

Machine Learning Algorithm to Estimate Cardiac Output Based On Less-Invasive Arterial Blood Pressure Measurements

Hamo, Alan; Ottenhof, Niki; Korstanje, Jan Wiebe H.; Dauwels, Justin

DOI

[10.1109/EMBC53108.2024.10781760](https://doi.org/10.1109/EMBC53108.2024.10781760)

Publication date

2024

Document Version

Final published version

Published in

46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2024 - Proceedings

Citation (APA)

Hamo, A., Ottenhof, N., Korstanje, J. W. H., & Dauwels, J. (2024). Machine Learning Algorithm to Estimate Cardiac Output Based On Less-Invasive Arterial Blood Pressure Measurements. In *46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2024 - Proceedings* (Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS). IEEE. <https://doi.org/10.1109/EMBC53108.2024.10781760>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

MACHINE LEARNING ALGORITHM TO ESTIMATE CARDIAC OUTPUT BASED ON LESS-INVASIVE ARTERIAL BLOOD PRESSURE MEASUREMENTS

Alan Hamo*

Niki Ottenhof†

Jan-Wiebe H Korstanje†

Justin Dauwels*

* Department of Microelectronics, Delft University of Technology, Delft, the Netherlands.

†Department of Anaesthesia, Erasmus University Medical Center, Rotterdam, the Netherlands.

ABSTRACT

Cardiac output (CO) is a vital hemodynamic parameter that reflects the blood volume pumped by the heart per minute. A less-invasive way to estimate CO is by analyzing arterial blood pressure (ABP) waveforms. However, the relationship between CO and blood pressure is unknown. This study uses machine learning and feature engineering techniques to discover the relationship between CO and ABP. We apply the sparse identification non-linear dynamics (SINDy) algorithm to discover features. Additionally, we investigate the optimum number of cardiac cycles required for feature extraction to achieve the best performance. The proposed approach achieves clinically acceptable performance regarding radial limits of agreement (RLOA) and bias (RBias). Further, the proposed approach is validated on an external dataset. Finally, similarities to the Navier-Stokes equations are presented.

Index Terms— Cardiac output, feature discovery, machine learning, Arterial blood pressure, SINDy algorithm

1. INTRODUCTION

Cardiac output (CO) is crucial for determining tissue oxygen delivery and indicates the heart's ability to meet the body's demands. Researchers have developed various methods to measure CO, including the invasive thermodilution technique, also known as the Swan-Ganz catheter [1]. Although the Swan-Ganz catheter technique is considered the gold standard for CO measurements, it only provides an approximation since measuring the actual CO is extremely hard in clinical practices. Alternatively, a less-invasive way to estimate CO is by analyzing the arterial blood pressure (ABP) waveform. Many models were developed over the years trying to relate CO with ABP via less invasive procedures. These models can be categorized into models derived from physical principles called hereafter classical models [2, 3, 4] and models obtained from data-driven techniques such as convolutional neural networks [5] and machine learning [6]. This study exploits machine learning techniques and feature engineering to develop an explainable algorithm for estimating CO from the ABP waveform. To achieve this goal, 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. We apply the SINDy [7] algorithm

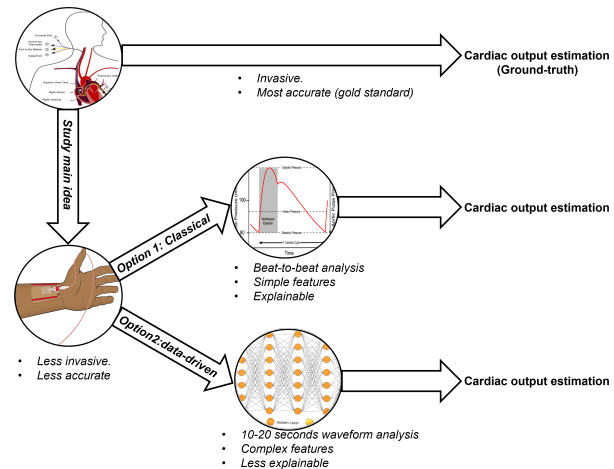


Fig. 1: Summary of the idea behind the project.

to discover a novel set of features and use these features as input to the algorithm. Furthermore, we examine the optimal duration of the ABP waveform used for feature extraction. Figure 1 summarizes the main ideas behind this study. The rest of the paper is organized as follows: the next section explains the methods, followed by the results and conclusion in section 3 and 4, respectively.

2. METHODS

The methodology pipeline starts with a signal abnormality index to examine the quality of the waveform. After that, we extract hemodynamic and waveform features for the next step (experiment 1: feature discovery). In experiment 2, we investigate the optimum number of cardiac cycles. We calculate clinically accepted metrics to evaluate the accuracy and precision using the gold standard for CO measurements as a reference. We apply leave-one-patient-out cross-validation to validate the model. Additionally, we investigate the performance of six machine learning models, including linear (linear, ridge, and lasso) and tree-based (decision-tree, XG-Boost, and random-forest) models [8]. Figure 2 shows the methodology pipeline.

2.1. Feature extraction

We extract hemodynamic and waveform features alongside demographic information. Figure 3 summarizes the extracted hemodynamic and waveform features.

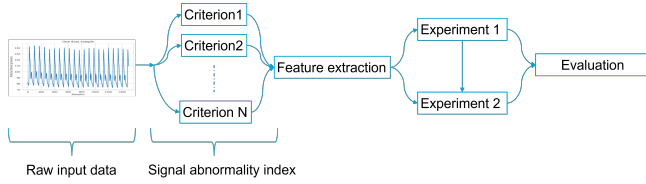


Fig. 2: Methodology pipeline.

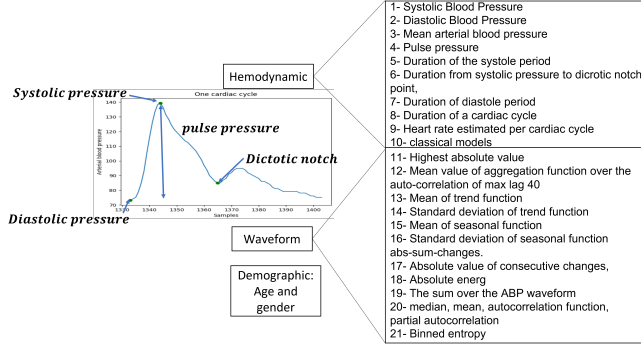


Fig. 3: Summary of the extracted features (Hemodynamic, waveform, and demographic).

Age and gender features were used as demographic information.

2.2. The SINDy algorithm

The SINDy [7] algorithm operates by constructing a library of potential candidate functions. Sparse regression techniques are applied to identify the active terms within the library matrix that characterize the system's dynamics. To this end, the SINDy algorithm solves the following optimization problem:

$$\arg \min_{\beta} \left\| \frac{dx}{dt} - \phi(x)\beta \right\|_2 + \gamma \|\beta\|_1. \quad (1)$$

In this equation, β represents the coefficient vector obtained from solving the constrained regression problem. The term $\frac{dx}{dt}$ denotes the derivative of the features with respect to time, capturing the system's dynamics. The matrix $\phi(x)$ represents the library of possible functions applied to the feature set. Finally, γ is the hyper-parameter. We apply the sequentially thresholded least squares algorithm (STLSQ) [9] method to solve the optimization problem with 0.005 as a threshold. We adopt a recursive approach to address the curse of dimensionality problem. Each iteration focuses on three features, and CO is always one of them. The augmentation of the features (the proposed library function) is based on the classical models. The procedure of feature discovery is described in Algorithm 1.

2.3. The number of cardiac cycles

We investigate the optimal time window or the number of cardiac cycles required for feature extraction to estimate CO accurately, explicitly addressing how far back in time we need to go to estimate CO accurately. To estimate cardiac

Algorithm 1 Feature discovery

```

 $\phi(x) = [\phi(x)_{\text{Logarithmic}}, \phi(x)_{\text{exponential}}, \phi(x)_{\text{polynomial}}]$ 
for #patients > patient iterator do
  for #feature index > feature iterator do
    for #feature index - feature iterator > feature iterator2 do
       $x = [CO, f_{\text{feature iterator}}, f_{\text{feature iterator2}}]$ 
      1)  $\arg \min_{\beta} \left\| \frac{dx}{dt} - \phi(x)\beta \right\| + \gamma \|\beta\|$ 
      2) store the names of non-zero  $\phi(x)$ 
      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  $\phi(x)$ 

```

output, we consider different numbers of cardiac cycles, ranging from one to eight.

2.4. Clinically acceptable algorithm

The precision level of the reference method is $\pm 20\%$ [10]. 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, Bland-Altman analysis and Trending ability analyses are used [10]. Metrics such as Radial limits of agreements (RLOA), Radial bias (RBias), Concordance rate (CR), and Percentage error (PE) should be evaluated. Table 1 summarizes the requirements of these analyses such that the developed model is clinically accepted [10, 11, 12].

Table 1: Summary of the clinically accepted algorithm requirements.

PE [10]	CR [11]	RBias [12]	RLOA [12]
< 30%	> 90–95%	< 5°	< 30°

3. RESULTS

3.1. Data preparation

We obtain the data from the Vital Data Bank [13]. 47 patients were selected from the Vital Recorder data based on the availability of the ABP waveforms and the corresponding Swan-Ganz catheter CO measurements. The waveform was sampled at a rate of 500 Hz. For further processing, this was down-sampled to 100 Hz. CO measurements appeared at 2-second intervals. 350 samples were extracted from each patient. 18900 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. CO values range from 1.6 to 11.8 with a mean of 6.45 and a standard deviation of $2.01 \frac{L}{min}$. 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.

3.2. The discovered features

Following the procedure described in Algorithm 1, 33 features were discovered and summarized in Table 2. The features are categorized into six types: hemodynamic, derivatives (representing changes between cardiac cycles), logarithmic, non-linear, waveform, and demographic features. 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 velocity are indicated with (**u**) in Table 2. The features that capture pressure-related characteristics are denoted with (**p**). Features that describe the power and energy are denoted by (**w**) and (**e**). Lastly, features that describe the randomness are denoted by (**r**).

Table 2: Summary of the discovered features (hemodynamics, waveform, and demographic).

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

3.3. The optimum number of cardiac cycles

Experimenting with different cardiac cycles 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. 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 CO 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 CO measurement than the one before. Ridge regression was the best-performing model among other linear and tree-based models. The results of different cardiac cycles can be seen in Table 3, and the comparison between models can be seen in Table 4.

Table 3: Ridge regression performance for different numbers of cardiac cycles (CC).

	RMS	MAE	R	R2	Bias	LOA	PE%	CR%	RLOA°	RBias°
2CC	1.13	0.96	0.80	0.63	-0.06	2.22,-2.35	35.38	52.30	28.54	-0.88
3CC	1.03	0.92	0.82	0.68	-0.05	2.08,-2.18	32.77	77.61	27.23	1.00
4CC	1.04	0.93	0.82	0.68	-0.07	2.07,-2.21	32.98	79.10	28.83	5.13
5CC	1.04	0.94	0.81	0.67	-0.07	2.10,-2.23	33.34	79.10	25.49	3.23
6CC	1.05	0.95	0.81	0.66	-0.07	2.12,-2.27	33.84	82.08	29.15	4.44

Table 4: Comparison of the performance of six regression models using discovered features and five cardiac cycles.

	RMS (L/min)	MAE (L/min)	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
RandForrest	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

3.4. Feature contribution

The contribution of four feature sets, namely hemodynamic (H), waveform (W), demographic, and SINDy features, is summarized in Table 5. It can be seen from Table 5 that the discovered features using the SINDy algorithm have a significant impact as the results were improved.

Table 5: Feature contribution analysis.

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

3.5. Model evaluation using MIMIC

To validate the model on an external dataset, we obtain data from MIMIC-II version 2 [14] dataset. ABP waveforms, corresponding CO measurements, and the gender of the patients are extracted. The ABP was sampled at 125 Hz in the MIMIC dataset, and the CO measurements appeared at irregular intervals. 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, and features were standardized using the mean and variance obtained from the Vitaldb. The validation test results are summarized in Table 6. The model achieved good results regarding MSE, MAE, and RLOA. The model achieved a clinically acceptable level regarding radial bias. This means that the results are accurate and precise when the model can track the changes in the reference values.

Table 6: Validation results: the model was trained on Vitaldb and tested on MIMIC dataset.

	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

4. DISCUSSION

4.1. Explanation and interpretation

To write the learned model in a symbolic equation, a vector **u** is defined that contains velocity-related features. Another vector **x** is also defined to include logarithm-related features. Finally, **R** and **v** vectors are defined to include randomness-related features and all other non-linear features. Using the defined vectors, the learned model can be written in short notation as a linear combination of these vectors:

$$\text{CO} = (\mathbf{u} + \Delta\mathbf{u} + \mathbf{p} + \Delta\mathbf{p} + \mathbf{x} + \log(\mathbf{x}) + \mathbf{R} + \mathbf{v})\mathbf{W}. \quad (2)$$

Here **W** are the learned weights from training the model. Upon closer look at the learned equation, similarity to the

Navier-Stokes equation is observed. The term $\mathbf{u} + \Delta \mathbf{u} + \mathbf{p} + \Delta \mathbf{p}$ 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 $\mathbf{x} + \log(\mathbf{x}) + \mathbf{R} + \mathbf{v}$ are related to turbulence and viscosity phenomena similar to the Navier-Stokes equations, leaving the CO as an external force (input) to the system. Modeling the problem as a fluid flow problem with CO as input strongly connects with the classical approach and the physiological nature of the problem.

4.2. Limitation

Although the model achieved clinically acceptable performance, it should be noted that the study was conducted within specific clinical settings where a limited range of hemodynamic situations was captured. Furthermore, it was unexpected that demographic information such as age and BMI would play no role in the model. Finally, the connection to Navier-Stokes equations is still not fully clear.

5. CONCLUSION

The proposed approach achieved a clinically acceptable level of performance regarding radial limits of agreement and bias. The learned model was validated on two different datasets and achieved comparable performance. Finally, an interpretation of the learned model was provided. For future research, we will benchmark MIMIC performance against other published methods. Additionally, we will explore the time-delay embedding coordinates to model blood flow dynamics with CO and demographic information as input to the system and explore more advanced techniques, such as auto-encoders, for dynamics discovery.

6. REFERENCES

- [1] Rahul Nanchal and Robert W Taylor, "Hemodynamic monitoring," *Chap*, vol. 41, pp. 471–486, 2007.
- [2] Joseph Erlanger, "An experimental study of blood-pressure and of pulse-pressure in man," *Bull Johns Hopkins Hosp*, vol. 12, pp. 145–378, 1904.
- [3] Nicholas T Kouchoukos, Louis C Sheppard, and DONALD A McDONALD, "Estimation of stroke volume in the dog by a pulse contour method," *Circulation Research*, vol. 26, no. 5, pp. 611–623, 1970.
- [4] 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," *Journal for all experimental medicine*, vol. 59, pp. 105–122, 1928.
- [5] Hyun-Lim Yang, Chul-Woo Jung, Seong Mi Yang, Min-Soo Kim, Sungho Shim, Kook Hyun Lee, Hyung-Chul Lee, 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," *JMIR medical informatics*, vol. 9, no. 8, pp. e24762, 2021.
- [6] Liao Ke, Armagan Elibol, Xiao Wei, Liao Cenyu, Wang Wei, and Nak Young Chong, "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.
- [7] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz, "Sparse identification of nonlinear dynamics with control (sindyc)," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 710–715, 2016.
- [8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [9] Brian de Silva, Kathleen Champion, Markus Quade, Jean-Christophe Loiseau, J. Kutz, and Steven Brunton, "Pysindy: A python package for the sparse identification of nonlinear dynamical systems from data," *Journal of Open Source Software*, vol. 5, no. 49, pp. 2104, 2020.
- [10] Peter M Odor, Sohail Bampoe, and Maurizio Cecconi, "Cardiac output monitoring: Validation studies—how results should be presented," *Current anesthesiology reports*, vol. 7, pp. 410–415, 2017.
- [11] 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," *Anesthesia & Analgesia*, vol. 111, no. 5, pp. 1180–1192, 2010.
- [12] Lester A Critchley, Xiao X Yang, and Anna Lee, "Assessment of trending ability of cardiac output monitors by polar plot methodology," *Journal of cardiothoracic and vascular anesthesia*, vol. 25, no. 3, pp. 536–546, 2011.
- [13] Hyung-Chul Lee, Yoonsang Park, Soo Bin Yoon, Seong Mi Yang, Dongnyeok Park, and Chul-Woo Jung, "Vitaldb, a high-fidelity multi-parameter vital signs database in surgical patients," *Scientific Data*, vol. 9, no. 1, pp. 279, 2022.
- [14] Mohammed Saeed, Christine Lieu, Greg Raber, and Roger G Mark, "Mimic ii: a massive temporal icu patient database to support research in intelligent patient monitoring," in *Computers in cardiology*. IEEE, 2002, pp. 641–644.