Dear Dr. Kiss,

Thank you for your letter and opportunity to revise our manuscript “The Inverse Problem for Cardiac Arrhythmias”. The comments made by the reviewers have been very helpful in constructing what we believe to be an improved version of the original manuscript. Notably, we have (i) conducted a statistical test using Akaike’s Information Criterion to assess the relative performance of each model and (ii) updated Figure 6 to include simulations from the other models, which helps to compare and contrast their output.

Attached is our response to the reviewer comments and description of the changes made to the manuscript.

Thank you again for considering our manuscript for publication in Chaos: An Interdisciplinary Journal of Nonlinear Science. We look forward to hearing from your office.

Best wishes,

Thomas Bury
Postdoctoral researcher
University of McGill

Referee: 1

This work presents techniques for developing mathematical models for premature ventricular complexes (PVCs). Analysis of data from long recordings is used to estimate the parameters of the models, and then the simulated rhythms of the model are compared against the data from patients. The manuscript illustrates the techniques with the analysis of data from a patient with frequent PVCs, many of them in the form of bigeminy.
The work is original and interesting. The manuscript is very well written. Only some comments need to be addressed, as described below.

Thank you for your helpful comments and taking the time to review our manuscript.

1. Page 2, first paragraph of Section II. Please mention if the identification of either normal beats or PVCs, which was performed automatically, had any manual supervision by a trained technician or physician. Also, please comment if all PVCs had only one morphology. If PVCs had more than one morphology, a brief comment in the Discussion section may be needed.

All ectopic beats were identified and categorised by a technician in concert with software template recognition software and only then automatically counted. This is reviewed by a physician and confirmed. The PVCs were of a single morphology. We have added this information on Line 115.

2. Page 3, second column, first paragraph. The curves reflecting dependencies on t_s(s) and VN(s) were performed with linear functions instead of exponential ones, as the authors mentioned. Could this choice of linear functions influence the performance of the models? Please comment on this in the manuscript.

It is true that an exponential function is often used to model the conduction time as a function of recovery time (see e.g. Sun et al. J. theor. Biol. 1995). However, given the remarkably linear relationship between the VV and VN intervals (R^2=0.98 in Fig 2d), we do not believe that using an exponential function here would improve the performance of the models. We have added a comment to this end on Line 200.

3. Page 3, second column, second paragraph. Does the noise term in equation 4 have a mean value of zero?

Yes it does, we have added this information.

4. Page 5, first column, last paragraph, says: "All parameters are as in the original model except for the following which are adjusted by a multiplicative factor to facilitate EADs". Then, the adjustments are described. Please mention how the authors decided on each of these adjustments.

The change in each parameter value is based on a physiological understanding of the effect of channel conductances on afterdepolarizations (Guo et al. J Cardiovasc Pharmacol 2011). We use parameters that were demonstrated by the original authors of the model (Tomek et al. ELife, 2020) to produce EADs. This information is added on Line 339.

5. Page 6, first column, last paragraph. The patient had an ablation procedure that eliminated the abnormal rhythm and its symptoms. Could you please comment on the follow-up time of the case regarding the outcome of the ablation procedure? In some patients, the arrhythmias may return not from the exact anatomical origin but from other susceptible regions, depending on the mechanism.

During the 45-minute waiting period at the conclusion of the procedure, 0 PVCs were noted at rest, with isoproterenol, or with up to triple ventricular extra-stimuli. This was a dramatic change compared to pre-ablation. This patient was lost to clinical follow up after the day of the procedure; however, this is irrelevant for the purposes of the manuscript. The important point is that we showed that there was a unifocal source for the PVCs that was confirmed by elimination of the source during ablation. We have clarified what we mean 'elimination of the abnormal rhythm' on Line 413.

Referee: 2

Summary
In this study by Bury et al., the authors develop three mathematical models for Premature Ventricular Complexes (PVCs) based on the morphological timings of a single patient's ECG. The proposed models are based on three hypothesized mechanisms for PVCs: Reentry, enhanced automaticity, and triggered activity. Each of these mechanisms are described and understood through different variables such as ECG event-timings, phase reduction, and action potential timings. These mechanisms are then rewritten in terms of the ECG event-timing variables to allow a fair comparison among the models. The mechanistic functional relationship is based on event-timing linear regressions of event correlations. As a result, the constructed models provide a qualitative and statistical description of PVC's events. Each of the models were found to fail 58%, 55%, and 53% of the time. The authors propose improving the model by considering time-dependent parameters, longer time series, and a better mechanistic description of PCVs.

Thank you for your helpful comments and taking the time to review our manuscript.

Important observations
1. Can the analysis be extended to more than one patient? It seems that the characteristics of this patient were not the average ones regarding PVCs and the authors constantly remark this fact to justify their analysis.

This analysis is applicable to patients who show a significant number of PVCs in a bigeminal rhythm (alternating sinus and ectopic beats). Patients who show very frequent PVCs (>30%) will typically fall into this category. In our cohort of 33 patients, 5 fall into this category, and so it is not particularly uncommon.

2. The authors mention they use least-squares for their linear regressions, yet they do not specify how they calculate the determination coefficient R2. When calculating the correlation plots and the functional relationships of the ECG event-timings, have the authors used the adjusted R2 coefficient? It adjusts for the number of predictor variables. As an alternative, have they applied a L1 regression as opposed to L2 regression (least-squares) to avoid any biased fits from outliers when considering long time series?

We are computing the R2 value as the square of the Pearson correlation coefficient r, where

$$r = \frac{\Sigma(x_i-\hat{x})(y_i-\hat{y})}{\sqrt{\Sigma(x_i-\hat{x})^2\Sigma(y_i-\hat{y})^2}},$$

x=VN and y=VV. We have noted this in Appendix A, Line 538. When we do linear regression, there is only a single predictor variable (VN), so using the adjusted R2 value would make no difference.

Thanks for the suggestion to consider L1 regression. We have opted to stick with L2 regression as the parameters can be obtained analytically without having to use local optimisation techniques, and our data is preprocessed to remove severe outliers.

3. Given that some values such as the coupling interval have an uncertainty, how do the authors determine the effects of error propagation in their models?

The sample rate of the Holter recording is 180 Hz, resulting in a measurement error of +/- 2.8 ms for the timing of each beat, an error that is 0.6% the size of the average coupling interval (467 ms). So this error is relatively small, and we do not expect it to have an effect on our findings. Moreover, there is no 'error propagation' as such, since the models are not iterative.

4. The assumption that the APD-DI restitution curve and the CV-DI (dispersion) curve are linear is a strong one for a general case. Do the authors justify this assumption with data or is this an assumption to simplify the mathematics?

The assumption of linearity for the relationship between the VV and VN interval and the relationship between the NV and t_s interval are based on what we observe in the data. The assumption of linearity between EAD latency and basic cycle length is based on observation from simulations of the ionic model. Indeed, these assumptions simplify the mathematics, but we feel they are also justified based on these observations.

5. The number of segments in the clinical data that are significantly different from the output of the model simulation are 58% (reentry), 55% (modulated parasystole), and 53% (triggered activity). These numbers are in the same order of magnitude, making the assertion that "the triggered activity model is the more successful" is somewhat questionable. Could the authors produce a statistically stronger conclusion be written by considering more patients or improving the model assumptions?

This is a good point - we acknowledge that there are a significant proportion of 5-minute segments that are not well explained by each of the models proposed. This is likely due to the non-stationarity of the data, and the fact that we fix the model parameters for the whole record. Using time-varying parameters would likely improve performance, but it is not clear which parameters in each model should be varying–something to investigate in future work.

Our first approach to evaluating the performance of the models does not compare the performance of each model on each given segment. In order to do this, we now use Akaike's Information Criterion (AIC), a statistical measure that determines the relative goodness between a set of models on a dataset. Upon computing the AIC scores for each model on each segment of the clinical data, we find that out of the 206 5-minute segments, the model for triggered activity is best fitting model in 138, the model for modulated parasystole in 48, and the model for reentry in 20—a demonstration that, out of the three simple models considered, triggered activity is the most plausible. We present this analysis starting on Line 385.

6. If the success probability of the models is 50% or less, would the model success be considered random? Please ask the authors to convince us otherwise.

That depends on what you mean by random! A random model would likely be successful on 0% of the segments. A model that is successful on 50% of the data is successful half the time. We believe this is impressive for a simple model with a few fixed parameters attempting to fit to 20 hours of Holter recording.

7. No figures for the first two models were presented that support the model's success. Could the authors show more evidence that supports their concluding remarks?

We have updated Figure 6 to show output from the first two models as well as the model for triggered activity, which helps demonstrate the qualitatively similar behaviour between the patient data and the triggered activity model, that is not displayed by the other two models.

8. The authors discuss the building block of the models and their accuracy. Can they show any of the model's predictive capabilities?

Testing the predictive capabilities of the model would be interesting, but not necessary for the purposes of this paper. This could be pursued in future work.

9. The models by construction are linear in VN (coupling interval). In the manuscript the authors suggest that the model could be improved by considering time-dependent parameters, but there is no discussion about including nonlinear relationships among the event-timings? Why are the authors pursuing the former strategy to improve the model?

From analysing the data, we have found remarkably linear relationships between VV and VN within a given 5-minute segment. However, this linear relationship (slope and intercept) appears to change in different

segments. It is for this reason that we believe time-dependent parameters will play an important role. That said, nonlinear relationships between these variables could also be a factor, which we now allude to on Line 504.

Minor errors
1. Figure 4 (blue) was never referenced.

Thanks, the blue trace in Figure 4 is now referenced.