Article in Press

Rapid calibration of atrial electrophysiology models using Gaussian process emulators in the ensemble Kalman filter

Mariya Mamajiwala, Cesare Corrado, Christopher W. Lanyon, Steven A. Niederer, Richard D. Wilkinson & Richard H. Clayton

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Rapid calibration of atrial electrophysiology models using Gaussian process emulators in the ensemble Kalman filter

Mariya Mamajiwala*[1], Cesare Corrado[2], Christopher W. Lanyon[1], Steven A. Niederer[2], Richard D. Wilkinson[1,*], and Richard H. Clayton[3,*]

[1]University of Nottingham, School of Mathematical Sciences, Nottingham, NG7 2RD, United Kingdom
[2]Imperial College London, National Heart and Lung Institute, London, SW3 6LY, United Kingdom
[3]University of Sheffield, School of Computer Science and Insigneo institute, Sheffield, S10 2TN, United Kingdom
[*]mariya.mamajiwala@nottingham.ac.uk, r.d.wilkinson@nottingham.ac.uk, r.h.clayton@sheffield.ac.uk

ABSTRACT

Atrial fibrillation (AF) is a common cardiac arrhythmia characterised by disordered electrical activity in the atria. The standard treatment is catheter ablation, which is invasive and irreversible. Recent advances in computational electrophysiology offer the potential for patient-specific models that can be used to guide clinical decisions. To be of practical value, we must be able to rapidly calibrate physics-based models using routine clinical measurements. We pose this calibration task as a static inverse problem, where the goal is to infer spatially homogenous tissue-level electrophysiological parameters from the available observations. To make this tractable, we replace the expensive forward model with Gaussian process emulators (GPEs), and propose a novel adaptation of the ensemble Kalman filter (EnKF) for static non-linear inverse problems. The approach yields parameter samples that can be interpreted as coming from the best Gaussian approximation of the posterior distribution. We compare our results with those obtained using Markov chain Monte Carlo (MCMC) sampling and demonstrate the potential of the approach to enable near-real-time patient-specific calibration, a key step towards predicting outcomes of AF treatment within clinical timescales. The approach is readily applicable to a wide range of static inverse problems in science and engineering.

Keywords: cardiac electrophysiology, atrial fibrillation, radiofrequency ablation, calibration, ensemble Kalman filter

Introduction

The heart is an electromechanical pump, where electrical activation acts to initiate and synchronise contraction. Disruption of the normal sequence of electrical activation and recovery results in a cardiac arrhythmia. Atrial fibrillation (AF) is a common type of arrhythmia that is characterised by irregular and often rapid electrical activity in the atria that is becoming increasingly prevalent[1]. The origin of fibrillatory activity is typically in the left atrium (LA). Radiofrequency (RF) catheter ablation of the left atrial tissue is often the recommended treatment for persistent AF.

Computational models of cardiac electrophysiology have become valuable tools for understanding the mechanisms that underlie normal and abnormal electrical activity in the heart[2]. In order to develop personalised computational models of patient-specific atrial electrophysiology, we must estimate unknown model parameters using patient data. These models have the potential to predict the success of an RF ablation, and hence can be used to guide clinical procedures[3]. These personalised atrial models require two key elements. The first is anatomy, which can be obtained from magnetic resonance imaging (MRI)[4], computed tomography (CT), or electroanatomical mapping systems. The second is a map of material properties that may vary across the cardiac tissue. Traditionally, rule-based methods[5–7] have been used to assign spatial parameter fields such as tissue excitability, conduction velocity, or action potential duration, based on anatomical coordinates, imaging-derived features, or regional differences in structure and function. While these approaches are computationally efficient and interpretable, they provide only point estimates. In our probabilistic framework, these rule-based maps can serve as informative priors, which can then be refined using calibration methods that infer patient-specific parameter distributions from clinical data such as voltage maps, conduction velocities, or pacing measurements[8]. This combination leverages the physiological realism of rule-based assignments while enabling uncertainty-aware, personalized model calibration.

When creating or validating models of the LA from electroanatomical mapping data, there are inherent challenges as the data can be noisy, sparse, and are typically limited to local activation times (LATs) recorded at a few locations within the LA during programmed pacing. Given these data, our objective is to calibrate an electrophysiology model on a patient-specific anatomy using only the measurements that are obtained in a typical routine clinical procedure, ideally in near real-time so that

timely guidance can be given during the procedure. Our focus is developing a fast and efficient method for a probabilistic calibration of the tissue parameters from sparse and noisy observations. This is a critical component in the broader goal of predicting the outcome of radio-frequency (RF) ablation through patient-specific electrophysiological modelling, combined with confidence on these predictions. The tissue parameters that define a patient-specific model vary across the LA and are thus represented as fields. In previous work we have used Gaussian process manifold interpolation[9] and latent Gaussian processes[35] to represent this spatial heterogeneity, and Markov chain Monte Carlo (MCMC) for calibration[10, 35]. However, these approaches are computationally expensive and require measurements of effective refractory period (as a surrogate for action potential duration (APD)) that are difficult to obtain within clinically acceptable timescales. Here, we adapt the ensemble Kalman filter (EnKF)[11], a widely used and robust algorithm for inverse problems and rapid system identification domains[12–14], to achieve this. While the EnKF was originally developed for dynamic systems with time-series data, it has also shown promise in static inference problems. We propose specific adaptations of the EnKF to enable its use when using Gaussian process emulators, and demonstrate its effectiveness on a non-linear exemplar problem and for estimating tissue parameters of the LA. We model the LA as homogeneous and isotropic tissue, and use the phenomenological modified Mitchell Schaeffer model (mMS)[15] to characterise the cell action potential. Phenomenological models such as mMS provide a compact yet physiologically meaningful representation of tissue-level excitability, enabling efficient parameter estimation while retaining key restitution and refractory properties relevant to atrial dynamics. The results are also compared with those from MCMC, obtained at much greater computational cost. Finally, we discuss some advantages and disadvantages of the proposed approach, related work, modifications of the approach, and some directions for future research.

Methods

Our contributions
We present a novel adaptation of the ensemble Kalman filter to solve a static inverse problem in close to real time using Gaussian process emulators. We give an expression for the ensemble Kalman gain matrix for a non-linear measurement function[16] in terms of the Gaussian process emulator (GPE) covariance function, and apply the method to calibrate tissue parameters in the left atrium (LA), and we demonstrate that the posterior mean recovers the ground truth parameter value with credible intervals.

Calibrating with an emulator
Given noisy measurements, $y \in \mathbb{R}^p$, of LATs and/or APDs at certain locations on the left atrial surface, we want to estimate unknown tissue parameters $\theta \in \mathbb{R}^5$, which we assume are constant across the atrium. We let $\pi(\theta)$ denote our prior belief about θ before observing the data, and we assume that

$$y = h(\theta) + \varepsilon, \tag{1}$$

where $h(\theta) \in \mathbb{R}^p$ is the simulated LAT/APD at the same set of spatial locations using tissue properties θ. We assume that $\varepsilon \sim \mathcal{N}(0, R)$, i.e., we have zero-mean Gaussian observation error. Our aim is to estimate the posterior distribution of the parameters given the data:

$$\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta) \tag{2}$$

where the likelihood $\pi(y|\theta)$ is determined by the observation equation (1).

However, evaluating $h(\theta)$ is computationally expensive, making many standard computational strategies unfeasible in the clinic. All inference must be done using a limited and precomputed ensemble of function evaluations $\mathscr{E} = \{\theta_m, h(\theta_m)\}_{m=1}^M$. Given $\mathscr{E}$, we can train a surrogate model for $h(\theta)$ that can then be used in any analysis [17, 18]. We use GPEs, and assume that $h(\cdot) \sim GP(m(\cdot), k(\cdot, \cdot))$, where $m(\cdot)$ and $k(\cdot, \cdot)$ are the prior mean and covariance functions. GPEs are closed under Bayesian conditioning, so that $h(.)|\mathscr{E} \sim GP(\bar{m}(\cdot), \bar{k}(\cdot, \cdot))$ where $\bar{m}$ and $\bar{k}$ are the posterior mean and covariance functions, which are easy to compute from $\mathscr{E}$[19]. Note that $m(\theta) \in \mathbb{R}^p$ and $k(\theta, \theta') \in \mathbb{R}^{p \times p}$. We found no benefit from modelling the covariance between different outputs, and so for computational ease assume independent outputs so that $k(\theta, \theta')$ is a diagonal matrix, see Supplementary Section SM2 for details on the training of GPEs.

Once we have replaced the simulator by an emulator, it is possible to sample from the posterior distribution $\pi(\theta|y, \mathscr{E})$ using sampling methods such as MCMC. However, due to the inherent serial nature of MCMC, this may still take too long for clinical purposes, as typically we must simulate long chains requiring of the order 10^5–10^6 iterations. Our aim is to estimate the posterior in near real-time during a clinical procedure; therefore, computing a fast and robust estimate of tissue properties is essential. To achieve this, we modify the ensemble Kalman filter (EnKF) to solve the inverse problem, and show that this provides an approach of sufficient speed for practical clinical use.

Background: Ensemble Kalman filter for dynamic problems

The ensemble Kalman filter (EnKF)[11] provides a Monte Carlo approximation to the Bayesian filtering problem, i.e., determine the posterior distribution of the state and parameters at time t given measurements upto time t, $\pi(X_t, \theta_t | y_{\{1...t\}})$. For combined state and parameter estimation, we assume the following equations hold:

$$\mathrm{d}X_t = f(X_t, \theta_t)\mathrm{d}t + \sigma_x \mathrm{d}B_t \tag{3}$$
$$\mathrm{d}\theta_t = \sigma_\theta \mathrm{d}W_t \tag{4}$$
$$Y_t = h(X_t, \theta_t) + \sigma_y Z_t \tag{5}$$

X_t is the state, Y_t an observation, and θ_t the unknown model parameter at time t. Equations (3) and (5) describe the process and measurement dynamics respectively, and equation (4) introduces artificial dynamics for the parameters. $\mathrm{d}B_t$ and $\mathrm{d}W_t$ are independent Brownian motion increments, Z_t is a sequence of independent standard Gaussian random variables, and σ_x, σ_θ and σ_y are noise intensity matrices (which may or may not be constant with time). To approximate this evolving posterior, the EnKF employs a Monte Carlo approach in which a finite ensemble (collection) of size N, consisting of independent samples of the joint state and parameter vectors, $[X_0^n, \theta_0^n]_{n=1}^N$, is propagated and updated over time. Starting from this ensemble, the EnKF proceeds iteratively through two stages as follows: Predict: $\begin{bmatrix} \tilde{X}_{t+1}^n \\ \tilde{\theta}_{t+1}^n \end{bmatrix} = \begin{bmatrix} X_t^n \\ \theta_t^n \end{bmatrix} + \begin{bmatrix} f(X_t^n, \theta_t^n)\Delta t \\ 0 \end{bmatrix} + \begin{bmatrix} \sigma_x \Delta B_t^n \\ \sigma_\theta \Delta W_t^n \end{bmatrix}$; $n \in \{1, \dots, N\}$

Update: $\begin{bmatrix} X_{t+1}^n \\ \theta_{t+1}^n \end{bmatrix} = \begin{bmatrix} \tilde{X}_{t+1}^n \\ \tilde{\theta}_{t+1}^n \end{bmatrix} + K_{t+1} \left[Y_{t+1} - h(\tilde{X}_{t+1}^n, \tilde{\theta}_{t+1}^n) \right]$; $n \in \{1, \dots, N\}$

The prediction step is performed using explicit Euler-Maruyama integration of the stochastic process dynamics and at the update stage, the particles are given a correction based on the Kalman gain matrix K (when $h(\cdot)$ is non-linear[16]) and the innovation vector $(Y_{t+1} - h(\tilde{X}_{t+1}^n, \tilde{\theta}_{t+1}^n))$.

Proposed method: EnKF for calibration of a static problem with an emulator

Equation (2) defines a static calibration problem, with a single observation, no state variable, and constant unknown parameter. The EnKF is not directly applicable, but can be modified for this situation [20]. We modify existing EnKF approaches to the scenario where the observation operator, h, is itself an uncertain surrogate for the deterministic simulator.

The absence of state dynamics is straightforward to deal with; it may be ignored and the only process model needed is for the evolution of parameters. Equation (4) is an artificial dynamics introduced to allow to exploration of the parameter space as we assimilate information from the data. If the difference between the posterior distribution and the prior is large, then assimilating the information in a single step (with the static measurement) is challenging as the particles in the ensemble would initially be in the wrong part of parameter space. Instead, we create an artificial time history of measurements by perturbing the original (already noisy) measurement at every time step. This allows for a smoother transition of the ensemble from prior to posterior. The perturbation of the data at each step is essential so as not to over-count the information available in y. The variance of the added noise is the original noise variance multiplied by the number of iterations of the algorithm (see Supplementary Section SM4 for details). Similar schemes for perturbing the measurements have been used in other work[21–23].

Finally, in order to account for the GPE uncertainty within the EnKF framework, we propose a modification of the ensemble Kalman gain matrix to reflect that we are using a stochastic emulator for the measurement function. Although there is work on using a GPE in the process equation[24] equation (3), to the best of our knowledge, there is no variant of the EnKF that accounts for emulator uncertainty in the measurement equation (5). The pseudo-code for our modified EnKF is provided in Algorithm 1. We show numerically that as accuracy of emulator improves, the posterior obtained via Algorithm 1 converges to the solution of the EnKF for a known non-linear function, demonstrating consistency. See Supplementary Fig. SM1. Note that when $\sigma_\theta = 0$, the emulator mean function is replaced with the known measurement function, and the emulator variance is set to 0, i.e. $\bar{m}(\cdot) = h(\cdot)$ and $\bar{k}(\cdot, \cdot) = 0$, Algorithm 1 reduces to the usual ensemble Kalman filter for static problems. In the case $\sigma_\theta = 0$, it has been established[25] that Algorithm 1, with a finite ensemble size and in the limit when time-step tends to 0, yields posterior mean and variance identical to the true posterior, consistent with the behavior of the standard EnKF. We provide an analysis for the effect of non-zero σ_θ in Supplementary Section SM1.

Electrophysiology model

We model the LA electrophysiology using the modified Mitchell-Schaeffer model [15]. We assume homogeneous isotropic tissue, characterised by a monodomain model given by the following set of coupled partial differential equations

---

Algorithm 1 EnKF for Static Calibration with a GPE measurement operator

---

Inputs: Noise covariance matrix R, number of iterations K, process covariance σ_θ, initial distribution $\mathcal{N}_d(\mu_0, \Sigma_0)$, observation vector $Y \in \mathbb{R}^p$.

Initialise: Sample parameters $\theta_0^n \sim \mathcal{N}_d(\mu_0, \Sigma_0)$ for $n = 1, \ldots, N$.

for $k = 0, \ldots, K-1$ **do**

    $\varepsilon_y \sim \mathcal{N}_p(\mathbf{0}, KR)$ ▷ 1. Perturb observation vector

    $Y_{k+1} = Y + \varepsilon_y$

    for $n = 1, \ldots, N$ **do**

        $\varepsilon_\theta \sim \mathcal{N}_d(\mathbf{0}, I)$

        $\tilde{\theta}_{k+1}^n = \theta_k^n + \sigma_\theta \varepsilon_\theta$ ▷ 2. Predict θ

    end for

    $P_{k+1} = \frac{1}{\sqrt{N-1}} \left[\tilde{\theta}_{k+1} - \langle \tilde{\theta}_{k+1} \rangle \right]$

    $H_{k+1} = \frac{1}{\sqrt{N-1}} \left[\bar{m}(\tilde{\theta}_{k+1}) - \langle \bar{m}(\tilde{\theta}_{k+1}) \rangle \right]$

    where $\tilde{\theta}_{k+1}, \bar{m}(\tilde{\theta}_{k+1})$ are matrices with n^{th} column $\tilde{\theta}_{k+1}^n$ and $\bar{m}(\tilde{\theta}_{k+1}^n)$

    $\langle A \rangle$ denotes the matrix with N identical columns, each being the mean of the columns of A

    for $n = 1, \ldots, N$ **do**

        $K_{k+1}^n = P_{k+1} H_{k+1}^T (H_{k+1} H_{k+1}^T + R + \bar{k}(\tilde{\theta}_{k+1}^n, \tilde{\theta}_{k+1}^n))^{-1}$ ▷ 3. Compute Kalman gain matrix

        $\theta_{k+1}^n = \tilde{\theta}_{k+1}^n + K_{k+1}^n [Y_{k+1} - \bar{m}(\tilde{\theta}_{k+1}^n)]$ ▷ 4. Update θ

    end for

end for

$$\frac{\partial v_m}{\partial t} = \nabla \cdot (D \nabla v_m) + h \frac{v_m (v_m - v_{gate})(1 - v_m)}{\tau_{in}} - (1-h)\frac{v_m}{\tau_{out}} + F_{stim} \tag{6}$$

$$\frac{\partial h}{\partial t} = \begin{cases} (1-h)/\tau_{open} & \text{if } v_m \leq v_{gate} \\ -h/\tau_{close} & \text{otherwise} \end{cases} \tag{7}$$

where $v_m \equiv v_m(x,t)$ denotes the normalised trans-membrane voltage field over the atrial surface (i.e. $v_m \in [0,1]$), $h \equiv h(x,t)$ is a gating variable that controls recovery, and $F_{stim} \equiv F_{stim}(x,t)$ is an externally applied forcing/stimulus. We wish to infer $\theta = (\tau_{in}, \tau_{out}, \tau_{open}, \tau_{close}, D)$, where $\tau_{in}, \tau_{out}, \tau_{open}, \tau_{close}$ are cell parameters that characterise the four stages of the trans-membrane potential, and D is the tissue conductivity. Prior parameter ranges are summarised in Supplementary Table SM1. Boundary conditions are no-flux across the tissue boundaries to reflect electrical isolation of the tissue, and we use initial conditions that set all cells to resting membrane potential, and we fix the cell excitation threshold at $v_{gate} = 0.1$. We use openCARP[26] to solve equations (6) and (7) using a finite element method, with a computational mesh generated from imaging data [27]. We chose a relatively large LA geometry, since such geometries are more likely to sustain AF, which is our focus, and in previous work[28] we showed that parameter sensitivities of tissue outputs are largely independent of geometry. We simulated a paced activation consisting of a stimulus applied in the rectangular region shown in Supplementary Fig. SM2 (b)–(c), located near the coronary sinus (CS). We use an S1S2 pacing protocol consisting of 3 S1 stimuli 800 ms apart, followed by a premature S2 stimulus at a 500 ms coupling interval.

We focus on local activation times (LATs), which correspond to the time of rapid depolarisation of the cardiac tissue, as they are one of the few electrophysiological markers that can be directly recorded during electroanatomical mapping. We define the LAT at atrial location x, to be the time at which the voltage at x, $v_m(x,\cdot)$, crosses a threshold of 0.75 from below (i.e., crosses with a positive time derivative). In addition to LAT, we also compute the action potential duration (APD) at each atrial location. We define APD to be the difference between Local Recovery Time (LRT) and LAT, where LRT is the time at which v_m drops back to 10% of the maximum amplitude. APD is difficult to measure in the clinical setting, but effective refractory period (ERP) can be considered as a surrogate although is not always a close approximation [29]. Furthermore, measuring ERP is time consuming as it requires repeating the S1S2 protocol for progressively decreasing values of the coupling interval S2 until failure to propagate, and so these measurements may not always be available.

Our rationale for including APD in the present study was to examine the potential benefit to calibration afforded by the use of ERP measurements in addition to LAT, so that these benefits can be weighed against the additional cost of making them.

Results

We assess the efficacy of the proposed EnKF algorithm to accurately and quickly calibrate in non-linear static settings in a toy problem (Supplementary Section SM3) and in the electrophysiology model, as discussed in the next subsection. The speed of the EnKF approach allows us to investigate the practical identifiability of the tissue parameters using different combinations of S1S2 measurements, enabling us to assess their relative value to guide clinical decisions. In the second part of our results, we conducted a pilot study to predict atrial fibrillation post-ablation using calibrated posterior parameters (using S1S2 observations), which represents the motivating clinical application.

Calibration of tissue parameters of the LA

We evaluate the proposed approach for rapid calibration of tissue parameters of the LA based on noisy measurements of LAT and APD. As a first step, we trained Gaussian process emulators (GPEs) of LATs and APDs as a function of the tissue parameters. Prior ranges for the parameters are given in Supplementary Table SM1[30]. Best practice [18] suggests the design $\{\theta_i\}_{m=1}^M$ used to generate training data $\mathscr{E} = \{\theta_m, h(\theta_m)\}_{m=1}^M$ should be space-filling, such as a maximin Latin hypercube [18]. However, the prior support contains parameter values that produce activation patterns that are not physiologically plausible (for example, APDs larger than 500 ms). Therefore, we designed a series of rejection steps to arrive at the prior support to train the GPE. We first trained a classifier to predict whether a cell model evaluated at θ results physiologically feasible outputs. Because the cell model (rather than the tissue model) is cheap, we can afford to do sufficient simulation to train a classifier with high predictive accuracy. We then created a space filling design in the prior support, and rejected parameters if the classifier predicts nonphysiological cell predictions. We then perform a second layer of rejection based on tissue simulations performed with openCARP on ARCHER2[31] with the S1S2 protocol described in the electrophysiology section (some parameter values gave plausible cell simulations, but implausible simulations across the LA). This resulted in an ensemble of $M = 202$ simulations. See Supplementary Section SM2 for full details.

From the simulations, we extracted the following information: (i) local activation time (LATs) of the third S1 beat; (ii) LAT of the S2 beat; and (iii) action potential duration (APD) of the S2 beat, which we refer to as S1, S2, and APD henceforth. In a clinical setting, activation maps can make recordings at 100s of points on the left atrium when using a fixed pacing, but S1S2 pacing protocols tend to only record LAT at a smaller number of locations. To reflect this we used measurements from 15 spatially well-distributed locations on the left atrium (Supplementary Fig. SM2c) , based on the universal atrial coordinate (UAC) system[32]. The output of the simulator, $h(\theta)$, can thus be considered to be a 45 dimensional vector (3 output types at 15 locations). We split the ensemble $\mathscr{E}$ into training and test sets, and trained independent Gaussian process emulators (GPEs) for each output variable. The resulting GPEs were validated by assessing their predictive accuracy on the unseen test data. In all 45 cases, the coefficient of determination, R^2, which computes the proportion of variance explained, was greater than 95%. See Supplementary Section SM2 for full details.

To assess the practical identifiability [33] of the parameters, and how that changes with the different observation types, as well as the efficacy of Algorithm 1, we conducted a series of 50 calibration experiments. In each, we selected a ground truth parameter and simulation from the ensemble $\mathscr{E}$, and added zero-mean Gaussian noise to the outputs to create synthetic measurements. Calibration was performed separately for each of the following measurement combinations: (i) S1 only, (ii) S1 and S2, and (iii) S1, S2 and APD. For each instance, the ensemble mean of the proposed method at the final iteration was taken as the calibrated parameter estimate. Fig. 2 shows posterior means of the model parameters obtained through calibration compared to ground truth values used to generate the data. It is clear that τ_{in} and conductivity, D, can be identified from the data - albeit with varying accuracy - irrespective of the type of measurements used for calibration. Other parameters, however, were difficult to estimate with S1 measurements alone. The posterior mean for these parameters remains the same as the prior mean, and the posterior variance is dominated by the parameter dynamics. They only become identifiable with the addition of more measurements. This is consistent with global sensitivity analyses of the model[28], where it was shown that τ_{in} and conductivity are the most influential on the S1 output, and other parameters become important for other outputs such as APD.

To quantitatively compare the estimated posterior uncertainty from the EnKF and MCMC, kernel density estimates (KDE) of the marginal posterior distributions are shown in Fig. SM3. Overall, the EnKF variances show good agreement with those of the largest MCMC mode for the identifiable parameters, τ_{in} and D, while for the less well identified parameters, the EnKF has broader posterior uncertainty reflecting weaker data constraints.

To assess the predictive accuracy of the calibrated parameters, we performed S1S2 simulations using the calibrated parameters corresponding to the 150 instances. From these simulations, we extracted the S1, S2 and APD outputs across the entire left atrial (LA) mesh. We then calculated the root-mean-square errors (RMSEs) for the three outputs over the LA mesh (~300,000) nodes. Figure 3 presents the distribution of these RMSEs. We now discuss the three plots in Fig. 3. The first plot presents the RMSE in the predicted S1 outputs. This error is highest when calibration is performed using only S1 measurements. Incorporating S2 measurements leads to a modest improvement, while the inclusion of APD measurements does not yield substantial additional gains. While the addition of S2 measurements enhances the calibration of other parameters

(see Fig. 2), it also improves the estimation of τ_{in} and the conductivity, both of which are influential in determining the S1 output. On the other hand, APD measurements primarily enhance the calibration of parameters that have limited impact on S1 output, explaining the minimal reduction in RMSE. The second figure shows the RMSE in the S2 output. Here, the RMSE is notably high when calibration is based solely on S1 measurements. This is expected, as S2 outputs are influenced by the S1S2 coupling interval, which is not captured by the S1 output alone. Once S2 measurements are included in the calibration process, the RMSEs decrease significantly as shown in the middle plot of this figure. As with the S1 case, the addition of APD measurements further refines the calibration of parameters that do not substantially affect S2 outputs, resulting in only a marginal improvement in RMSE. Finally, the third plot displays the RMSE in the APD output. In this case, a clear and consistent reduction in RMSE is observed as more measurement types are incorporated. The APD is governed primarily by τ_{in}, τ_{out}, and τ_{close}, which are not fully captured when using only S1 data. The inclusion of S2 outputs enhances the calibration of these parameters, and the direct use of APD measurements further improves their estimation. This progressive refinement is reflected in the observed reduction in RMSE.

For clarity, we show the LAT maps corresponding to one of the 50 instances using S1+S2+APD measurements in Fig. 1.

It may be observed that the S2 output (4th column) in the first row has the largest error - this is not surprising as this row corresponds to calibration against only S1. All other errors are negligible. For 2 of the 50 ground truths, we show comparison of results from the proposed approach with those via MCMC, see Fig. 4. Results for other ground truth instances compare similarly. Metropolis-Hastings is implemented using the emcee package in Python with GPEs. The prior distribution used in MCMC is the same as the initial distribution for the propsed approach. We run 10 independent chains, each with 40,000 samples; removing 10,000 samples as burn-in for each chain and then and applying thinning with a factor of 10, we finally have 30,000 samples for each ground truth. Pair-wise scatter plots are shown in Fig. 4. Compared to MCMC or other sequential Monte Carlo methods, the quality of performance of the proposed approach scales more effectively with increasing measurement data. Contrary to expectation, MCMC methods often struggle as more data becomes available. This trend is clearly observed in Fig. 4, where the performance of MCMC deteriorates while that of the proposed method improves as more measurements are used. This can be understood intuitively: as the quantity of data increases, the posterior distribution becomes more concentrated, making it harder for proposal samples to fall within high-probability regions. Consequently, a greater proportion of proposals are rejected, leading to poor mixing and requiring significantly longer chains to adequately explore the posterior. In contrast, the proposed approach applies additive corrections to all ensemble members based on the available measurements, enabling efficient assimilation of information and improved convergence.

The computational cost of the algorithm is $O(NK)$, but the computational time is $O(K)$ if we parallelise. We varied the hyperparameters in Algorithm 1 to assess robustness to these choices. Specifically, we varied the ensemble size N from 200 to 500, the measurement noise intensity from 0.01 to 1, and the number of steps K from 20 to 100. For hyperparameter values in this range, the change in the posterior mean estimate was indistinguishable from the Monte Carlo variability. The computational cost per sample for MCMC also increases with more data because evaluating the likelihood (often via the forward model) is more expensive. On the other hand, adding more measurements does not significantly increase cost per update for the proposed method since the update step involves linear algebra operations (matrix-vector multiplications and inversions), whose cost grows modestly with more measurements. The number of ensemble members usually stays fixed (e.g., 100–500), and all members may be updated in parallel. Furthermore, as more measurements are added, EnKF often converges in fewer iterations because the system is better constrained. Overall, potential for parallelisation and linear update structure make EnKF computationally more scalable with increased data.

Prediction of AF - pilot study

Our objective is to predict whether atrial fibrillation (AF) will be sustained or terminated in a particular patient. AF is a complex phenomenon, and it is difficult to collect directly relevant electrophysiology data. Usually, we only have data from S1S2 or similar protocols, and on the basis of this, need to estimate tissue parameters before making a calibrated prediction about AF termination. AF initiation in the model can be achieved using various methods, such as cross-field stimulation[34] or multiple spiral waves. We used a four-spiral wave initiation protocol based on universal atrial coordinates (UACs) and performed electrophysiology simulations accordingly[32]. From the simulation results, we classify each parameter vector, θ, as either leading to sustained AF or to successful termination. Averaging with respect to the posterior distribution for the parameters, then gives us an estimate of whether AF will sustain for a particular patient, given the available data. Explicitly, if $g(\theta)$ is an indicator function taking the value 1 if AF sustains in the simulation, and is otherwise 0, then

$$\mathbb{P}(\text{AF sustain} \mid y) = \int g(\theta)\pi(\theta \mid y)\mathrm{d}\theta.$$

The challenge is that $h(\theta)$ (the simulator prediction of the S1S2 outcomes) and $g(\theta)$ may have very different sensitivites to θ. Thus, the parameters we can identify from the S1S2 data may not be the parameters we need to know to predict AF termination.

To investigate this, we conducted a pilot study to assess whether parameters calibrated using S1S2 measurements can reliably predict AF behaviour. Two representative cases from $\mathscr{E}$ were selected, and for each, we drew 200 samples from both the prior and posterior distributions. AF simulations were then performed for all samples as well as for the ground-truth parameters. The classification results are shown in Figures 4c and 4d. Figure SM2a illustrates one representative case, where AF is sustained for the ground truth and all posterior samples but terminates for all prior samples, highlighting the predictive capability of the calibrated parameters. Across both cases, at least 94 % of the posterior samples correctly predicted AF termination in agreement with the ground truth, suggesting that despite limitations around parameter identifiability, the proposed approach holds promise for clinical applications such as guiding ablation therapy.

Discussion

This work builds on and extends our previous approaches [9, 10, 35], as well as several established frameworks for Bayesian parameter inference in computationally intensive settings. In particular, our approach relates closely to methods such as Ensemble Kalman Inversion (EKI) [20, 36], the Ensemble Kalman Sampler (EKS)[37], and the broader "Calibrate–Emulate–Sample"[38] paradigm. Below, we situate our approach in the context of these methods, highlighting key differences and contributions.

The proposed algorithm can be viewed as a generalisation of Ensemble Kalman Inversion (EKI) [21], differing in three important aspects. First, it incorporates parameter dynamics via a pseudo-time evolution, whereas standard EKI corresponds to the special case where this evolution noise σ_θ is zero. Second, our method modifies the measurement perturbation strategy by scaling the measurement noise with the number of iterations ($K \times R$), an idea explored in prior work [20, 36]. Third, it replaces the deterministic forward model with a GPE, treating the emulator mean $\bar{m}(\cdot)$ as a surrogate for the true measurement operator and including emulator uncertainty through the kernel $\bar{k}(\cdot, \cdot)$. We tested simplified variants of our method that closely resemble EKI by incorporating the first two modifications, but these typically yielded less accurate posterior means and underestimated posterior variances—consistent with known behavior of EKI. Additionally, we experimented with omitting the GPE covariance term in the Kalman gain; while this works well when emulator uncertainty is negligible as compared to the measurement uncertainty, i.e. $k(\cdot, \cdot) \ll R$, it may introduce a bias when the GPE is less accurate.

Our method also diverges from the "Calibrate–Emulate–Sample" framework [38], where emulator training, calibration (typically using the exact forward model), and sampling (often via MCMC) occur sequentially after data acquisition. This workflow, while flexible, is computationally intensive and not well-suited for time-critical applications. In contrast, our approach can be viewed as "Emulate–Calibrate"—we shift the expensive GPE training phase to an offline stage before measurements are available. The inference step then proceeds quickly using GPE evaluations alone, producing approximate posterior samples directly through Algorithm 1, without a separate sampling phase. This design makes our method practical for real-time or near-real-time applications.

Finally, our work is conceptually related to the Ensemble Kalman Sampler (EKS) [37], which also aims to approximate posterior distributions using interacting ensemble updates. However, EKS is designed for settings where the forward model $h(\theta)$ is available but its derivatives are intractable. In our case, the forward model itself is expensive to evaluate, motivating the use of a surrogate model. We address this challenge by incorporating a GPE into the inference framework, enabling efficient updates while still targeting a Gaussian approximation of the posterior. Despite the different assumptions about model accessibility, both methods aim to sample from the same target distribution, i.e. a Gaussian approximation of the posterior.

Our proposed approach is particularly well-suited for applications requiring real-time or near-real-time predictions. Most of the computational burden is handled offline, where GPEs are trained over the relevant parameter space. Once trained, the inference step becomes computationally inexpensive, enabling rapid predictions as soon as measurements become available. This makes the method ideal for time-sensitive or clinical scenarios. For the problems considered in this work, a single finite element simulation on ARCHER2 takes approximately 15 minutes. Therefore, calibrating with 1,000 simulations (with an ensemble size of 500 over 20 steps) requires not only approximately 250 hours of HPC compute time, but also significant pre- and post-processing effort. This is because the 500 initial simulations must be completed and post-processed before the next 500 sets of parameters can be generated (with the calibration algorithm) pre-processed and run—introducing substantial overhead costs. In contrast, calibrating with GPEs via Algorithm 1 takes less than one minute on a standard CPU (see Tables SM2 and SM3 for detailed timings).

Another strength of the approach is that the ensemble is embarrassingly parallel. Ensemble members can be updated independently, allowing efficient use of computational resources. Convergence is typically achieved within 20–30 iterations, and theoretically, in static settings with a single observation, the algorithm can converge in just one step. This efficient convergence further supports its practical utility in demanding environments. In terms of scalability, the method performs robustly as parameter dimensionality increases. Unlike Sequential Monte Carlo methods, which may suffer from slow mixing and low acceptance rates in high-dimensional spaces—especially in the absence of informative priors—the ensemble Kalman Filter (EnKF) remains computationally feasible. This is due to its additive updates, which do not rely on an accept-reject mechanism.

Although MCMC provides asymptotically exact samples from the posterior, it often faces practical challenges such as poor mixing or entrapment in regions of high GPE uncertainty (see Fig. 4b). Our numerical experiments show that the proposed approach consistently avoids these pitfalls, converging to the correct mode more reliably. A key limitation, however, is that the method assumes the posterior distribution to be approximately Gaussian. In cases where the true posterior is bimodal, the estimated posterior will converge to one of the modes depending on the random initialisation. While the EnKF does not yield exact posterior samples but rather samples from a Gaussian approximation, this is often sufficient for parameter estimation tasks—our primary focus in this work. Another limitation of the present work is the use of a phenomenological model which, unlike full biophysical models, does not capture complex non-linearities. However, it has been shown [39–41] that phenomenological models can capture most important behaviours of cardiac tissue electrophysiology, and we believe that calibration of these simplified models is a reasonable starting point. A logical extension of this work would be to evaluate our method in a biophysically detailed model, using a subset of parameters that are deemed to have the most influence on model behavior[42, 43] , however this is beyond the scope of the present work.

It is important to emphasise that the accuracy of the calibration process (via any method) is fundamentally tied to the fidelity of the GPE. The proposed algorithm identifies parameters such that the discrepancy between GPE-predicted outputs and the observed measurements aligns with the assumed measurement noise. As illustrated in Supplementary Fig. SM1, even if the inferred parameters exhibit some bias relative to the true values, the measurements predicted by the GPE at those parameters remain unbiased—confirming that the calibration is consistent with the trained emulator. Notably, measurements predicted using the original forward model (not shown) do show bias, attributable to GPE approximation error. This bias diminishes as GPE accuracy improves, reinforcing the need for high-quality surrogate modeling in the offline phase.

Finally, note that the current workflow requires us to generate a surrogate model for each new patient, incurring a substantial computational cost to create the patient specific training set. Lanyon et al.[44]show that by transferring information between cardiac patients it is possible to significantly reduce the size of the dataset required to train surrogate models, which would then be straightforward to integrate into the calibration framework described here. However, it is would also be possible to include geometric and tissue heterogeneity directly into the GPE as inputs. For example, atrial heterogeneity can be represented by subdividing the left atrium into five regions corresponding to distinct image intensity ratio (IIR) ranges[27]. Each region can be characterized by its own parameter values, yielding approximately 25 parameters in total to describe the left atrial tissue. For a Gaussian process emulator, we would typically require 250–500 training simulations for parameters of this dimension[45]. In a typical clinical workflow, the time interval between pre-procedural MRI and ablation is approximately 24–72 hours[4], thus making it feasible to perform the several hundred simulations required to enable a surrogate model to be created within clinically relevant timescales.

Conclusions

We have proposed an adaptation of the ensemble Kalman filter to solve a static inverse problem of the type $y = h(\theta) + \varepsilon$. Given noisy measurements, y, the objective is to determine the posterior distribution of θ where $h(\theta)$ is only available as a GPE. With the help of a toy problem, we have shown that Algorithm 1 converges to the original EnKF solution as the emulator uncertainty goes to zero. We then showed via a sketch proof, that the effect of the pseudo-parameter dynamics on the posterior variance is an inflation by $\sigma_\theta \sigma_\theta^T t$ relative to the case $\sigma_\theta = 0$ when the ensemble size goes to infinity and integration time-step goes to zero. The inflation for a finite ensemble size is therefore bound above by $\sigma_\theta \sigma_\theta^T t$. In practice, however, we have shown through various examples, that this effect is negligible.

We applied the proposed approach to a series of synthetic problems of calibration of tissue parameters of the left atrium and summarised the performance in Figures 2 and 3. A specific case is shown in Fig. 1 for clarity. For a couple of instances, we show a comparison of the posterior distributions obtained via Algorithm 1 with those obtained from MCMC, treating it as a gold standard in Fig. 4; the ground truth parameter is also shown for reference. We then performed a pilot study to determine whether AF is sustained or not given measurements only from the S1S2 pacing protocol. As demonstrated with the pilot study, the proposed method shows promise to enable this prediction with reasonable accuracy (about 94% posterior samples predicted the AF behaviour correctly) in real-time as we move the expensive training of GPEs offline.

In this paper we have focused on calibrating homogeneous tissue parameters. A natural next step is to model spatial variation in the tissue parameters. However, there is a trade-off between expressivity of the parameter representation, and our ability to estimate the parameters. This will need to be investigated using pilot clinical data. Such fields can be represented using methods like Gaussian process manifold interpolation[9] or basis expansions informed by imaging, though the optimal choice remains an open question. For example, high-resolution activation time measurements can be used to infer spatially varying coefficients, enabling the estimation of local conduction heterogeneities[46]. Moreover, region-specific conduction slowing—such as in Bachmann's bundle observed even in sinus rhythm[47]—can be incorporated as spatially varying conductivities to guide parameter estimation. Once parameterised, the coefficients can be inferred using Algorithm 1, with only a moderate increase in dimensionality. This is manageable with richer measurement data—clinical settings often provide hundreds of data points,

making the extension both feasible and promising.

The current pseudo-dynamics, modeled as Brownian motion, do not inherently enforce physiological constraints on parameters (e.g., positivity or bounded ranges). While the EnKF update often corrects violations, proposals outside the trained GPE region—such as negative values—can lead to high emulator uncertainty and weak updates, slowing convergence. A promising direction is to incorporate constrained dynamics (e.g., reflecting or truncated Brownian motion) to ensure physically plausible proposals and improve numerical stability. Approaches include Doob's h-transform[48], killed diffusions [49], or manifold-based Brownian motion with tailored metrics[50, 51]. Although the current EnKF adaptation yields rapid and robust parameter estimates with credible intervals for the homogeneous phenomenological case, future efforts should explore a formal convergence analysis and assess more deeply the reliability of uncertainty quantification. This could include examining sensitivity to ensemble size, hyperparameter choices (such as σ_θ), or extensions to heterogeneous tissues.

Another key direction for future work is the design of clinical experiments[52] to improve calibration quality. This includes selecting informative measurements, influenced by stimulation protocols, sensor placement, and tissue properties. In our context, the goal is to identify a pacing strategy that best informs AF-related parameters. Optimising sensor number and spatial distribution—especially in heterogeneous tissue—can enhance parameter identifiability, with methods like sensitivity analysis or information-theoretic criteria offering guidance. Additionally, anatomical factors such as atrial shape, fiber orientation, and wall thickness affect measurements; incorporating insights from virtual cohorts or statistical shape models [53] can improve robustness and generalisability. Although demonstrated primarily in the context of cardiac electrophysiology, the proposed method may be readily adapted to a broad class of static inverse problems across diverse application domains.
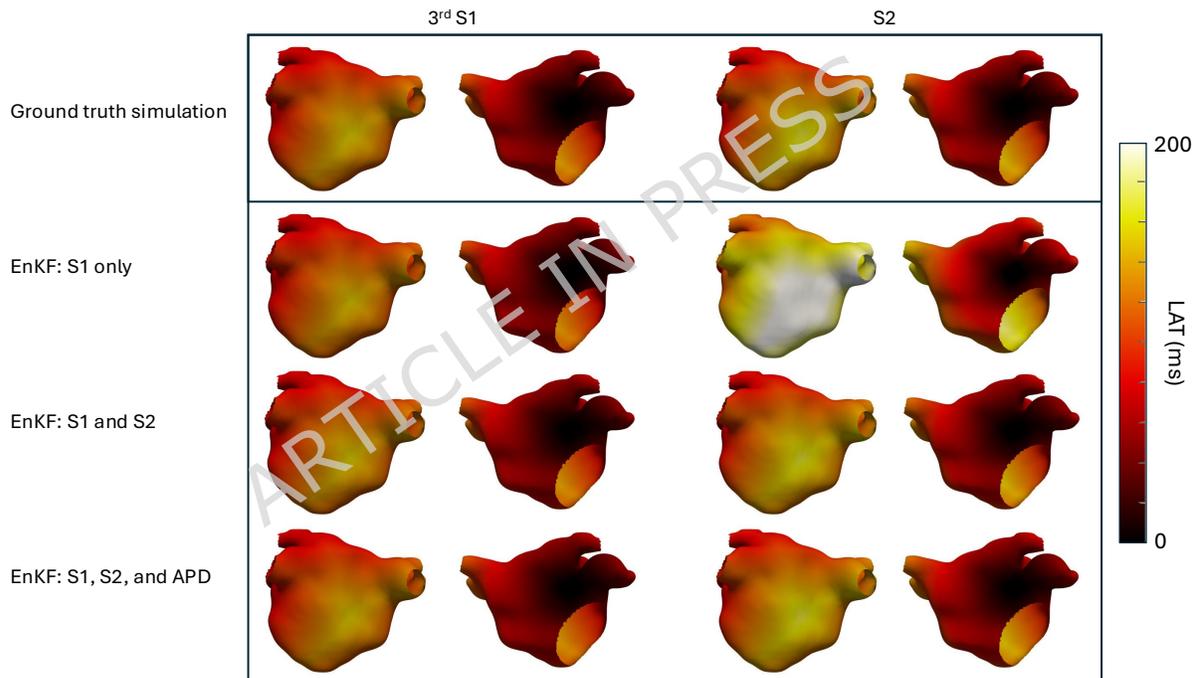

Figure 1. Example local activation time (LAT) maps for the S1S2 pacing protocol. The first two columns show LAT maps for the final S1 beat, and the next two colums show LAT for the S2 beat. The top row shows LAT maps obtained from a simulation with a singla parameter sample; the subsequent rows show LAT maps obtained from simulations using the posterior point estimate of each parameter (i.e., with the ensemble mean after the final iteration of the EnKF) for different combinations of measurements as shown on the figure.
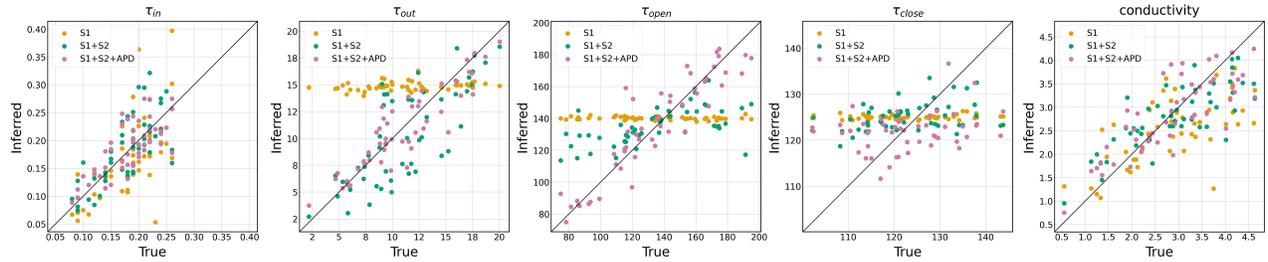
Figure 2. Comparison of ground truth and calibrated parameters for 150 instances of calibration; 50 for each of the datasets used for calibration (see text for details). Each subfigure corresponds to one of the tissue parameters. In each subfigure, the true parameter values are plotted on the x-axis, and the corresponding inferred (calibrated) values on the y-axis. Different measurement types are indicated by distinct colors.
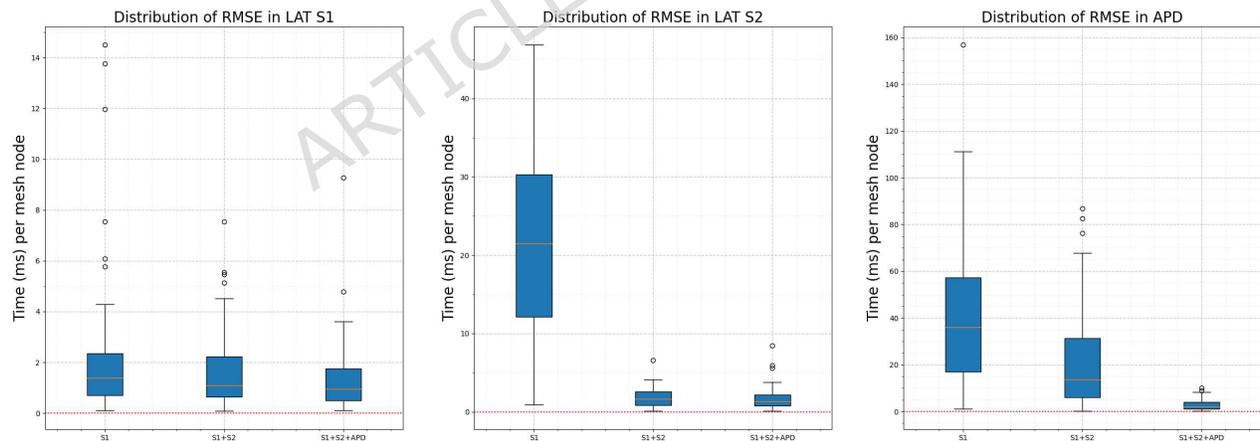

Figure 3. The Figure shows the distribution (over 50 simulations) of RMSEs (over the LA mesh with ~300,000 nodes) of different outputs, viz. RMSEs in LAT S1, LAT S2 and APD in the left, middle and right subfigures respectively. In each subfigure, the three boxplots correspond to 3 calibration scenarios viz., when only S1 (LAT of third S1 beat) measurement is used for calibration, when S2 (LAT of the S2 beat) is used in addition to S1 and when S1+S2+APD (LAT of S1 and S2 beats in addition to the last APD) are used.
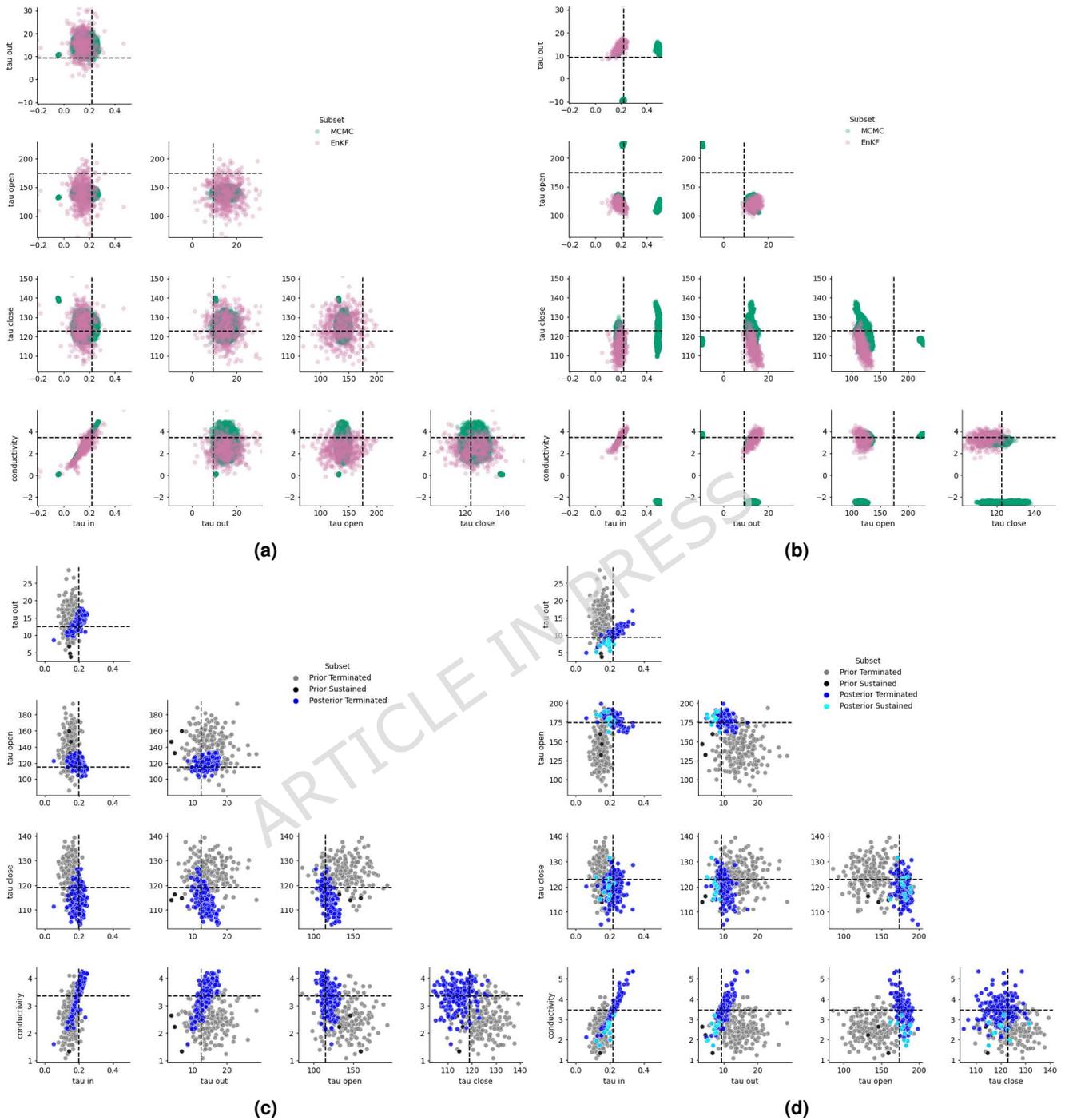
Figure 4. Top row: Comparison of posterior distribution of parameters obtained from MCMC and the proposed approach. Subfigure (a) shows results of calibration against only S1 measurements, while (b) shows those against S1+S2+APD measurements. Bottom row: Results from the AF pilot study. (c), and (d): Two instances for a pilot study on predicting AF from parameters calibrated against the S1S2 protocol. (c) AF is terminated for the ground truth as well as all of the posterior samples; (d) AF is terminated for the ground truth, and all but 11 of the posterior samples.

Data availability

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

References

1. Linz, D., Gawalko, M., Betz, K., Hendriks, J. M., Lip, G. Y. H., Vinter, N., Guo, Y. and Johnsen, S. Atrial fibrillation: epidemiology, screening and digital health. *Lancet Reg. Health Eur.* **37** (2024).

2. Clayton, R. H., Bernus, O., Cherry, E. M., Dierckx, H., Fenton, F. H., Mirabella, L., Panfilov, A. V., Sachse, F. B., Seemann, G. and Zhang, H. Models of cardiac tissue electrophysiology: progress, challenges and open questions. *Prog. Biophys. Mol. Biol.* **104**, 1–3, 22–48 (2011).

3. Niederer, S. A., Lumens, J. and Trayanova, N. A. Computational models in cardiology. *Nat. Rev. Cardiol.* **16**, 2, 100–111 (2019).

4. Boyle, P. M., Zghaib, T., Zahid, S., Ali, R. L., Deng, D., Franceschi, W. H., Hakim, J. B., Murphy, M. J., Prakosa, A. and Zimmerman, S. L. Computationally guided personalized targeted ablation of persistent atrial fibrillation. *Nat. Biomed. Eng.* **3**, 11, 870–879 (2019).

5. Bayer, J. D., Blake, R. C., Plank, G. and Trayanova, N. A. A novel rule-based algorithm for assigning myocardial fiber orientation to computational heart models. *Ann. Biomed. Eng.* **40**, 10, 2243–2254 (2012).

6. Roney, C. H., Solis Lemus, J. A., Lopez Barrera, C., Zolotarev, A., Ulgen, O., Kerfoot, E., Bevis, L., Misghina, S., Vidal Horrach, C. and Jaffery, O. A. Constructing bilayer and volumetric atrial models at scale. *Interface Focus* **13**, 6, 20230038 (2023).

7. Jaffery, O. A., Melki, L., Slabaugh, G., Good, W. W. and Roney, C. H. A review of personalised cardiac computational modelling using electroanatomical mapping data. *Arrhythmia & Electrophysiology Review* **13**, e08 (2024).

8. Nairn, D., Eichenlaub, M., Mueller-Edenborn, B., Huang, T., Lehrmann, H., Nagel, C., Azzolin, L., Luongo, G., Figueras Ventura, R. M. and Rubio Forcada, B. Differences in atrial substrate localization using late gadolinium enhancement-magnetic resonance imaging, electrogram voltage, and conduction velocity: a cohort study using a consistent anatomical reference frame in patients with persistent atrial fibrillation. *Europace* **25**, 9, euad278 (2023).

9. Coveney, S., Corrado, C., Roney, C. H., O'Hare, D., Williams, S. E., O'Neill, M. D., Niederer, S. A., Clayton, R. H., Oakley, J. E. and Wilkinson, R. D. Gaussian process manifold interpolation for probabilistic atrial activation maps and uncertain conduction velocity. *Phil. Trans. R. Soc. A* **378**, 2173, 20190345 (2020).

10. Coveney, S., Corrado, C., Oakley, J. E., Wilkinson, R. D., Niederer, S. A. and Clayton, R. H. Bayesian calibration of electrophysiology models using restitution curve emulators. *Front. Physiol.* **12**, 693015 (2021).

11. Evensen, G. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367 (2003).

12. Moradkhani, H., Sorooshian, S., Gupta, H. V. and Houser, P. R. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Adv. Water Resour.* **28**, 2, 135–147 (2005).

13. Huang, R., Diao, R., Li, Y., Sanchez-Gasca, J., Huang, Z., Thomas, B., Etingov, P., Kincic, S., Wang, S. and Fan, R. Calibrating parameters of power system stability models using advanced ensemble Kalman filter. *IEEE Trans. Power Syst.* **33**, 3, 2895–2905 (2017).

14. Wu, M. and Smyth, A. W. Application of the unscented Kalman filter for real-time nonlinear structural system identification. *Struct. Control Health Monit.* **14**, 7, 971–990 (2007).

15. Corrado, C. and Niederer, S. A. A two-variable model robust to pacemaker behaviour for the dynamics of the cardiac action potential. *Math. Biosci.* **281**, 46–54 (2016).

16. Tang, Y., Ambandan, J. and Chen, D. Nonlinear measurement function in the ensemble Kalman filter. *Adv. Atmos. Sci.* **31**, 551–558 (2014).

17. Kennedy, M. C. and O'Hagan, A. Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63**, 3, 425–464 (2001).

18. Gramacy, R. B. *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences*. Chapman and Hall/CRC (2020).

19. Williams, C. K. I. and Rasmussen, C. E. *Gaussian Processes for Machine Learning*. MIT Press (2006).

20. Iglesias, M. A., Law, K. J. H. and Stuart, A. M. Ensemble Kalman methods for inverse problems. *Inverse Problems* **29**, 4, 045001 (2013).

21. Kovachki, N. B. and Stuart, A. M. Ensemble Kalman inversion: a derivative-free technique for machine learning tasks. *Inverse Problems* **35**, 9, 095005 (2019).

22. Raveendran, T., Gupta, S., Vasu, R. M. and Roy, D. Pseudo-time particle filtering for diffuse optical tomography. *J. Opt. Soc. Am. A* **28**, 10, 2070–2081 (2011).

23. Raveendran, T., Gupta, S., Vasu, R. M. and Roy, D. A pseudo-time EnKF incorporating shape based reconstruction for diffuse optical tomography. *Med. Phys.* **39**, 2, 1092–1101 (2012).

24. Lin, Z., Sun, Y., Yin, F. and Thiéry, A. H. Ensemble Kalman filtering meets Gaussian process SSM for non-mean-field and online inference. *IEEE Trans. Signal Process.* (2024).

25. Schillings, C. and Stuart, A. M. Analysis of the ensemble Kalman filter for inverse problems. *SIAM J. Numer. Anal.* **55**, 3, 1264–1290 (2017).

26. Plank, G., Loewe, A., Neic, A., Augustin, C., Huang, Y.-L., Gsell, M., Karabelas, E., Nothstein, M., Sánchez, J., Prassl, A., Seemann, G. and Vigmond, E. The openCARP simulation environment for cardiac electrophysiology. *Comput. Methods Programs Biomed.* **208**, 106223 (2021).

27. Roney, C. H., Sim, I., Yu, J., Beach, M., Mehta, A., Alonso Solis-Lemus, J., Kotadia, I., Whitaker, J., Corrado, C., Razeghi, O., Vigmond, E., Narayan, S. M., O'Neill, M., Williams, S. E. and Niederer, S. A. Predicting atrial fibrillation recurrence by combining population data and virtual cohorts of patient-specific left atrial models. *Circ. Arrhythm. Electrophysiol.* **15**, 2, e010253 (2022).

28. Mamajiwala, M., Corrado, C., Niederer, S., Wilkinson, R., and Clayton, R. Global sensitivity analysis of left atrial electrophysiology models. In *Computing in Cardiology*, 1–4 (2024).

29. Bode, F., Kilborn, M., Karasik, P. and Franz, M. R. The repolarization-excitability relationship in the human right atrium is unaffected by cycle length, recording site and prior arrhythmias. *J. Am. Coll. Cardiol.* **37**, 3, 920–925 (2001).

30. Corrado, C., Williams, S., Karim, R., Plank, G., O'Neill, M. and Niederer, S. A. A work flow to build and validate patient specific left atrium electrophysiology models from catheter measurements. *Med. Image Anal.* **47**, 153–163 (2018).

31. Beckett, G., Beech-Brandt, J., Leach, K., Payne, Z., Simpson, A., Smith, L., Turner, A. and Whiting, A. ARCHER2 service description. *Zenodo* (2024).

32. Roney, C. H., Pashaei, A., Meo, M., Dubois, R., Boyle, P. M., Trayanova, N. A., Cochet, H., Niederer, S. A. and Vigmond, E. J. Universal atrial coordinates applied to visualisation, registration and construction of patient specific meshes. *Med. Image Anal.* **55**, 65–75 (2019).

33. Wieland, F.-G., Hauber, A. L., Rosenblatt, M., Tönsing, C. and Timmer, J. On structural and practical identifiability. *Curr. Opin. Syst. Biol.* **25**, 60–69 (2021).

34. Pertsov, A. M., Davidenko, J. M., Salomonsz, R., Baxter, W. T. and Jalife, J. Spiral waves of excitation underlie reentrant activity in isolated cardiac muscle. *Circ. Res.* **72**, 3, 631–650 (1993).

35. Coveney, S., Roney, C. H., Corrado, C., Wilkinson, R. D., Oakley, J. E., Niederer, S. A. and Clayton, R. H. Calibrating cardiac electrophysiology models using latent Gaussian processes on atrial manifolds. *Sci. Rep.* **12**, 1, 16572 (2022).

36. Iglesias, M. and Yang, Y. Adaptive regularisation for ensemble Kalman inversion. *Inverse Problems* **37**, 2, 025008 (2021).

37. Garbuno-Inigo, A., Hoffmann, F., Li, W. and Stuart, A. M. Interacting Langevin diffusions: gradient structure and ensemble Kalman sampler. *SIAM J. Appl. Dyn. Syst.* **19**, 1, 412–441 (2020).

38. Cleary, E., Garbuno-Inigo, A., Lan, S., Schneider, T. and Stuart, A. M. Calibrate, emulate, sample. *J. Comput. Phys.* **424**, 109716 (2021).

39. Fenton, F. and Karma, A. Vortex dynamics in three-dimensional continuous myocardium with fiber rotation: filament instability and fibrillation. *Chaos* **8**, 1, 20–47 (1998).

40. Fenton, F. H., Cherry, E. M., Hastings, H. M. and Evans, S. J. Multiple mechanisms of spiral wave breakup in a model of cardiac electrical activity. *Chaos* **12**, 3, 852–892 (2002).

41. Bueno-Orovio, A., Cherry, E. M. and Fenton, F. H. Minimal model for human ventricular action potentials in tissue. *J. Theor. Biol.* **253**, 3, 544–560 (2008).

42. Corrado, C., Roney, C. H., Narayan, S. M., Giles, W. R. and Niederer, S. A. The effect of clinically relevant changes in extracellular electrolyte concentrations on human atrial arrhythmias. *Commun. Med.* **6**, 7 (2025).

43. Strocchi, M., Longobardi, S., Augustin, C. M., Gsell, M. A. F., Petras, A., Rinaldi, C. A., Vigmond, E. J., Plank, G., Oates, C. J. and Wilkinson, R. D. Cell to whole organ global sensitivity analysis on a four-chamber heart electromechanics model using Gaussian processes emulators. *PLOS Comput. Biol.* **19**, e1011257 (2023).

44. Lanyon, C. W., Rodero, C., Quayyum, A., Baptiste, T. M. G., Niederer, S. A. and Wilkinson, R. D. Weaving the digital tapestry: methods for emulating cohorts of cardiac digital twins using Gaussian processes. *Ann. Biomed. Eng.* (to appear).

45. Loeppky, J. L., Sacks, J. and Welch, W. J. Choosing the sample size of a computer experiment: a practical guide. *Technometrics* **51**, 4, 366–376 (2009).

46. van Schie, M. S., Heida, A., Taverne, Y. J. H. J., Bogers, A. J. J. C. and de Groot, N. M. S. Identification of local atrial conduction heterogeneities using high-density conduction velocity estimation. *EP Europace* **23**, 11, 1815–1825 (2021).

47. Heida, A., van Schie, M. S., van der Does, W. F. B., Taverne, Y. J. H. J., Bogers, A. J. J. C. and de Groot, N. M. S. Reduction of conduction velocity in patients with atrial fibrillation. *J. Clin. Med.* **10**, 12, 2614 (2021).

48. Roy, D. and Gorti, G. V. *Stochastic Dynamics, Filtering and Optimization*. Cambridge University Press (2017).

49. Kappen, H. J. Path integrals and symmetry breaking for optimal control theory. *J. Stat. Mech. Theory Exp.* **2005**, 11, P11011 (2005).

50. Hsu, E. P. *Stochastic Analysis on Manifolds*. American Mathematical Society (2002).

51. Mamajiwala, M. and Roy, D. Stochastic dynamical systems developed on Riemannian manifolds. *Probabilistic Eng. Mech.* **67**, 103179 (2022).

52. Rainforth, T., Foster, A., Ivanova, D. R. and Bickford Smith, F. Modern Bayesian experimental design. *Statist. Sci.* **39**, 1, 100–114 (2024).

53. Piazzese, C., Carminati, M. C., Pepi, M. and Caiani, E. G. Statistical shape models of the heart: applications to cardiac imaging. In *Statistical Shape and Deformation Analysis*, 445–480. Elsevier (2017).

Acknowledgements

Funding

Author contributions statement

M.M. conceived and implemented the method. C.C. contributed scripts for atrial electrophysiology simulations, and C.W.L. contributed scripts for MCMC. R.H.C. and S.A.N. advised on cardiac modelling. R.D.W. advised on the mathematical formulation and design of numerical experiments. All authors reviewed the manuscript.

Competing Interests

The authors declare no competing interests.