Cross-Subject Prediction in ECGI using a MARS-Based Regression Framework

Amaël Mombereau[1], Nicolas Montagne[1], Ayoub El Ghebouli[1], Yesim Serinagaoglu[2*] & Laura Bear[1*]

[1]Université de Bordeaux, INSERM, CRCTB, U 1045, IHU Liryc, Bordeaux, France
[2]Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara,Turkey
*These authors contributed equally as last authors

Abstract

Noninvasive electrocardiographic imaging (ECGI) reconstructs epicardial electrical activity from body surface potentials (BSP). In previous work, we demonstrated that a regression framework based on Multivariate Adaptive Regression Splines (MARS) trained within a subject achieved higher fidelity than Tikhonov regularization. In this study, we extend this framework to assess cross-subject generalization. A leave-one-pig-out protocol was applied on in-vivo data from four anesthetized pigs with torso and epicardial recordings (184 BSP and 239 EGM electrodes). To enable cross-subject training, BSP and EGM were standardized to 2D electrode maps. Results showed that the generalized MARS model maintained accuracy close to subject-specific training, and in some cases outperformed it when training data within the test subject were limited. These findings highlight the potential of data-driven ECGI frameworks to move beyond subject-specific models and towards clinically transferable solutions.

1. Introduction

Electrocardiographic imaging (ECGI) aims to reconstruct epicardial electrical activity from noninvasive body surface potentials (BSP). The classical formulation is an ill-posed inverse problem, typically solved with Tikhonov regularization [1]. Although stabilizing the solution, such approaches tend to oversmooth electrograms and activation time (AT) maps, are highly sensitive to errors in geometry or electrode registration, and may introduce artifacts such as spurious lines of block [2-3].

To overcome these limitations, data-driven approaches have recently been proposed. By directly learning the BSP-EGM mapping, they avoid explicit dependence on forward models and may be more robust to anatomical variability. Neural networks, hybrid physics-informed methods, and nonparametric regressions have shown promising results [4-6]. In particular, we previously introduced a regression framework based on Multivariate Adaptive Regression Splines (MARS), trained within a single subject (M1). This subject-specific model improved EGM and AT fidelity compared to Tikhonov [7], but required subject-specific training data, limiting clinical translation.

Achieving cross-subject generalization is a critical step for ECGI. A generalized model could reduce the need for patient-specific imaging, improve robustness to inter-individual variability, and allow pooling of heterogeneous datasets. Recent studies in ECGI have highlighted the potential of cross-subject training, provided that signals are standardized across geometries [8]. Standardization strategies such as cylindrical unwrapping of BSP and bullseye projection of EGMs enable alignment between subjects while preserving spatial topology.

In this study, we extend the MARS framework to a generalized setting (M2) using a leave-one-subject-out protocol on in-vivo pig data. We hypothesize that this approach, combined with standardized 3D-to-2D projections of BSP and EGM signals, can maintain reconstruction fidelity comparable to M1 and Tikhonov, while enabling a more robust and clinically transferable solution.

2. Methods

.2.1. Data Sets

Experimental data were obtained from anesthetized, closed-chest, pigs (n=4, 30-40 kg). Epicardial and torso potentials were recorded simultaneously using an elastic "sock" (239 unipolar electrodes, Auckland Uniservices Ltd, New Zealand) and flexible strips attached to the body surface (184 electrodes, BioSemi, the Netherlands). For each pig, recordings were made during sinus rhythm, and pacing left and right endo- , and epicardium sites with the number of pacing locations varying between 5 and 21 depending on the animal. Overall, 70 records were obtained. Upon completion, the heart was arrested and MRI performed. The heart was excised and perfusion-fixed. Epicardial electrode locations were captured with a multi-axis digitizing arm (FARO Technologies, FL). MRI

ISSN: 2325-887X DOI: 10.22489/CinC.2025.367

contrast markers placed on the "sock" and body surface strips were used to registration.

2.2. Inverse Problem Formulation

In this work, we utilized EPs (i.e., electrograms (EGM)) as the bioelectric source model. The relationship between EPs and BSPs is described by a linear equation:

$$y(t) = Ax(t) + n(t)$$

Here, $x(t) \in \mathbb{R}^{N\times T}$ and $y(t) \in \mathbb{R}^{M\times T}$ are the EPs and BSPs vectors at time t, where M , N and T are the number of ECG electrodes, epicardial nodes and number of time samples, respectively. $A \in \mathbb{R}^{M\times N}$ is the forward operator, and $n(t) \in \mathbb{R}^{M\times T}$ represents the measurement noise. The forward matrix A was calculated by solving the forward ECG problem using the boundary element method. For the inverse problem, zero-order Tikhonov regularization was employed as a benchmark for comparison with the MARS method. The regularization parameter was computed using the L-curve method

2.3. Multivariate Adaptive Regression Splines (MARS)

MARS method is a non-parametric regression technique designed to model complex relationships between a dependent variable and multiple predictors. It achieves this by using spline functions to divide the data into different regions and fit local models in each interval. The central idea of MARS is to construct the model through a series of basis functions formed around knots, which act as breakpoints in the predictor space. These basis functions are defined as pairs of reflected splines and can capture both non-linearities and interactions between variables. This part already have been explain in [6-7] .

Several MARS libraries are available. We implemented our methods with the help of "Earth: Multivariate Adaptive Regression Splines R package" and its standalone C version, which can be used in MATLAB.

.2.4. Geometric Standardization and Interpolation

A major challenge for cross-subject training in ECGI lies in the inter-individual variability of electrode configurations and cardiac-torso geometries. Directly pooling BSP and epicardial data across subjects without alignment would introduce inconsistencies in spatial correspondence, preventing the regression model from learning transferable mappings. To overcome this, we applied a spatial standardization procedure to project all subject-specific measurements into a common 2D reference domain.

To enable cross-subject training, torso electrode positions had to be expressed in a common coordinate frame, independent of subject-specific geometry. We implemented a cylindrical projection and unwrapping procedure to convert 3D body surface electrode coordinates into a standardized 2D representation.

First, a cylinder was fitted to the 3D torso surface using principal component analysis to determine its central axis and least-squares estimation to compute its center and radius (figure 1). This fitted cylinder provided a geometrical reference aligned with the longitudinal axis of the torso.

The cylinder and the projected electrodes were then rotated and translated so that their main axis aligned with the global Z-axis and their origin coincided with the manually defined cardiac apex. The rotation matrix was computed using the axis–angle formulation, aligning the PCA-derived vector *V* with the vertical unit vector [0 0 1].

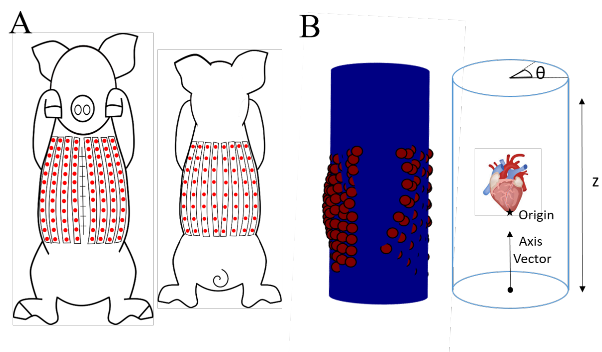

Figure 1: Cylindrical fitting and alignment of body surface electrodes. (A) Experimental torso electrode layout (Red dots) on the pig model. (B) Cylinder fitted to the torso surface, then aligned with the global Z-axis and positioned with its origin at the manually defined cardiac apex.

Next, each 3D electrode position was converted into cylindrical coordinates (θ,ρ,z), where the angular component θ represents the circumferential position around the torso and *z* the vertical coordinate along the heart-torso axis.

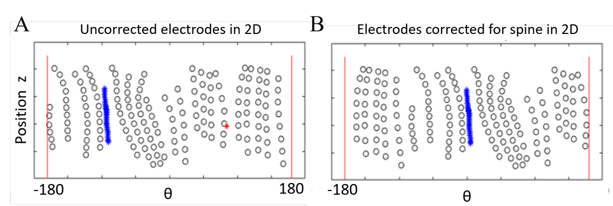

Figure 2: Cylindrical unwrapping and angular correction. (A) Initial 2D projection. (B) Electrodes recentred with the spine as angular origin. White dots:electrode ; the blue line: sternum; red dot: initial spine position; red line: corrected spine position.

Finally, electrodes were unwrapped into a 2D map defined by [θ,z] coordinates and recentered so that the spine corresponded to the angular origin (θ=0), ensuring consistent orientation across subjects (figure 2).

For EGMs, epicardial electrode positions were projected onto a standardized 2D bullseye representation, following an adapted UNISYS approach [8–9]. To preserve spatial topology during flattening, the 3D ventricular surface was first mapped onto a virtual cone whose apex was positioned along the ventricular axis. This intermediate conical projection maintains relative geodesic distances between neighboring electrodes before unfolding the surface into the 2D bullseye map.

After projection of BSP and EGM electrodes into their respective 2D standardized domains, the signals were interpolated onto a uniform grid using ordinary kriging (figure 3).

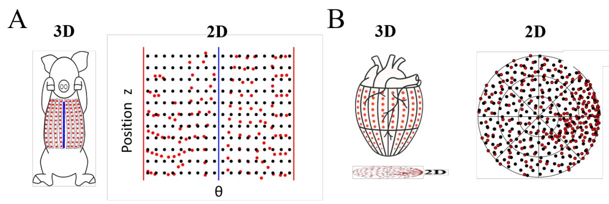

Figure 3: Standardization of electrode positions for cross-subject training. (A) BSP projected to a 2D cylindrical map. (B) EGM projected to a 2D bullseye map. Red: original positions; black: after standardization.

This method was chosen to preserve spatial smoothness and to generate regularly sampled potential maps consistent across subjects.

In ordinary kriging, each unknown potential value ẑ(u) at a grid node u is estimated as a weighted sum of the measured electrode signals $z(s_i)$:

$$\hat{z}(u) = \sum_i w_i(u) z(s_i),$$

where the weights w_i(u) depend on the spatial distance between electrodes and satisfy the unbiasedness constraint $\sum_i w_i = 1$.

Pairwise Euclidean distances were computed between all known electrode positions and between known and target grid points in the 2D domain. These distances were used to form the kriging system, where the covariance between points decreases with distance. Solving this system yields a geometry-specific weight matrix W that can be reused for all time samples and beats $\hat{Z} = WZ$ with Z being the matrix of electrode signals.

.2.5. Data Training

Five beats per pacing site were used for training to balance accuracy and computational efficiency [1].

Two training configurations were tested. In the **s**ubject-specific model (M1)**,** training and testing were performed within the same animal using a leave-one-pacing-site-out protocol. In the generalized model (M2)**,** a leave-one-pig-out strategy was applied, training on three animals and testing on the remaining one.

Training was restricted to the QRS complex of each beat to focus on ventricular activation and avoid interference from baseline or repolarization components. All BSP and EGM signals were normalized beat by beat to zero mean and unit variance prior to model fitting. The MARS implementation (*earth* package, version 5.3.3) automatically selected the most relevant basis functions through internal cross-validation.

.2.6. Data Analysis

Data analysis was performed on all beats for each record. Correspondence between measured and reconstructed electrograms (EGMs) was quantified using the Pearson's correlation coefficient (CC).
The significance of differences among the means of the inverse methods were examined using Friedman test.

3. Results

Preliminary observations suggest that the generalized MARS model (M2) may offer consistent performance across subjects (figure 4).

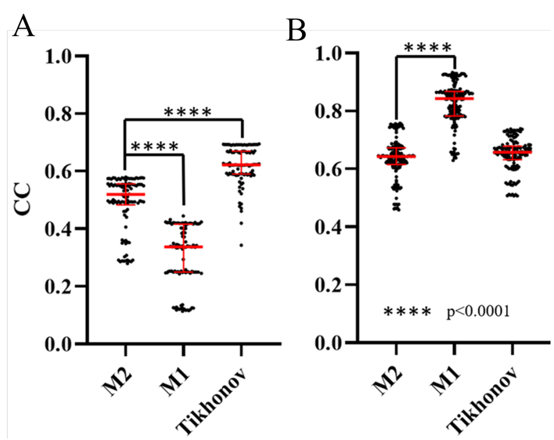

Figure 4: Comparison of electrogram reconstruction accuracy across methods. (A) Pig with limited training data for M1. (B) Pig with extensive training data. **** $p < 0.0001$.

In the case where the subject-specific model (M1) performed poorly due to limited training data (Pig 2), M2 achieved higher accuracy than M1 (median CC = 0.52 [0.48–0.56] vs 0.34 [0.25–0.42], $p < 0.0001$), although it remained significantly lower than Tikhonov (0.62 [0.59–0.67], $p < 0.0001$). Conversely, for the animal with a richer pacing dataset (Pig 3), M1 reached the best reconstruction accuracy (0.84 [0.77–0.89]), while M2 was slightly lower (0.64 [0.60–0.67]) and comparable to Tikhonov (0.66 [0.62–0.70]).

4. Discussion

These results represent preliminary but encouraging evidence that a data-driven regression framework such as MARS can generalize across subjects in ECGI. While the present study included a limited number of animals, the consistent behavior of the generalized model (M2) across different datasets suggests that data diversity, rather than subject-specific optimization, can drive stable reconstruction performance.

Several improvements could further enhance model robustness and generalization. In particular, the spatial standardization step could be refined. In the current implementation, epicardial and BSP potentials were interpolated in the 2D standardized domain, implicitly assuming that electrode distances are preserved after projection. This is not strictly true, and geometric distortions may affect local spatial relationships between electrodes. A more accurate approach would be to perform the interpolation in 3D prior to 3D-to-2D transformation, preserving geodesic distances and minimizing spatial distortion before model training.

Future work will focus on optimizing the spatial standardization procedure, particularly by improving the 3D-to-2D transformation and interpolation steps. This will also include estimating the remaining animals using the generalized framework and potentially augmenting the training set with simulated data to enhance model robustness and variability coverage.

5. Conclusion

This study introduces a generalized, cross-subject version of the MARS-based ECGI framework. Preliminary results demonstrate that the model can reconstruct epicardial electrograms with accuracy comparable to subject-specific training and, in data-limited cases, superior performance. These findings highlight the potential of data-driven approaches for developing geometry-independent, patient-agnostic ECGI methods.

Acknowledgments

This research was financially supported by the National Research Agency (ANR-10-IAHU-04) and (ANR-22-CE17-0023), and the PHC BOSPHORE “INVERSE” grant (47978NM), and Scientific and Technological Research Council of Türkiye (TÜBİTAK-221N175).

References

[1] A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-posed problems*. Washington, D.C.: John Wiley & Sons, New York: V. H. Winston & Sons, 1977.

[2] M. J. M. Cluitmans *et al.*, « In Vivo Validation of Electrocardiographic Imaging », *JACC Clin. Electrophysiol.*, vol. 3, nº 3, p. 232-242, mars 2017, doi: 10.1016/j.jacep.2016.11.012.

[3] L. R. Bear *et al.*, « Forward Problem of Electrocardiography: Is It Solved? », *Circ. Arrhythm. Electrophysiol.*, vol. 8, nº 3, p. 677-684, june 2015, doi: 10.1161/CIRCEP.114.001573.

[4] K.-W. Chen, L. Bear, and C.-W. Lin, « Solving Inverse Electrocardiographic Mapping Using Machine Learning and Deep Learning Frameworks », *Sensors*, vol. 22, nº 6, p. 2331, mars 2022, doi: 10.3390/s22062331.

[5] X. Meng, Z. Li, D. Zhang, and G. E. Karniadakis, « PPINN: Parareal Physics-Informed Neural Network for Time-Dependent PDEs », *Comput. Methods Appl. Mech. Eng.*, vol. 370, p. 113250, oct. 2020, doi: 10.1016/j.cma.2020.113250.

[6] O. Onak, T. Erenler, and Y. Serinagaoglu, « A Novel Data-Adaptive Regression Framework Based on Multivariate Adaptive Regression Splines for Electrocardiographic Imaging », *IEEE Trans. Biomed. Eng.*, vol. 69, nº 2, p. 963-974, fevr. 2022, doi: 10.1109/TBME.2021.3110767.

[7] A. Mombereau, Y. Serinagaoglu Dogrusoz, R. Dubois, and L. R. Bear, « Improved Performance of Data-Adaptive Regression Framework Based on Multivariate Adaptive Regression Splines for Electrocardiographic Imaging », presented at 2024 *Computing in Cardiology* Conference, déc. 2024. doi: 10.22489/CinC.2024.294.

[8] J. Stoks, U. Chau Nguyen, R. Peeters, P. Volders, and M. Cluitmans, « An Open-Source Algorithm for Standardized Bullseye Visualization of High-Resolution Cardiac Ventricular Data: UNISYS », presented at 2020 *Computing in Cardiology* Conference, déc. 2020. doi: 10.22489/CinC.2020.160.

[9] N. Montagne, A. Mombereau, O. Bernus, and L. R. Bear, «Estimating Septum Rotation to Improve Accuracy of Cardiac Biventricular Bullseye Representation Using the UNISYS Algorithm», *Computing in Cardiology*, accepted for publication, 2025. Available: https://cinc.org/2025/Program/accepted/348.html

Address for correspondence:

Amaël Mombereau
Av. du Haut Lévêque, 33600 Pessac
Amael.mombereau@ihu-liryc.fr