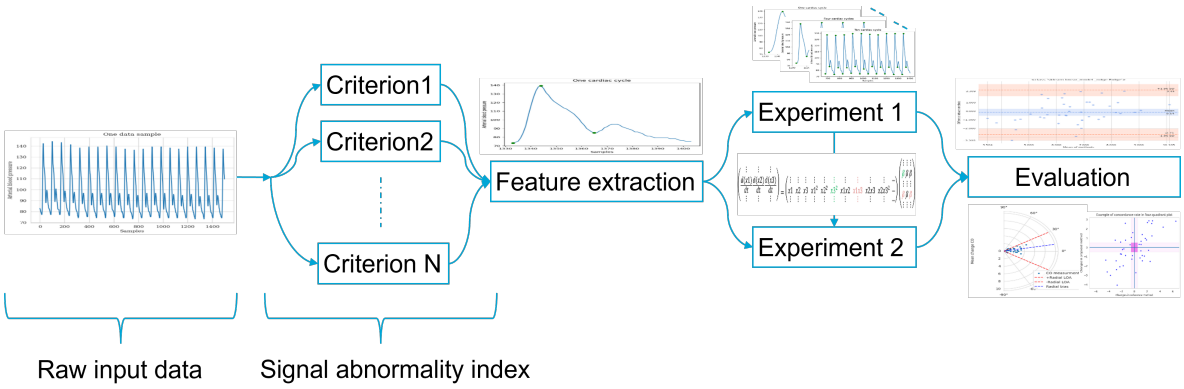


Figure 3.1: Thesis methodology pipeline

3.1 Signal processing

This section focuses on the processing steps in preparing the arterial waveform to serve as input for the model and post-processing the estimated values. The first step is downsampling the arterial blood pressure waveform from 500 Hz to 100 Hz to reduce the number of data points in the waveform. After that, the signal abnormality index is used to assess the quality of the arterial blood pressure waveform. Such criteria are based on physiologic, noise/artifact, and beat-to-beat variation values and were proposed in [45] to remove waveforms that may result in undesirable results. The criteria are summarized in the table 3.1.

Another processing step is performed at the end of the methods pipeline. After training the model, a low-pass filter will be applied to the estimated values to reduce

the fluctuation. The CO does not change rapidly, mainly during anesthesia when the patients are at rest. Therefore, a low pass filter will mitigate any fluctuation that does not correspond to physiological events. This filter will be applied per patient so that patient trackability is unaffected.

Table 3.1: Summary of signal abnormalities criteria. Any waveform that satisfies these criteria was deleted

No.	Feature	Abnormality criteria
1	Systolic pressure	$P_s > 300$mmHg
2	Diastolic pressure	$P_d < 20$mmHg
3	Mean pressure	$P_m < 30$mmHg or $P_m > 200$mmHg
4	Pulse pressure	$P_p < 20$mmHg
5	Heart rate	HR < 20 or HR > 200 bpm
6	Mean of negative slopes	$\omega <$ -40mmHg/100ms
7	$P_s[k]$-$P_s[k-1]$	$\lvert\Delta P_s\rvert > 20$ mmHg
8	$P_d[k]$-$P_d[k-1]$	$\lvert\Delta P_d\rvert > 20$ mmHg
9	T[k]-T[k − 1] (duration of cardiac cycle)	$\lvert\Delta T\rvert > \frac{2}{3}$ sec

In this study, the following criteria were used next to the previous ones summarized in table 3.2:

Table 3.2: Summary of signal abnormalities criteria added in this study. Any waveform that satisfies these criteria was deleted

No.	Feature	Abnormality criteria
11	Difference between max and min in the waveform	max(ABP)-min(ABP)< 30mmHg
12	Abrupt change	diff(ABP) > 30

3.2 Feature engineering

In this section, the process of feature extraction is described. Following [21, 32], we extract hemodynamic and waveform features alongside demographic information. The feature extraction will be performed on an arterial blood pressure waveform long enough to capture at least eight cardiac cycles. After extracting the features, a feature matrix will be constructed for further analysis.

3.2.1 Hemodynamic features extraction

Hemodynamic features provide valuable information about blood pressure characteristics during cardiac cycles as they capture essential indicators of cardiac output, such as systolic, diastolic, and diacritic peaks. Extracting these three peaks will allow extracting other features related to the duration of each phase, estimation of the heart rate (using the duration of the cardiac cycle), and calculation of classical estimators. The extracted hemodynamic features are described in table 3.3, and the most important hemodynamic features can be seen in figure 3.2.

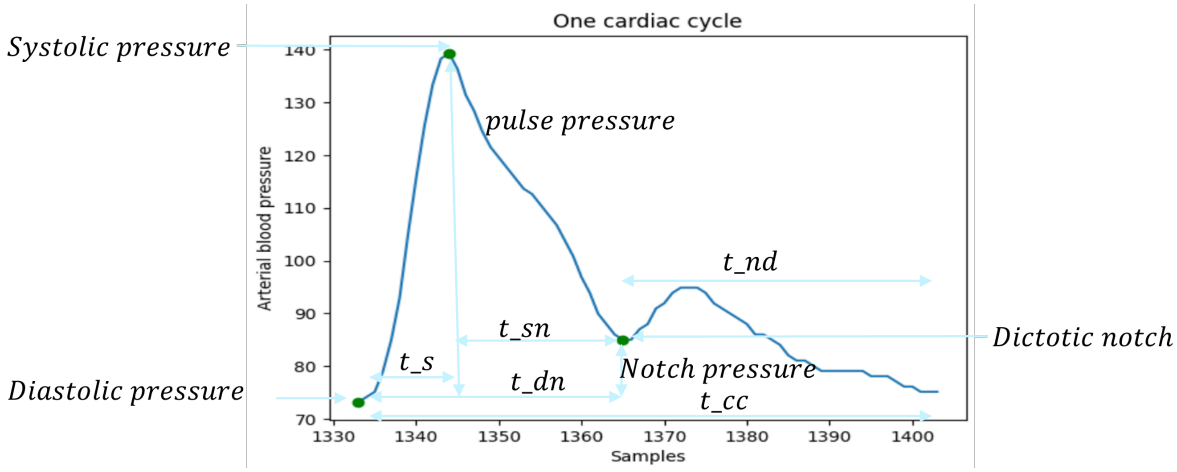


Figure 3.2: Example showing one cardiac cycle with most important hemodynamic features

Table 3.3: Summary of extracted hemodynamic features

Feature	**Description**
SBP	Systolic Blood Pressure. The maximum pressure reached during systole
DBP	Diastolic Blood Pressure. The pressure at the end of diastole
MAP	Mean arterial blood pressure
PP	Pulse pressure (SBP − DBP)
T_{SB}	Duration of the systole period
T_{SN}	Duration from systolic pressure to dicrotic notch point
T_{DP}	Duration of diastole period
T_C	Duration of a cardiac cycle
HR	Heart rate estimated per cardiac cycle
Classical estimators	This includes Windkessel, Windkessel with RC, Herd, Liljestrand-Zender, systolic area, a systolic area with correction, systolic area with corrected impedance, and root mean square models.

3.2.2 Waveform features extraction

The second type of feature extracted in this study is waveform features. Waveform features capture the shape and characteristics of the arterial blood pressure waveform. They provide additional information about the dynamics and patterns within the cardiac cycles. The extracted waveform features are summarized in table 3.4.

Table 3.4: Summary of extracted waveform features

Feature	**Description**
abs_{max}	Highest absolute value
agg_{acc}	Mean value of aggregation function over the auto-correlation of max lag 40
$\text{trend}_{\text{mean}}$	Mean of trend function
$\text{trend}_{\text{std}}$	Standard deviation of trend function
$\text{seasonal}_{\text{mean}}$	Mean of seasonal function
$\text{seasonal}_{\text{std}}$	Standard deviation of seasonal function
abs-sum-changes	Absolute value of consecutive changes $\sum_{i=1,..,n} \lvert x_{i+1} - x_i \rvert$
abs-energy	Absolute energy: $E = \sum_{i=1,..,n} x_i^2$
sum-value	The sum over the arterial pressure waveform
media	The median of the arterial pressure waveform
mean	The mean of the arterial pressure waveform
autocorrelation	The value of the first nine lags of the autocorrelation function
Partial-autocorrelation	The value of the first nine lags of the partial autocorrelation function
Fourier-Entropy	Binned entropy of the power spectral density

3.3 Experiment1: Feature discovery

This experiment will be conducted to augment the features and select the relevant ones based on sparse techniques. Answering the following research question:

RQ 1: *Which arterial blood pressure features should be used as input for the algorithm?*

After extracting hemodynamics and waveform features, the SINDy algorithm [4] will be applied to the feature matrix to augment and select relevant features. We apply the leave-one-patient-out technique to avoid any biases in the feature selection. For each iteration, one patient will be excluded, the SINDy algorithm will be used, and the designer will select features manually. A pseudo-algorithm is provided in 1. This algorithm was repeated twice, once for hemodynamics and once for waveforms features.

The SINDy algorithm

The SINDy algorithm is a powerful technique to discover sparse representations of dynamical systems based on measured data. Its primary objective is to identify the governing equations that describe the dynamics of a system by employing sparsity-promoting techniques.

The algorithm operates by constructing a library of potential candidate functions, which can be selected based on prior knowledge of the system or systematically generated. This library contains a set of functions relevant to the system's dynamics under investigation.

Once the library is established, the SINDy algorithm applies sparse regression techniques to identify the active terms within the library matrix that characterize the system's dynamics. In other words, it aims to determine the most significant functions of the library that contribute to the system's dynamics. To achieve this, the SINDy algorithm solves the following optimization problem:

$$\arg\min_{\beta} \quad \left\| \frac{\mathrm{dx}}{\mathrm{dt}} - \phi(\mathrm{x})\beta \right\| + \gamma|\beta|. \tag{3.1}$$

In this equation, β represents the coefficient vector obtained from solving the constrained regression problem. The term $\frac{\mathrm{dx}}{\mathrm{dt}}$ denotes the derivative of the features with respect to time, capturing the system's dynamics. The matrix $\phi(\mathrm{x})$ represents the library of possible functions applied to the feature set. Finally, γ is a hyper-parameter that needs to be selected by the designer and influences the sparsity of the solution.

By solving the optimization problem, the SINDy algorithm identifies the most relevant terms and their corresponding coefficients in the library matrix, effectively capturing the underlying dynamics of the system. The resulting sparse representation provides valuable insights into the relationship between variables and enables the discovery of the governing equations that describe the system's behavior.

This experiment uses the SINDy algorithm to discover new features that capture the cardiac output and arterial blood pressure dynamics. We extend the feature set by including the derivative of the features, representing how these features change over cardiac cycles.

We adopt a recursive approach to address the curse of dimensionality problem and facilitate the analysis. Each iteration focuses on three features, where cardiac output is always one of them, and the goal is to investigate how cardiac output is related to the other two features. Thus, we construct the following matrix equation for polynomial features as an example:

$$\begin{pmatrix} \frac{\Delta \mathrm{CO}}{\Delta \mathrm{t}} & \frac{\Delta \mathrm{Feature1}}{\Delta \mathrm{t}} & \frac{\Delta \mathrm{Feature2}}{\Delta \mathrm{t}} \\ \frac{\Delta \mathrm{CO}}{\Delta \mathrm{t}} & \frac{\Delta \mathrm{Feature1}}{\Delta \mathrm{t}} & \frac{\Delta \mathrm{Feature2}}{\Delta \mathrm{t}} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \mathrm{CO} & \mathrm{Feature1} & \mathrm{CO}\cdot\mathrm{Feature1} & \mathrm{Feature1}^2 & \dots \\ \mathrm{CO} & \mathrm{Feature1} & \mathrm{CO}\cdot\mathrm{Feature1} & \mathrm{Feature1}^2 & \dots \\ \mathrm{CO} & \mathrm{Feature1} & \mathrm{CO}\cdot\mathrm{Feature1} & \mathrm{Feature1}^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots \\ \beta & \beta & \beta \\ \vdots & \vdots & \vdots \end{pmatrix}. \tag{3.2}$$

In each iteration, the values of Feature1 and Feature2 are varied while CO remains constant. This approach lets us learn the relationship between cardiac output and the proposed features. Let us consider an example of one iteration:

$$\begin{aligned} \frac{\Delta \mathrm{CO}}{\Delta \mathrm{t}} &= \mathrm{CO} + \mathrm{Feature1}^2 + \mathrm{Feature1}\cdot\mathrm{Feature2}. \\ \frac{\Delta \mathrm{Feature1}}{\Delta \mathrm{t}} &= \mathrm{CO} + \mathrm{Feature2}^2. \end{aligned} \tag{3.3}$$

From the first equation in 3.3, we observe that cardiac output is related quadratically to Feature1, and a non-linear term exists between cardiac output and Feature1 $\cdot$ Feature2. From the second equation, we discover cardiac output is related to the derivative of Feature1 and the square of Feature2. Finally, these identified features are added to the original feature set and serve as input to the model.

Algorithm 1 Feature discovery

$\phi(\mathrm{x}) = [\phi(\mathrm{x})_{\mathrm{Logarithmic}}, \phi(\mathrm{x})_{\mathrm{exponential}}, \phi(\mathrm{x})_{\mathrm{polynominal}}]$
for #patients > patient iterator **do**
  for #feature index > feature iterator **do**
    for #feature index − feature iterator > feature iterator2 **do**
      $\mathrm{x} = [\mathrm{CO}, \mathrm{f}_{\text{feature iterator}}, \mathrm{f}_{\text{feature iterator2}}]$
      1) $\arg\min_{\beta} \quad \|\frac{\mathrm{dx}}{\mathrm{dt}} - \phi(\mathrm{x})\beta\| + \gamma|\beta|$
      2) store the names of non-zero $\phi(\mathrm{x})$
      3) feature iterator2 = feature iterator2 + 1
    end for
    feature iterator = feature iterator + 1
  end for
  patient iterator = patient iterator + 1
end for
Select overlapped features in $\phi(\mathrm{x})$

3.4 Experiment 2: The number of cardiac cycles

This experiment aims to investigate the optimal time window or the number of cardiac cycles required to accurately estimate cardiac output, explicitly addressing how far back in time we need to go to estimate cardiac output accurately. Answering the posed research question:

RQ 2: *What is the optimal number of cardiac cycles required for extracting features that yield the best performance?*

For each sample, we systematically progressively analyzed cardiac cycles, starting from the most recent and moving back in time. To estimate cardiac output, we considered different numbers of cardiac cycles, ranging from one to eight. This approach allowed us to investigate how including additional cardiac cycles affects the accuracy and reliability of cardiac output estimation.

After data collection and segmentation (a segment is equivalent to a sample where a sample is 15 seconds of arterial blood pressure waveform with its corresponding cardiac output value). After that, features were extracted for eight cardiac cycles for each

sample, and the feature matrix was formed:

$$x_{features} = \begin{pmatrix} Feature1_1 & \dots & Feature1_{10} & \dots & Feature2_1 & \dots & FeatureN_{10} \\ Feature1_1 & \dots & Feature1_{10} & \dots & Feature2_1 & \dots & FeatureN_{10} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{3.4}$$

Only features with subscript $_1$ of all features are used to investigate the performance of one cardiac cycle. To examine the performance of two cardiac cycles, the average was taken. As an example, in the two cardiac cycles test, the first feature becomes $\frac{\text{Feature1}_1+\text{Feature1}_2}{2}$, and for three cardiac cycles: $\frac{\text{Feature1}_1+\text{Feature1}_2+\text{Feature1}_3}{3}$, Etc. Only for the eight cardiac cycle tests the whole matrix is used. Figure 3.3 shows an example of different numbers of cardiac cycles, namely, 10, 7, 4, and 1.

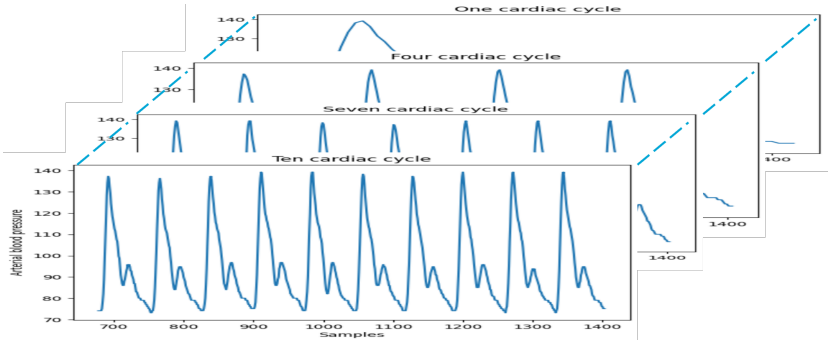


Figure 3.3: An image showing a different number of cardiac cycles waveforms 10, 7, 4, and 1 cardiac cycles

The estimation performance was assessed using well-established metrics commonly employed for evaluating clinically accepted algorithms, including agreement and trending-ability analysis. The performance of the estimation will be compared with the gold standard for measuring cardiac output (the Swan-Ganz catheter).

Statistical analysis and visualization techniques will be employed to identify trends as the number of cardiac cycles increased. The objective is to determine the optimal number of cardiac cycles at which the incremental improvement in estimation accuracy and precision becomes negligible or reaches a clear maximum value.

3.5 Models

This section will discuss the machine learning models employed in this study. The aim is to explore and compare the performance of various algorithms, including linear regression models such as linear, lasso regression, and ridge regressions. Tree-based regression models such as decision trees, XGBoost, and random forests will also be tested.

Linear models

Linear regression models are simple and easy to understand since they calculate the weighted sum of the features. Linear models aim to establish a linear relationship between the dependent variable (cardiac output) and the independent variables (discovered features). This study will use linear regression models to develop a linear equation that predicts cardiac output based on the input features. The linear regression equation can be represented as follows:

$$\mathrm{CO} = \beta_0 + \beta_1 * \mathrm{X}_1 + \beta_2 * \mathrm{X}_2 + ... + \beta_\mathrm{f} * \mathrm{X}_\mathrm{f}. \tag{3.5}$$

Here CO is the estimated cardiac output, X_1, X_2, X_3,...,X_f represent the features used where f is the total number of features used, and β_1, β_2, β_3, ... β_f are the coefficients to be learned during model training. On the other hand, lasso regression incorporates a penalty term in the linear regression model by adding an $l1$ regularization term. The lasso regression equation can be represented as:

$$\mathrm{CO} = \beta_0 + \beta_1 * \mathrm{X}_1 + \beta_2 * \mathrm{X}_2 + ... + \beta_\mathrm{f} * \mathrm{X}_\mathrm{f} + \lambda * \sum |\beta|. \tag{3.6}$$

Here, λ is the regularization parameter that controls the sparsity of the model. The last linear model that will be used is the ridge regression. Ridge is a similar technique to the lasso regression but utilizes a different regularization term, namely, the $l2$ norm. Ridge regression aims to mitigate multicollinearity/ non-orthogonality between the input variable. The ridge regression equation is given by:

$$\mathrm{CO} = \beta_0 + \beta_1 * \mathrm{X}_1 + \beta_2 * \mathrm{X}_2 + ... + \beta_\mathrm{f} * \mathrm{X}_\mathrm{f} + \lambda * \sum |\beta|_2. \tag{3.7}$$

Tree-based models

On the other hand, tree-based regression models are more advanced techniques that capture more complex relationships between the features and still provide interpretability through visualizing the tree. Tree-based models, such as decision trees, are individual models that use a hierarchical structure of nodes and branches to make predictions. Each node in the tree represents a decision based on a specific feature, and each leaf node represents the predicted outcome. Tree-based models can capture complex relationships and non-linear patterns in the data.

The decision tree recursively partitions the feature space to find the parameter that minimizes the cost function:

$$\theta^* = \arg\min_{\theta} G(\Theta_\mathrm{m}, \theta). \tag{3.8}$$

Here, Θ_m is the data at node m and θ= $(\mathrm{j}, \mathrm{t}_\mathrm{m})$ is the candidate split consisting of feature j and threshold t_m. For the regression problem, the cost function is either mean square error RMS or mean absolute error MAE. On the other hand, random forest and XGBoost are ensemble learning techniques that combine multiple decision trees to make predictions

3.6 Evaluation metrics

In this section, the evaluation metrics of a clinically accepted algorithm will be discussed. The results of the trained model must be compared with a reference method, such as the Swan-Ganz catheter method, to evaluate a new algorithm that estimates cardiac output. Such a comparison should investigate the accuracy and precision of the proposed algorithm. The accuracy of the new algorithm describes how close the estimated value is to the actual/reference value, and the precision describes the spread of repeated values due to random error. Evaluating the accuracy and precision of the model will answer the posed research question:

> **RQ 3:** *How does the developed model compare to the traditional invasive method (Swan-Ganz catheter) regarding accuracy and precision?*

In the following subsections, these analyses will be explained, and elaboration on what parameter to be measured and how will be given.

Bland-Altman analysis

The recognized standard statistical method of assessing the agreement between two serial measurements of the same clinical variable is the Bland-Altman plot [2]. The Bland-Altman analysis determines the bias as a measure of accuracy and the limits of agreement as a measure of precision. The Bland-Altman plot is presented as a scatter plot where the x-axis is the average of the two methods, and the y-axis is the difference between the methods. Figure 3.4 shows an example of such a plot. The limits of the agreement are calculated as follows:

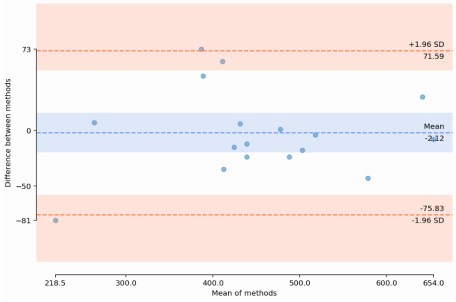


Figure 3.4: Example of Bland-Altman plot.[17]

$$\text{LOA} = (\text{bias}) \pm \text{t}_{\text{a,n}-1} \cdot (\text{SD}). \tag{3.9}$$

Here, SD is the standard deviation of the differences, and $\text{t}_{\text{a,n}-1}$ is the t-value corresponding to $\text{n} - 1$ degrees of freedom and a type I error a of 0.05. A good algorithm should have a bias as close as possible to zero and as tight as possible limits of agreements. An acceptable limit of the agreement should be defined beforehand to justify the acceptance of the new algorithm. However, this is hard to determine. Therefore, it is suggested that the percentage error of the Bland-Altman analysis should be reported [39]. The percentage error is calculated as follows:

$$\text{PE}(\%) = 100\% \cdot \text{t}_{\alpha,\text{n}-1} \cdot \frac{\text{SD}}{\text{meanCO}}. \tag{3.10}$$

A percentage error PE of ±30 has been suggested as a guide to determine if the proposed method is a good alternative to the reference method.

Trending ability analysis

The Bland-Altman analysis shows the agreement between two measurements for the same variable. However, such an analysis does not provide information about the trend in the data since a trend analysis should involve the change in the cardiac output ΔCO. To assess the trending ability, it is suggested to use the four-quadrant analysis to calculate the concordance rate and polar plot analysis to calculate the radial bias and radial limits of agreement [7, 14]. In the following, an explanation of these two analyses will be provided, and how they are calculated.

Four-quadrant analysis

The Four-quadrant analysis is an analysis to assess the trending ability of an algorithm compared to a reference method. In such an analysis, the change in the tested algorithm ΔCO_{tested} is plotted against the change in the reference method $\Delta CO_{reference}$. Such a plot shows whether the data are randomly distributed or follow the line of identity y = x. In figure 3.5, an example of such a plot can be seen. When both methods change in the same direction, either positive or negative, this is considered agreement. This is considered disagreement when one method changes positively and the other changes negatively or the opposite. Data points around the origin describe small changes in CO and do not reflect any trending ability.

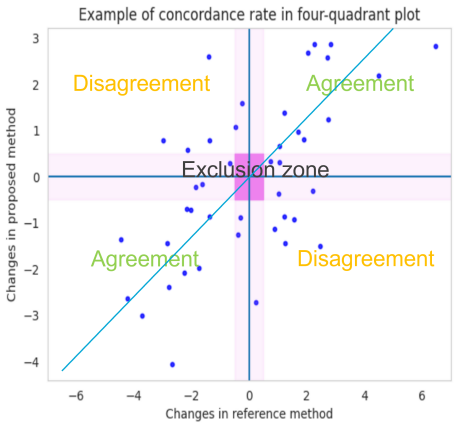


Figure 3.5: Example of Four-Quadrant plot

Such changes are randomly distributed, unpredictable, and could be originated from noise or measurement errors [6]. Therefore, an exclusion zone is defined around the origin to exclude these small changes. From such analysis, the concordance rate can be calculated. The conventional concordance rate is defined as follows:

$$CR = \frac{\#\text{agreement} - \#\text{agreement}_{\text{exclusion}}}{\#\text{points} - \#\text{exclusion}}. \tag{3.11}$$

Concordance rate between $90 \sim 95$ are considered acceptable clinically

Polar plot analysis

One drawback of the Four-quadrant analysis is that it ignores information regarding the size of the change in CO. In addition, only one property of agreement is tested: whether the change is positive or negative. As a solution for these drawbacks, the polar plot analysis is suggested. The idea is to treat each point from the Four-quadrant analysis as a vector originating from the origin with an angel and magnitude. This angle and radius (length) are represented as a polar plot. The points are rotated by a 45°, making the line of identity at angel 0°(which is easier for visual inspection).

Data with negative change are rotated 180° since the change direction is unimportant. The data in the exclusion zone are also excluded from the polar plot analysis since they have large random errors. Finally, the radial limits of agreement and radial bias are calculated in the same way as in the Bland-Altman analysis:

$$\text{RLOA} = \text{mean}(\theta) \pm 1.96\text{SD}(\theta). \tag{3.12}$$

Where RLOA is the radial limits of agreement, SD is the standard deviation, and θ is the angle of each point in the polar plot calculated as follows:

$$\text{angle} = \arctan(\frac{\Delta\text{CO}_{\text{predicted}}}{\Delta\text{CO}_{\text{tested}}}). \tag{3.13}$$

$$\text{Bias} = \frac{\Delta\text{CO}_{\text{tested}} + \Delta\text{CO}_{\text{predicted}}}{2}. \tag{3.14}$$

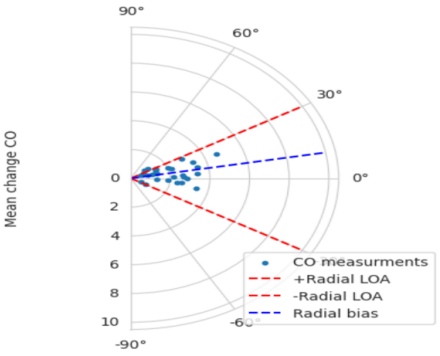


Figure 3.6: Example of polar plot analysis used to assess the trending ability of the proposed algorithm

Figure 3.6 shows an example of the polar plot analysis where the two red lines are at angle 30° and the blue line is the bias of all points. **A RLOA of $\leq 30°$ and RBias of ≤ 5 is accepted clinically**

Table 3.5 summarises the requirements of a clinically acceptable algorithm.

Table 3.5: Summary of the clinically accepted algorithm requirements

Percentage error	**Concordance rate**	**Radial bias**	**Radial limits of agreements**
30%	90∼95	5°	30°

Additionally, in this study, other metrics are also reported. These additional metrics are summarised in table 3.6

Table 3.6: Summary of the additional metrics used in this study

Percentage error	**Concordance rate**
Pearson correlation R	$r = \frac{\sum(y-m_y)(\hat{y}-m_{\hat{y}})}{\sqrt{\sum(y-m_y)^2 \sum(\hat{y}-m_{\hat{y}})^2}}$
coefficient of determination R^2	$R^2 = 1 - \frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y}_i)^2}$
coefficient of determination R^2	$R^2 = 1 - \frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}$
Root mean square error RMSE	$RMSE = \sqrt{\frac{\sum(\hat{y}_i-y_i)^2}{N}}$
Mean absolute error MAE	$MAE = \frac{1}{n_{samples}} \sum \lvert y_i - \hat{y}_i \rvert$

In table 3.6, y is the tested variable, $\hat{y}$ is the estimated value, N is the number of the samples, and $\bar{y}$ is the mean of the measured variable.

4 Results

In this chapter, the implementation results will be presented. The prime objective of this chapter is to evaluate the performance and effectiveness of the implemented solution. This chapter starts with an explanation of the dataset used for model development. This will be followed by the results of pre-processing, feature extraction, feature discovery, the number of cardiac cycles, and model evaluation, where results will be shown, and an explanation will be provided. The evaluation of the models will be based on clinically accepted metrics and leave-one-patient-out cross-validation. Finally, the learned model and the proposed approach will be validated on an external dataset to assess the generalizability of the proposed approach.

4.1 Dataset

The data utilized in this thesis were obtained from the Vital Data Bank, a freely accessible public database comprising biosignal waveforms and vital signs. The Vital Recorder program captured time-synchronized, high-resolution data from various anesthesia devices [26, 27]. Patients were selected from the Vital Recorder data based on the following measurements:

- Arterial pressure wave: The arterial pressure waveforms were collected using the Tram-Rac 4A (SNUADC) device, a patient monitor device manufactured by GE Healthcare. The waveform was sampled at a rate of 500 Hz.
- Cardiac output and stroke volume: The Vigilance II cardiac output monitor from Edwards Lifescience collected data on cardiac output and stroke volume. The data were recorded at intervals of 2 seconds. This monitor utilizes the Swan-Ganz catheter and pulmonary artery thermodilution technique, widely recognized as the gold standard for measuring cardiac output and stroke volume.

We used the open-source Python package provided by the dataset called vitaldb 1.4.1 to extract arterial blood pressure waveform, corresponding cardiac output, and demographic information. Forty-seven cases were identified with cardiac output and arterial blood pressure measurements.

This study defines a sample as a 15-second waveform extracted from the arterial blood pressure starting from the time t_0 when a cardiac output was received till (t_{0-15}). From each patient, 350 samples were extracted. Eighteen thousand nine hundred samples were collected from all patients before any pre-processing or signal quality check. After pre-processing and signal quality check, the number becomes 15460 samples with 43 cases.

The dataset comprised 27 females and 16 males with the following demographic characteristics: a height of 161.59 ± 7.54 cm, a weight of 60.45 ± 12.6 kg, BMI (Body Mass Index) of 20.07 ± 4.04, and an age of 54.58 ± 14.75 years. Figure 4.1 displays a box plot visualizing the distribution of the demographic information. A wide range of patient demographic information is available in the dataset, varying between men and women.

All patient IDs and surgical operations used in this study are in the appendix C.1. The reference cardiac output measurements have the following characteristics: mean of 6.45 $\frac{\text{L}}{\text{min}}$ and standard deviation of 2.01 $\frac{\text{L}}{\text{min}}$. Figure 4.2 show the distribution of the reference cardiac output measurements covering values ranging from 1.6 $\frac{\text{L}}{\text{min}}$ to 11.8 $\frac{\text{L}}{\text{min}}$.

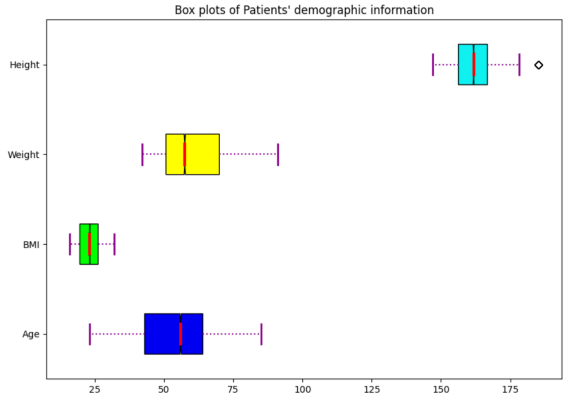


Figure 4.1: Patient's demographic information box plots height, weight, BMI (body mass index), and Age

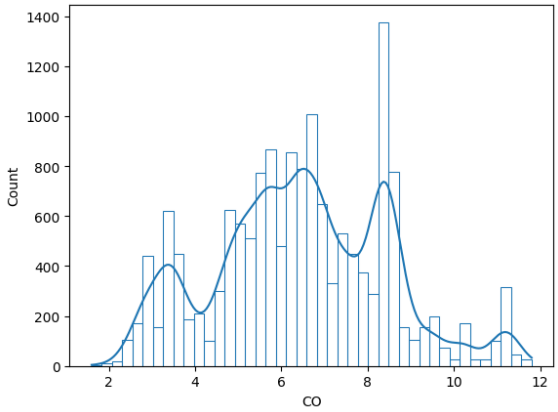


Figure 4.2: Reference cardiac output distribution

4.2 Pre-processing and feature extraction

As mentioned in section 3.1, the first step for pre-processing the waveform is to down-sample it from 500 Hz to 100 Hz. To this end, we utilized the Python package vitaldb 1.4.1, which includes a function to load a specific case and allows for downsampling the waveform.

After down-sampling the waveforms, we conducted a signal abnormality check to assess the waveform's quality. The criteria outlined in Tables 3.1 and 3.2 were implemented in Python, and any waveforms that satisfied the criteria were removed. Figure 4.3 and 4.4 show signal distortion where the signal shape changes. Patient movements likely cause these distortions. Additionally, Figures 4.5 and 4.6 display a different type of distortion where the signal increases instead of reaching a maximum during systole. This pattern persists for approximately 6 to 10 seconds before returning to the standard shape. It occurs when doctors draw blood from the patient, leaving the arterial line

sensor open. Figure 4.5 also demonstrates that more than one criterion could be met in one sample, showing two vertical blue lines with a difference between less than 30 mmHg and the waveform exceeding 200 mmHg abnormality criterion. Lastly, Figure 4.7 exhibits a misleading waveform where the signal appears suitable for feature extraction. However, physiologically, pulse pressure below 30 mmHg should be disregarded.

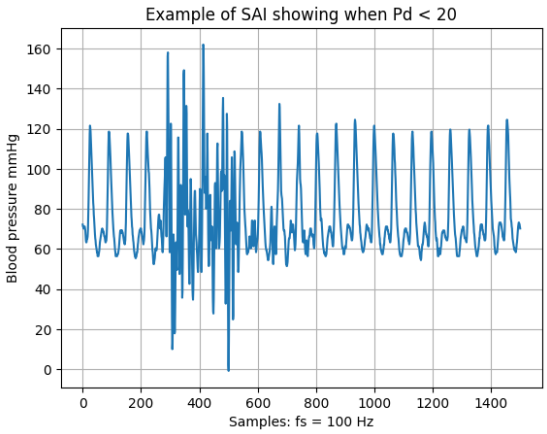


Figure 4.3: Results of SAI criteria 2 from table 3.1. Diastolic blood pressure is less than 20 mmHg, mainly caused by patient movements

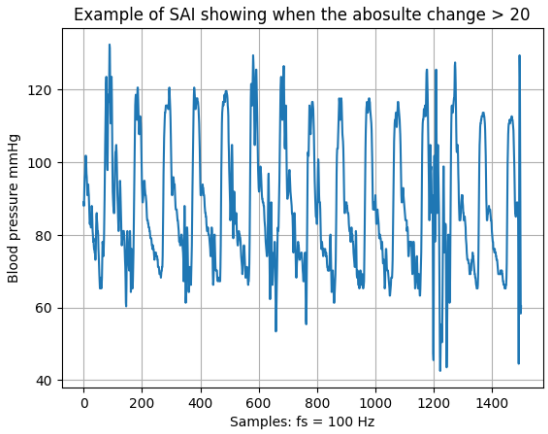


Figure 4.4: Results of SAI criteria 12 from table 3.2. Absolute change in the waveform is bigger than 20, mainly caused by patient movements

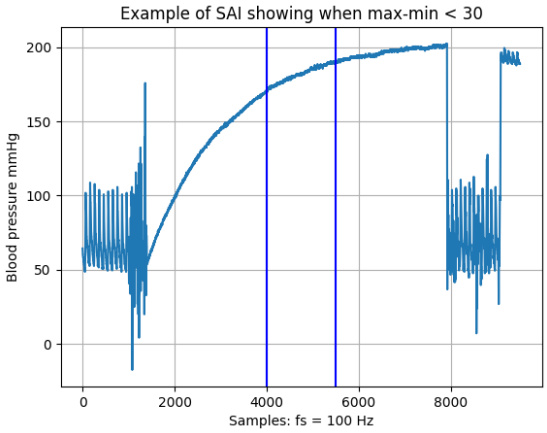


Figure 4.5: Results of SAI criteria 11 from table 3.2 criteria is met between the two vertical blue lines. The difference between min and max is less than 30 in 2 seconds. mainly cause when the sensor is open

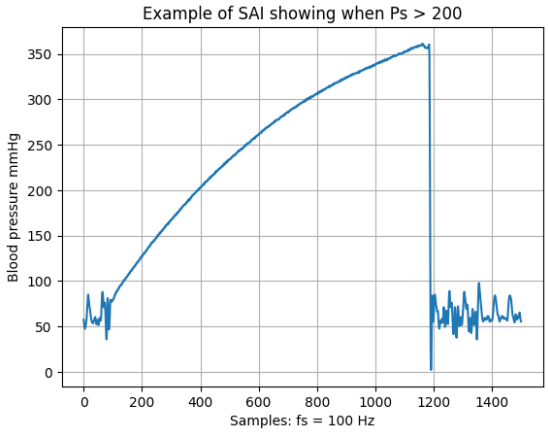


Figure 4.6: Results of SAI criteria 1 from table 3.1. Systolic blood pressure is bigger than 200 mmHg. mainly cause when the sensor is open

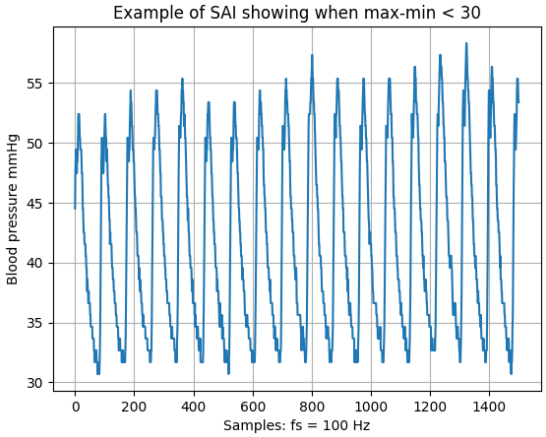


Figure 4.7: Results of SAI criteria 11 from table 3.2. The difference between systolic and diastolic is less than 30 $mmHg$.

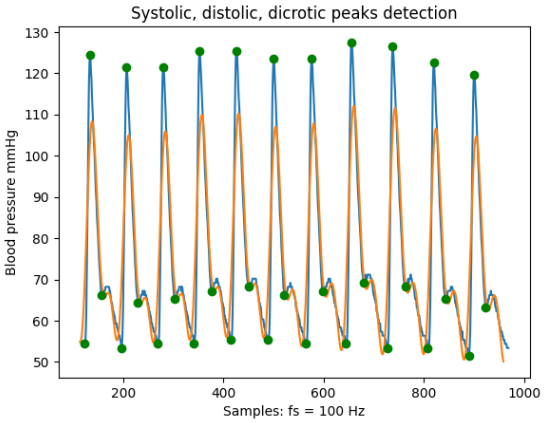


Figure 4.8: Hemodynamic feature extraction. The blue waveform is the original waveform. The Orange one is the processed one to extract features. Green dots are the features. The blue waveform is used for further processing

After downsampling, and signal abnormality checks, we extracted hemodynamics and waveform features. For this purpose, we utilized the Neurokit Python package [31], which includes a function to process electrodermal activity EDA signals. Although this library is not explicitly designed for arterial blood pressure waveforms, the two waveforms share a similar shape, and the library provides valuable information about the signal, such as rise time, recovery time, and critical peaks. Figure 4.8 exemplifies an arterial pressure waveform with the detection of essential hemodynamic peaks, including systolic, diastolic, and dicrotic peaks. These peaks are indicated as green points. The orange waveform represents the signal after processing by the EDA algorithm, while the blue waveform represents the original arterial blood pressure waveform after filtering. The algorithm introduces attenuation and phase shifts in the signal and hence can not be used for further processing. Therefore, the peaks are mapped to the original waveform (blue) by utilizing a window search of 20 samples around the detected peaks by the algorithm (orange). This allows us to use the original signal without attenuation or phase shift. Extracting these peaks allowed us to calculate all other features mentioned in table 3.3. We extracted waveform features listed in Table 3.4 using the open-source package tsfresh, designed for processing time series signals. All the features were utilized to construct a feature matrix, which will be employed for further development.

4.3 Feature discovery

As discussed in section 3.3, the SINDy algorithm will be utilized to discover features related to the dynamics of arterial blood pressure and cardiac output. We manually constructed the feature matrices and followed a recursive approach to discover these

features as described in algorithm 1. The Python implementation of the SINDy algorithm [20, 43] was employed for this purpose. We tested polynomial, exponential, and logarithmic features. Table 4.1 summarizes the 33 discovered features, categorized into six types: hemodynamic, derivatives (representing changes between cardiac cycles), logarithmic, non-linear, waveform, and demographic features. The implementation code of this procedure can be found on GitHub

Upon closer examination of these features, it can be inferred that they describe five physical phenomena: velocity, pressure, power, energy, and randomness. Features related to duration describe the velocity of blood flow. For example, a longer cardiac cycle duration indicates slower blood movement. These features are indicated with (**u**) in table 4.1. Additionally, the features that capture pressure-related characteristics of the blood flow are denoted with (**p**). Features that describe the power and energy of the blood flow are denoted by (**w**) and (**e**), respectively. Lastly, features that describe the randomness are denoted by (**r**).

Table 4.1: Discovered features using SINDy algorithm

Hemodynamic	Derivatives	Logarithmic	Non-linear	Waveform	Demographic
Heart rate (**u**)	Heart rate (**u**)	Heart rate (**u**)	1/systolic pressure (**p**)	Absolute energy (**e**)	Gender
Systolic duration (**u**)	Systolic duration (**u**)	Cardiac cycles duration (**u**)	1/diastolic pressure (**p**)	Fourier entropy 5 (**r**)	
Systolic notch duration (**u**)	Systolic notch duration (**u**)	Pulse pressure (**p**)	1/mean pressure (**p**)	Fourier entropy 10 (**r**)	
Diastolic duration (**u**)	Diastolic duration (**u**)	Absolute energy (**e**)	1/systolic duration (**u**)	Fourier entropy 100 (**r**)	
Cardiac cycles duration (**u**)	Cardiac cycles duration (**u**)		systolic*pulse (pressure) (**p**)		
systolic area model	Systolic pressure (**p**)		systolic/diastolic (pressure) (**p**)		
Liljestrand-Zander model (**p**)			systolic/pulse (pressure) (**p**)		
RMS pressure (**w**)			systolic/diastolic (duration) (**u**)		
Area with correction (**w**)			systolic/cardiac cycle (duration) (**u**)		

4.4 The number of cardiac cycles

Before conducting the experiment explained in section 3.4 to determine the optimum number of cardiac cycles. All regression models were tested using the discovered features mentioned in table 4.1 and one cardiac cycle. The aim was to select the best-performing model to save training time. The evaluation results can be seen in table 4.2. This table shows that ridge regression is the best-performing model regarding all evaluation metrics except for the concordance rate CR. The best-performing model for the concordance rate is the lasso regression. This indicates that the lasso regression is the most precise model. However, it is not accurate as the ridge. Also, linear regression has comparable performance to ridge regression. However, ridge regression is slightly better than linear regression. Therefore, ridge regression will be used for further development as it is the most accurate and precise model.

Table 4.2: Regressors comparison

	RMS	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
Linear	1.21	**1.01**	**0.76**	**0.57**	-0.05	2.43,-2.51	38.21	50.9	**32.96**	**-2.19**
Lasso	1.30	1.13	0.67	0.44	0.02	2.85,-2.81	43.53	**53.38**	43.93	6.17
Ridge	**1.20**	**1.01**	**0.76**	**0.57**	**-0.04**	**2.43,-2.50**	**37.96**	50.6	32.94	-2.18
DecisionTree	1.97	1.78	0.36	-0.21	0.02	4.22,-4.17	64.55	49.63	26.85	-6.29
RandomForrest	1.45	1.27	0.61	0.36	-0.06	2.98,-3.10	46.79	50.63	38.35	0.46
XGBoost	1.44	1.25	0.62	0.38	-0.12	2.87,-3.11	45.98	50.36	36.63	2.49

After selecting ridge regression, the experiment described in section 3.4 was conducted using the features discovered through the SINDy algorithm mentioned in table 4.1 and ridge regression. The implementation results can be seen in figures 4.9, 4.10 4.11, 4.12, 4.13, 4.14, 4.15, 4.16,4.17, and 4.18. It can be observed that three cardiac cycles yield the best results regarding limits of agreement of 2.2 figure 4.11, mean absolute error of 0.96 figure 4.12, percentage error of 35.4 figure 4.13, and radial limits of agreements of 28.5 figure 4.17 by achieving a minimum. Three cardiac cycles were also the best number of cardiac cycles regarding root mean square error of 1.12 figure 4.18, correlation R of 0.8 figure 4.14, and R2 of 0.63 figure 4.15 by achieving a maximum.

Regarding radial bias, the best performing cardiac cycles is at five cardiac cycles by achieving a minimum, as can be seen in figure 4.16. However, it should be noted that the radial bias is within the clinically acceptable level for all cardiac cycle numbers. Regarding the concordance rate, a maximum is achieved at six cardiac cycles, as shown in figure 4.10, indicating that six cardiac cycles provide the most precise estimation. On the other hand, three cardiac cycles provide the most accurate solution, as observed from the other metrics. Therefore, choosing the best number of cardiac cycles is a trade-off between accuracy (three cardiac cycles)and precision (six cardiac cycles).

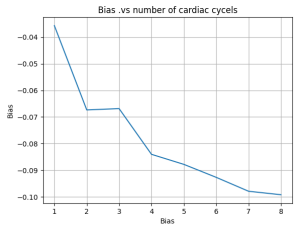


Figure 4.9: #cardiac cycles vs. bias

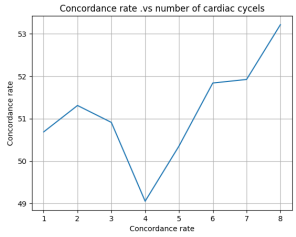


Figure 4.10: #cardiac cycles vs. concordance rate

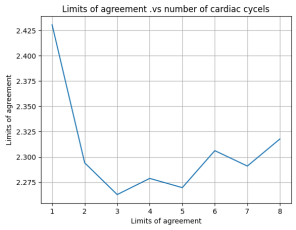


Figure 4.11: #cardiac cycles vs. limits of agreement

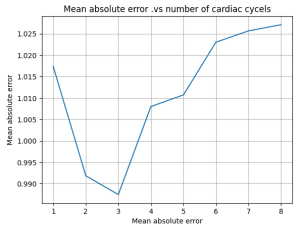


Figure 4.12: #cardiac cycles vs. mean absolute error

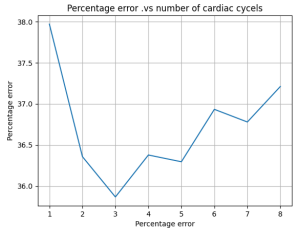


Figure 4.13: #cardiac cycles vs. percentage error

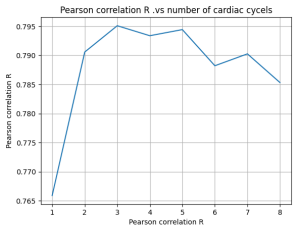


Figure 4.14: #cardiac cycles vs. correlation

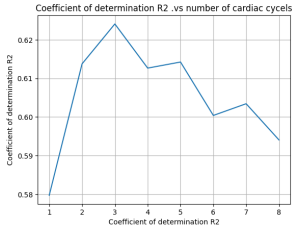


Figure 4.15: #cardiac cycles vs. R2

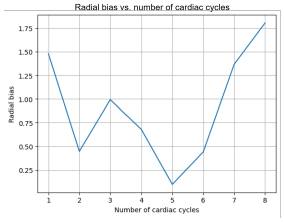


Figure 4.16: #cardiac cycles vs. radial bias

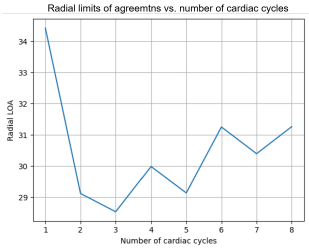


Figure 4.17: #cardiac cycles vs. Radial limits of agreement

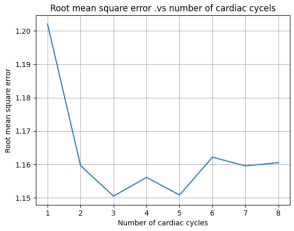


Figure 4.18: #cardiac cycles vs. RMS

4.5 Post processing

In this section, we present the results of the post-processing stage, where we applied a low-pass filter to the estimated values to address fluctuations in the model output. The input for the filter comprises the CO estimation obtained through discovered features and ridge regression. Additionally, we investigated the impact of the number of cardiac cycles on the filtering process. To achieve this, we tested three, four, five, and six cardiac cycles, each with different cut-off frequencies. The chosen filter was a Butterworth low-pass filter of order 5, selected based on experimentation and implemented in Python using the SCIPY library [49].

Figures 4.21, 4.23, 4.24, 4.25, and 4.27,4.22 and 4.28 present the performance of the ridge regression model using three cardiac cycles with varying cut-off frequencies. The performance of the ridge regression with four, five, and six cardiac cycles, along with their respective cut-off frequencies, is available in Appendix A.3, A.4, A.5, andA.6. The cut-off frequencies tested were: 0.25, 0.20, 0.15, 0.10, 0.05, 0.025, 0.015, and 0.005 Hz. Upon analyzing the results, it was found that a cut-off frequency of 0.025 Hz yields the best overall performance, displaying superior results in terms of limits of agreement, percentage error, correlation, R2, radial limits of agreement, root mean square error, and mean absolute error, where either a maximum or minimum is achieved. Although the best cut-off frequency for radial bias and bias is not at 0.025 Hz, it still falls within clinically acceptable performance levels. Thus, 0.025 Hz was the most appropriate cut-off frequency for the filtering process.

Table 4.3 provides a numerical comparison of the performance of three cardiac cycles using different cut-off frequencies. Table 4.4 further showcases the best-performing cut-off frequency (0.025 Hz) performance for different cardiac cycles. Three cardiac cycles demonstrate the most accurate results, while six have the best precision. Therefore, selecting the appropriate number of cardiac cycles represents a trade-off between accuracy and precision. Upon closer examination of Table 4.4, five cardiac cycles strike the best balance between accuracy and precision. As a result, five cardiac cycles will be utilized for further development.

To visualize the impact of the filter, Figure 4.29 demonstrates the filter's application (in blue) to the estimated values (in orange). Conversely, Figure 4.30 showcases the model's performance without the filter. The filter effectively reduces fluctuations in the estimated values, consequently enhancing the model's performance.

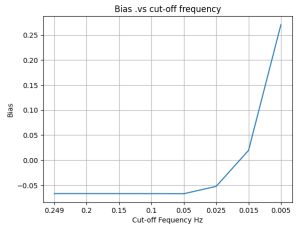


Figure 4.19: Cut-off frequency vs. bis

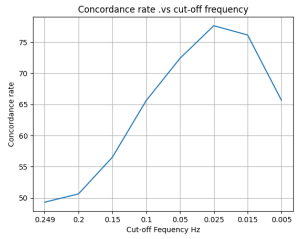


Figure 4.20: Cut-off frequency vs. concordance rate

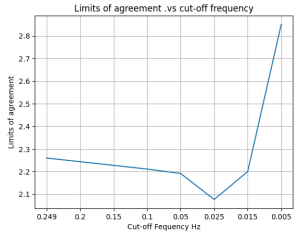


Figure 4.21: Cut-off frequency vs. limits of agreement

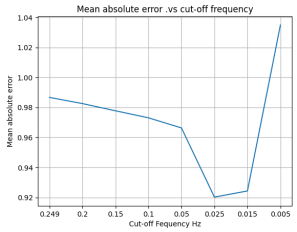


Figure 4.22: Cut-off frequency vs. mean absolute error

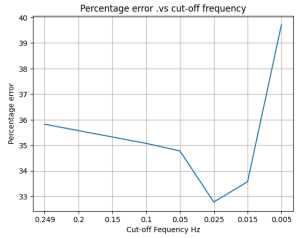


Figure 4.23: Cut-off frequency vs. percentage error

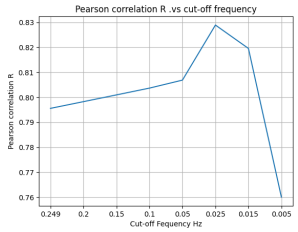


Figure 4.24: Cut-off frequency vs. correlation

Figure 4.25: Cut-off frequency vs. R2

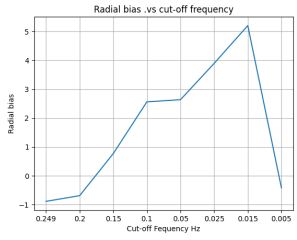


Figure 4.26: Cut-off frequency vs. radial bias

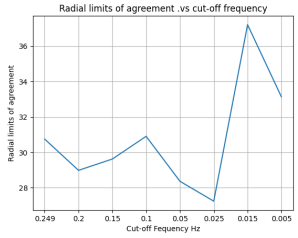


Figure 4.27: Cut-off frequency vs. radial limits of agreement

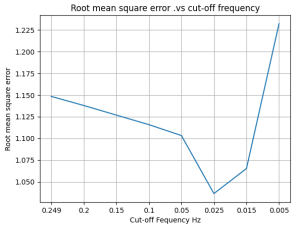


Figure 4.28: Cut-off frequency vs. RMS

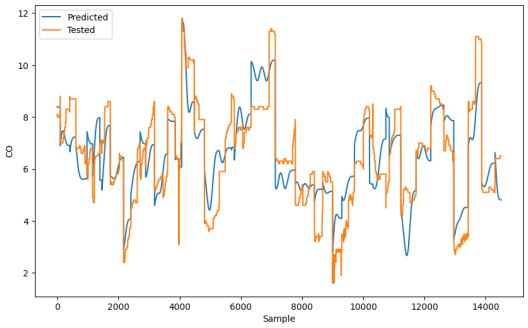


Figure 4.29: Estimated CO in blue vs. tested CO in orange (with filter)

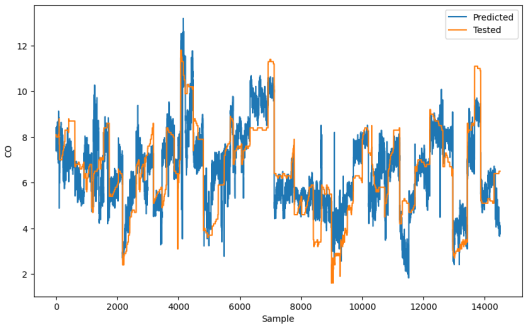


Figure 4.30: Estimated CO in blue vs. tested CO in orange (without filter)

Table 4.3: Low-pass filter cutt-off frequencies comparison

Ḣz/	**RMS**	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
0.25	1.14	0.98	0.79	0.62	-0.07	2.26,-2.39	35.82	49.30	30.74	-0.88
0.2	1.13	0.98	0.79	0.63	-0.07	2.24,-2.38	35.57	50.63	28.98	-0.69
0.15	1.12	0.97	0.80	0.63	-0.07	2.23,-2.36	35.32	56.50	29.62	0.76
0.1	1.11	0.97	0.80	0.64	-0.07	2.21,-2.35	35.07	65.59	30.90	2.56
0.05	1.10	0.96	0.80	0.64	-0.07	2.19,-2.33	34.77	72.36	28.36	2.63
0.025	**1.03**	**0.92**	**0.82**	**0.68**	**-0.05**	**2.08,-2.18**	**32.77**	**77.61**	**27.23**	3.89
0.015	1.06	0.92	0.81	0.67	0.02	2.2,-2.16	33.57	76.11	37.21	5.20
0.005	1.23	1.03	0.76	0.52	0.27	2.85,-2.31	39.71	65.67	33.13	**-0.42**

Table 4.4: Ridge regression performance w.o.f (without filter) compared to w.f.x (with filter number of cardiac cycles)

	RMS	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
w.o.f	1.13	0.96	0.80	0.63	-0.06	2.22,-2.35	35.38	52.30	28.54	-0.88
w.f.3	1.03	0.92	0.82	0.68	-0.05	2.08,-2.18	32.77	77.61	27.23	1.00
w.f.4	1.04	0.93	0.82	0.68	-0.07	2.07,-2.21	32.98	79.10	28.83	5.13
w.f.5	1.04	0.94	0.81	0.67	-0.07	2.10,-2.23	33.34	79.10	25.49	3.23
w.f.6	1.05	0.95	0.81	0.66	-0.07	2.12,-2.27	33.84%	82.08%	29.15	4.44

4.6 Models evaluation

Upon finalizing the system design, the regression models were subjected to an evaluation process. This involved utilizing the discovered features, five cardiac cycles, and applying a low pass filter with a cut-off frequency of 0.025 Hz. This evaluation aimed to assess whether the proposed approach enhances the performance of all regression models. The evaluation results are presented in Table 4.5. A comparison with the results from Table 4.2 reveals a significant improvement in the performance of all models, particularly in the case of linear and ridge regression. These models now achieve a clinically acceptable level of performance concerning radial limits of agreement and bias.

On the other hand, concordance rate and percentage error still fall short of clinically acceptable performance. The specified requirements were a percentage error of less than 30% and a concordance rate higher than 90%. Nevertheless, it is worth mentioning that these parameters can be enhanced by considering various factors, such as adjusting the number of cardiac cycles or employing different cut-off frequencies, depending on the specific application of the algorithm. However, the best possible concordance rate with this model was 82.06%, and the best possible percentage error was 32.77%.

Figure 4.31 presents a visual representation of the performance of the six regression models through bar graphs. It is evident from this figure that both linear and ridge regression outperforms the other models. Furthermore, the tree-based models exhibit a reduced radial bias compared to the linear models.

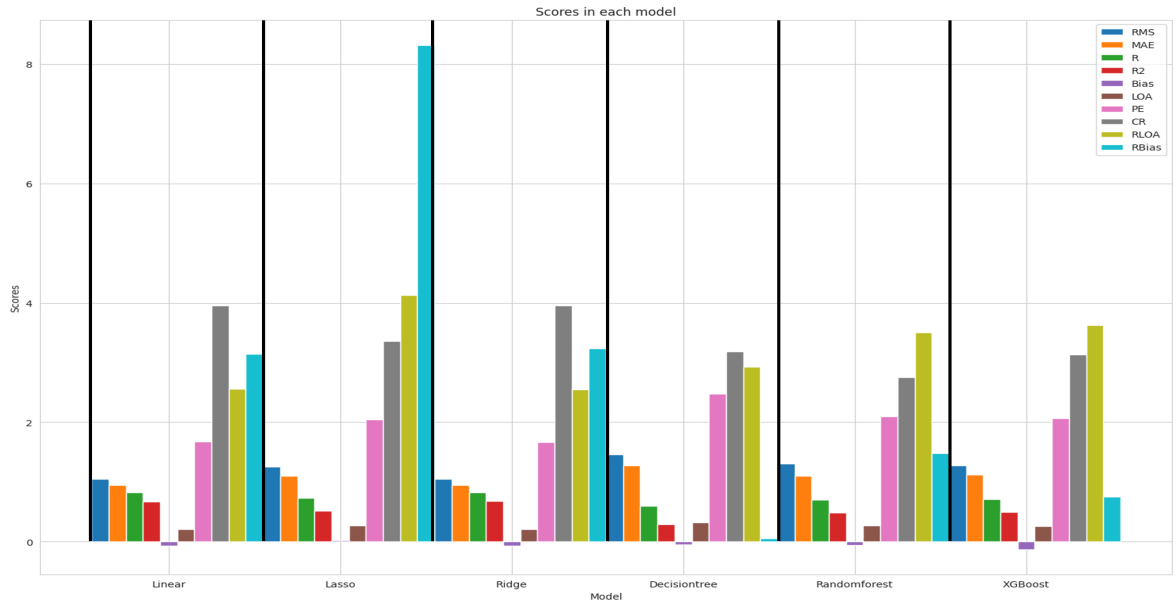


Figure 4.31: Bar-graph showing the results of the re-evaluation of six regression models

Table 4.5: Regressors comparison re-evaluation

	RMS	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
Linear	1.05	0.95	**0.82**	**0.67**	-0.07	2.11,-2.25	33.50	**79.10**	25.57	**3.14**
Lasso	1.25	1.09	0.72	0.51	**0.03**	2.68,-2.62	40.81	67.16	41.28	8.31
Ridge	**1.04**	**0.94**	**0.82**	**0.67**	-0.07	**2.10,-2.23**	**33.34**	**79.10**	**25.49**	3 .23
DecisionTree	1.45	1.27	0.59	0.28	-0.05	3.16,-3.26	49.43	63.76	29.29	0.05
RandomForrest	1.30	1.10	0.70	0.48	-0.06	2.66,-2.78	41.84	55.07	34.99	1.47
XGBoost	1.27	1.11	0.70	0.49	-0.13	2.56,-2.82	41.40	62.68	36.26	0.75

Figure 4.6 illustrates the four-quadrant plot of the best-performing model, which includes ridge regression, five cardiac cycles, and a cut-off frequency of 0.025 Hz. Notably, many estimates lie within the agreement region (quadrants 1 and 3), with fewer points fluctuating around the origin. These points are further visualized in the polar plot depicted in figure 4.6, where most cluster around the angle zero. This clustering indicates both low radial limits of agreement and bias.

Moreover, the low percentage error is reflected in the Bland-Altman analysis showcased in figure 4.6. The low limits of agreement in this analysis reflect the precision of the model, while the low bias reflects its accuracy. Lastly, it is essential to note that

cardiac output values remain relatively constant within each patient, with variations primarily occurring between patients. As a result, most changes fall within the exclusion zone 12%, 0.624 [L/min]. This explains the limited number of points observed in the figures. These findings suggest that the presented results demonstrate the model's effectiveness in distinguishing between different patients effectively.

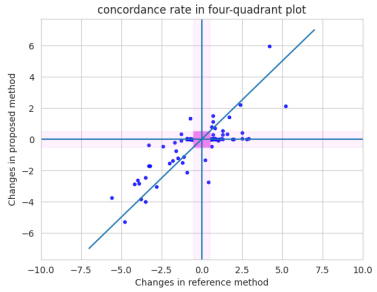


Figure 4.32: Best-performing model: four-quadrant analysis

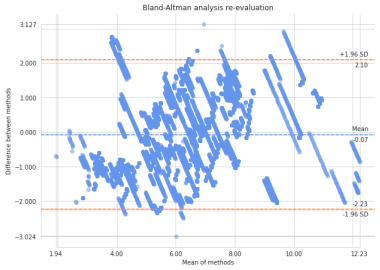


Figure 4.33: Best-performing model: Bland-Altman analysis

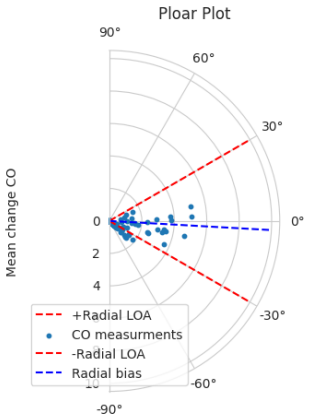


Figure 4.34: Best-performing model: polar plot analysis

4.7 Feature contribution

This section shows the results of analyzing the contribution of the features in the developed model. Evaluating feature importance is crucial for assessing performance and understanding the model's effectiveness. The contribution of four feature sets, namely hemodynamic features, waveform features, and demographic/SINDy features, is demonstrated in table 4.6. Detailed descriptions of these features can be found in table 4.1. Analysis of table 4.6 reveals that the discovered features using the SINDy algorithm have a significant impact, leading to notable improvements in radial limits of agreement and bias, reaching clinically acceptable levels. Moreover, these features improve performance regarding percentage error and concordance rate.

Furthermore, we calculated and visualized the mutual information and F-score of the features with the ground truth in Figures 4.36 and 4.35, respectively. These analyses provide additional insights into the contributions of individual features to the model's performance. The results indicate that heart rate and the logarithm of heart rate have the highest contribution, as evidenced by their high mutual information and F-scores. Duration-related features also exhibit a high F-score, while the inverse of systolic and diastolic pressure demonstrate high mutual information with the ground truth. Additionally, non-linear features show robust mutual information with the ground truth variable, emphasizing their significance. Conversely, derivative-related features demonstrate lower F-scores and mutual information. In summary, the feature contribution

analysis provides valuable insights into the relative importance of different features in the model. The SINDy-discovered features prove to be particularly impactful, significantly improving performance metrics.

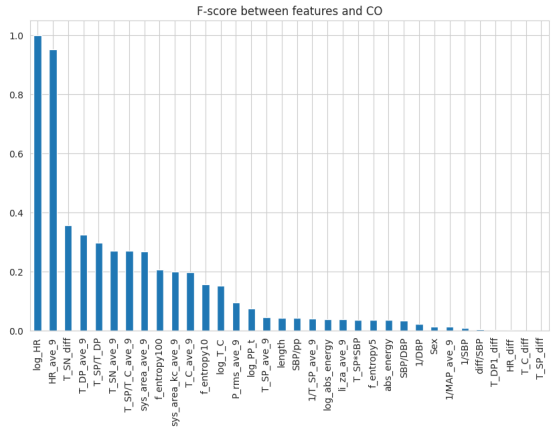


Figure 4.35: Features F-score with cardiac output

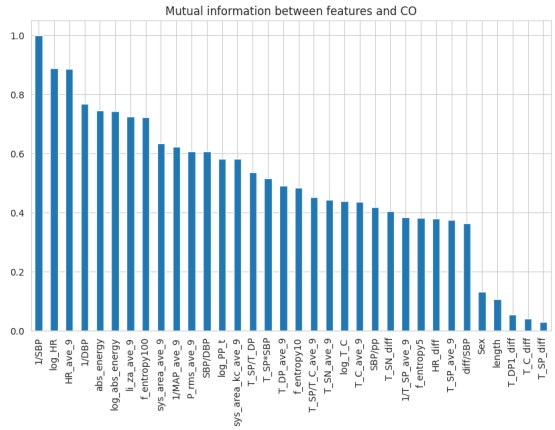


Figure 4.36: Mutual information between cardiac output and the features

Table 4.6: Regressors comparison re-evaluation

	RMS	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
Hemodynamics	1.23	1.05	0.72	0.52	0.02	2.67,-2.64	40.31	64.17	37.67	3.29
Hemodynamics + Waveform	1.22	1.02	0.73	0.54	**0.02**	2.61,-2.58	39.71	71.64	42.44	6.02
Hemodynamics + Waveform + SINDy	1.04	0.94	0.82	0.67	-0.03	2.25,-2.31	33.35	77.61	29.15	5.50
Hemodynamics + Waveform + SINDy + gender	**1.04**	**0.94**	**0.82**	**0.67**	-0.07	**2.10,-2.23**	**33.34**	**79.10**	**25.49**	**3.23**

4.8 Model validation using MIMIC dataset

In this section, we validate the proposed approach utilizing an external dataset, namely, the MIMIC-II Waveform Database, version 2 dataset [40]. In the previous sections, it was found that ridge regression is the best model. Five cardiac cycles are the optimum number, and a cut-off frequency of 0.0025 Hz gives the best performance. In this section, these findings will be validated using an external dataset.

MIMIC II dataset

The MIMIC (Multiparameter Intelligent Monitoring in Intensive Care) dataset contains de-identified health records from patients admitted to intensive care units. It includes various clinical data, such as vital signs, laboratory results, medication administration records, Etc. To validate the model, arterial blood pressure waveforms, corresponding cardiac output measurements, and the gender of the patients are needed. Patients were selected based on the availability of the needed measurements. The arterial blood pressure was sampled at 125 Hz in the MIMIC dataset, and the cardiac output was measured at one-minute intervals. However, cardiac output measurements were unavailable for each minute and appeared at irregular intervals (from one minute up to two hours). One hundred three patients were found, with 1112 samples, 64 male and 39 female.

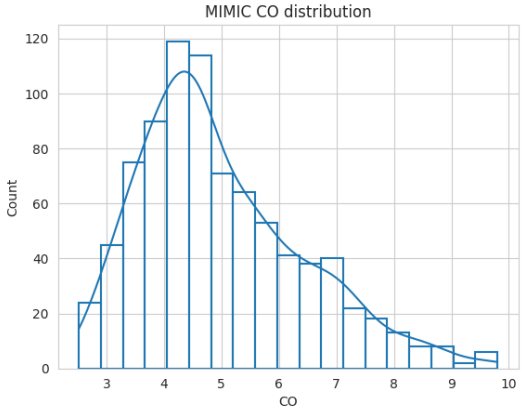


Figure 4.37: Reference cardiac output in eh MIMIC dataset distribution

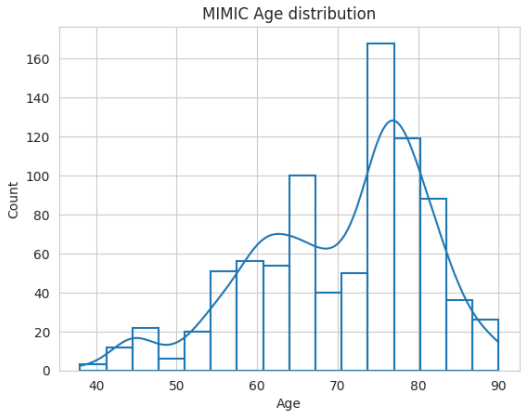


Figure 4.38: Patient's age distribution

Signal processing

Before testing the model, the quality of the waveform was assessed manually to ensure that only suitable waveforms were used in further processing. The patient's ID of the deleted waveform and the used patient's ID can be found in the appendix C.1. After the quality assessment check, 737 samples were left for further testing. These samples are from 87 patients, 56 male and 31 female. Additionally, the waveform was down-sampled to 100 Hz such that the sampling frequency of both datasets matched. However, nothing could be done regarding the measurement interval of the cardiac output as it was measured at irregular intervals in the MIMIC dataset. In contrast, it was measured at two seconds intervals in the Vitaldb dataset. After cleaning, the waveforms, hemodynamics, and waveform features were extracted similarly for the Vitaldb dataset in section 4.2.

After that, the features were standardized using the mean and variance obtained from the Vitaldb. This differed from the standard procedure, where the features should be

standardized using mean and variance from the MIMIC dataset using k-fold cross-validation. However, the MIMIC dataset is 20 times smaller than the Vitalddb, and certain hemodynamic situations were not measured in the MIMIC dataset. While in the Vitaldb, a wider range of hemodynamic situations was captured. For example, before standardization, the heart rate in the Vitaldb is between 41.5 and 145.7, with a mean value of 76.4, as shown in figure 4.39, the orange distribution. After standardization, this mean value will be mapped to zero, as seen in figure 4.40, the orange distribution.

On the other hand, before standardization, the heart rate in the MIMIC dataset is between 53.2 and 205.4, with a mean of 95.5 as seen in figure 4.39, the blue distribution. Therefore, when standardizing using the mean and variance from the MIMIC dataset, the heart rate below 95.5 will be negative relative to the mean. However, after standardization, the model learned from the Vitaldb that a heart rate less than 76.4 gets a negative value. Therefore, since the features have a physiological meaning, keeping the features' range relative to Vitaldb's mean makes more sense as a heart rate of 200 should be mapped into an extremely high heart rate that was not learned in the Vitaldb and not to a region of high heart rate that Vitaldb had learned.

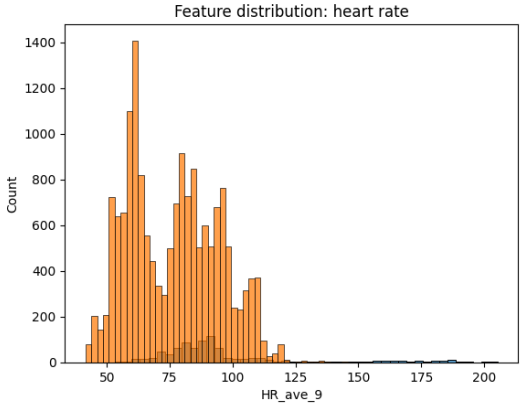


Figure 4.39: Heart rate feature distribution before standardization

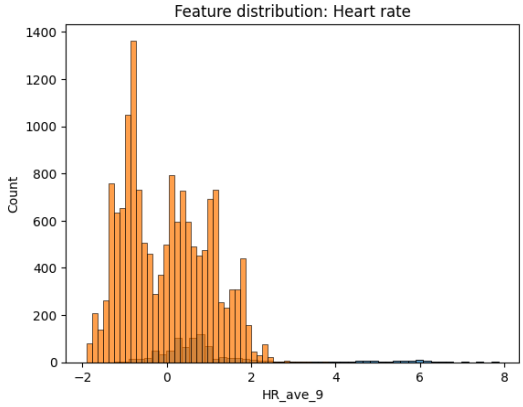


Figure 4.40: Heart rate feature distribution before standardization

Validation results

This section will present the validation results of testing the learned model from the Vitaldb dataset on the MIMIC dataset. The validation test results are summarized in table 4.7. The model achieved good results regarding RMS, MAE, and RLOA. Remarkably, the model achieves a clinically acceptable level regarding radial bias RBias. This means that the results are accurate and precise when the model can track the changes in the reference values. This can be seen in figure 4.43 when most points are distributed between 30° and - 30°. The model's ability to track changes is reflected by the concordance rate of 54.74 %. This can be seen in figure 4.41 where half of the points are in the agreement region (quarters 1 and 3). From the figure 4.42, it can be seen that the model underestimates high cardiac output values as the maximum mean

of methods is around 8.1 and overestimate the low cardiac output values leading to a high percentage error of 45.41% as most of the estimation distributed between 4 and 7 [L/min].

The reference cardiac output was measured at irregular intervals ranging from one measurement per ten minutes to one measurement per two hours. Therefore, the ground truth cardiac output fluctuates a lot such that no two sequential measurements are close to each other, as seen in figure 4.45. This has led to a negative R2 of -0.15, meaning that an estimator that always predicts the mean value is almost similar to the learned model. However, this is not the case, as seen in figure 4.45, where the estimated values (orange signal) do not only predict the mean but, on average, have almost similar performance to an estimator that always predicts the mean. From figure 4.44, it can be seen that the error distribution is centered around 0.21, and most of the errors made are around that low error region. However, fewer values have large errors leading to high limits of agreements of 2.49,-2.91. These errors are mainly due to new hemodynamic situations that appeared in the MIMIC dataset and were not in the Vitalddb dataset. Therefore, the model will be re-trained using the MIMIC dataset in the next section to test if the proposed approach can learn different hemodynamic situations.

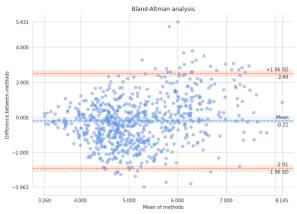


Figure 4.41: Linear regression: Four-quadrant analysis

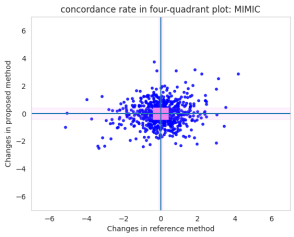


Figure 4.42: Linear regression: Bland-Altman analysis

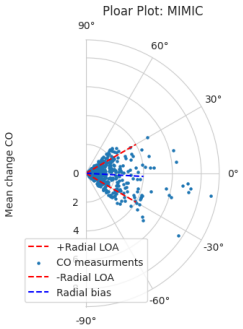


Figure 4.43: Linear regression: Polar plot analysis

Table 4.7: Validation results: the model was trained on Vitaldb and tested on MIMIC datases

	RMS	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
MIMIC	1.39	1.09	0.29	-0.15	-0.21	2.49,-2.91	45.41	54.74	30.07	3.15
Vitaldb	1.04	0.94	0.82	0.67	-0.07	2.10,-2.23	33.34	79.10	25.49	3.23

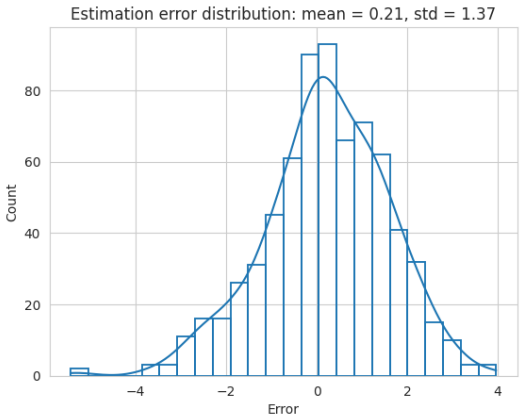


Figure 4.44: Estimation error distribution

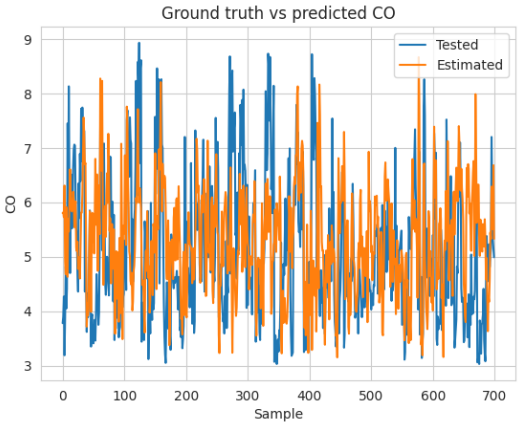


Figure 4.45: Image showing the fluctuation in the ground-truth CO. Test CO in blue .vs estimated value in orange

Re-train on MIMIC

This section presents the results of re-training the model on the MIMIc dataset. The procedure of this test is similar to the previous section 4.8 except that the model will be trained on the MIMIC dataset instead of the Vitaldb to assess the proposed approach's ability to learn different hemodynamic situations. Additionally, different models were tested, including linear models such as linear, lasso, and ridge regression and tree-based regression models such as decision tree, random forest, and XGBoost.

Using leave-one-patient-out cross-validation, six models were trained and tested. The results are summarized in table 4.8. Using the same approach, ridge regression is again the best-performing model alongside linear regression, as it was found using the Vitaldb. However, the performance is not as good as the training on the Vitaldb. This is mainly due to the limited amount of data in the MIMIC dataset.

Figures 4.46. 4.47, and 4.48 show the performance of the ridge regression re-trained on the MIMIC dataset, including Bland-Altman analysis, four-quadrant analysis, and polar plot analysis, respectively. As can be seen from the results, the model achieved better performance when re-trained using the MIMIC, indicating that the proposed approach can learn different hemodynamic situations.

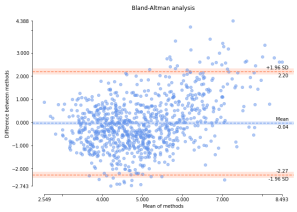


Figure 4.46: Linear regression: Four-quadrant analysis

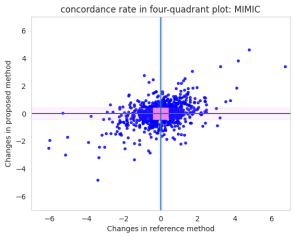


Figure 4.47: Linear regression: Bland-Altman analysis

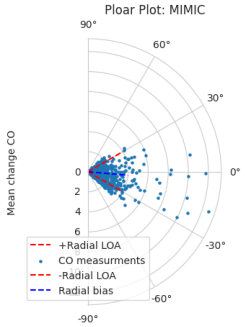


Figure 4.48: Linear regression: Polar plot analysis

Table 4.8: Regressors comparison re-train using MIMIC dataset

	RMS	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE%**	**CR%**	**RLOA°**	**RBias°**
Linear	**1.13**	**0.89**	**0.60**	**0.35**	**-0.02**	2.21,-2.26	**44.77**	67.49	31.54	5.47
Ridge	**1.13**	**0.89**	**0.60**	**0.35**	-0.04	**2.20,-2.27**	**44.77**	**67.08**	**31.51**	**5.43**
Lasso	1.16	0.89	0.57	0.33	-0.02	2.25,-2.30	45.58	66.47	31.13	5.63
DecisionTree	1.58	1.19	0.40	-0.24	-0.15	2.94,-3.24	61.93	58.08	23.83	-5.62
RandomForrest	1.20	0.93	0.54	0.27	-0.1	2.26,-2.46	47.27	62.04	33.72	7.78
XGBoost	1.23	0.95	0.53	0.23	-0.12	2.29,-2.54	48.40	58.94	31.46	6.11

5 Discussion

In this chapter, we will present a discussion of the obtained results. This study aimed to design an algorithm for estimating cardiac output from arterial blood pressure measurements. The study aimed to determine the optimal number of cardiac cycles used for feature extraction to achieve the highest level of accuracy and precision. Furthermore, the study employed the SINDy algorithm to discover new features. Finally, the proposed method was validated using the MIMIC II dataset. The findings of this study offer valuable insights into the system's dynamics, considering the optimal number of cardiac cycles and the newly identified features.

5.1 Evaluation

Using the optimum number of cardiac cycles (five) alongside the discovered features mentioned in table 4.1, the proposed solution achieved remarkable results outperforming other literature in terms of RMS (root means square error), MAE (mean absolute error), R2, LOA (limits of agreement), PE (percentage error), RLOA (radial limits of agreement), and RBias (radial bias) when training and testing using the Vitlaldb dataset. Additionally, our model is the best-performing model regarding MAE, LOA, RLOA, and RBias when training and testing on the MIMIC II dataset. Finally, the model achieved the best-performing model regarding RLOA and RBias when trained on the Vitaldb and tested on the MIMIC dataset. Table 5.1 compares this study's performance with other studies. Regarding Bias, CR (concordance rate), and R (correlation), the proposed solution in this study was close to the best literature.

It should be noted that in [21], the data was split into training and test datasets based on samples. This implies that some of the test patients were seen during training. Especially that age and gender were essential features in their study and model. Additionally, it should be mentioned that the dataset used in [21] is twice as big as the MIMIC data used in this study. They used versions II and III, while we used only version II due to data accessibility issues.

It is also important to highlight that the results achieved in [25] are in comparison to a commercially available cardiac output device. Such devices utilize arterial blood pressure, hence less invasive measurements, to estimate cardiac output. Inaccuracies were reported for such devices. Data was split into training and testing datasets; hence, the test was not performed on all patients. A more proper way to evaluate the system was using leave-one-patient-out cross-validation, as we did in this study.

As a result of this comparison with the literature, this is the only study that has used proper leave-one-patient-out cross-validation and reported all clinically acceptable metrics. Additionally, this study is the only study that validated the learned model on an external dataset. Finally, this study is the only study that achieved a clinically acceptable level regarding radial limits of agreements (RLOA) and radial bias (RBias).

Table 5.1: Comparison with literature. Cells with (-) indicate that the study did not report the metric. * The study estimated stroke volume instead of cardiac output. Therefore, the comparison is complex since stroke volume should first be converted to cardiac output. ** reference CO was not the gold standard.

	Vitaldb	MIMIC II	MIMIC II validation	[21] (MIMIC II+III)	[55] (Vitaldb)	[25]** (Vitaldb)	[34]
RMS	**1.04**	1.13	1.39	1.192	-	1.45	-
MAE	**0.94**	**0.89**	1.03	0.94	-	1.01	-
R	0.82	0.60	0.29	-	0.64	0.95**	**0.84**
R2	**0.67**	0.35	-0.15	0.49	-	-	-
Bias	-0.07	-0.04	-0.21	**-0.01**	*	-0.85	*
LOA	**2.10,-2.23**	**2.20,-2.27**	2.49,-2.91	2.32,-2.35	*	0.72,-2.88	*
PE	**33.34%**	44.77	45.41	39.44%	-	-	-
CR	79.10%	67.08	54.74	**79.91%**	53%	96.26% **	77.74
RLOA	**25.49**	**31.51**	**30.07**	31.6	-	-	-
RBias	**3.23**	**5.43**	**3.15**	16.3	-	-	-

5.2 Explanation

In this section, an interpretation of the achieved results will be addressed. This section explains why the five cardiac cycles yield the best performance. After that, the interpretation of the learned features and elaboration on the learned model will be provided.

Upon investigation, it was revealed that three cardiac cycles yielded the most accurate results, while six cardiac cycles provided the highest precision. Further analysis indicated that five cardiac cycles struck the best balance between accuracy and precision. The average duration of a cardiac cycle is approximately 0.8 seconds. On the other hand, cardiac output measurements were taken at 2-second intervals, as illustrated in figure 5.1. Based on this figure, we can conclude that three cardiac cycles

align most closely with the 2-second intervals. Similarly, six cardiac cycles align with the two cardiac output measurements taken within a 4-second interval. This finding suggests that the model utilizes the data from the last two seconds, equivalent to the previous three cardiac cycles that were used to measure the ground truth for the most accurate results. Additionally, to provide the most precise results, the model considers the average of the last two cardiac output measurements, equivalent to the previous six cardiac cycles. Finally, five cardiac cycles aim to strike the best balance between accuracy and precision by relying more on the last cardiac output measurement than the one before.

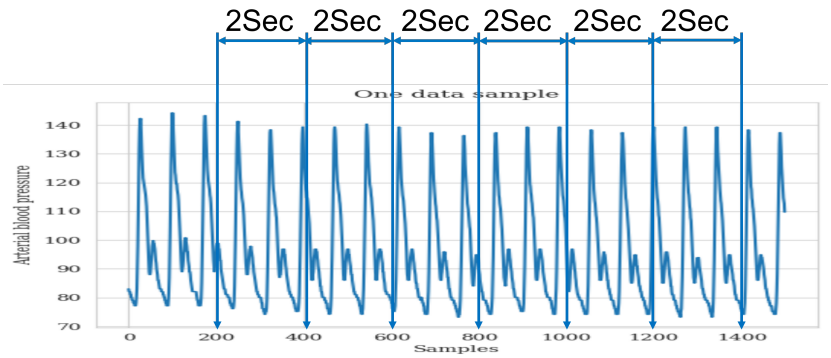


Figure 5.1: Arterial blood pressure waveform with arrows indicating cardiac output measurement moments (2 seconds interval)

Furthermore, applying the SINDy algorithm allowed us to discover new features for cardiac output estimation. The algorithm effectively identified 33 features that offer insight into the underlying dynamics and relationships between arterial blood pressure and cardiac output. These features were related to five physical phenomena: velocity, pressure, power, energy, and randomness. A linear combination of these features describes the learned model.

To write the leaned model in a symbolic equation, a vector u is defined that contains velocity-related features: heart rate, systolic duration, systolic notch duration, diastolic duration, and duration of the cardiac cycle. Upon closely examining the feature set, it can be observed that the vector u also defines the derivative or change-related features as Δu.

$$\mathrm{u} = [\mathrm{HR}, \text{systolic duration}, \text{systolic notch duration}, \text{diastolic duration}, \text{cardiac cycle duration}]. \quad (5.1)$$

Another vector x is also defined to include logarithm-related features such as heart rate, cardiac cycle duration, pulse pressure, and absolute energy. It should be noted that x and u share heart rate and cardiac cycle duration. Therefore, the logarithmic features can be written as $\mathrm{x} + log(\mathrm{x})$ since the features defined in x are also included with their logarithm in the feature matrix. Finally, R and v vectors are defined to include randomness-related features and all other non-linear features, respectively.

$$\mathrm{x} = [\mathrm{HR}, \text{systolic duration}, \text{pulse pressure}, \text{absolute energy}]. \quad (5.2)$$

$$\mathrm{R} = [\text{Fourier entropy5}, \text{Fourier entropy10}, \text{Fourier entropy100}]. \tag{5.3}$$

Using the defined vectors, the learned model can be written in short notation as a linear combination of these vectors.

$$\mathrm{CO} = (\mathrm{u} + \Delta\mathrm{u} + \mathrm{p} + \Delta\mathrm{p} + \mathrm{x} + log(\mathrm{x}) + \mathrm{R} + \mathrm{v})\mathrm{W}. \tag{5.4}$$

Here W is the learned weights from training the model $W = [\mathrm{w}_1, \mathrm{w}_2, ...\mathrm{w}_n]$. Upon closer look at the learned equation, similarity to the Navier-Stokes equation is observed. The term $\mathrm{u} + \Delta\mathrm{u} + \mathrm{p} + \Delta\mathrm{p}$ holds information about the velocity and pressure terms with their derivatives similar to the Navier-stokes equations that describe the flow of a fluid. The Navier-Stokes equations also have other terms related to the density and deviatoric stress tensor. These terms describe turbulence and viscosity in the flow of a fluid. This study suggests that the other terms in the learned model $\mathrm{x} + log(\mathrm{x}) + \mathrm{R} + \mathrm{v}$ are related to turbulence and viscosity phenomenons similar to the Navier-Stokes equations.

Therefore, the learned model is interpreted as a differential equation that shares similarities with the Navier-Stokes equations, where cardiac output acts as an external force to the system. Modeling the cardiac output as an external force is connected to the classical approach, where cardiac output is modeled as a current source delivering power to a circuit consisting of resistors and capacitors. However, modeling the problem as fluid flow does make more sense than modeling the problem from a circuit theory perspective, as the nature of the problem at hand is fluid (volume of blood).

Additionally, the fact that ridge regression was the best-performing model also says something about the nature of the problem. Originally, ridge regression was designed to solve non-orthogonal systems of equations where $\mathrm{A}^T\mathrm{A} \neq \mathrm{A}\mathrm{A}^T$ [15]. Where A is the features matrix in this study. This was verified to be the case by multiplying $\mathrm{A}^T\mathrm{A}$ and $\mathrm{A}\mathrm{A}^T$. Systems that obey non-orthogonal systems of equations are well known in fluid dynamics as non-normal/degenerate systems. The more significant the difference between $\mathrm{A}^T\mathrm{A}$ and $\mathrm{A}\mathrm{A}^T$, the more non-normal the system is. Non-normal systems are described by $\mathrm{te}^{-\mathrm{t}\lambda}$ where such a system seems to be unstable at the beginning; however, after a certain time, the system goes back to a stable state. Such behavior is shown in figure 5.2. Similarities were also observed in the behavior of such a system and the continuous blood pressure waveform shown in figure 5.3. Therefore, we suggest that part of the arterial blood pressure system in a non-orthogonal system was captured by the terms $\mathrm{x} + log(\mathrm{x})$ in the feature matrix. Having the cardiac output as a power of an exponent has a solid connection to the classical approach for cardiac output estimation.

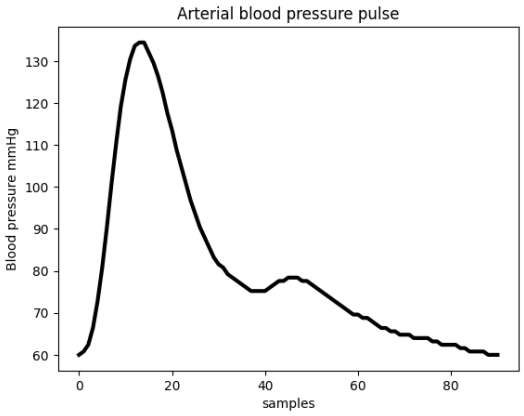


Figure 5.2: Simulation of 100 points of $te^{t\lambda}$

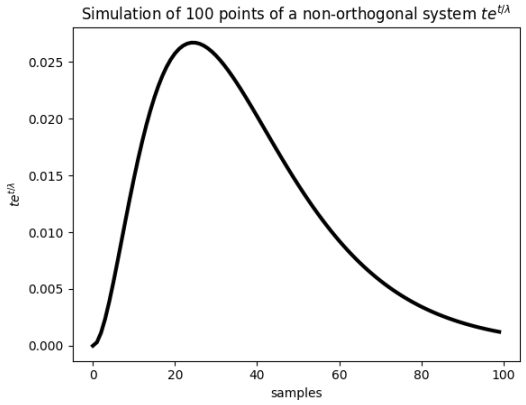


Figure 5.3: One cardiac cycle of arterial blood pressure

5.3 Limitations

In this section, the limitations associated with this study are highlighted. Although the model achieved clinically acceptable performance, it should be noted that the study was conducted within a specific population and clinical settings where a limited range of hemodynamic situations was captured. This was evident when the model was tested on the external dataset leading to performance degradation.

Furthermore, our anesthesiologists from the Erasmus MC told us that the model's velocity and pressure-related features are logical. Additionally, having non-linear terms is understandable because the relation between cardiac output and blood pressure is not linear. Therefore, having logarithmic or rational features is acceptable. However, it was unexpected that demographic information such as age and body mass index (BMI) would play no role in the model. In the learned model, only gender was used as a demographic feature, as there is a difference between men and women regarding cardiac output. However, the relationship between age and BMI with cardiac output seems non-linear; therefore, a linear combination of age and BMI will not be a solution. Further investigation is needed to determine how to add these features to the system.

It should also be mentioned that the quality of the waveform hindered the feature extraction process. This was evident when the quality of the MIMIC waveforms was assessed manually. It was noticed that feature extraction does not work well in certain situations. For example, when the dicrotic notch disappears from the waveform. Such a phenomenon has a physiological interpretation, meaning the heart did not relax well. We did not have any method to detect such a situation, which led to false pressure and time-related feature calculation. However, since the waveforms in the MIMIC dataset were assessed manually, a labeled dataset is now available to design an algorithm that can detect such situations.

Additionally, the interpretability of the model should be further investigated to describe the similarity between the model and the Navier-Stokes equation, as the connection is still not fully clear. Also, perhaps, an expert in differential equations and fluid dynamics can help understand the model more.

Finally, the hyper-parameter of the ridge regression was hard to determine as by tuning this parameter, an assumption is made about the prior distribution of the weights, which may not be true. From a Bayesian perspective, ridge regression is the solution of the Bayesian linear regression when the prior of the weights is Guissian.

6 Conclusion and future work

In this study, our primary objective was to design an algorithm for less invasive estimation of cardiac output based on continuous arterial blood pressure measurement. We aimed to develop a method that could offer comparable performance to the standard gold method for measuring cardiac output, namely the Swan-Ganz catheter, and to be as explainable as possible. Using machine learning techniques and feature engineering, we successfully designed an algorithm that achieved clinically accepted performance regarding radial limits of agreements and angular bias. The learned model was validated on an external dataset and achieved comparable performance. Furthermore, the developed model was interpreted as a differential equation related to the Navier-Stokes equations.

The core focus of our system design revolved around identifying the input features and determining the optimum number of cardiac cycles for accurate estimation of cardiac output. These factors play a crucial role in the overall performance and reliability of the system. We captured essential information necessary for the explainable estimation of cardiac output by engineering relevant features. Additionally, investigating the optimum number of cardiac cycles allowed us to optimize the efficiency and practicality of the system.

Furthermore, the interpretability of the developed model is a notable strength of our system. By providing valuable insights into the system's interpretation, we enhance our understanding of the underlying physiological processes involved in cardiac output estimation. This contributes to the scientific knowledge in the field and facilitates medical professionals' acceptance and adoption of the system.

Future work

This research contributes to the ongoing efforts to develop less invasive and accessible methods for cardiac output estimation, ultimately leading to improved patient care and enhanced diagnostic capabilities. In this chapter, an elaboration on future work will be provided.

For future work, it is recommended to design an algorithm to assess the quality of the arterial blood pressure waveform. Such an algorithm will enhance the model's performance and will allow the model to detect a broader range of hemodynamic situations. We propose utilizing the manually labeled dataset available in the appendix C.4 to achieve this.

An additional area of research could be estimating the time delay between arterial blood pressure measurements using an arterial line and the corresponding moment when the heart pumps the blood. In addition to the time delay, the attenuation in the blood pressure waveform should be estimated to reconstruct the original shape of the blood pressure as it was pumped by the heart. This valuable information can be extracted from the ECG signals. When the heart pumps, the ECG signal exhibits a peak, which may also reflect the strength of the pump.

The underlying concept is to align the blood pressure signal with the heart's activity. Achieving this alignment is crucial to accurately capture the relationship between the heart's pumping and the corresponding blood pressure variations. We hypothesize that estimating cardiac output from pressure information measured next to the heart would give better results.

In cases where the ECG signals alone cannot fully reflect the attenuation in the blood pressure waveform, an alternative approach could involve leveraging machine learning techniques to learn this relationship from invasively measured data. The Vitaldb dataset presents an opportunity for this, as it provides measurements of SPB (Systolic Blood Pressure) and DBP (Diastolic Blood Pressure), which were acquired invasively, and these measurements can be used as ground truth for the model.

By delving into this area of research and exploring the alignment of ECG and blood pressure signals, we could unlock a novel method for estimating cardiac output and potentially improve the understanding of cardiovascular dynamics.

Additionally, diversifying the training approach to encompass distinct models for estimating cardiac output holds potential advantages. As demonstrated in this study, cardiac output can be categorized into three ranges: low, mid, and high. The evaluation of the MIMIC dataset revealed that the model tended to underestimate high values and overestimate low values, displaying a preference for estimating values in the mid-range. This discrepancy may indicate that the relationship between high and low values differs slightly between blood pressure and cardiac output.

Consequently, exploring the development of three separate models could prove fruitful. Each model would be tailored to a specific range (low, mid, and high), allowing for more accurate estimations across the spectrum. However, it is essential to acknowledge a potential challenge of this approach—the scarcity of data points corresponding to high and low cardiac output values. Since such instances are infrequent, ensuring sufficient data for each model could be problematic. However, exploring the creation of multiple models for cardiac output estimation, each designed for a specific range, could yield accuracy improvements and enhance the findings' clinical relevance.

Moreover, it is recommended that the model is applied using data from the Erasmus MC as part of future work. By fitting the model to the specific data from the Erasmus

MC and conducting live tests on-site, we can take significant strides toward the practical implementation of this research. An option worth exploring would be merging the Vitaldb or MIMIC data with the Erasmus MC data, given that the distributions of the features are similar. This data fusion would provide a more diverse and comprehensive dataset, allowing for further fine-tuning and validation of the model.

For future research, we also aim to utilize the time-delay embedding coordinates to describe the system's dynamics and its relation to cardiac output. Since we learned from this study that the system is non-orthogonal, and an orthogonal basis exists that describes the system enabling linear combination instead of non-linear relationships. This is also in line with the Koopman operator theory that state that a non-linear system could be linearized in a higher dimension (in theory, infinity dimension), assuming the existence of hidden variables that are not measured.

This operator is extremely hard to find in practice. However, an alternative exists known as the Henkel alternative that utilizes the eigenvectors of a Henkel matrix obtained from a time-series signal. Therefore, the idea is to construct a Henkel matrix as shown in the matrix below 6.1, and by each raw shift, the hidden variable will interact. This interaction could be captured by taking the singular value decomposition of the matrix.

$$\mathrm{x}_{\mathrm{ABP}} = \begin{pmatrix} \mathrm{ABP}_1 & \dots & \mathrm{ABP}_{10} & \dots & \mathrm{ABP}_{20} & \dots & \mathrm{ABP}_N \\ \mathrm{ABP}_2 & \dots & \mathrm{ABP}_{11} & \dots & \mathrm{ABP}_{21} & \dots & \mathrm{ABP}_{N+1} \\ \mathrm{ABP}_3 & \dots & \mathrm{ABP}_{12} & \dots & \mathrm{ABP}_{22} & \dots & \mathrm{ABP}_{N+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{6.1}$$

The assumption is that the system's behavior could be described using a linear combination of these eigenvectors. The SINDy algorithm could be used again to capture this relationship, and we learned from this study that the cardiac output is an input of such a system. Therefore, we use the cardiac output as input.

A model based on the eigenvectors of a Henkel matrix and accepting inputs is very desirable. Since we could also use demographic information as input to the system and anesthesia drug dosages such as propofol, this could be very helpful in estimating cardiac output and determining the depth of anesthesia based on arterial blood pressure waveform.

Bibliography

[1] Tatsuya Arai et al. "Estimation of changes in instantaneous aortic blood flow by the analysis of arterial blood pressure". In: *Journal of Applied Physiology* 112.11 (2012), pp. 1832–1838.

[2] J Martin Bland and DouglasG Altman. "Statistical methods for assessing agreement between two methods of clinical measurement". In: *The lancet* 327.8476 (1986), pp. 307–310.

[3] Maurice J Bourgeois et al. "Characteristics of aortic diastolic pressure decay with application to the continuous monitoring of changes in peripheral vascular resistance". In: *Circulation research* 35.1 (1974), pp. 56–66.

[4] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. "Sparse identification of nonlinear dynamics with control (SINDYc)". In: *IFAC-PapersOnLine* 49.18 (2016), pp. 710–715.

[5] Gari D Clifford, Daniel J Scott, and Mauricio Villarroel. "User guide and documentation for the MIMIC II database (version 2, release 1)". In: (2010).

[6] Lester A Critchley, Anna Lee, and Anthony M-H Ho. "A critical review of the ability of continuous cardiac output monitors to measure trends in cardiac output". In: *Anesthesia & Analgesia* 111.5 (2010), pp. 1180–1192.

[7] Lester A Critchley, Xiao X Yang, and Anna Lee. "Assessment of trending ability of cardiac output monitors by polar plot methodology". In: *Journal of cardiothoracic and vascular anesthesia* 25.3 (2011), pp. 536–546.

[8] Nader Jafarnia Dabanloo, Fatemeh Adaei, and Ali Motie Nasrabadi. "The performance of neural network in the estimation of cardiac output using arterial blood pressure waveforms". In: *2011 Computing in Cardiology*. IEEE. 2011, pp. 145–148.

[9] Kate E Drummond and Edward Murphy. "Minimally invasive cardiac output monitors". In: *Continuing Education in Anaesthesia, Critical Care & Pain* 12.1 (2012), pp. 5–10.

[10] Joseph Erlanger. "An experimental study of blood-pressure and of pulse-pressure in man". In: *Bull Johns Hopkins Hosp* 12 (1904), pp. 145–378.

[11] Qi Guo and Xiaomei Wu. "Measuring cardiac output through thermodilution based on machine learning". In: *Journal of Mechanics in Medicine and Biology* 21.05 (2021), p. 2140003.

[12] Feras Hatib et al. "Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis". In: *Anesthesiology* 129.4 (2018), pp. 663–674.

[13] J Alan Herd, NORMAN R Leclair, and WILLIAM Simon. "Arterial pressure pulse contours during hemorrhage in anesthetized dogs". In: *Journal of Applied Physiology* 21.6 (1966), pp. 1864–1868.

[14] Mayu Hiraishi, Kensuke Tanioka, and Toshio Shimokawa. "Concordance rate of a four-quadrant plot for repeated measurements". In: *BMC Medical Research Methodology* 21 (2021), pp. 1–16.

[15] Arthur E Hoerl and Robert W Kennard. "Ridge regression: Biased estimation for nonorthogonal problems". In: *Technometrics* 12.1 (1970), pp. 55–67.

[16] R Jacob, B Dierberger, and G Kissling. "Functional significance of the Frank-Starling mechanism under physiological and pathophysiological conditions". In: *European heart journal* 13.suppl_E (1992), pp. 7–14.
[17] jaketmp and Lee Tirrell. *jaketmp/pyCompare:* June 2021.
[18] Young-Seob Jeong et al. "Prediction of blood pressure after induction of anesthesia using deep learning: A feasibility study". In: *Applied Sciences* 9.23 (2019), p. 5135.
[19] Max M Jonas and Suzie J Tanser. "Lithium dilution measurement of cardiac output and arterial pulse waveform analysis: an indicator dilution calibrated beat-by-beat system for continuous estimation of cardiac output". In: *Current opinion in critical care* 8.3 (2002), pp. 257–261.
[20] Alan A. Kaptanoglu et al. "PySINDy: A comprehensive Python package for robust sparse system identification". In: *Journal of Open Source Software* 7.69 (2022), p. 3994. DOI: 10.21105/joss.03994. URL: https://doi.org/10.21105/joss.03994.
[21] Liao Ke et al. "Machine Learning Algorithm to Predict Cardiac Output Using Arterial Pressure Waveform Analysis". In: *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE. 2022, pp. 1586–1591.
[22] Samir Kendale et al. "Supervised machine-learning predictive analytics for prediction of postinduction hypotension". In: *Anesthesiology* 129.4 (2018), pp. 675–688.
[23] Jack W Kern and William C Shoemaker. "Meta-analysis of hemodynamic optimization in high-risk patients". In: *Critical care medicine* 30.8 (2002), pp. 1686–1692.
[24] Nicholas T Kouchoukos, Louis C Sheppard, and DONALD A McDONALD. "Estimation of stroke volume in the dog by a pulse contour method". In: *Circulation Research* 26.5 (1970), pp. 611–623.
[25] Hye-Mee Kwon et al. "Estimation of stroke volume variance from arterial blood pressure: Using a 1-D convolutional neural network". In: *Sensors* 21.15 (2021), p. 5130.
[26] Hyung-Chul Lee and Chul-Woo Jung. "Vital Recorder—a free research tool for automatic recording of high-resolution time-synchronised physiological data from multiple anaesthesia devices". In: *Scientific reports* 8.1 (2018), pp. 1–8.
[27] Hyung-Chul Lee et al. "VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients". In: *Scientific Data* 9.1 (2022), p. 279.
[28] G Liljestrand and E Zander. "Comparative determination of the minute volume of the heart in humans using the nitric oxide method and blood pressure measurement". In: *Journal for all experimental medicine* 59 (1928), pp. 105–122.
[29] Nehemiah T Liu et al. "Blood pressure and heart rate from the arterial blood pressure waveform can reliably estimate cardiac output in a conscious sheep model of multiple hemorrhages and resuscitation using computer machine learning approaches". In: *Journal of Trauma and Acute Care Surgery* 79.4 (2015), S85–S92.
[30] S Lorsomradee et al. "Continuous cardiac output measurement: arterial pressure analysis versus thermodilution technique during cardiac surgery with cardiopulmonary bypass". In: *Anaesthesia* 62.10 (2007), pp. 979–983.
[31] Dominique Makowski et al. "NeuroKit2: A Python toolbox for neurophysiological signal processing". In: *Behavior Research Methods* 53.4 (Feb. 2021), pp. 1689–1696. DOI: 10.3758/s13428-020-01516-y. URL: https://doi.org/10.3758%5C%2Fs13428-020-01516-y.
[32] James F Martin et al. "Application of pattern recognition and image classification techniques to determine continuous cardiac output from the arterial pressure waveform". In: *IEEE transactions on biomedical engineering* 41.10 (1994), pp. 913–920.

[33] LJ Montenij et al. “Methodology of method comparison studies evaluating the validity of cardiac output monitors: a stepwise approach and checklist”. In: *British Journal of Anaesthesia* 116.6 (2016), pp. 750–758.

[34] Young-Jin Moon et al. “Deep learning-based stroke volume estimation outperforms conventional arterial contour method in patients with hemodynamic instability”. In: *Journal of clinical medicine* 8.9 (2019), p. 1419.

[35] Ramakrishna Mukkamala et al. “Continuous cardiac output monitoring by peripheral blood pressure waveform analysis”. In: *IEEE Transactions on Biomedical Engineering* 53.3 (2006), pp. 459–467.

[36] Rahul Nanchal and Robert W Taylor. “Hemodynamic monitoring”. In: *Chap* 41 (2007), pp. 471–486.

[37] Dmitri Nepogodiev et al. “Global burden of postoperative death”. In: *The Lancet* 393.10170 (2019), p. 401.

[38] Michael F O’Rourke and Toshio Yaginuma. “Wave reflections and the arterial pulse”. In: *Archives of internal medicine* 144.2 (1984), pp. 366–371.

[39] Peter M Odor, Sohail Bampoe, and Maurizio Cecconi. “Cardiac output monitoring: Validation studies–how results should be presented”. In: *Current anesthesiology reports* 7 (2017), pp. 410–415.

[40] Mohammed Saeed et al. “MIMIC II: a massive temporal ICU patient database to support research in intelligent patient monitoring”. In: *Computers in cardiology*. IEEE. 2002, pp. 641–644.

[41] M Sanders, S Servaas, and C Slagt. “Accuracy and precision of non-invasive cardiac output monitoring by electrical cardiometry: a systematic review and meta-analysis”. In: *Journal of clinical monitoring and computing* 34.3 (2020), pp. 433–460.

[42] B Saugel et al. “Noninvasive continuous cardiac output monitoring in perioperative and intensive care medicine”. In: *British journal of anaesthesia* 114.4 (2015), pp. 562–575.

[43] Brian de Silva et al. “PySINDy: A Python package for the sparse identification of nonlinear dynamical systems from data”. In: *Journal of Open Source Software* 5.49 (2020), p. 2104. DOI: 10.21105/joss.02104. URL: https://doi.org/10.21105/joss.02104.

[44] Non-Invasive Sphygmomanometers—Part. *2: Clinical validation of automated measurement type*. Tech. rep. ANSI/AAMI/ISO Standard 81060-2, 2009.

[45] JX Sun, AT Reisner, and RG Mark. “A signal abnormality index for arterial blood pressure waveforms”. In: *2006 Computers in Cardiology*. IEEE. 2006, pp. 13–16.

[46] JX Sun et al. “Estimating cardiac output from arterial blood pressurewaveforms: a critical evaluation using the MIMIC II database”. In: *Computers in Cardiology, 2005*. IEEE. 2005, pp. 295–298.

[47] A Vakily et al. “A system for continuous estimating and monitoring cardiac output via arterial waveform analysis”. In: *Journal of biomedical physics & engineering* 7.2 (2017), p. 181.

[48] JJ Van Lieshout and KH Wesseling. *Editorial II: Continuous cardiac output by pulse contour analysis?* 2001.

[49] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[50] Homer R Warner et al. “Quantitation of beat-to-beat changes in stroke volume from the aortic pulse contour in man”. In: *Journal of Applied Physiology* 5.9 (1953), pp. 495–507.

[51] Thomas B Watt Jr and Charles S Burrus. "Arterial pressure contour analysis for estimating human vascular properties". In: *Journal of applied physiology* 40.2 (1976), pp. 171–176.

[52] KH Wesseling. "A simple device for the continuous measurement of cardiac output". In: *Adv Cardiovasc Phys* 5 (1983), pp. 16–52.

[53] Wikipedia contributors. *Arterial line*. https://en.wikipedia.org/w/index.php?title=Arterial_line&oldid=1139831322. Accessed: NA-NA-NA. Feb. 2023.

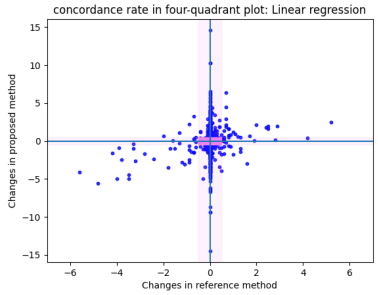
[54] Wikipedia contributors. *Pulmonary artery catheter*. https://en.wikipedia.org/w/index.php?title=Pulmonary_artery_catheter&oldid=1153235339. Accessed: May 2023.

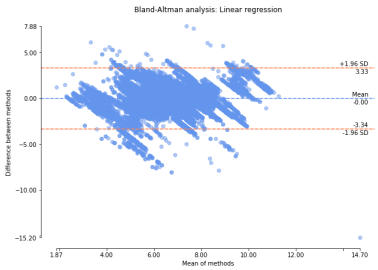
[55] Hyun-Lim Yang et al. "Development and validation of an arterial pressure-based cardiac output algorithm using a convolutional neural network: retrospective study based on prospective registry data". In: *JMIR medical informatics* 9.8 (2021), e24762.

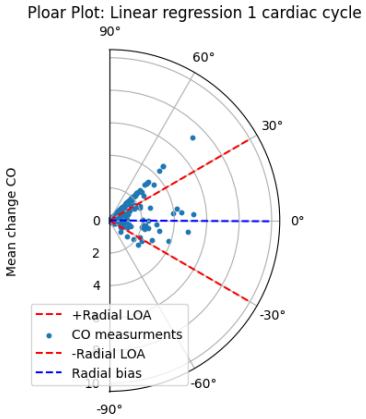
Results

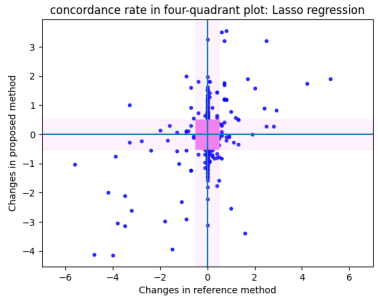
A

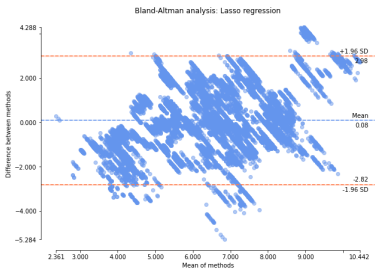
A.1 Results of comparison between regressors

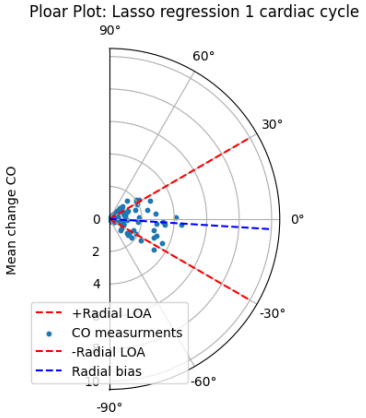


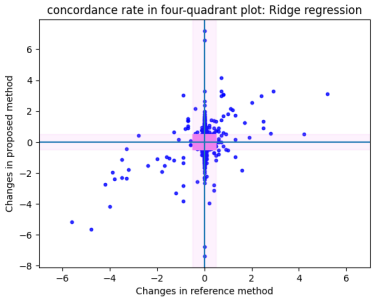


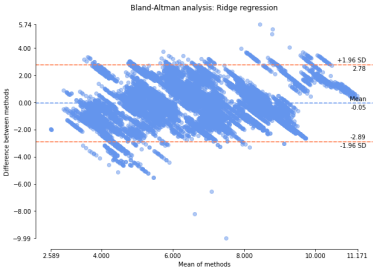


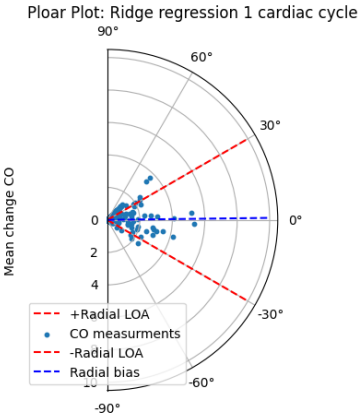


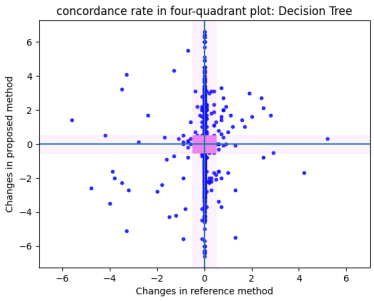




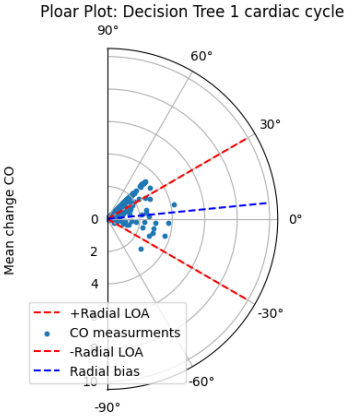

concordance rate in four-quadrant plot: Ridge regression
Changes in proposed method
Changes in reference method

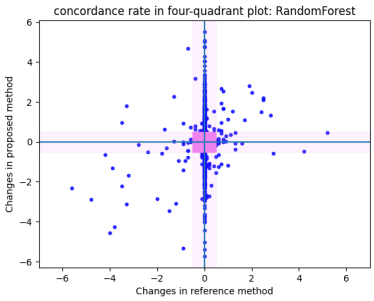

Bland-Altman analysis: Ridge regression
+1.96 SD
2.78
Mean
-0.05
-2.89
-1.96 SD
Difference between methods
Mean of methods

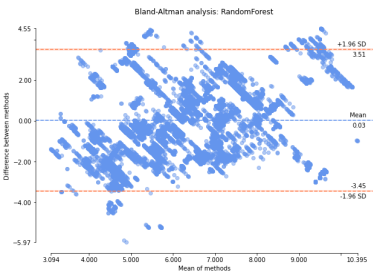

Ploar Plot: Ridge regression 1 cardiac cycle
Mean change CO
+Radial LOA
CO measurments
-Radial LOA
Radial bias

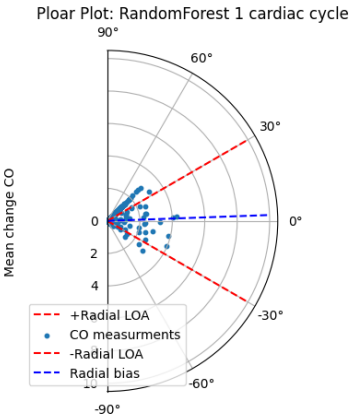

concordance rate in four-quadrant plot: Decision Tree
Changes in proposed method
Changes in reference method

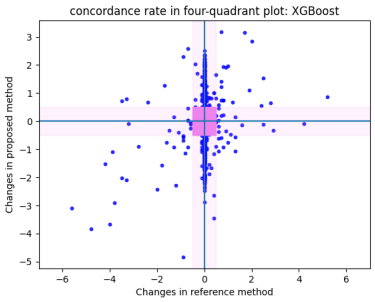

Bland-Altman analysis: Decision Tree
+1.96 SD
4.47
Mean
-0.01
-4.48
-1.96 SD
Difference between methods
Mean of methods

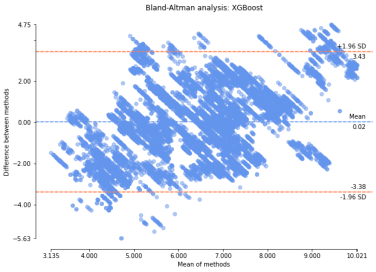

Ploar Plot: Decision Tree 1 cardiac cycle
Mean change CO
+Radial LOA
CO measurments
-Radial LOA
Radial bias

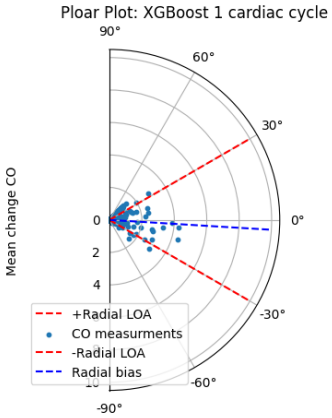

concordance rate in four-quadrant plot: RandomForest
Changes in proposed method
Changes in reference method


Bland-Altman analysis: RandomForest
+1.96 SD
3.51
Mean
0.03
-3.45
-1.96 SD
Difference between methods
Mean of methods


Ploar Plot: RandomForest 1 cardiac cycle
Mean change CO
+Radial LOA
CO measurments
-Radial LOA
Radial bias

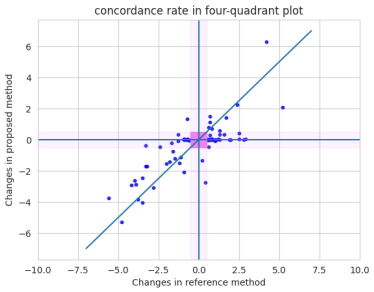


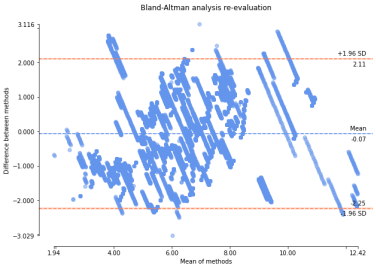


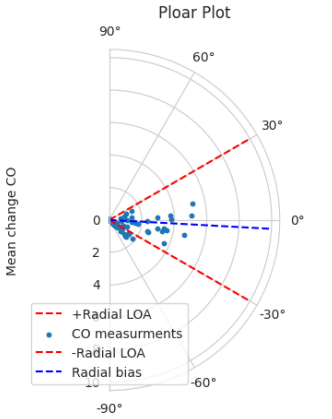


A.2 Results of comparison between regressors (reevaluation)

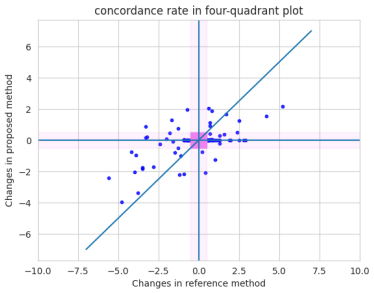
Linear regression

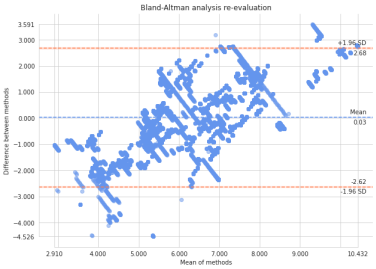


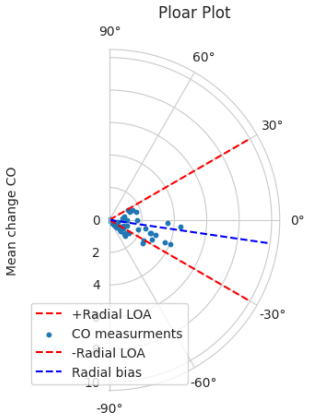




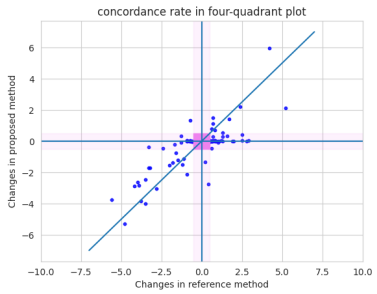
Lasso regression

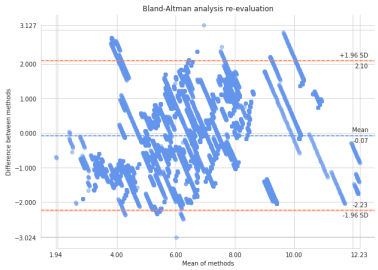


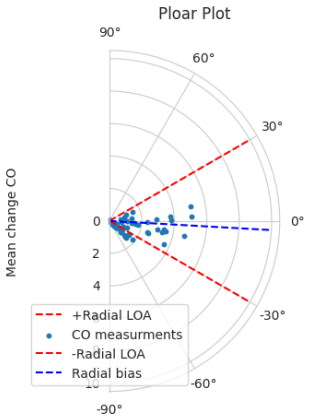




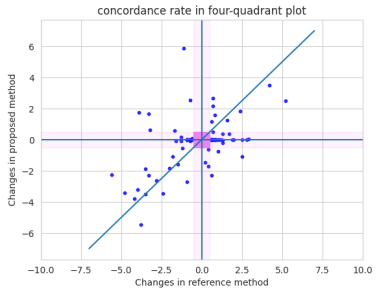
Ridge regression

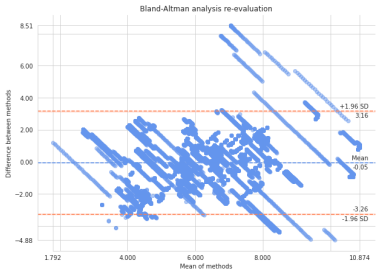


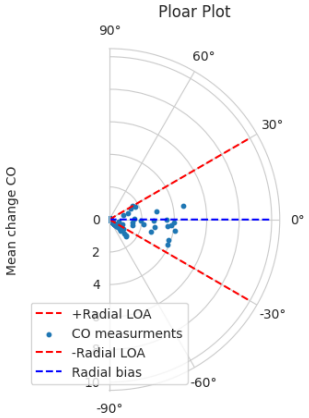




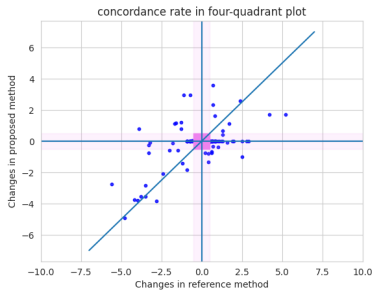
Decision tree

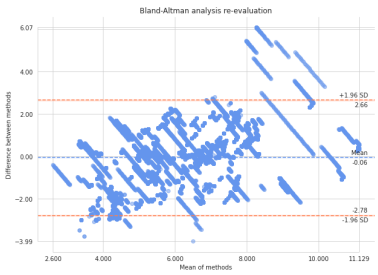


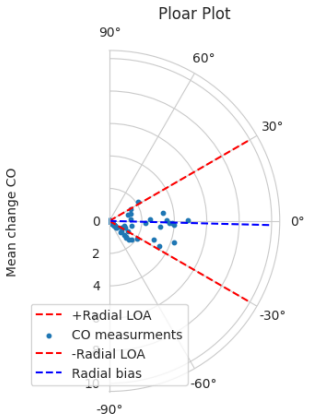




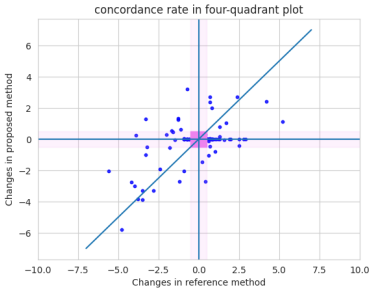
Random forest

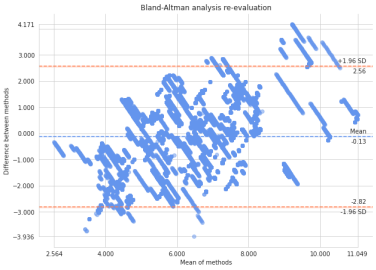


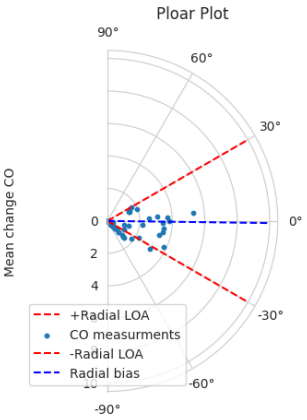




XGBoost







A.3 Filter design three cardiac cycles

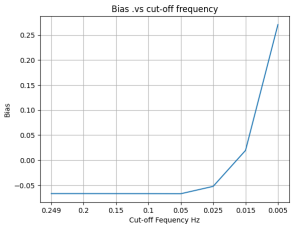


Figure A.1: Cut-off frequency vs. bis

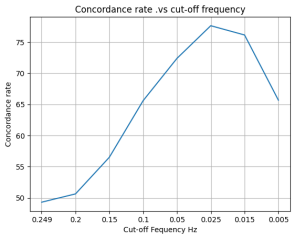


Figure A.2: Cut-off frequency vs. concordance rate

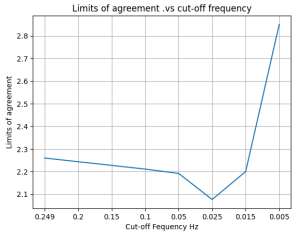


Figure A.3: Cut-off frequency vs. limits of agreement

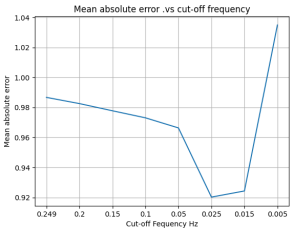


Figure A.4: Cut-off frequency vs. mean absolute error

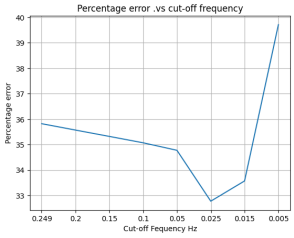


Figure A.5: Cut-off frequency vs. percentage error

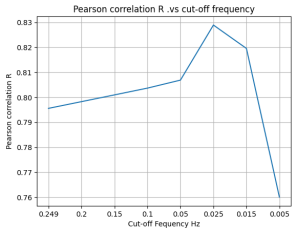


Figure A.6: Cut-off frequency vs. correlation

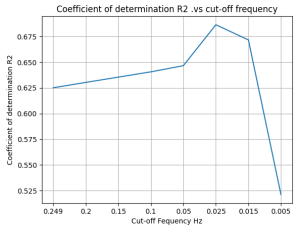


Figure A.7: Cut-off frequency vs. R2

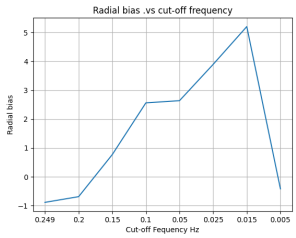


Figure A.8: Cut-off frequency vs. radial bias

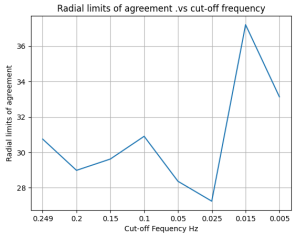


Figure A.9: Cut-off frequency vs. radial limits of agreement

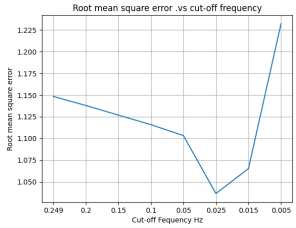


Figure A.10: Cut-off frequency vs. RMS

Table A.1: Regressors comparison

Ḣz/	**RMS**	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE**	**CR**	**RLOA**	**RBias**
0.25	1.148	0.986	0.795	0.625	-0.07	2.26,-2.39	35.82%	49.30%	30.74°	-0.88°
0.2	1.137	0.982	0.798	0.630	-0.07	2.24,-2.38	35.57%	50.63%	28.98°	-0.693°
0.15	1.126	0.977	0.800	0.635	-0.07	2.23,-2.36	35.32%	56.50%	29.62°	0.765°
0.1	1.115	0.973	0.803	0.640	-0.07	2.21,-2.35	35.07%	65.59%	30.90°	2.560°
0.05	1.103	0.966	0.806	0.646	-0.07	2.19,-2.33	34.77%	72.36%	28.36°	2.636°
0.025	1.036	0.920	0.828	0.686	-0.05	2.08,-2.18	32.77%	77.61%	27.23°	3.895°
0.015	1.065	0.924	0.819	0.671	0.02	2.2,-2.16	33.57%	76.11%	37.21°	5.205°
0.005	1.232	1.034	0.760	0.521	0.27	2.85,-2.31	39.71%	65.67%	33.13°	-0.420°

A.4 Filter design four cardiac cycles

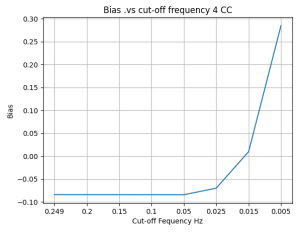


Figure A.11: Cut-off frequency vs. bis

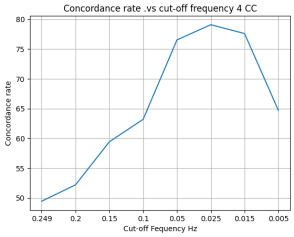


Figure A.12: Cut-off frequency vs. concordance rate

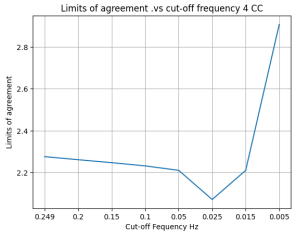


Figure A.13: Cut-off frequency vs. limits of agreement

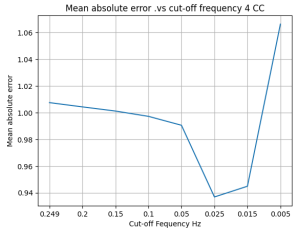


Figure A.14: Cut-off frequency vs. mean absolute error

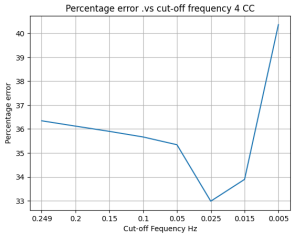


Figure A.15: Cut-off frequency vs. percentage error

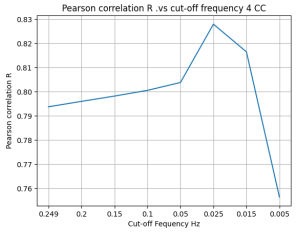


Figure A.16: Cut-off frequency vs. correlation

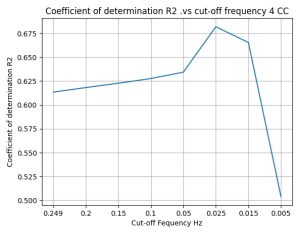


Figure A.17: Cut-off frequency vs. R2

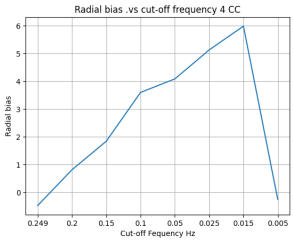


Figure A.18: Cut-off frequency vs. radial bias

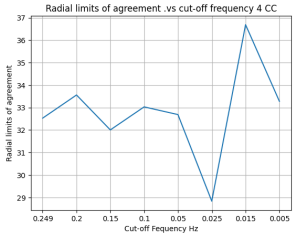


Figure A.19: Cut-off frequency vs. radial limits of agreement

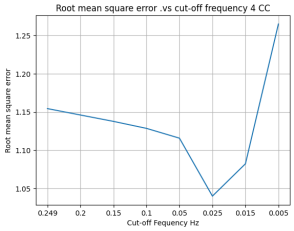


Figure A.20: Cut-off frequency vs. RMS

Table A.2: Regressors comparison

Ḣz/	**RMS**	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE**	**CR**	**RLOA**	**RBias**
0.25	1.15	1.00	0.79	0.61	-0.08	2.28,-2.44	36.34%	49.45%	32.53°	-0.47°
0.2	1.14	1.00	0.79	0.61	-0.08	2.26,-2.43	36.11%	52.20%	33.55°	0.81°
0.15	1.13	1.00	0.79	0.62	-0.08	2.25,-2.42	35.90%	59.42%	32.00°	1.84°
0.1	1.12	0.99	0.80	0.62	-0.08	2.23,-2.40	35.66%	63.21%	33.03°	3.60°
0.05	1.11	0.99	0.80	0.63	-0.08	2.21,-2.38	35.34%	76.54%	32.68°	4.08°
0.025	1.04	0.93	0.82	0.68	-0.07	2.07,-2.21	32.98%	79.10%	28.83°	5.13°
0.015	1.08	0.94	0.81	0.66	0.01	2.21,-2.19	33.89%	77.61%	36.68°	5.98°
0.005	1.26	1.06	0.75	0.50	0.28	2.91,-2.34	40.36%	64.70%	33.28°	-0.25°

A.5 Filter design five cardiac cycles

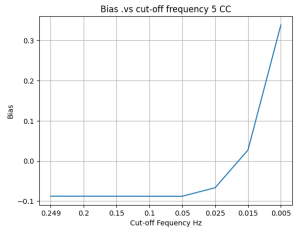


Figure A.21: Cut-off frequency vs. bis

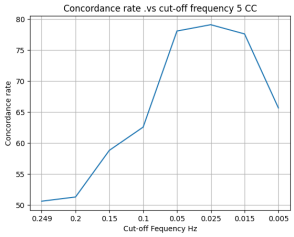


Figure A.22: Cut-off frequency vs. concordance rate

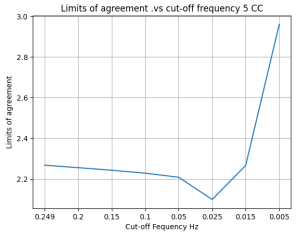


Figure A.23: Cut-off frequency vs. limits of agreement

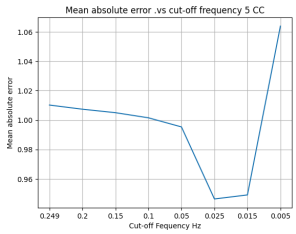


Figure A.24: Cut-off frequency vs. mean absolute error

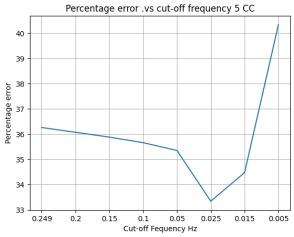


Figure A.25: Cut-off frequency vs. percentage error

Figure A.26: Cut-off frequency vs. correlation

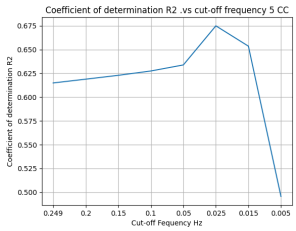


Figure A.27: Cut-off frequency vs. R2

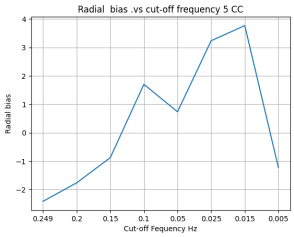


Figure A.28: Cut-off frequency vs. radial bias

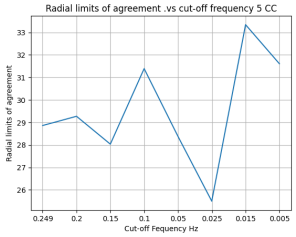


Figure A.29: Cut-off frequency vs. radial limits of agreement

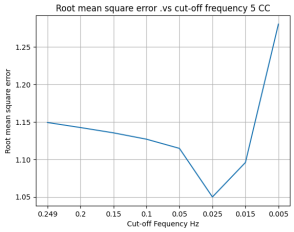


Figure A.30: Cut-off frequency vs. RMS

Table A.3: Regressors comparison

Ḣz/	**RMS**	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE**	**CR**	**RLOA**	**RBias**
0.25	1.14	1.01	0.79	0.61	-0.09	2.27,-2.44	36.26%	50.61%	28.86°	-2.41°
0.2	1.14	1.00	0.79	0.61	-0.09	2.26,-2.43	36.07%	51.29%	29.27°	-1.76°
0.15	1.13	1.00	0.79	0.62	-0.09	2.24,-2.42	35.88%	58.82%	28.03°	-0.88°
0.1	1.12	1.00	0.80	0.62	-0.09	2.23,-2.40	35.66%	62.57%	31.38°	1.70°
0.05	1.11	0.80	0.79	0.63	-0.09	2.21,-2.38	35.35%	78.08%	28.37°	0.73°
0.025	1.04	0.82	0.81	0.67	-0.07	2.10,-2.23	33.34%	79.10%	25.49°	3.23°
0.015	1.09	0.80	0.80	0.65	0.03	2.27,-2.21	34.47%	77.61%	33.34°	3.76°
0.005	1.28	0.75	0.76	0.49	0.34	2.96,-2.28	40.33%	65.67%	31.60°	-1.22°

A.6 Filter design six cardiac cycles

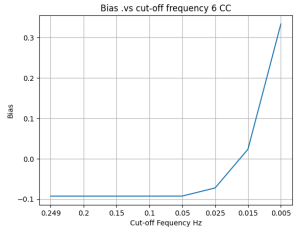


Figure A.31: Cut-off frequency vs. bis

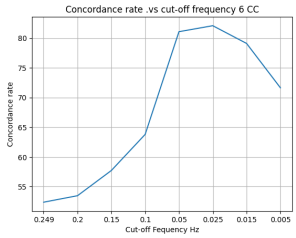


Figure A.32: Cut-off frequency vs. concordance rate

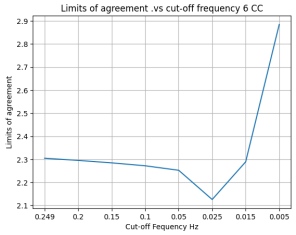


Figure A.33: Cut-off frequency vs. limits of agreement

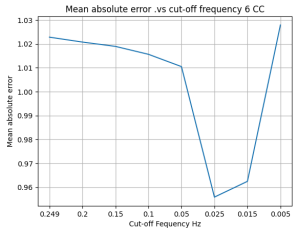


Figure A.34: Cut-off frequency vs. mean absolute error

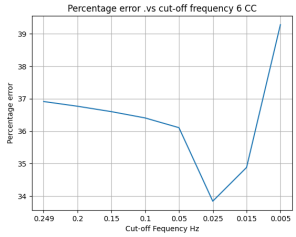


Figure A.35: Cut-off frequency vs. percentage error

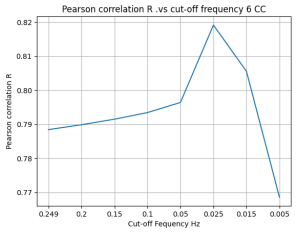


Figure A.36: Cut-off frequency vs. correlation

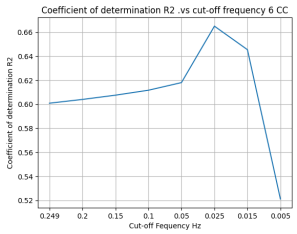


Figure A.37: Cut-off frequency vs. R2

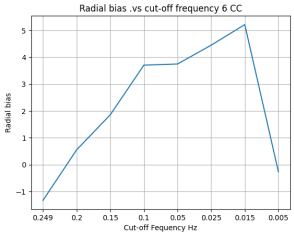


Figure A.38: Cut-off frequency vs. radial bias

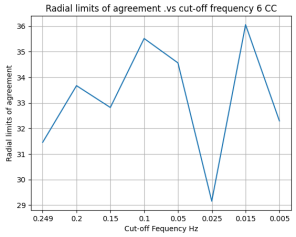


Figure A.39: Cut-off frequency vs. radial limits of agreement

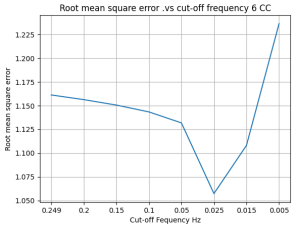


Figure A.40: Cut-off frequency vs. RMS

Table A.4: Regressors comparison

Ḣz/	**RMS**	**MAE**	**R**	**R2**	**Bias**	**LOA**	**PE**	**CR**	**RLOA**	**RBias**
0.25	1.16	1.02	0.78	0.60	-0.09	2.30,-2.49	36.91%	52.38%	31.45°	-1.33°
0.2	1.15	1.02	0.78	0.60	-0.09	2.30,-2.48	36.76%	53.47%	33.66°	0.54°
0.15	1.15	1.01	0.79	0.60	-0.09	2.28,-2.47	36.59%	57.71%	32.81°	1.85°
0.1	1.14	1.01	0.79	0.61	-0.09	2.27,-2.46	36.40%	63.79%	35.51°	3.70°
0.05	1.13	1.01	0.79	0.61	-0.09	2.25,-2.44	36.10%	81.08%	34.55°	3.74°
0.025	1.05	0.95	0.81	0.66	-0.07	2.13,-2.27	33.84%	82.08%	29.15°	4.44°
0.015	1.10	0.96	0.80	0.64	0.02	2.29,-2.24	34.88%	79.10%	36.05°	5.21°
0.005	1.23	1.02	0.76	0.52	0.33	2.88,-2.22	39.27%	71.64%	32.29°	-0.26°

A.7 Features weights

/

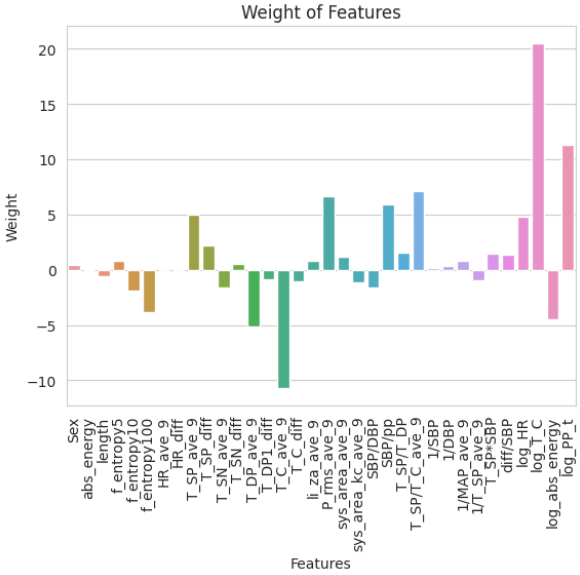

Weight of Features
Weight
20
15
10
5
0
−5
−10
Sex
abs_energy
length
f_entropy5
f_entropy10
f_entropy100
HR_ave_9
HR_diff
T_SP_ave_9
T_SP_diff
T_SN_ave_9
T_SN_diff
T_DP_ave_9
T_DP1_diff
T_C_ave_9
T_C_diff
li_za_ave_9
P_rms_ave_9
sys_area_ave_9
sys_area_kc_ave_9
SBP/DBP
SBP/pp
T_SP/T_DP
T_SP/T_C_ave_9
1/SBP
1/DBP
1/MAP_ave_9
1/T_SP_ave_9
T_SP*SBP
diff/SBP
log_HR
log_T_C
log_abs_energy
log_PP_t
Features

Table A.5: Summary of extracted hemodynamic features

No	Feature	weight
1	*Gender*	4.16824346e-01
2	*Absolute energy*	3.76669128e-07
3	*Fourier entropy*5	-5.64677346e-01
4	*Fourier entropy*10	8.14264005e-01
5	*Fourier entropy*100	-1.92112057e+00
6	*Heart rate*	-3.79286509e+00
7	$Diff$ *heart rate*	5.80360575e-03
8	*Systolic duration*	4.98815153e+00
9	$Diff$ *Systolic duration*	2.21602988e+00
10	*Systolic notch duration*	-1.62738187e+00
11	$Diff$ *Systolic notch duration*	5.51504244e-01
12	*Distolic duration*	-5.17908499e+00
13	$Diff$ *Distolic duration*	-8.50643685e-01
14	*Cycle duration*	-1.07255053e+01
15	$Diff$ *Cycle duration*	-1.04201849e+00
16	$Liljestrand - Zender$	7.85049739e-01
17	$RMS_{pressure}$	6.66613057e+00
18	*Systolic area*	1.13170352e+00
19	*Systolic area correction*	-1.12962825e+00
20	*Systolicpressure/Distolicpressure*	-1.58442166e+00
21	*Systolic pressure/pulse pressure*	5.94630432e+00
22	*Systolic duration/Distolic duration*	1.52133793e+00
23	*Systolic duration/Cycle duration*	7.07841294e+00
24	1 / *Systolic pressure*	1.16715166e-01
25	1 / *Distolic pressure*	3.71947205e-01
26	1 / *Mean pressure*	8.39398828e-01
27	1 / *Systolic duration*	-9.93281845e-01
28	*Systolic duration*pressure*	1.47326325e+00
29	$Diff$ *Systolic pressure/Systolic pressure*	1.39106183e+00

Extra project literature review

B

The use of arterial blood pressure to monitor cardiac output was proposed more than a century ago. Researchers developed models to represent the arterial tree such as the Windkessel model which was assumed well representative of the arterial tree by two parameters. According to this model, the arterial blood pressure should decay like a pure exponential [35]. However, the contour of the arterial pulse changes significantly in different locations of the arterial tree [38]. To overcome the site dependence issue, researchers added an improvement to the model such as in [35], and [1] where they made the model frequency dependent. They estimated the impulse response of the measured arterial blood pressure waveform and took the average time constant of the fitted exponential functions over 6 minutes intervals by computing the arithmetic mean and median in [35], and 17-beat moving window size in [1]. The overall RMSNE was 22.8%. Such an approach provided more accurate results than previous studies, is less site dependent, and took into account possible distortion involved in the arterial blood pressure waveform. However, the cardiac output error depends on the heart rate. And many sources of error were assumed to be negligible such as that the frequency-dependent parameter may not be constant (non-linear parameter) over the pressure range considered in the study which may lead to underestimating or overestimating the cardiac output. Also, the viscous component of the arterial wall was assumed to be negligible which may potentially affect the Windkessl time constant. The contribution of peripheral venous pressure (PVP) to the pressure gradient was also assumed to be negligible. Under most circumstances this could be negligible, however, under extreme hypertensive conditions, PVP should not be neglected because this will overestimate the cardiac output measurement. It should also be noticed that this study was conducted on animals and human testing is still needed.

In the mathematical analysis approach, only one feature was used to analyze the waveform (the decay rate of pure exponential), and non-linear patterns were challenging. Data-driven models or learning techniques can provide powerful tools for solving such problems. One of the first attempts to use learning techniques for cardiac output estimation was to use a machine learning-based conjunctive rule to calculate the average weights of two models [29]. This study showed that arterial blood pressure waveform can be used to reliably estimate blood pressure and heart rate and that such an approach is integrable into decision support systems. However, this study contained only 10 records, the techniques were tested on animals, and heart rate and blood pressure were derived from arterial blood pressure rather than vital signs monitors. In another study, [47], the area under the systolic portion of the arterial pressure waveform from the end-diastole to the end of the ejection phase was used to estimate the stroke volume. This is an implementation of the theory that suggests that stroke

volume is related to blood pressure in an artery and vascular resistance. The system was evaluated using 7 signals of arterial pressure and the corresponding COs. The CO values were estimated using thermodilution measurements. The system has an average error of 6.5% in estimating CO. In this study, only the stroke volume was estimated. However, CO is also related to vascular resistance that varies from patient to patient. Therefore, an individual calibration was needed for each patient. This parameter was estimated from the first few seconds of the wave recordings. Recently, in [55], deep learning and transfer learning techniques were used to develop an algorithm that mimicked the commercial APCO algorithm [55]. The model was built using CNN to extract features from the data and a regressor to predict the stroke volume values. This study showed more accurate results than a commercial APCO device for estimating cardiac output. However, more data is needed for further development and implementation since the data used in this study is from a single-center registry, therefore, there may be a limited range of CO. In addition, the ground-truth CO used in this data may be less accurate in certain situations compared with the thermodilution technique. The data collected in this study were from two different devices and the delay was not fully revealed. Also, there was a statistical difference in age between the training and testing sets, and finally, there was no visualization with explainable artificial intelligence that explains why the proposed model produce the results. The raw data set of this study was disclosed for any other improvement from other researchers to achieve the global optimum model.

Other studies used learning techniques to predict cardiac output-related values in the future such as blood pressure [18] or hypo-tensions[22], [12]. In [12], logistic regression was used to predict hypo-tension 15 minutes before its occurrence with a sensitivity of 88% and specificity of 87%. In this study, even though it predicts the hypo-tension event before its actual occurrence it is not applicable in real-time predictions because it was assumed that the hypotension event does not occur again within 20 minutes. In another study, recurrent neural nets RNN were used to predict blood pressure 3 minutes ahead. The mean blood pressure and the systolic blood pressure errors were 9mmHg and 11 mmHg respectively. The standard error for automatic arterial pressure monitoring should not be greater than 5 mmHg and the standard deviation not greater than 8 mmHg as was defined by the association for the advancement of medical instrumentation [44]. It is clear that the study did not meet the requirements of the acceptable error, however, data used in this study were obtained from only 102 operations which is not enough to capture diverse patterns of future blood pressure. Therefore, having more data may improve the results. Another minor limitation of this study is the first prediction. To get a prediction, data must be collected and this takes time. Therefore, data collection should start before the induction of anesthesia. Supervised machine learning techniques were also used to predict post-induction hypo-tension events [22]. In this study, they tested multiple techniques from which the gradient boosting machine performs the best 0.76. 10000 cases were incorporated into machine learning in this study, however, all cases were extracted over six months in a single institution. A data set over a longer period may improve the results. Also, the lack of transparency was a limitation in this study where the algorithm is seen as a black box. This may

be inadequate if a complete understanding of the clinical implication is needed for any practice modification.

From this literature review, approaches can be categorised into two. The first one being predicting blood pressure or hypo-tensions in the future. This approach can support decision making and is important since it provides the ability for early interventions. However, practical limitations are associated with this approach such as that non-practical assumptions should be made or results have bigger error than the standard acceptable error. Having more data may solve the problem and reduce the error, however, after predicting the blood pressure in the future, cardiac output still needs to be estimated from this predictions. This may reduce the accuracy more since the algorithm for CO estimation will also have some estimation errors. In addition, future hear rate should also be estimated. In other words, the ground-truth for CO estimation in the future may be unreliable. Therefore, predicting CO in the future is not feasible in one project since many steps should be taken: predicting blood pressure and heart rate in the future, estimating stroke volume from predictions, and calculating the CO. The second approach in this literature review assumes that blood pressure and hear rate are available which is the case in current time. This approach is more feasible since only one algorithm should be developed and is needed in practice since the current methods are either invasive and complex or non-invasive and inaccurate.

In [55], the second approach was used were learning techniques were used (deep learning and transfer learning) and they achieved more accurate results than the commercially available APCO devices. However, two major limitations were associated with this study. The first one being the small amount of data. Having more data could improve the results even more such that the algorithm achieves the accuracy of the PAC method (the gold standard). The second limitation was that not all arterial blood pressure wave-forms had corresponding CO estimation, therefore, they needed to be estimated rather than getting the values from CO monitor. The model in the mentioned study is promising and performs the best ,and therefore, it is suitable to be chosen. However, to add improvement on the study, the associated limitations should be tackled.

Regarding the first limitation (small data-set), the Erasmus MC hospital will provide access to their data registry. This data will be added to the data from the study conducted in [55]. Regarding the second limitation, in this paper, an algorithm will be implemented to give rough CO estimation for wave-forms without reference CO label through mathematical analysis of the wave-forms as it was done in [47]. After that, the model from [55] can be trained and tuned using transfer learning and data labeled from PAC monitor. Having done these steps, it is expected to achieve more accuracy and wider range of CO values will be achieved.

B.1 Design requirements

The most important aspect of designing the requirements is the reference ground truth that the model is trained with. Although, PAC is considered as the gold standard and generally accepted as the clinical standard method, there is no consensus that such a technique is the appropriate comparator when testing less/non-invasive devices. The frequently used method to compare two methods is correlation analysis. It should be noted that correlation analysis does not measure agreement between two methods, but their relationship. Statistical methods to assess a system are accuracy, precision, and trending ability [42].

Accuracy is defined as the bias between the results from this study and the reference data (ground truth). Precision is defined as the standard deviation (SD) of the bias. Limits of agreement (LoA) defined as [bias +/- 1.96*SD]. Mean percentage error (MPE) is derived from the SD and mean CO [41]. The bland-Altman analysis is the accepted statistical approach for the evaluation of agreement (accuracy and precision). When using Bland-Altman analysis, the mean difference (bias), and the limits of the agreement reflect accuracy and precision. There are no generally accepted cut-off values for the acceptable mean difference and the corresponding standard deviation assessed by the Bland-Altman analysis. For the acceptable agreement 30% is usually applied. However, this can only be applied when the ground truth and the method to be tested have a precision of +/- 20 % [41], [42].

In addition, the percentage error can be calculated as 2 times the standard deviation of the mean difference divided by the mean of the measurements. Also, the ability to accurately follow changes in CO is crucially important, 4-quadrant plot can be used to assess this. Usually, small changes of CO are excluded from the 4-quadrant plot by using an exclusion zone (10% or 0.5 $\frac{L}{min}$) [42].

To provide improvement to the study conducted in [55], their achieved results will be used as requirements in this study:

- The error difference between the estimated value and the ground truth should be less than -0.5 mL with a standard deviation of 2.1 mL.
- The absolute error between the estimated value and the ground truth should be less than 1.8 mL and a standard deviation of 0.3 mL.
- The percentage error between the estimated value and the ground truth should be less than 1.2 mL and a standard deviation of 1.7 mL.
- The absolute percentage error between the estimated value and the ground truth should be less than 2.2 mL and a standard deviation of 0.2.

An additional requirement for the ML system is the estimate update time. Arterial blood pressure waveform is sampled at 100 Hz frequency over a period of 20 s. Therefore, the CO values should be estimated and updated at 20 s intervals [30].

Table B.1: Summary of design requirements

ID	**Requirement**	**Verification method**
DA.01	The error should be less than -0.5 mL with standard deviation of 2.1 mL.	Test
DA.02	The absolute error difference should be less than 1.8 mL and standard deviation of 0.3 mL.	Test
DA.03	The percentage error difference should be less than 1.2 mL and standard deviation of 1.7 mL.	Test
DA.04	The absolute percentage error difference should be less than 2.2 mL and standard deviation of 0.2.	Test
DA.05	Limits of agreement should be 30 %	Test
DA.06	Exclusion zone in 4-quadrant plot is (10%)	Test
DA.07	The estimate and update time of the CO measurement should be done at 20 s intervals	Test

C Dataset

C.1 Patient's IDs

4034.0, 349.0, 4959.0, 251.0, 2252.0, 4721.0, 5693.0, 5698.0, 5018.0, 2872.0, 5682.0, 2252.0,5607.0, 5442.0, 3188.0,2185.0, 6227.0, 4800.0, 1083.0, 4573.0, 2700.0, 3719.0, 553.0, 5837.0, 3113.0,6297.0, 1900.0, 1959.0, 3967.0, 3849.0, 2453.0, 985.0, 4398.0,4283.0, 5687.0, 945.0, 1730.0, 397.0, 2326.0,4897.0, 1785.0, 783.0, 1018

C.2 surgical operations

Among these cases, the following surgical operations were performed: 1 Cesarean section, 2 Total gastrectomies, 3 Pylorus preserving pancreaticoduodenectomies, 1 Radical prostatectomy, 2 Excisions, 1 Simple mastectomy, 2 Lung wedge resections, 3 Liver transplantations, 2 Cholecystectomies, 4 Breast-conserving surgeries, 1 Radical hysterectomy, 1 Ovarian cystectomy, 1 Kidney transplantation, 1 Thyroglossal duct cyst excision, 1 Total thyroidectomy, 1 Hysterectomy, 1 Closure of wound, 2 Metastasectomies, 2 Anterior resections, 1 Bleeding control, 3 Distal gastrectomies, 1 Extended hemihepatectomy, 3 Hemicolectomies, 2 Exploratory laparotomies, 1 Hemihepatectomy, 1 Thyroid lobectomy, 1 Lung segmentectomy, 1 Spleen preserving distal pancreatectomy, and 1 Proximal gastrectomy.

C.3 MIMIC patients ID

The following patients were deleted because of the low waveform quality:
deleted-patient = [3613920, 3814000, 3299549, 3162867, 3689670, 3654372, 3402722, 3695934, 3911888, 3582566, 3284836, 3632160]

The following IDs are the patient's IDs from the MIMIC dataset that was used to train and test:
patient-ID-MIMIC = [3544749, 3689670, 3027112, 3701096, 3162867, 3697467,3767318, 3607858, 3721513,3220385, 3715271, 3790349, 3171184, 3964904, 3255734, 3609675, 3980460, 3548811, 3000126, 3112718, 3600781, 3295369, 3574553, 3458812, 3455496, 3294756, 3299549, 3475959, 3280565, 3628710, 3475592, 3123796, 3676005, 3814000, 3574889, 3613920, 3828483, 3882370, 3161169, 3126479, 3950676, 3079821, 3041277, 3523851, 3722858, 3157770, 3347680, 3548075, 3215863, 3555576, 3145459, 3903980, 3667327, 3367125, 3425244, 3188562, 3326633, 3000105, 3222216, 3573321, 3617554, 3599408, 3345464, 3871964, 3344874, 3072283, 3386452, 3989828, 3599099, 3546834, 3239573, 3057224, 3922725, 3447088, 3676483, 3551695, 3265081, 3046261, 3956395,

3632160, 3284836, 3318909, 3127390, 3686173, 3043670, 3603259, 3551259, 3769651, 3021157, 3402722, 3654372, 3695934, 3289691, 3911888, 3582566, 3069909, 3803050, 3760748, 3182265, 3926530, 3344926, 3321106, 3702369, 3955830]

C.4 Data for waveform quality assessment

3544749: 2, 4, 6, 11
3027112: 26, 30, 31
3701096: 34, 36, 41, 42
3697467: 60, 61, 63
3607858: 77, 80, 82, 84, 86, 88, 90, 91
3790349: 107, 108, 109, 110
3964904: 119
3609675: 125,126,127,128,129,130,131,132,135,136,137,138,139,140
3548811: 146,147
3000126: 149,154,155
3295369: 175,176,177,178,180,182
3458812: 198,201
3455496: 208,209,210,213,215
3299549: 227,232,236,237
3280565: 247,250
3628710: 252,264,265,266
3123796: 307,316
3676005: 320,321,324,325,326,329
3613920: 342,347
3828483: 351,352,353
3161169: 364,365,368,370,371,373,374,375,377,378
3126479: 379,380,381,384,385
3950676: 389,390
3079821: 396,397,398,399,400
3041277: 404,411
3523851: 414,415,419,420,422,423,424
3157770: 433,436,437,439
3347680: 440,442
3215863: 450,451,452
3903980: 479,482,483,484,487,489
3667327: 495,496,497,498,499,500,501,509
3367125: 513,516,518,521,524,525,526,527,528,529,530,536,539
3326633: 556,561,563,564,565,566
3222216: 569,570,575
3573321: 578,579,580,583,587,588,589
3617554: 592,593
3599408: 600,601,602,603
3871964: 608,609,610,611,612,613,614,615,617,618,620,621
3072283: 627,628,629,630,631,632,633,634,635,637

3386452: 640,641,643,644,645,646,647,652,654,656,666,668,672,674,661
3989828: 675,676,677,678,679,680,681,682,683,684,685,686,687,688,694,700,701,703
3599099: 704,705
3546834: 711,712,715
3057224: 719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,739
3922725: 742,743,744,745,746,747,748,749,752,753,756,759
3676483: 772,773,775,777,778,781,782,783
3265081: 788,789,790,791
3046261: 793
3956395: 798,799
3632160: 804,805,808
3284836: 812,813,814,819
3127390: 829,830,832,834,836,837,838,839,840
3043670: 848,849,851,852
3021157: 873,874
3402722: 880,881,882,883,884,885,886
3654372: 897,898,899,900,902,903,908,910,911,912,913,915,916
3695934: 922,923,928,929
3289691: 934
3911888: 940,944
3069909: 952,954,955,957
3803050: 959,960,964,965,968,969,970,971,973,974,975,976
3760748: 980,983,989,994,1000
3182265: 1010
3926530: 1013,1015,1017,1020
3344926: 1022,1024,1025,1026,1028,1029,1032
3702369: 1045,1056,1059
delete full: 3689670, 3162867, 3767318, 3220385, 3171184, 3255734, 3980460, 3112718, 3600781, 3574553, 3294756, 3475959, 3475592, 3574889, 3882370, 3188562, 3000105,3239573,3686173,3603259,3769651,3402722,3582566, 3321106, 3955830