Dear editor CharLotte Krawczyk,

Thank you very much for your comments. In order to include your observations, we include the two paragraphs in the discussion section. Moreover we change the name of subsection 2.2.
We are very grateful to you and the reviewers for helping us improve the quality of our work.

Following we include the two inserted paragraphs.

1. The results are sensitive to the size of the domain. An exhaustive parametric analysis using machine learning techniques to classify the synthetic series as function of the input parameters (the size N, P, and πfrac) was carried out in Monterrubio-Velasco et al. (2018a). In Figure 17 (taken from Monterrubio-Velasco et al. (2018a)), we show the mean error of three different ML classification algorithms (Random Forest, Supported vector machine, and Flexible discriminant analysis), as a function of the domain (grid) size. The figure shows that as the grid size is increased, the classification error decreases, meaning that large grid sizes allow us to distinguish among the different properties. In other words, for small grid size, the difference is indistinguishable, while larger grid sizes are able to capture the differences. We observe the results using as classification two input parameters P (in red) and πfrac (in blue). When we use the P parameter, we observe that the size domain has to increase in order to reduce the mean classification error, and it becomes minimum for $N \geq 300$. On the other hand, if we want to classify the synthetic catalogs considering frac, the figure shows that the error classification reaches a minimum value for lower grid sizes $N \geq 200$. So, if we consider the case of $P = 0$, and the classification is based on πfrac then a proper grid sizes used to model aftershocks, including faults, is for $N \geq 200$. We can confirm that an optimization of the parametric search using classification machine learning techniques can be very useful in this stochastic model.

2. Considering the example of Northridge our results suggest that the best combination of parameters to approximate to real cases, depends on the minimum magnitude of the real catalogues, as shown in Table 4. Related with the completeness magnitude, Davidsen and Baiesi (2016), define the Short Term Aftershock Incompleteness (STAI) as a phenomenon that arises from overlapping wave-forms and /or detector saturation, such as events that are missed in the coda of preceding ones. One important consequence of STAI is an increase in the local magnitude of completeness, since small events are not well recorded. It is worth noting that in this work we are not analyzing the STAI phenomena because we are not explicitly modelling this process. We use the Northridge catalog obtained by the Southern California Seismic Network (SCSN), and we analyze it as a "final" catalog. In our statistics and analysis applied to the real catalog, we consider different magnitude cut-offs, as shown in Table 3. The cut-off magnitude is not related with the time. On the other hand, it is noteworthy that our model is not affected by the STAI, because this phenomenon arises from overlapping wave-forms, and in our approach we are not considering explicitly this physical process. To modify the minimum magnitude in the synthetic catalogs we only filter the events with small rupture areas.