# Comments on ”Probabilistic Evaluation of Climate Models For Seasonal Forecasting in the MENA Region”

**Title**: as the evaluation carried out is both deterministic and probabilistic, I suggest the a more appropriate title would be just “Evaluation of Climate Models For Seasonal Forecasting in the MENA Region”

**Authors**

The affiliation “Hassania School of Public Works” should be after Nohayla BERRAHMOUCH, Mohamed EL-BADRI and before the Supervisors.

Please note that for formal reports etc, my name is Nicholas (Nick is short form of my name)

Please add Wafae’s and my affiliations as well.

I think that for Wafae this should be “Direction Generale de la Meteorologie, Morocco”

My affiliation is “Met Office, Exeter, UK”

**Acknowledgments**

Please can you also add   
Nicholas Savage’s time was funded via the WISER MENA project. The Weather and Climate Information Services (WISER) Programme is funded with UK International Development from the UK government and led by the Met Office in the UK. This work has been partially supported by UK International Development from the UK government; however, the views expressed do not necessarily reflect the UK government’s official policies.

“Special attention is given to the models’ skill in forecasting extreme weather events” – this is not a correct description of the report at the moment. The focus in the report is on tercile forecasts which are not extremes.

**2.2.2 Impact-Based Evaluation**

The paper you cite is now published open access, so you should use this link instead: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2024EF004936>

and give the authors.

Please state that this study is global and not specific to MENA and that it focuses on **extreme** precipitation. A key aspect of the study is the use of regions which were vulnerable to wildfire and flooding

**3.2.1 Spearman rank correlation**

The title and the metric described is Spearman rank correlation but the plots state they are Anomaly Correlation Coefficient – please make sure this is consistent.

Please provide more description of the maps – where do models have better or worse skill? Are they consistent with each other in the areas they have high skill for or does this depend from one model to another?

I find it very hard to read the values in the heatmaps. Please increase the text size

The heatmaps say they are for “the mena region” – what is the definition of this region used here? Which correltation coefficient is used (Perason, ACC, Spearman?). Also explain what is meant by period (I assume this is the forecast range)

The color scale used in this figure needs to be consistent for all models

Please give more description of these results. Which models are best, how does skill change with period? Are some seasons better for most models or particularly bad?

This request for more descriptions and changes to the heatmaps colour scales also apply to all following sections

**3.2.2 RMSE and 3.2.3 Coefficient of Determination (R 2 )**

As above – please describe the results in more detail. Same comments on colour scale

**3.3 Probabilistic Evaluation Metrics**

At some point in the text, you need to explain what a tercile based forecast is and explain here that all these probabilistic scores are for tercile based forecasts (I assume they are but you don’t say this)

**3.3.2 Reliability**

Please add more explanations of the reliability diagram – what does it show and why is a perfect model has a 45 degree slope

Figure 13: The Reliability Score – please replace this with a table of results

Reliability diagram – explain what the slope below 45 degrees for almost all of the models tells us.