Point-by-point response to Reviewer #2

Florian Ehmele on behalf of the co-author

July 09, 2018

Thank you very much for your work and the useful and valuable comments that helped to improve the scientific quality of our manuscript. Please find below our reply to the individual points.

This is a difficult paper to read – I think some rewriting and tightening would help exposition and the reader understand what the main contributions are.

We understand the reviewer's point that the manuscript may be difficult to follow. In the revised version of the manuscript, we try to improve the readability by re-organizing the sections and by tightening / deleting details that are not that important. We will also add a flow chart showing the different components of the model as suggested by Reviewer #1 to better understand the model's different components and their links.

In terms of the scientific problems, my major concerns can be summarized by the following comments:

1) How can you be sure the model isn't overfit? There are numerous parameters and features, and a seemingly exhaustive parameter estimation method is used, but shouldn't there be a cross-validation study where training data are used to fit the model, and held-out testing data are used to validate the goodness-of-fit?

We don't really see a conflict with overfit or "overengineering". All four components of the model (orographic, background, frontal, embedded-convection) are physically-based processes related to vertical lifting on different scales. The parameter estimation mainly is required for the dynamical core (wave dynamics) and the microphysics. For clarification, we will add a comment on this in the conclusion section. Considering the question of cross-validation, we fully agree. However, this was already done. The model is trained for a set of approx. 100 events and then driven stochastically over 10.000 events. The stochastic simulations, the main purpose of the model, are evaluated against the observations using different statistical quantities. To avoid any confusion, we will highlight this in the manuscript.

2) Lack of comparison against a simpler model. There are many moving pieces in this model; which components are giving the most improvement? In particular, it would be helpful to consider simpler versions of the model and compare their relative performance in simulation, this would help the reader understand which contributions are the most important and where future research may focus.

We agree that adding a comparison to simpler versions or to other models would highlight the potential and skill of our approach (this point was also recommended by reviewer #1). As to the best of our knowledge there is no comparable stochastic model available, we will use COSMO-CLM reanalysis instead, but focusing on historic events. We will split the four-part Figure 12 into two Figures (new 13 and 14), one for the median and one for the 90^{th} percentile and add the corresponding statistics for both cases using the reduced stochastic model (rSPM, basic model) and reanalysis data performed with the COSMO-CLM model. We same will apply for Figure 13 (new Fig. 15). With those additions the improvements of the model should become clearer.

3) How well are spatial correlations maintained in the model? Spatially aggregated statistics like max, min and means are validated, but what about raw correlations?

We are not sure what the reviewer exactly means with “raw correlations”. The presented analysis of maximum and minimum values in Figure 13 is not spatially aggregated, but show max/min values at each grid point in the model domain. The median and percentile values presented in Figure 12 as well as the difference between model and observations for different return periods shown in Figure 14 are grid point-based statistics and not spatially aggregated. As the SPM is event-based and not designed for continuous simulations of extreme events it is not possible to correlate “time series” between simulations and observations for the different grid points. Considering wave dynamics in combination with an FFT algorithm, a spatial conjunction is already given.