The authors here present an interesting combination of statistical approaches to compare the probability density functions of observed and simulated chlorophyll across biomes and Longhurst provinces. I congratulate the authors on the work that has gone into this manuscript. It is extremely well written and the rationale behind their approach is cogently explained. I thoroughly hope that it encourages the ocean biogeochemical community to move beyond the use of correlation and RMSD as standard measures of model performance. I am happy to recommend the manuscript be accepted as it is but have a few very minor recommendations given below.

Perhaps the authors could discuss how the approach might be extended to provide greater insight on model-observation mismatch? Where coincident physical observations are available, one could presumably subset model and observation data to access for example the ability of models to capture the impact of regional heat waves on chlorophyll distributions.

As most biogeochemical models do not produce a Chl\textsubscript{Rrs} output, perhaps the authors could offer some thoughts/recommendations on the use of the approach when only Chl\textsubscript{mod} output is available. The approach still seems to offer plenty of insight and this should perhaps be emphasized a bit more.

L37 Should RMSD be standard deviation? While the RMSD is sometimes also given in Taylor plots this doesn’t seem to be given in Figure 2.

L80 Suggest giving the model name here.
Fig 1. Suggest increasing the label font size for the regions identified.

Fig 2. A point color legend for these Taylor plots would be helpful.

Fig 4a. The colorbar should have a label.

Fig 4b Suggest adding a point color legend.