

Application of Order and Sample Selection in Uncertainty Quantification of Cardiac Models

Anna Busatto^{1,2,3}, Lindsay C R Tanner^{1,2,3}, Karli Gillette⁵, Akil Narayan^{1,4}, Gernot Plank⁵, Rob S MacLeod^{1,2,3}

¹ Scientific Computing and Imaging Institute, University of Utah, SLC, UT, USA

² Nora Eccles Cardiovascular Research and Training Institute, University of Utah, SLC, UT, USA

³ Department of Biomedical Engineering, University of Utah, SLC, UT, USA

⁴ Department of Mathematics, University of Utah, SLC, UT, USA

⁵ Institute of Biophysics, Medical University of Graz, Graz, Austria

Abstract

Simulating the electrical behavior of the heart requires accounting for parameter errors, model inaccuracies, and individual variations in settings, which can all be influenced by user choices or disease conditions. To map the effects of parameter uncertainty, we built on previous findings employing bi-ventricular activation simulations and robust uncertainty quantification (UQ) techniques based on polynomial chaos expansion (PCE) that maps variability in propagation simulations. The PCE approach offers efficient stochastic exploration with reduced computational demands. To ensure reliable results, we focused here on the importance of testing for polynomial order and sample size, aiming to obtain accurate outcomes with minimal computational burden. Order testing involves determining the polynomial degree used for calculating statistics, whereas sample testing pertains to identifying the necessary number and values of the parameters from which the UQ model is estimated. The guide for both steps was to ensure consistency in the results, roughly emulating a convergence analysis. We applied this approach to a bi-ventricular activation simulation using UncertainSCI and quantified the effects of physiological variability in conduction velocity. Our results show that the selection of the appropriate polynomial degree order and sample dataset influences the outcomes of simulations and should be a required step before performing a UQ analysis.

variability, which exists naturally but can also be linked to disease states. We have previously shown the importance of uncertainty quantification (UQ) in evaluating the effects of parameter variability in simulations using the polynomial chaos expansion (PCE) approach. [1, 2] PCE facilitates efficient stochastic exploration with reduced computational demands by employing an emulator for the underlying forward model.

However, there are considerations that need to be taken into account to be confident in the results of UQ; the order of the underlying polynomials and the selection of samples are of particular importance. Order testing refers to the order of the polynomial degree used to fit the model uncertainty and from which to calculate the statistics, whereas sample testing refers to the number and values of points needed to fit the polynomial. One computational goal in all UQ is to achieve accurate results with the smallest possible polynomial degree order and with the smallest sample size. Optimization strategies exist for both aspects, and we have set out to test their utility in the specific context of modeling cardiac propagation.

Building on previous findings, we tested the order and sampling effects in the UQ of bi-ventricular activation simulations based on PCE via UncertainSCI. [3, 4] We evaluated the specific impact of physiological variability in conduction velocity (CV) on simulations of cardiac activation and then tested the impact of polynomial order and sample size on the UQ results.

1. Introduction

Accuracy in simulating the electrical behavior of the heart can be affected by parameter and model errors, as well as individual variability in parameter settings. This variability stems from either user choices or physiological

2. Methods

Uncertainty Quantification: Order and Sample Testing

To conduct this study, we analyzed uncertainty in activation patterns due to variability in CV using the PCE approach. UQ was implemented via UncertainSCI, an open-

Table 1: Conduction Velocity Ranges

Parameter	Minimum Value	Maximum Value
Myocardial Longitudinal (M_L)	50 cm/s	80 cm/s
Myocardial Transverse (M_T)	$1/3 * M_L$	$2/3 * M_L$
Myocardial Sheet (M_S)	$1/3 * M_L$	$2/3 * M_L$
Endocardial Longitudinal (E_L)	200 cm/s	400 cm/s
Endocardial Transverse (E_T)	$1/2 * E_L$	$1 * E_L$
Endocardial Sheet (E_S)	$1/2 * E_L$	$1 * E_L$

source software we have developed. [3, 4] The PCE approach builds a polynomial emulator of the system under study from a discrete number of cases computed from predetermined values of the input parameters. The user controls the order of the polynomial, which, in turn, guides, but does not dictate, the number of sample points required. In this case, we tested orders 5 to 10 for the polynomial emulator and report here the results of orders 5 and 6 because they were sufficient to reach consistent results. For each order, we explored the response to sample selection, using two different numbers of points, which were selected from uniformly distributed values within the boundaries of Table 1. For each order, the first set of points was automatically determined by UncertainSCI based on the number of parameters (in our case, six). We then manually doubled the number of points to see if the performance changed. For both the single and double sample sizes, we tested three randomly selected sample sets. For each sample set, we used 472 points for order 5, 944 for order 5-doubled, 934 for order 6, and 1868 for order 6-doubled. The computed statistical parameters included the mean and standard deviation (STD) of the resulting activation times on the volume. We also tested the residuals, which are a metric of the difference between the values of the dependent variable predicted by the emulator and the observed values.

Simulation Model and Protocol

To compute the propagation sequences, we employed a bi-ventricular geometric model of a porcine heart that we created by segmenting a computerized tomography (CT) scan acquired from the explanted heart from one of our experiments. The result was a finite-element mesh with an average edge length of 800 μm . [5] To mimic the effect of the Purkinje network, we implemented a fast endocardial layer with CV set to four to five times the longitudinal value of the rest of the myocardium. Table 1 contains a summary of the resulting endocardial and myocardial CVs. The activation patterns were determined using the Eikonal technique implemented in the *Cardiac Arrhythmia Research Package* (CARP). [6] We stimulated each beat from a single location on the LV endocardial surface, which was assigned using universal ventricular coordinates

of the geometric model. [7, 8]

3. Results

Fig 1 (top) shows the results from the UQ analysis with order 5 and three sets of 472 (single sample size) and 944 points (double sample size). The results for the mean showed consistency, with the mean maximum activation time, a measure of total ventricular activation, being virtually the same (74 ms) for all cases. However, the amplitudes of STD varied greatly between the single and double-sized samples, as did the spatial patterns of STD. STD values also varied across the different sample sets of the same size, with minimal values from 0.0–4.2 ms and maximal values between 12–15 ms. Spatial patterns of STD also varied, most notably across sample sets in the single-sized samples (second row of Fig 1). The three sample sets of the double sample size are visually more similar, even though sample set one had slightly higher STD values (lighter green) compared to sample sets two and three. The residuals (third row of Fig 1) were lower for the single sample size, with a maximum of 4.6×10^{-4} ; however, they were still low for the double sample size, with a maximum of 3.8×10^{-2} . The mean residuals for the single sample size were small but variable at 1.1×10^{-7} for sample set one, 4.2×10^{-5} for sample set two, and 1.1×10^{-4} for sample set three, with spatial patterns that varied mostly in amplitude. The corresponding values for the double sample size were larger at 1.3×10^{-2} for sample set one, 7.6×10^{-3} for sample set two, and 8.3×10^{-3} for sample set three. The spatial patterns varied only slightly for the double sample size.

We also tested the performance of order 6 with 934 (single sample size) and 1868 points (double sample size), with three sets of samples each, as shown in Fig 1 (lower panel). For the mean, the single sample size showed very small variation, with minimum values ranging between -1.1 ms and 1.9 ms and a maximum between 74 ms and 75 ms. The double sample size yielded similar amplitudes and very consistent spatial patterns.

The STD values and patterns showed clear differences between the single and double sample sizes. For the single

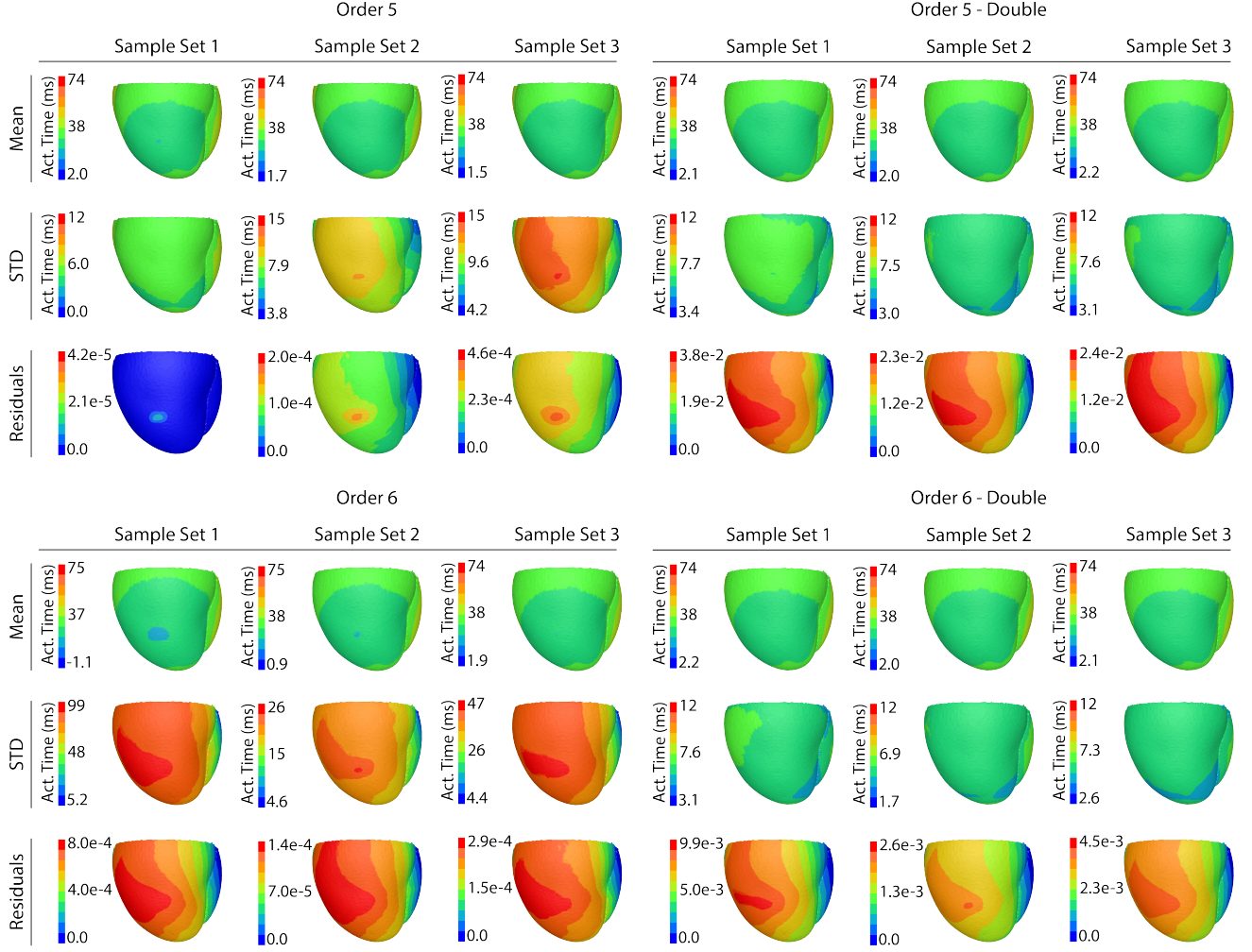


Figure 1: Examples of order and sample testing for left ventricular endocardial stimulation: order 5 (top) and order 6 (bottom). The standard statistics (three columns on the left) were calculated from a total of 472 points for order 5 and 934 points for order 6. The statistics for double sample sizes (three columns on the right) were calculated from a total of 944 points for order 5 and 1868 points for order 6. The first row shows the mean, the middle row shows the standard deviation (STD), and the bottom row shows the residuals. The results are visualized on the left ventricular epicardial surface.

sample size, the minimum STD ranged between 4.4 ms and 5.2 ms; however, the maximum was highly variable, ranging between 26 ms and 99 ms; visual differences between the three sample sets are evident in Fig 1 (bottom). In particular, even though the color pattern appears to be similar across the three, the highly varying maxima indicate a large variation in activation pattern. For the double sample size, the STD had a minimum ranging between 1.7 ms and 3.1 ms and a maximum of 12 ms. Visually, they follow a similar pattern, as shown in Fig 1 (bottom). The residuals, once again, were overall lower for the single sample size, with a maximum of 8.0×10^{-4} ; however, they were still low for the double sample size, with a maximum of 9.9×10^{-3} . The mean residuals for the single sample

size were 2.8×10^{-4} for sample set one, 4.8×10^{-5} for sample set two, and 1.1×10^{-4} for sample set three. For double sample size, they were 3.0×10^{-3} for sample set one, 6.9×10^{-4} for sample set two, and 1.4×10^{-3} for sample set three.

4. Discussion and Conclusions

The aim of this study was to show the importance of selecting the correct polynomial order and sample size to obtain accurate and consistent results in UQ. We explored the specific application of cardiac propagation, but we expect the results to be generally applicable. [3] We explored a large range of polynomial orders, but here we focused on

orders 5 and 6 because they were deemed to provide a good representation of the overall results, with particular interest in their precision and computational overhead. Other applications will likely yield different specific values but will also require a similar exploration of parameters. Our results indicated complex relationships between sample size and stability of the results. In particular, inconsistencies in results appeared for both orders 5 and 6 with single sample sizes, suggesting that the minimal number of points generated by UncertainSCI was inadequate in this type of simulation. In both orders, increasing the number of points yielded reduced STD values and made them more stable across different sample sets. The benefits came at a small cost in the amplitudes of the residuals. Although the residuals increased with a higher number of points, they were lower with the double sample size of order 6 compared to the double sample size of order 5.

Our findings suggest that, in this specific simulation setup, order 6 with double the points is the best option to obtain consistent results across the statistics analyzed. The mean varied by only 0.2 ms across the three sample sets, whereas the STD varied by 1.4 ms. Both mean and STD also showed consistent spatial distributions across the three repetitions (sample sets). Order 6 showed slightly better consistency of the spatial patterns over order 5, mostly when comparing it with the performance observed with superior orders, suggesting it was the more trustworthy option. Furthermore, lower residuals for order 6 with double sample size could indicate better performance compared with order 5 with double sample size.

The results also differed among sample sets, *i.e.*, across random selections. The variations for both orders were small as long as they came from the doubled number of points. Hence, one could select any of the three sample sets and be confident in the results. Given the opportunity, the best selection would be the one that produces the lowest residuals, sample set two for order 6 double, in this study.

Limitations of this study include using a single cardiac geometry and a single stimulus site in the analysis. The specific results are also tied to the underlying model and the parameters targeted in the UQ analysis, in this case, the six anisotropic CVs.

Other applications of UQ will require adjustments in the polynomial order and testing sample selections, which will add computational time to the UQ analysis depending on the number of points and parameters, but our findings underscore the importance of carefully selecting these parameters to ensure reliable and robust results.

Acknowledgments

Support for this research came from the Center for Integrative Biomedical Computing (www.sci.utah.edu/cibc), NIH/NIGMS grants P41 GM103545 and R24 GM136986, NIH/NIBIB grant U24EB029012, and the Nora Eccles Harrison Foundation for Cardiovascular Research.

References

- [1] Busatto A, Rupp LC, Gillette K, Narayan A, Plank G, MacLeod RS. Capturing the influence of conduction velocity on epicardial activation patterns using uncertainty quantification. In 2023 Computing in Cardiology (CinC), volume 50. IEEE, 2023; 1–4.
- [2] Rupp LC, Busatto A, Bergquist JA, Gillette K, Narayan A, Plank G, MacLeod RS. Uncertainty quantification of fiber orientation and epicardial activation. In 2023 Computing in Cardiology (CinC), volume 50. IEEE, 2023; 1–4.
- [3] Narayan A, Liu Z, Bergquist JA, Charlebois C, Rampersad S, Rupp L, Brooks D, White D, Tate J, MacLeod RS. UncertainSCI: Uncertainty quantification for computational models in biomedicine and bioengineering. *Comput Biol Med* 2023;152:106407.
- [4] Tate J, Liu Z, Bergquist JA, Rampersad S, White D, Charlebois C, Rupp L, Brooks DH, MacLeod RS, Narayan A. Uncertainsci: A python package for noninvasive parametric uncertainty quantification of simulation pipelines. *JOSS* 2023;8(90):4249.
- [5] Zenger B, Good W, Bergquist J, Burton B, Tate J, Berkenbile L, Sharma V, MacLeod R. Novel experimental model for studying the spatiotemporal electrical signature of acute myocardial ischemia: a translational platform. *J Physiol Meas* Feb 2020;41(1):015002.
- [6] Neic A, Campos FO, Prassl AJ, Niederer SA, Bishop MJ, Vigmond EJ, Plank G. Efficient computation of electrograms and ECGs in human whole heart simulations using a reaction-eikonal model. *J Comput Phys* 2017;346:191–211.
- [7] Bayer J, Prassl AJ, Pashaei A, Gomez JF, Frontera A, Neic A, Plank G, Vigmond EJ. Universal ventricular coordinates: A generic framework for describing position within the heart and transferring data. *Med Image Anal* 2018;45:83–93.
- [8] Gillette K, Gsell MA, Prassl AJ, Karabelas E, Reiter U, Reiter G, Grandits T, Štern Darko, Urschler M, Bayer J, Augustin CA, Neic AV, Pock T, Vigmond EJ, Plank G. A framework for the generation of digital twins of cardiac electrophysiology from clinical 12-leads ECGs. *Med Image Anal* 2021;.

Address for correspondence:

Anna Busatto
University of Utah
72 Central Campus Dr, Salt Lake City, UT 84112
annabusatto@sci.utah.edu