

A Python diagnostics package for evaluation of MJO-Teleconnections in S2S forecast systems

Cristiana Stan, Saisri Kollapaneni, Andrea M. Jenney, Jiabao Wang, Zheng Wu, Cheng Zheng, Hyemi Kim, Chaim I. Garfinkel, and Ayush Singh

This has the potential to be a very useful package for the model evaluation community and a detailed description of the capabilities and applications will be paramount to the users of this package.

Overall the manuscript and in particular the summary at the end is very brief and short on details. I would expect the summary to include a discussion of how this package will assist users in evaluating model skill. Ideally, this section should also include a paragraph on how these diagnostics have helped improve models in the past and how they were applied in those cases.

Another aspect that is missing from the manuscript in present form how the different diagnostics discussed in the manuscript fit together to give a comprehensive view of MJO skill and/ or what aspects are still missing and are not touched on by these metrics. For example, one useful application that is currently not available is the capability to compare verification and multiple sets of models, this should be highlighted in the summary (I think it is mentioned briefly in section 2.2).

Comments

1. Page 2, lines 44-50: I don't think this is necessary in the introduction, especially since the author's don't refer to the phases again in the introduction. It feels out of place and should maybe go into a separate section on the MJO and which aspects are relevant for prediction skill and which of those to diagnostics aim to address. This detailed description feels out of place here. This is also one of only a few places in the manuscript that go into great detail on the MJO, in contrast most of the manuscript does not discuss MJO dynamics or thermodynamics in detail. This feels out of balance. The authors could consider adding a section describing the relevant aspects of the MJO in detail and then refer back to that when discussing each diagnostic in detail to highlight which aspects of the MJO each diagnostic addresses.
2. Showing the comparison to the model is appreciated. It would be useful to have a short discussion on why the exact model is not specified. I assume it is because the authors don't want to focus on the model evaluation, but rather the package capabilities. Even so, I think it is still important, when showing each diagnostic and comparison between ERAI/IMERG and model, to include how the application of the diagnostic helps inform forecasts and/ or users. This is currently not done. Given that the main point of diagnostics is to improve models and characterize their skill compared to other models these sections need more details on how this is achieved.

Minor comments

1. Page 1, line 21: "...S2S *time scales*."
2. Page 2, line 52: It may be better to be more specific and say "barotropic Rossby waves".
3. Page 8, line 165: ". Pacific North America *sector* and *the* North Atlantic."
4. Page 14, lines 277 - 282: I would add equation numbers for each of these.