Authors Response to Referee #1

We thank referee #1 for all the time they invested in reading and understanding this paper and in particular, for their constructive comments and suggestions to the manuscript we have submitted. We took all the comments into consideration for generating a revised version of the manuscript. We respond to the original comments (italic) point-by-point.

This emulator, a new addition to the MESMER family, produces fields of monthly precipitation as a function of corresponding fields of monthly average surface temperature. It does so by modeling precipitation at each grid point as the product of a mean component, driven by temperature in a neighborhood of the location, and a residual component whose variability is estimated as constant along the length of the simulation.

The results, for both the case when temperature is the output of the ESM being emulated, and the case when it is in turn a product of emulation by the mean component of MESMER-M, are validated according to numerous metrics that address first and higher order behavior, trends, spatial and temporal coherence within the precipitation output and between precipitation and temperature, with satisfactory results.

I enjoyed reading this paper, which is clearly written and presents a wide range of validation results creatively conceived and displayed. The authors have a clear-eyed take on the outcomes, which are not perfect but constitute an important step forward in emulating impact relevant variables. This is especially true for an emulator that can generate novel realizations, akin to additional initial condition ensemble members, and therefore can help characterize statistics of extremes (within the limits of what ESMs and their emulator can represent, through monthly average quantities). As the authors recognize, there is room for improvement, especially for the highest quantiles of precipitation which often appear overestimated when the validation is done within the confines of an individual model statistics, but these errors pale compared to the inter-model variability/differences. It is also good that the authors are able to offer an hypothesis about the sources of the emulator shortcomings (having to do with the variability term being modeled as stationary rather than allowing it to vary with temperature changes as well) and therefore can point at directions for further development.

AR: Thank you for this very accurate summary of our work.

My comments are in the spirit of creating a slightly more comprehensive picture of the emulator-ESM comparisons (again recognizing the nice effort in providing a multidimensional depiction of the emulator performance, always hard to do). What I would have liked to see are:

- *Maps of gridded trends and patterns of change for a few time slices along the century, comparing emulated and ESM projections, maybe as seasonal averages. These could give an additional sense of the mean/trend performance at the grid-point level, performance evaluation which is currently - for these aspects -- limited to aggregated regional (SREX) scales*.

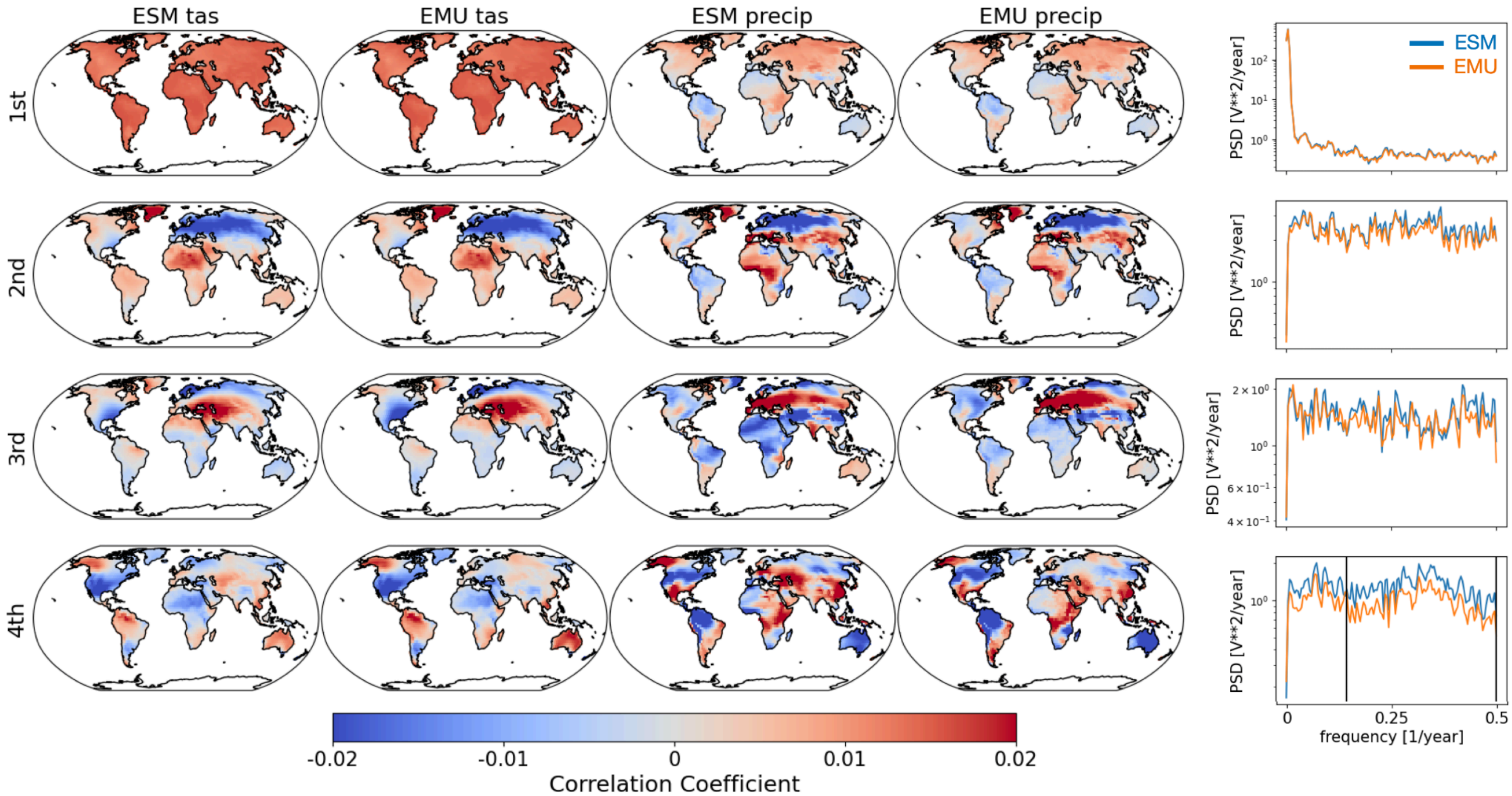
AR: We think showing seasonal estimates of gridded output would be very useful. We have included Figure 2 (and Figures B1 and B2) in the revised manuscript.

- *Additional analyses of spatial and temporal characteristics of the generated precipitation fields. These could be performed for a set of locations, chosen to represent different latitudes/climate conditions. These analyses could include variograms, which are the tool of choice for analyzing the spatial correlation characteristics of a field from spatial statistics, and could be computed for the regions around each location. Similarly, for time series of output at specific locations one could look at the entire autocorrelation structure, or spectrum, and have a full picture of what the emulator does well or misses (for example, ENSO-driven oscillations in precipitation at points known to have strong teleconnection, and of course compared to what the ESM does, since it is not a given that ESMs represent those signals well. Hopefully at least one of these three ESMs has a decent ENSO behavior and it would be interesting to see if the emulator can represent it).*

AR: We agree with all the suggestions. As we want to avoid displaying too many verification metrics in the main of the paper, we have addressed the suggestions in Appendix B3 called “Additional Spatio-temporal Validation Metrics” (L434 ff.). The main analysis steps we took are the following:

1. Geo-spatial statistics: Figure B10 now displays variograms that focus on 4 locations and their 300 closest grid-points. The locations were selected at random within the regions SCA, NEU, CAF, SEA. These regions are consistent with the validation approach employed throughout the entire paper and represent a wide range of possible precipitation behaviors. We discuss the results in L440-446.
2. Temporal characteristics: For the same four locations, we generated the full frequency-periodograms and included the results in Figure B11. Periodograms decompose the power of the precipitation signal into individual frequency components. The findings are described in L446-454.

3. Lastly, we also tried to incorporate your suggestions on ENSO behavior. The main difficulty being that most established ENSO indices we are aware of rely on ocean data while we emulate land data only. We nevertheless used MPI-ESM1-2-LR, a model that depicts ENSO feedback reasonably well [1], to perform experiments identifying ENSO signals over land. To this end, we jointly decomposed the gridded temperature and precipitation fields into their principal components independently for ESM and EMU data. Below, we are showing the correlation coefficients between the first four principal components (top to bottom) and January ESM/EMU temperature (1st/2nd column) and ESM/EMU precipitation (3rd/4th column). We identify the 4th component to resemble ENSO behavior as expected in December, January and February (e.g. warming over North-Western North America, Eastern and South Asia, Eastern South Africa, North-West South America, Madagascar, Australia; cooling over Eastern North America; Dry conditions over Eastern South Africa, Madagascar; Wet conditions in Western, Central and Eastern North America, North Central America). In addition, we performed a frequency-analysis on the first four principal components. The results are displayed in the fifth column (ESM: blue lines, EMU: orange lines). The El Nino frequency bands (2-7 years) are highlighted for the third component using black vertical lines. The signal strength peaks at an occurrence of 0.33 (once every 3 years). The signal is present in the emulated data as well, although it is less pronounced. We appreciate that this analysis, despite showing some encouraging results, is limited as we do not have ocean data available as part of our emulation. Because of the limitations, we are hesitant to include the analysis on ENSO behavior in the main manuscript, but now mention it as an area for future development (L415-417).
    a. [1] Tian, Q., Ding, R., & Li, J. (2023). Simulations of the North Tropical Atlantic Mode–ENSO connection in CMIP5 and CMIP6 models. *Journal of Geophysical Research: Atmospheres*, *128*(16), e2023JD039018.



My other overarching feedback is that some of the figures are very small and it is hard to distinguish the different points/color hues. Either expanding the area of the page dedicated to them, or reducing the range of some of the axis to focus on the region of the plot where something is happening could help (even if I realize that the same ranges may be needed for ease of comparison). But all this said, I appreciate the graphics that the authors have created to display these results.

AR: We agreed and adjusted all graphics for better visibility. In particular, we adjusted all tick and label sizes and hope that the graphics are more easily accessible now.

I am aware of another emulator proposed recently, for daily precipitation, that may be relevant as a citation (if only for its use of a Gamma distribution): Kemsley, S. W., Osborn, T. J., Dorling, S. R., & Wallace, C. (2024). Pattern scaling the parameters of a Markov-chain gamma-distribution daily precipitation generator, International journal of climatology, 44(1), 144–159. https://doi.org/10.1002/joc.8320

AR: Thank you for catching this one. We have included it in our paper now.

These are my only substantial comments, and I would be pleased to see a slightly revised version addressing them, but I consider this work in very good shape already so I would characterize my requests as minor revisions and I hope the authors can meet them with ease. Nice work.

AR: Thank you again for taking the time to look at our paper so thoroughly and for the, overall, very supportive and positive feedback. We hope we could address your suggestions sufficiently in the revised manuscript.