Inferring Cardiovascular Biomarkers with Hybrid Model Learning

Ortal Senouf EPFL
Jens Behrmann Apple
Jörn-Henrik Jacobsen Apple
Pascal Frossard EPFL

Emmanuel Abbé EPFL & Apple
Antoine Wehenkel Apple

Abstract

Wearable devices offer continuous monitoring of biomarkers, presenting an opportunity to diagnose cardiovascular diseases earlier, potentially reducing their fatality rate. While machine learning holds promise for predicting cardiovascular biomarkers from sensor data, its use often depends on the availability of labeled datasets, which are limited due to technical and ethical constraints. On the other hand, biophysical simulations present a solution to data scarcity but face challenges in model transfer from simulation to reality due to inherent model simplifications and misspecifications. Building on advancements in hybrid learning, we introduce a method that combines a pulse-wave propagation model, rooted in biophysical simulations, with a correction model trained with unlabeled real-world data. This generative model transforms cardiovascular parameters into real-world sensor measurements and, when trained as an auto-encoder, also provides the inverse transformation, mapping measurements to cardiovascular biomarkers. Notably, when assessed using real pulse-wave data, our hybrid method appears to outperform models based solely on simulations in inferring cardiovascular biomarkers, opening new avenues for inferring physiological biomarkers in data-limited scenarios.

1 Introduction

Cardiovascular (CV) diseases are the leading cause of mortality worldwide [1, 2]. Arterial pulse waves (PW) propagate from the heart to peripheral locations through the arterial network and hold valuable information regarding the underlying cardiac health. The development of wearable devices, which continuously measure arterial PWs at peripheral locations, holds several promises for the early diagnostic of CV diseases and drastic reduction of their fatality rates [3, 4]. Nevertheless, inferring important cardiac biomarkers, e.g., stroke volume or left ventricular ejection time, from these measurements has remained a challenge, owing to the complex nature of this inverse problem [5].

Machine learning (ML) has emerged as a promising approach for predicting biomarkers from wearables' measurements [6, 7]. Typically, the labeled data in this domain come from specialized medical tests, primarily involving individuals with health concerns. As a result, labeled data are scarce and not representative of the general population, hindering the scaling of ML solutions beyond proof-of-concept. Moreover, the usual strategies to address data scarcity, e.g., data augmentations or self-supervised learning, are not directly applicable, as the nature of symmetries in the inverse problem remains unclear.

Alternatively, biophysical models, such as whole-body 1D hemodynamics simulators [8, 9], can describe the forward process that generates PWs based on a set of physiological parameters for the CV system. These models implicitly offer an inverse mapping from peripheral PW measurements to CV

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

biomarkers. Owing to this, they have recently gained attention as a solution to real-world data scarcity. Specifically, they are used to train models that learn the inverse mapping using simulated samples [5, 10, 11, 12, 13]. However, the *simulation-to-reality* gap of hemodynamics models challenge the success of these approaches. Inspired by recent advances in hybrid learning [14, 15, 16, 17], we jointly address labelled-data scarcity and biophysical models' misspecification by learning, and inverting, a conditional generative model, mapping biomarkers to PWs.

More specifically, we develop a biophysics-inspired neural surrogate of the 1D hemodynamics simulations from [9]. Then, we jointly solve the inference and misspecification's modelling problems by learning a hybrid auto-encoder. Finally, we provide a preliminary empirical validation of the hybrid learning method's ability to model the PWs generative process. Our approach holds potential in enhancing biomarker inference on real-world data by bridging the gap between simulation-based and data-driven inference.

2 Method

Inference of CV parameters from PWs represents a classic inverse problem, where the objective is to deduce the underlying system parameters given measurements. Consider $g_p(\mathbf{z}_p, \mathbf{z}_t) = \mathbf{x}$, where g_p is the forward generative model describing the relationship between cardiac parameters of interest (POI) $\mathbf{z}_p$, the nuisance parameters $\mathbf{z}_t$ and the measured PW, $\mathbf{x}$. The objective of the inverse problem lies in determining $\mathbf{z}_p, \mathbf{z}_t$, such that the model's output matches the observed data as closely as possible. We use the L2 norm to compare reconstructed and original PWs, leading to the following optimization problem:

$$\hat{\mathbf{z}}_p, \hat{\mathbf{z}}_t = \arg\min_{\mathbf{z}_p, \mathbf{z}_t} ||g_p(\mathbf{z}_p, \mathbf{z}_t) - \mathbf{x}||^2.$$

The challenge becomes particularly pronounced when the forward model g_p does not fully encapsulate the true physical process, potentially due to unaccounted parameters $\mathbf{z}_c$. This is a common scenario since models are typically simplifications of the actual physics. This deviation, represented as $\epsilon(\mathbf{z}_p, \mathbf{z}_t, \mathbf{z}_c)$, then modifies the objective of the inverse problem, from merely inferring $\mathbf{z}_p$,$\mathbf{z}_t$ to doing so while modelling ϵ. This task can be framed as:

$$\hat{\mathbf{z}}_p, \hat{\mathbf{z}}_t, \hat{\mathbf{z}}_c = \arg\min_{\mathbf{z}_p, \mathbf{z}_t, \mathbf{z}_c} ||g_p(\mathbf{z}_p, \mathbf{z}_t) + \epsilon(\mathbf{z}_p, \mathbf{z}_t, \mathbf{z}_c) - \mathbf{x}||^2 \tag{1}$$

2.1 Differentiable and efficient pulse-wave model

Following the approach of [5], our study employs the whole-body 1D hemodynamics simulator presented by [9] as the forward model g_p. This model simulates PW propagation across 116 significant arterial segments encompassing the thorax, limbs, and head. A balance between real-world fidelity and manageable complexity makes this model especially relevant. It can simulate measured arterial pressure (APWs) and photoplethysmograms (PPGs) waveforms at various body locations, given physiological parameters representing the cardiovascular system's bio-physical traits. Here the POI, $\mathbf{z}_p$, are set to be the heart rate (HR), left-ventricular ejection time (LVET), and stroke volume (SV), which characterize the cardiac pulse, as depicted in Fig.1A.

The simulator encompasses two key components: the cardiac pulse generation model $g_{p1}(\mathbf{z}_p)$ and the pulse propagation model $g_{p2}(g_{p1}, \mathbf{z}_t)$, as depicted in Fig.1A. The latter, g_{p2}, requires solving differential equations, making it computationally demanding. Consequently, directly using the simulator g_p for iterative tasks, such as training a neural network (NN) to solve Eq.1, becomes computationally prohibitive. To navigate this, we implement g_{p1} in a differentiable manner, introducing a substantial bias with respect to the role of the POI in PW dynamics. We then employ a surrogate model $\tilde{g}_p(;\theta_p)$ for pulse propagation modelling. This surrogate, defined as a NN with parameters θ_p, serves as a lightweight, differentiable approximation of g_{p2}, reducing computation times from minutes (simulator) to mere seconds (surrogate). It is trained using triplets of POI parameters $\mathbf{z}_p$, nuisance parameters $\mathbf{z}_t$, and the corresponding APW $\mathbf{x}_p$ from the simulator, as illustrated in Fig.1B. The cardiac pulse, generated by $g_{p1}(\mathbf{z}_p)$, alongside $\mathbf{z}_t$, becomes the surrogate input. The training objective is to minimize the mean squared error (MSE) between the surrogate's predicted wave $\hat{\mathbf{x}}_p$ and the simulator-generated wave $\mathbf{x}_p$.

2.2 End-to-End modelling and inference

We train a physics encoder, $f_{p,e}(;\phi_p)$, parameterized by a NN with weights ϕ_p, to infer $\mathbf{z}_p$ from simulated observed APW, $\mathbf{x}_p$ (Fig.1C). This training employs a supervised approach using triplets of the same simulated data utilized for the surrogate training. The loss function is the MSE between the predicted parameters $\hat{\mathbf{z}}_p$, $\hat{\mathbf{z}}_t$, and the actual ones that generated the simulated APW, $\mathbf{x}_p$.

Given a real measured wave $\mathbf{x}$ as input, the encoder predicts the parameters $\mathbf{z}_p, \mathbf{z}_t$, which are then fed into the differentiable cardiac pulse generator and surrogate simulator $\tilde{g}_p(;\theta_p)$ to generate a reconstructed wave $\hat{\mathbf{x}}$. The discrepancy between reconstructed and real waves is used to measure the accuracy of the parameters inference. To adjust for deviations between real wave $\mathbf{x}$ and simulated $\mathbf{x}_p$, a correction module that models $\epsilon(\mathbf{z}_p, \mathbf{z}_t, \mathbf{z}_c)$, is integrated in a hybrid pipeline with the surrogate physical model. The correction module (Fig.1D, bottom) consists of encoder $f_{c,e}(;\phi_c)$ identifying generative parameters $\mathbf{z}_c$ possibly missed by the physics simulator and decoder $g_c(;\theta_c)$ determining how these parameters influence APW propagation and defines the residual PW, $\mathbf{x}_c$. The optimization task is formulated as:

$$\hat{\mathbf{z}}_p, \hat{\mathbf{z}}_t, \phi_c, \theta_c = \arg \min_{\mathbf{z}_p, \mathbf{z}_t, \phi_c, \theta_c} ||\tilde{g}_p(f_{p,e}(\tilde{x}_p)) + g_c(f_{c,e}(x;\phi_c);\theta_c) - \mathbf{x}||^2 \tag{2}$$

where, $\tilde{\mathbf{x}}_p = \mathbf{x} - \mathbf{x}_c$, approximates the simulated APW $\mathbf{x}_p$ by subtracting the residual $\mathbf{x}_c$ from the real APW, $\mathbf{x}$.

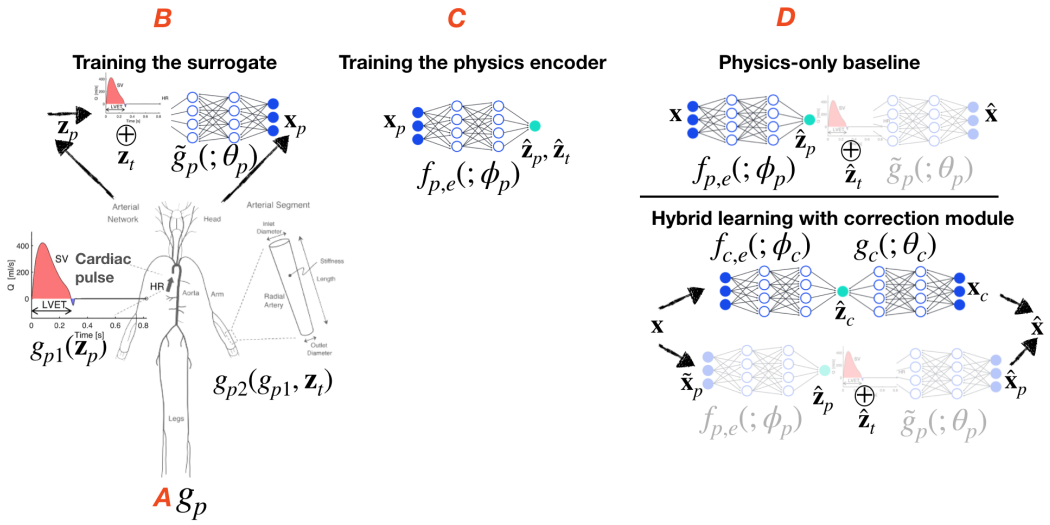

Figure 1: Hybrid model training scheme. The translucent elements denote fixed checkpoints or models. (A) The 1D pulse wave simulator [9] generates samples $\{\mathbf{z}_p, \mathbf{z}_t, \mathbf{x}_p\}$, which are used in (B) to train the surrogate. The cardiac pulse model g_{p1} is implemented explicitly and together with $\mathbf{z}_t$ serves as the input to the surrogate $\tilde{g}_p$. (C) The physics encoder is trained to invert the simulator using the same data as in A. (D) Top: Physics-only baseline (no correction model), the encoder is fine-tuned to infer parameters from real APW. Bottom: the hybrid approach, a correction encoder-decoder, $f_{c,e}, g_c$, is trained to account for the discrepancy between real and simulated APW.

3 Experiments

Simulations dataset. We train our surrogate model, $\tilde{g}_p(;\theta_p)$, on a dataset sourced from the database created by [9]. This database comprises 4,374 virtual healthy subjects aged between 25 and 75 years. The dataset has 6 degrees of freedom (age, diameter, HR, LVET, SV, mean blood pressure, pulse-wave velocity), which dictate all relevant parameters of the model. Subjects are simulated by sampling clinically meaningful values for each degree of freedom and simulating the corresponding APW.

Real pulse-waves (PulseDB) [18]. A large dataset for benchmarking cuff-less blood pressure estimation methods using ECG and PPG signals. It contains over 5 million waveform segments from

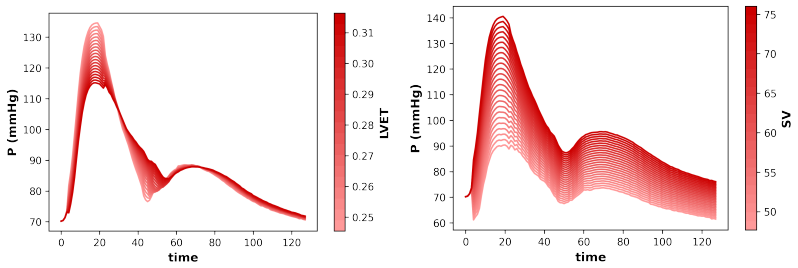

Figure 2: Visual sensitivity analysis of the surrogate model with respect to LVET and SV. The shade of the predicted APWs scales with the values of the inspected parameter

5,361 subjects. For evaluation, each APW sequence in the dataset is paired with HR and LVET labels, obtained by analyzing the associated ECG segments. It is important to note that these labels are utilized solely for assessment, not for training purposes.

Training & model selection process. Our study employs a multi-stage training approach, as illustrated in Fig.1. We begin by training the surrogate model, then advance to the physics encoder once optimal. For rigorous model selection, we use a nested cross-validation, dividing the data into 5 external folds with 3 internal subdivisions each. The surrogate model's validation for each external fold uses average scores from its internal folds. The chosen architecture for the surrogate model is a masked auto-regressive neural network (MAN) [19], since it introduces an inductive bias that facilitates time-series modeling. After finalizing the surrogate model, we qualitatively assess its sensitivity to the POI as shown in Fig.2. By tweaking parameters and maintaining others, we evaluate the surrogate's alignment with the simulator's ground truth APW for similar parameters. With the optimal surrogate configuration identified and fixed for each external fold, we then train and validate the physics encoder, designed as a convolutional NN, across the external 5 folds. Next, the hybrid model is trained using PulseDB APWs, by anchoring the physics encoder and surrogate model. As illustrated in Fig. 1.D (bottom), only the correction module undergoes training. The optimal training checkpoint is chosen based on the lowest reconstruction MSE.

Experimental settings. The model's efficacy is compared against a physics-only baseline (Fig. 1.D, top) wherein solely the physics encoder, $f_{p,e}(;\phi_p)$ is fine-tuned using actual APWs. Relative mean absolute error (RMAE) assesses the predicted parameters accuracy with respect to the actual labels. Two experimental settings are considered: one where HR is inferred and evaluated (HR in Table 1), and another where given HR, LVET inference gets evaluated ($\neg$ HR in Table 1). This distinction arises as the model often struggles to jointly infer multiple parameters possibly due to their interplay.

Results & conclusion. As shown in Table 1, the hybrid model exhibits lower RMAE, compared to the physics-only baseline. This is evident also when the reconstruction MSE are similar (0.072 vs. 0.057), suggesting that modelling the deviation between real and simulated waves is significant in solving the inverse problem and inferring the cardiac biomarkers. Nonetheless, it is important to note that these findings are preliminary, and our current approach has certain limitations. Throughout the training process, while validation MSE decreases, there comes a stage where the error in parameter inference grows. This suggests a need for model regularization, echoing findings from prior research [15]. In future work, we plan to explore regularization for hybrid learning for inverse problems. Additionally, we will examine the probabilistic approach to quantify uncertainty in parameter inference, potentially addressing the parameters interplay issue.

References

[1] Gregory A Roth, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi, Hedayat Abbastabar, Foad Abd-Allah, Jemal Abdela, Ahmed Abdelalim, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392(10159):1736–1788, 2018.

Model	Physics Only		Hybrid	
	HR	¬HR	HR	¬HR
Reconstruction MSE	0.23	0.072	0.032	0.057
Parameters RMAE	10.88%	14.21%	6.84%	7.45%

Table 1: Comparison of the physics only and hybrid learning approaches with two experimental settings. "HR": HR is inferred and evaluated, "¬ HR": HR is given and LVET is inferred and evaluated.

[2] Salim S Virani, Alvaro Alonso, Emelia J Benjamin, Marcio S Bittencourt, Clifton W Callaway, April P Carson, Alanna M Chamberlain, Alexander R Chang, Susan Cheng, Francesca N Delling, et al. Heart disease and stroke statistics—2020 update: a report from the American Heart Association. *Circulation*, 141(9):e139–e596, 2020.

[3] Tor Biering-Sørensen, Gabriela Querejeta Roca, Sheila M Hegde, Amil M Shah, Brian Claggett, Thomas H Mosley Jr, Kenneth R Butler Jr, and Scott D Solomon. Left ventricular ejection time is an independent predictor of incident heart failure in a community-based cohort. *European journal of heart failure*, 20(7):1106–1114, 2018.

[4] Denisse Castaneda, Aibhlin Esparza, Mohammad Ghamari, Cinna Soltanpur, and Homer Nazeran. A review on wearable photoplethysmography sensors and their potential future applications in health care. *International journal of biosensors & bioelectronics*, 4(4):195, 2018.

[5] Antoine Wehenkel, Jens Behrmann, Andrew C Miller, Guillermo Sapiro, Ozan Sener, Marco Cuturi, and Jörn-Henrik Jacobsen. Simulation-based Inference for Cardiovascular Models. *arXiv preprint arXiv:2307.13918*, 2023.

[6] Shing-Hong Liu, Zhi-Kai Yang, Kuo-Li Pan, Xin Zhu, and Wenxi Chen. Estimation of Left Ventricular Ejection Fraction Using Cardiovascular Hemodynamic Parameters and Pulse Morphological Characteristics with Machine Learning Algorithms. *Nutrients*, 14(19):4051, 2022.

[7] Young-Jin Moon, Hyun S Moon, Dong-Sub Kim, Jae-Man Kim, Joon-Kyu Lee, Woo-Hyun Shim, Sung-Hoon Kim, Gyu-Sam Hwang, and Jae-Soon Choi. Deep learning-based stroke volume estimation outperforms conventional arterial contour method in patients with hemodynamic instability. *Journal of clinical medicine*, 8(9):1419, 2019.

[8] Alessandro Melis. Gaussian process emulators for 1D vascular models. *PhD Thesis*, 2017.

[9] Peter H Charlton, Jorge Mariscal Harana, Samuel Vennin, Ye Li, Phil Chowienczyk, and Jordi Alastruey. Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes. *American Journal of Physiology-Heart and Circulatory Physiology*, 317(5):H1062–H1085, 2019.

[10] Neeraj Kavan Chakshu, Igor Sazonov, and Perumal Nithiarasu. Towards enabling a cardiovascular digital twin for human systemic circulation using inverse analysis. *Biomechanics and modeling in mechanobiology*, 20(2):449–465, 2021.

[11] Weiwei Jin, Philip Chowienczyk, and Jordi Alastruey. Estimating pulse wave velocity from the radial pressure wave using machine learning algorithms. *Plos one*, 16(6):e0245026, 2021.

[12] Vasiliki Bikia, Marija Lazaroska, Deborah Scherrer Ma, Méline Zhao, Georgios Rovas, Stamatia Pagoulatou, and Nikolaos Stergiopulos. Estimation of Left Ventricular End-Systolic Elastance From Brachial Pressure Waveform via Deep Learning. *Frontiers in Bioengineering and Biotechnology*, 9:754003, 2021.

[13] Eugenia Ipar, Nicolás A Aguirre, Leandro J Cymberknop, and Ricardo L Armentano. Blood Pressure Morphology as a Fingerprint of Cardiovascular Health: A Machine Learning Based Approach. In *Applied Informatics: Fourth International Conference, ICAI 2021, Buenos Aires, Argentina, October 28–30, 2021, Proceedings 4*. Springer, 2021.

[14] Naoya Takeishi and Alexandros Kalousis. Physics-integrated variational autoencoders for robust and interpretable generative modeling. In *Advances in Neural Information Processing Systems*, 2021.

[15] Yuan Yin, Vincent Le Guen, Jérémie Dona, Emmanuel de Bézenac, Ibrahim Ayed, Nicolas Thome, and Patrick Gallinari. Augmenting physical models with deep networks for complex dynamics forecasting. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12):124012, 2021.

[16] Emmanuel De Bézenac, Arthur Pajot, and Patrick Gallinari. Deep learning for physical processes: Incorporating prior scientific knowledge. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12):124009, 2019.

[17] Antoine Wehenkel, Jens Behrmann, Hsiang Hsu, Guillermo Sapiro, Gilles Louppe, and Jörn-Henrik Jacobsen. Robust Hybrid Learning With Expert Augmentation. *Transactions on Machine Learning Research*, 2023.

[18] Weinan Wang, Pedram Mohseni, Kevin L Kilgore, and Laleh Najafizadeh. PulseDB: A large, cleaned dataset based on MIMIC-III and VitalDB for benchmarking cuff-less blood pressure estimation methods. *Frontiers in Digital Health*, 4:1090854, 2023.

[19] Mathieu Germain, Karol Gregor, Iain Murray, and Hugo Larochelle. Made: Masked autoencoder for distribution estimation. In *International Conference on Machine Learning*, 2015.