Geosci. Model Dev. Discuss., 8, C2730–C2732, 2015
www.geosci-model-dev-discuss.net/8/C2730/2015/

Geoscientific Model Development Discussions
Open Access

GMDD
8, C2730–C2732, 2015

Interactive Comment

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 #2

Received and published: 29 October 2015

Summary Statement:

The main purpose of the technical note by Paja et al. is to re-evaluate the climate model failure data reported in Lucas et al. (2013). In particular, Paja et al. use a feature selection technique based on random forests, instead of sensitivity analysis, to identify parameters that influence simulation failures. Their results largely agree with those in the original paper. Lucas et al. determined that 4 parameters account for most of the variance in the failures (about 90%), which are the same four parameters identified by Paja as has having the largest feature scores. Paja et al. also show that the feature scores of the less influential parameters (i.e., those ranked lower than the top 4) depend on the train/test split. Their results are reasonable and not surprising because the

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

raw data displayed in figure 2 of Lucas et al. shows that the relationship between failures and parameter values degrades significantly going from higher to lower ranked parameters. It is less clear how much value the geoscientific community can take from the Paja et al. study because it re-evaluates an existing paper and reaches similar conclusions. I am not inclined to recommend the paper for publication in GMD as an original manuscript, but as a technical note it could suffice after addressing the items and comments listed below. I leave it to the discretion of the editor to decide if it passes this bar.

Item 1. The presentation of the material is still rough around the edges in terms of readability and language. I recommend that the authors work with someone to improve the readability.

Item 2. There is a mistake on page 5420 line 23. It should say 540 simulations, not 480 simulations. On the same line, "randomly" is probably a better word than "systematically".

Item 3. On page 5421 line 24, the authors state that the setup used by Lucas et al. "precludes estimation of statistical uncertainty". This is not strictly true, as bootstrapping estimates the distribution of failure probability due to different train/test splits and as a function of input parameter values. Even though they did not report the uncertainty in their sensitivity indices, Lucas et al. used bootstrapping for the sensitivity analysis.

Item 4. Page 5423 describes the general random forest algorithm, but doesn't provide the values used for the control parameters. One potential problem with random forests is the tradeoff between bias and variance during fitting. Can the authors comment on how they determined the values of the control parameters, whether they controlled for bias or variance, and what the impact of their choice is on the feature ranking?

Item 5. As shown in figure 1, the importance scores using the Boruta algorithm have values that range from about -10 to +120. How do these translate into sensitivity indicies? The latter are fractions between 0 and 1, and thus define the amount of variance

GMDD

8, C2730–C2732, 2015

explained by the parameters. Can a similar interpretation be made for the Boruta scores?

Item 6. Random forests can also have difficulties with correlated features, whereas polynomial chaos expansions, by design, explicitly decompose the sensitivities into various combinations of features. The authors should make some assessment of the potential effects of correlated features on their results. Furthermore, it is important to point out that a climate model parameter may be considered to be important even if it has a low feature score by itself but is correlated with parameters having high scores. This situation is analogous to parameters in figure 10 that have relatively small nodes but large edges.

Item 7. Paja et al. should also be aware that some of the co-authors of the Lucas et al paper were co-authors on a related paper that computed sensitivity information using random forest feature scores (doi:10.1002/2014JD022507).

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