Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-65-RC1, 2019

Geoscientific Model Development Discussions

Open Access

EGU

Interactive comment on "Evaluation of Unified Model Rainfall Forecasts over the Western Ghats and North East states of India" *by* Kuldeep Sharma et al.

Anonymous Referee #1

Received and published: 22 June 2019

Review of "Evaluation of Unified Model Rainfall Forecasts over the Western Ghats and North East states of India" by Kuldeep Sharma, Sushant Kumar, Raghavendra Ashrit, Sean Milton, Ashis K. Mitra, and Ekkattil N. Rajagopal

Summary of manuscript:

The authors use well-established verification metrics and a newly developed satellite-gauge merged rainfall product to evaluate global Met Office Unified Model forecasts over India, with a particular focus on the Western Ghats and North-Eastern states, which are regions with steep orography. The authors use deciles (80th and 90th percentile) to identify heavy rain and very heavy rainfall events in their forecasts and obser-

Printer-friendly version

Discussion paper

vations, which are upscaled to a 0.5-degree grid. The authors find significant improvements in the forecasts over a 11-year period, during which various model improvements have taken place that could have led to such improvements.

Summary of review:

The paper is generally well written: although the grammar could be improved throughout, there are only a few places where the scientific argument is difficult to follow due to grammar. The figures are generally of good quality. However, the paper does not contain any obvious scientific advances in terms of methodology, novelty in terms of research focus, or impact of its findings on future research or model development. It does not appear to be in scope of GMD, as it comes closest to "full evaluations of previously published models", but the results presented are far removed from a "full evaluation". It is noted that a previous review indicated similarities with the Sharma et al. (2017) paper and, indeed, these similarities are significant, with only the region of interest having shifted from central India to the more mountainous terrains. Furthermore, there are flaws in the experiment design that seriously affect the impact of this paper. Taking all of this into consideration, the recommendation is to reject this manuscript for publication.

Major comments:

1. Modelling experiment design

While the authors are conscientious listing the numerous model configuration changes that have taken place during the period of study, the variety of configurations makes the intercomparison practically meaningless for scientific model development. Tracking forecast skill is a useful endeavour, but could be done by modelling centres and published through reports. In order to attribute improvements in forecast skill to a particular change in configuration – whether it is change in resolution, data assimilation practice, dynamical core, or parameterisation practice – one would expect that different configurations are run for the same period of interest. Many of the configurations studied by the authors are no longer available, which makes findings related to those configurations practically useless for model development. One solution could be to focus on only one or two recent model configuration changes, such as a change in resolution and a change in parameterisation, and perhaps limit the period of study to the most recent monsoon seasons. Given that these forecasts are produced by another modelling centre, it is likely that these flaws in the study cannot be overcome.

2. Focus on orography

This has been highlighted in an earlier review and it remains a serious flaw of this study. Changes in orography – especially steep slopes as experience in the Western Ghats – will take place over scales much smaller than 0.5 degrees. How does a rainfall extreme (or decile) on the 0.5-degree scale relate to rainfall extremes at much smaller scales related to orography (valleys, catchments)? An observation-based rainfall product at such a coarse resolution will not be able to capture the detailed variations of rainfall related to orography on those scales that matter for flood forecasts, which appeared to be the motivation for this study based on the introduction. The GPM-derived IMERG rainfall product has a 0.1-degree resolution and could be more suitable for this purpose, at least for the most recent years. If the authors were to resubmit their paper, they should consider how their results are affected by the choice of observation for validation.

3. Deciles

This relates to the first point, but highlights the issue with experiment design. The authors consider the full period of study to derive the 80th and 90th percentile of rainfall from the observations, and do the same for the models. However, each model configuration will have its own climatological biases, leading to different deciles of rainfall. The model-derived deciles could therefore be skewed so that a configuration with a high bias will produce the model-averaged 80th percentile too frequently, while a configuration with a low bias would produce the 80th percentile very infrequently (for example). The authors demonstrate that the interannual variability leads to different frequencies of the multi-annual 80th decile in individual years, which complicates model evaluation further. On a similar note, it is unclear how the deciles are determined. Given that there are 475 grid point in the Western Ghats region and that there are 122 days in the JJAS period, I would expect 57950 values of daily rainfall, so 11590 counts of the 80th percentile, but the counts are around 3000. Do the authors consider the 80th percentile only conditional on there being rain measured? Looking at Figure 5, I would expect the mean counts to follow a linear relationship (each decile having approximately the same number of counts when summed over the 11-year period), but this does not appear to be the case. Why is there a change in relationship around the 60th percentile for the Western Ghats and the 40th percentile for the North Eastern states? If the authors were to resubmit the paper, ideally focusing on only two or three model configurations with changes that are relevant to the representation of rainfall over orography, they should: (1) Clearly state how the deciles are determined. (2) Determine the deciles for each model configuration and each year (the latter is already done for the observations). And (3) Report the actual values of these deciles in mm/day.

4. Bias

The authors should note that given the definition of bias = (a+b)/(a+c) and given that when using deciles, e.g. the 80th percentile (a+b)/(a+b+c+d) = 0.2 and (a+c)/(a+b+c+d) = 0.2, the bias should be 1 by definition. Perhaps interannual variability and changes in model configuration (see point 3) could affect the bias, and it is peculiar that it has no discernible effect. Nevertheless, it seems uninformative to use the bias when considering deciles.

Minor comments:

Line 14-21: This is very focused on methodology and does not explain at all what is novel and what is found in this study. The authors should use the abstract to describe their results more clearly, how these results are related to model configuration choices, and how their results could lead to future model development and improvement.

Line 33, 39: "heavy orography rainfall" and "heavy rainfall". Please specify actual amounts.

Line 45: "This" – does this refer to "complexity" (line 42) or "observational data sets" (line 44)?

Line 56-59: Please remove the abbreviations LPS and MD, these are used only once.

Line 68: A rainfall amount for the NE is mentioned. Please include a value for the WG.

Line 74: "report improved skill" – compared to what? Also line 75 "document", what did they find?

Line 72-89: The introduction requires a brief overview of rainfall observations over India, which are clearly important for understanding orography rainfall. The first paragraph of Section 2 should be moved to the introduction. Similarly, the rationale for using quantile-based thresholds belongs in the introduction. The first paragraph of Section 3 should be moved to the Introduction.

Section 2.1: It appears that the authors do not use the same observational data consistently throughout their study. It should be made more clear to the reader that for 2007-2011, the IMD gridded data are used. Are these on a 0.5-degree grid as well? For 2012-2015, the merged data (with TRMM) are used and for 2016-2018, the merged data (with GPM) are used. How do these different products compare in terms of deciles? Are the different products available for the same period of time at any point of the period of study? This is actually quite a concern, again, for the scientific quality of this analysis.

Line 162: How are these "improvements" evaluated? These are actually changes to the parameterization schemes, but how is it determined that these are improvements to the model?

Line 164: "Daily rainfall" – what is the period of accumulation, 00-24 UTC or some other standard time?

Line 218: “from observations” – the authors should clarify that three different observational products are used.

Line 233: “fine scale features” – the data are on a 0.5-degree grid so it is unclear what authors mean by these features.

Line 233: “successfully predict” – how “success” determined? Subjective eyeballing of precipitation maps?

Line 235: “over prediction” – this phrase has a different meaning. Use “overestimates”.

Line 242: The reference to a paper from 2011 is not appropriate to describe later model versions that are considered here (2013, 2015, 2018). The figure S2 is not clear either. How is systematic error calculated? Against analyses? How good are the analyses?

Line 245 and Figure 3: These figures would be easier to interpret if the authors reproduced two separate figures, one zoomed in to the WG and one for the NE.

Line 251: “false alarms” – this is not an appropriate phrase to use when comparing seasonally aggregated information.

Line 258-264: None of these attributions are justified without performing a systematic comparison of different model configurations for the same period of observations. See major comments.

Line 289-291: Remove “This . . . 2014)” – The authors are describing what a high POD means.

Line 297-302: Remove these lines as the Bias is not very informative when considering deciles.

Line 317: This should refer to Figure 8 and 9, not 7 and 8.

Line 320: Why is the magnitude of SEDI compared to other metrics, e.g. CSI?

Line 321-326: These claims are not supported by the findings due to the flaws in

Printer-friendly version

Discussion paper

experiment design.

Line 334: Remove “improved”.

Line 338: “identify and quantify the impact of” – this is not done in this paper due to the flaw in experiment design.

Line 341-342: What are the “large-scale monsoon rainfall features”?

Line 349: Rephrase to “Following increased grid resolution. . .”

Line 350: How is the “improved synoptic variability” determined?

Line 371: Is this 0.25-degree rainfall product used in the present study?

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-65, 2019.