Baseline Wander as a Limiting Factor in Heart Vector Space Modeling: Improving Neural ODE and PINN Predictions via Learned Decomposition

Richárd Tuhári[1], Gergő Bognár[1]

[1] Department of Numerical Analysis, Eötvös Loránd University, Budapest, Hungary

Abstract

Physics-informed models assume that observed motion is compatible with their governing dynamics. Baseline wander violates this assumption in vectorcardiography by adding structured low-frequency motion generated outside the modeled cardiac system. We evaluated a physics-informed recurrent predictor and a hybrid Neural ODE on one synthetic and three measured ECG datasets, before and after wavelet correction. Baseline removal consistently improved the PINN–GRU in the principal pointwise metrics on measured data, whereas Neural ODE performance depended on the dataset and metric. On synthetic data, the models behaved oppositely: cleaning strongly improved the Neural ODE but degraded the PINN–GRU, suggesting a mismatch between the synthetic generator and the physics encoded by the recurrent model. We also fitted a continuous neural baseline representation to a Gaussian-smoothed target. Its translation-invariant synthetic error was 27.52, comparable with the best wavelet result of 26.90. The current decomposition is offline because its symmetric target uses future samples, but it establishes that a learned model can attain wavelet-level baseline separation in heart vector space.

1. Introduction

This work studies how baseline wander affects the prediction of cardiac trajectories in vectorcardiographic space. Vectorcardiography (VCG) represents the electrical activity of the heart as a time-dependent three-dimensional vector and provides spatial information complementary to conventional electrocardiogram (ECG) analysis [1]. We reconstruct its X, Y, and Z components from the eight independent ECG leads using the inverse Dower transformation [2]. Forecasting this trajectory means predicting its future position from a preceding observation window.

Baseline wander is commonly treated as an ECG preprocessing artifact, but it can obscure clinically relevant morphology, while its removal can itself distort components such as the ST segment [3]. In VCG, baseline wander becomes an additional low-frequency displacement of the three-dimensional trajectory. The resulting motion is not produced by the cardiac dynamics that the prediction model is intended to describe. Consequently, a model may receive an observed trajectory that is incompatible with its assumed physical system.

This mismatch is particularly relevant to physics-guided prediction. A physics-informed recurrent model is penalized when its predictions violate a prescribed dynamical residual, while a Neural ordinary differential equation (Neural ODE) generates future trajectories by integrating a learned vector field. Neither model explicitly represents baseline wander. The models must therefore either interpret the artifact as cardiac motion or reject it through their learned dynamics, potentially reducing prediction accuracy.

Our primary goal is to determine whether removing baseline wander improves heart-vector prediction by making the observed trajectory more consistent with the modeled physics. We evaluate a physics-informed recurrent predictor (PINN–GRU) and a hybrid Neural ODE on raw and wavelet-corrected VCG trajectories. Synthetic data provide a clean cardiac reference. The measured evaluation uses two recordings from the Large-Scale 12-Lead ECG Arrhythmia Database [4] and one recording from the St Petersburg INCART 12-Lead Arrhythmia Database [5], for which the true baseline components are unknown.

Our second goal is to determine whether baseline wander can be represented by a learned continuous function rather than a fixed signal-processing procedure. We therefore fit a neural baseline representation and evaluate both its geometric agreement with wavelet decomposition and its effect on downstream prediction. The current learned decomposition is non-causal and is used only for offline evaluation; predictive real-time decomposition remains future work.

2. Methods

The methodology consists of three stages. First, the eight independent ECG leads are decomposed into cardiac and nuisance components and transformed into a three-dimensional VCG trajectory. Second, a physics-informed recurrent model and a hybrid Neural ODE are trained to predict future VCG positions from either uncorrected or baseline-corrected observations. Third, a learned continuous baseline representation is compared with wavelet removal in terms of both trajectory reconstruction and downstream prediction performance.

2.1. Signal model and wavelet reference

To examine how baseline wander distorts the reconstructed cardiac trajectory, we first model its contribution in ECG lead space and then remove it using a wavelet-based reference method. Let $\mathbf{s}_{\text{obs}}, \mathbf{s}_{\text{cardiac}}, \mathbf{n} \in \mathbf{R}^{\mathbf{8}}$ denote the measured ECG, its cardiac component, and the combined baseline-wander and residual-noise component, respectively:

$$\mathbf{s}_{\text{obs}} = \mathbf{s}_{\text{cardiac}} + \mathbf{n}. \tag{1}$$

With the inverse Dower matrix $\mathbf{D} \in \mathbf{R}^{\mathbf{3\times 8}}$ [2],

$$\mathbf{x}_{\text{obs}} = \mathbf{D}\mathbf{s}_{\text{obs}} = \mathbf{x}_{\text{cardiac}} + \mathbf{D}\mathbf{n}, \tag{2}$$

where $\mathbf{x}_{\text{obs}}, \mathbf{x}_{\text{cardiac}} \in \mathbf{R}^{\mathbf{3}}$ denote the observed and cardiac VCG vectors, respectively. The nuisance terms cannot be separated exactly in measured recordings. In the synthetic data, the added baseline $\mathbf{b}$ is known and linearity gives $\mathbf{D}(\mathbf{s}_{\text{cardiac}} + \mathbf{b}) = \mathbf{x}_{\text{cardiac}} + \mathbf{D}\mathbf{b}$.

The wavelet reference was computed independently in each ECG lead with a multilevel `sym10` transform and a 0.5-Hz target cutoff using PyWavelets [6]. Wavelet decomposition is well suited to baseline removal because it separates signal components across both time and scale, and previous comparisons have demonstrated its effectiveness on ECG signals with synthetically added baseline wander [7]. Approximation coefficients reconstructed the baseline; subtraction was followed by the inverse Dower transform. Reversing this order gave similar synthetic errors, so lead-space removal was fixed for the prediction experiments.

2.2. Physics-guided prediction

To test whether baseline wander interferes with the modeled cardiac dynamics, we predict future VCG trajectories using two complementary physics-guided approaches: a recurrent model regularized by a physical residual and a hybrid Neural ODE.

Physics-informed neural networks incorporate governing equations into the learning objective through a differential-equation residual [8]. Following this principle, the PINN–GRU predicts the next VCG position from a context of positions and finite-difference velocities. Its output corrects a constant-velocity extrapolation. A discrete forced damped-oscillator residual regularizes the prediction:

$$\mathbf{r_{t+1}} = \widehat{\mathbf{a}}_{\mathbf{t+1}} + \mathbf{c}\widehat{\mathbf{v}}_{\mathbf{t+1}} + \mathbf{k}(\mathbf{x_t} - \mathbf{x_0}) - \mathbf{p_t}\mathbf{d_t}, \tag{3}$$

where $c, k > 0$, equilibrium $\mathbf{x_0}$, driving magnitude p_t, and unit direction $\mathbf{d_t}$ are learned. The loss combines data, physics-residual, and drive-penalty terms. Multi-step predictions recursively reuse the fixed one-step model without parameter updates.

Neural ODEs represent the hidden-state dynamics of a neural network through a continuous-time differential equation [9]. Our hybrid Neural ODE encodes position, velocity, and acceleration histories into an initial physical and latent state, then integrates

$$\dot{\mathbf{x}} = \mathbf{v}, \tag{4}$$

$$\dot{\mathbf{v}} = -\gamma\mathbf{v} - \omega_{\mathbf{0}}^{\mathbf{2}}(\mathbf{x} - \mathbf{x_0}) + \alpha\mathbf{f_p}(\mathbf{x}, \mathbf{v}, \mathbf{z}, \mathbf{t}), \tag{5}$$

$$\dot{\mathbf{z}} = f_z(\mathbf{x}, \mathbf{v}, \mathbf{z}, \mathbf{t}). \tag{6}$$

This system represents a forced, damped second-order oscillator around the learned equilibrium position. Positive damping, natural frequency, and forcing scale are learned. Fixed-step fourth-order Runge–Kutta integration produces the complete prediction horizon. The equilibrium center in both physics-guided predictors was retained from an earlier central-potential formulation; the current oscillator form provided a simpler trainable system.

2.3. Learned decomposition

To determine whether baseline wander can be represented without a fixed signal-processing rule, we learn it directly as a continuous function of time and compare the resulting decomposition with the wavelet reference.

The learned baseline is a continuous coordinate network $\widehat{\mathbf{b}}(t) = f_\theta(t)$ with three 128-unit tanh hidden layers and a three-dimensional output. It is fitted separately to each VCG trajectory using mean squared error against a 101-sample Gaussian-smoothed target. The corrected trajectory is

$$\widehat{\mathbf{x}}_{\text{cardiac}}(t) = \mathbf{x}_{\text{obs}}(\mathbf{t}) - \widehat{\mathbf{b}}(\mathbf{t}) + \mathbf{c}, \tag{7}$$

where $\mathbf{c}$ restores the target centroid. This centering changes only translation, not trajectory shape. Because the Gaussian window is symmetric, the target contains future information. The current method is therefore an offline decomposition, not yet a valid predictive or real-time estimator.

Decomposition quality was measured by the translation-invariant RMS distance

$$d_{\text{TI}}(\mathbf{a}, \mathbf{b}) = \sqrt{\frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{i}} \left\| (\mathbf{a_i} - \bar{\mathbf{a}}) - (\mathbf{b_i} - \bar{\mathbf{b}}) \right\|_{\mathbf{2}}^{\mathbf{2}}}. \tag{8}$$

Table 1. Prediction errors on three measured ECG datasets and one synthetic dataset. Lower is better.

Data	Model	Input	MSE	RMSE	MAE	ADE	FDE
JS01025	H-NODE	raw	21 651.1	147.1	100.6	201.2	209.8
	H-NODE	BWR	**13326.0**	**115.4**	**54.5**	**115.4**	**164.3**
	P-GRU	raw	34 338.8	185.3	108.3	236.9	477.9
	P-GRU	BWR	**29476.8**	**171.7**	**97.9**	**220.0**	**336.8**
JS00064	H-NODE	raw	**12340.0**	**111.1**	**40.2**	**79.7**	**108.6**
	H-NODE	BWR	16 795.5	129.6	46.4	92.6	141.6
	P-GRU	raw	13 928.5	118.0	43.1	88.8	**166.2**
	P-GRU	BWR	**8313.4**	**91.2**	**42.1**	**86.1**	198.9
I0074	H-NODE	raw	**10312.9**	**101.6**	70.5	145.5	179.6
	H-NODE	BWR	12 231.3	110.6	**53.7**	**102.6**	**110.2**
	P-GRU	raw	34 485.9	185.7	91.5	177.4	299.1
	P-GRU	BWR	**20298.8**	**142.5**	**67.8**	**130.8**	**192.0**
Synthetic	H-NODE	raw	1877.5	43.3	25.2	50.0	79.0
	H-NODE	BWR	589.7	24.3	14.9	29.6	33.9
	H-NODE	clean	**543.2**	**23.3**	**11.3**	**22.3**	**29.3**
	P-GRU	raw	**965.2**	**31.1**	**19.7**	**39.3**	76.9
	P-GRU	BWR	1322.0	36.4	21.8	43.0	**75.9**
	P-GRU	clean	3719.7	61.0	35.1	71.8	155.6

3. Results

The evaluation addresses two questions. First, we examine how wavelet-based baseline removal changes PINN–GRU and Neural ODE prediction performance across one synthetic and three measured datasets. Second, we compare the learned decomposition with the wavelet reference in terms of trajectory reconstruction and downstream prediction accuracy.

3.1. Prediction across four datasets

Table 1 reports the four experiments. “BWR” denotes wavelet-corrected input; “clean” is available only for the synthetic data. ADE and FDE are mean and endpoint Euclidean displacement errors.

For measured data, BWR reduced PINN–GRU MSE, RMSE, MAE, and ADE on every record. FDE also improved for JS01025 and I0074, but increased for JS00064. The Hybrid Neural ODE was less uniform: all metrics improved for JS01025, all worsened for JS00064, and I0074 showed higher MSE and RMSE but lower MAE, ADE, and FDE after BWR. The latter combination is consistent with improved typical or local accuracy accompanied by occasional large deviations. It motivates testing higher-order terms in the differential equation, but does not by itself establish that missing higher-order dynamics are the cause.

The synthetic experiment reversed the PINN–GRU trend: cleaning degraded every metric except FDE after wavelet correction, whereas Neural ODE performance improved strongly and approached the ground-truth-clean result. One plausible explanation is that the recurrent physical residual better approximates the measured cardiac dynamics than the synthetic generator, making the PINN–GRU more sensitive to synthetic model mismatch. Alternative explanations include optimization variability and distributional differences; repeated runs are needed to distinguish them.

3.2. Learned Baseline Decomposition

Table 2 compares the learned and wavelet-based decompositions using translation-invariant distance. On synthetic data, the learned method performed comparably to the wavelet reference. On measured recordings, where no clean ground truth was available, the reported distances measure agreement between the two decomposition methods rather than absolute accuracy.

Table 2. Translation-invariant distances between VCG trajectories.

Data	Trajectories compared	d_{TI}
Synthetic	Clean vs. wavelet before VCG	34.91
Synthetic	Clean vs. wavelet after VCG	26.90
Synthetic	Clean vs. learned	27.52
Synthetic	Wavelet vs. learned	23.39
JS01025	Wavelet vs. learned	71.91
JS00064	Wavelet vs. learned	30.60
I0074	Wavelet vs. learned	27.75

Figure 1 compares the instantaneous reconstruction errors of the wavelet-based and learned decompositions on the synthetic recording. For this comparison, both corrected trajectories were expressed relative to the center learned by the decomposition network. Applying the same constant translation to the wavelet-corrected trajectory ensures that the pointwise comparison reflects differences in trajectory shape rather than an arbitrary spatial offset.

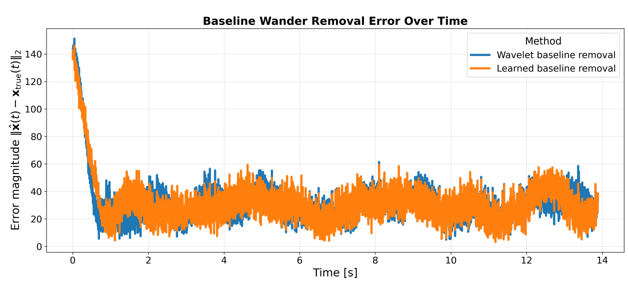

Figure 1. Instantaneous reconstruction error on the synthetic recording after wavelet-based and learned baseline removal. Both trajectories were aligned using the learned center before computing their Euclidean errors relative to the clean VCG trajectory.

After the initial transient, the two error curves remain in a similar range and follow comparable temporal variations. Neither decomposition is uniformly superior: their relative performance changes over time, and both exhibit occasional local deviations. Together with the translation-invariant distances in Table 2, this supports the conclusion that the learned network achieves wavelet-level offline decomposition on the synthetic trajectory. The comparison does not establish predictive or real-time performance because the Gaussian-smoothed training target is non-causal.

Table 3 shows that geometric agreement with the wavelet reference does not imply identical downstream prediction performance. The learned decomposition improved all reported metrics for both models on JS01025 and improved the Hybrid Neural ODE on JS00064. Its performance was comparable to wavelet removal for the P-GRU on JS00064, but was worse for both models on I0074. On synthetic data, it improved the Hybrid Neural ODE while substantially degrading the P-GRU. These results indicate that decomposition quality is model- and record-dependent and cannot be characterized by trajectory agreement alone. Nevertheless, the strong improvements on JS01025 and for the Neural ODE on JS00064 demonstrate that a learned decomposition can match or outperform the wavelet reference in downstream forecasting. Because the present Gaussian target is non-causal, these findings establish offline feasibility rather than real-time predictive decomposition.

Table 3. Prediction performance after wavelet-based (W) and learned (L) baseline decomposition. Lower values are better. The learned decomposition is evaluated offline because its Gaussian-smoothed target is non-causal.

Data	Model	RMSE W/L	MAE W/L	ADE W/L	FDE W/L
Synthetic	H-NODE	24.3/20.3	14.9/12.4	29.7/24.6	33.9/28.9
	P-GRU	36.4/124.6	21.8/39.5	43.0/79.8	75.9/135.6
JS01025	H-NODE	115.4/68.5	54.5/34.9	115.4/74.8	164.3/92.9
	P-GRU	171.7/129.5	97.9/67.6	220.0/151.2	336.8/266.6
JS00064	H-NODE	129.6/81.4	46.4/35.1	92.6/70.2	141.6/124.7
	P-GRU	91.2/91.4	42.1/41.8	86.1/87.4	198.9/206.1
I0074	H-NODE	110.6/171.2	53.7/57.5	102.6/110.6	110.2/250.5
	P-GRU	142.5/211.4	67.8/74.3	130.8/140.1	192.0/260.0

4. Discussion and Conclusion

We investigated whether baseline wander limits physics-guided prediction in heart vector space. A physics-informed recurrent predictor and a hybrid Neural ODE were evaluated on raw, wavelet-corrected, and learned-corrected VCG trajectories from one synthetic and three measured datasets.

Baseline removal generally improved the recurrent predictor on measured data, whereas the Neural ODE response depended on the recording and metric. Together with the opposite synthetic behavior, this indicates that performance depends not only on noise magnitude, but also on whether the observed trajectory is compatible with the dynamics encoded by the model. The results are preliminary single-run comparisons and do not yet identify the physical terms responsible for these differences.

The learned network achieved wavelet-level synthetic decomposition and improved prediction on selected measured recordings. This demonstrates the feasibility of integrating baseline estimation with downstream forecasting, although the current non-causal target restricts the method to offline use. Overall, baseline wander is a relevant source of model–observation mismatch, and decomposition should be evaluated jointly with the physical predictor. Future work will develop causal real-time decomposition, repeat the experiments across seeds and datasets, and investigate alternative physical dynamics.

Code and complete reproducibility materials are available at `https://github.com/9g77yfgb4h-spec/Baseline-Wander-as-a-Limiting-Factor-in-Heart-Vector-Space-Modeling`.

5. Acknowledgements

The project was supported by the University Excellence Fund of Eötvös Loránd University, Budapest, Hungary (ELTE).

References

[1] Vondrak J, Penhaker M. Review of processing pathological vectorcardiographic records for the detection of heart disease. Frontiers in Physiology 2022;13:856590.

[2] Edenbrandt L, Pahlm O. Vectorcardiogram synthesized from a 12-lead ecg: Superiority of the inverse dower matrix. Journal of Electrocardiology 1988;21(4):361–367.

[3] Lenis G, Pilia N, Loewe A, Schulze WHW, Dössel O. Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: A simulation study. Computational and Mathematical Methods in Medicine 2017;2017:9295029.

[4] Zheng J, Guo H, Chu H. A large-scale 12-lead electrocardiogram database for arrhythmia study. PhysioNet 2022;Version 1.0.0.

[5] Yakushenko E. St petersburg INCART 12-lead arrhythmia database, 2008. Version 1.0.0.

[6] Lee GR, Gommers R, Wasilewski F, Wohlfahrt K, O'Leary A. Pywavelets: A python package for wavelet analysis. Journal of Open Source Software 2019;4(36):1237.

[7] Chen CC, Tsui FR. Comparing different wavelet transforms on removing electrocardiogram baseline wanders and special trends. BMC Medical Informatics and Decision Making 2020;20(Suppl 11):343.

[8] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 2019; 378:686–707.

[9] Chen RTQ, Rubanova Y, Bettencourt J, Duvenaud D. Neural ordinary differential equations. In Advances in Neural Information Processing Systems, volume 31. 2018; 6571–6583.