Interactive comment on "Climate SPHINX: evaluating the impact of resolution and stochastic physics parameterisations in climate simulations" *by* Paolo Davini et al.

Anonymous Referee #2

Rhis paper describes the model configuration and experiment setup of the project Climate SPHINX (Stochastic Physics HIgh resolutioN eXperiments), which explores the impact of stochastic physics in an ensemble of 30-year climate integrations at five different atmospheric horizontal resolutions (from 125km up to 16km) within the EC-Earth model. This suite of simulations is impressive and, the setup thoughtful, in particular that of the coupled simulations, and the description of the experimental details excellent.

After the experimental description in the first 10 pages of the manuscript, preliminary results are summarized in the following 4 pages. They are comprised of an analysis of Tropical rainfall variability, an analysis of the Madden-Julian Oscillation variability

and an analysis of Mid-latitude atmospheric blocking variability. Frankly, I was disappointed by the evaluation of the simulations: while these are important topics, the manuscript did not go far beyond merely stating the results without a deeper analysis or understanding.

This is only the second time I am reviewing a GMD paper, and the focus of this journal is stated as "the public discussion of the description, development, and evaluation of numerical models of the Earth system and its components" rather than a more traditional science journal focusing on scientific findings. The current manuscript would falls under the subcategory "model experiment descriptions, including experimental details and project protocols". In short, it is unclear to me to which degree the evaluation of the numerical model is necessary and or it is sufficient in itself to describe the model experiments.

In the following I will address how the evaluation section could be improved and extended.

0. I would have liked to see some mean field comparison of temperature, wind, precipitation and in particular their variabiliyy.

Tropical Rain: 1. The manuscript interprets the difference between the TRMM and GPCP datasets as observational uncertainty. The TRMM data should be added to the right panels displaying the frequency fractions with regard to GPCP.

2. Since in the observational datasets the differences in the rainfall rates of 40mm/day are larger than the differences between the model simulations, a crude significance test would help with the interpretation of the results. This could be as simple as displaying rainfall rates for the first and second half of the simulation period.

3. I would suggest a second figure showing of rainfall rates and frequency fractions for the small rainfall rates only, so that the differences in the curves can be seen better.

MJO 4. Figure 6: Please specify how the patterns for the 4 phases of the MJO are

computed. Are they based on all data, or are they calculated for the observations only? It seems the comparison of the frequency of occurrence and the amplitude of the MJO should be done in a common basis system. Otherwise a case must be made, why different basis system are used and the patterns for each experiments (or at least some) shown and the change in patterns discussed.

5. Please refer to Weisheimer et al. 2014, who report that in particular SPPT improves MJO in System 4.

6. Are these results significant? Again a dot for each first half and second half of the simulations would be sufficient.

Blocking 7. Dawson et al. 2015 is in the Bibliography, but I couldn't find a reference in the text. This paper should be clearly cited in the blocking section.

8. Other work has suggested a positive impact of stochastic physics on blocking. Why is this not the case here? Maybe it only helps at horizontal resolutions of less than T159? Is there really no benefit even at resolutions of T159?