Geoscientific
Model Development
Discussions

Open Access

Interactive comment on "Application of all relevant feature selection for failure analysis of parameter-induced simulation crashes in climate models" *by* W. Paja et al.

Anonymous Referee #1

Received and published: 25 September 2015

This manuscript describes a study that used machine learning methods to analyze climate simulation failure (model crash) caused by perturbations in uncertain model parameters. The dataset and the goal of analysis were the same as those of Lucas et al. (2013), but different methods were applied which led to the conclusion that some of the parameters deemed important by the analysis of Lucas et al. (2013) were redundant or irrelevant.

Simulation failure analysis is a relevant topic for climate model development, and a more accurate identification of important parameters is beneficial. Hence the results of this study are potentially useful. On the other hand, I would recommend a serious revision so that the manuscript can be made more informative - and the messages

more convincing - for climate model developers and users.

My main difficulty with the manuscript is that it might not have been written with geoscientific model developers as the target audience. As such, I wonder whether the manuscript is more suited for a statistics or computer science journal.

From the perspective of a climate modeler, I think the manuscript does not provide sufficiently detailed descriptions of the methods and analysis procedure to allow many readers of GMD to reproduce the results or apply the same methods to analyze other datasets.

Comparing this study's results to those of Lucas et al. (2013), it is worth noting that the 3 strongly relevant parameters identified in this study were listed among the top-4 important parameters by Lucas et al. The benefit of this study's methods thus seems marginal. It would also be interesting to know what price one has to pay, in terms of algorithm complexity and computing time, in order to do the cross-validation and estimate the statistical uncertainty of the results.

Interactive comment on Geosci. Model Dev. Discuss., 8, 5419, 2015.