

Improving the Potential Accuracy and Usability of EURO-CORDEX Estimates of Future Rainfall Climate using Frequentist Model Averaging

Stephen Jewson, Giuliana Barbato, Paola Mercogliano, Jaroslav Mysiak, and Maximiliano Sassi

Response to Reviewer 1

Dear Reviewer 1,

Many thanks for taking the time to read our paper and make helpful comments, which have definitely improved it. I have pasted your comments below, greyed them out, and responded to them in black. Please note that I have made some additional changes based on reviewer 2's comments, and based on some additional literature on the topic of model averaging that I have found recently (in statistics and economics journals). In particular, as a result of what I've found in the statistics and economics literature, I've changed the names of the methods so that MMA becomes PMA, which means that SMMA becomes SPMA, and BMMA becomes BPMA.

The paper presents two methods to adjust ensemble mean of variables projected by climate models (CM) and compares their performances against two other adjusting approaches (i.e., conventional Akaike model averaging and statistical testing) and unadjusted mean, considering change along different future time frames, seasons, precipitation variables and RCP scenarios over the whole Europe.

The two proposed methods (MMA) have a common derivation based on minimisation of the predictive mean squared error.

The paper discusses the relative advantages of all the considered methods and shows that the application of MMA is particularly advantageous when the uncertainty of a given change is high due to small predicted changes and large spread among the CM signals (in such cases rejection of a null hypothesis of no change is usually the outcome of statistical tests).

As a general comment it is my opinion that the paper is timely, and results are of interest for NPG readers. However, the readability of the paper is not fluent and can be improved by a careful proofreading, since there are many parts of the manuscript that I needed to read and read again to understand the underlying message.

I have taken this comment to heart, and rewritten various sections. In particular:

- a) the introduction
- b) the section on probabilistic prediction (3.2.2)
- c) the section on Bayesian prediction (3.3)
- d) I've added a section on probabilistic skill scores (3.4)

Apart from these aspects I have only some minor issues, that are listed below.

The two presented methods (MMA) have been previously published in technical reports by the first author ((Jewson & Hawkins, 2009a, b; see reference list in the manuscript), as credited in Section 3.5. I suggest anticipating this information by providing proper credits in previous sections (e.g. sections 3.2 and 3.3) and to remove Section 3.5.

Thanks for this suggestion. I have moved these sections to 3.2 and 3.3, which I agree is the more logical place for them.

Line 243. “the scale parameter” should be “the square of the scale parameter”, “numbers” is “number”

Thanks for noticing those mistakes. I have corrected both.

Lines 251-256. Clarify if the objective prior is adjusted or unadjusted. Moreover the implementation of the Bayesian approach should be better explained.

The objective prior is the standard objective prior for this distribution, with no adjustment. I have significantly lengthened section 3.3, to explain the Bayesian approach in more detail, added some equations to make it clearer, and added another citation.

Line 263. Please explain the rationale of statistical testing and conventional AICc model averaging.

I have now explained the interpretation of the statistical testing in section 3.2.1, and the first paragraph in section 3 now explains AIC model averaging in more detail, and gives citations to two text-books that cover the topic.

Line 295. Similarly to previous comment: Please explain the rationale of Predictive mean log-likelihood

Based on this comment, and comments from reviewer 2, I have now added a paragraph in section 3.4 which explains the choice of score for evaluating the probabilistic forecasts. The score we are using is very commonly used, but has many different names in the literature (unfortunately). After reading around I've concluded that the name 'log-score' is more commonly used than 'mean log-likelihood', and so have changed the name from PMLL to predictive log-score (PLS) throughout.

Line 327. “Fig. 4d” I do not understand what is plotted in such subplot.

I have over-hauled figures 1 and 4, and I think they are much easier to understand now. Fig 4d now shows the absolute (unsigned) change in the ensemble mean change, in mm. So it shows the impact of applying the SMMA/SPMA method. The impact, when measured in mm, is largest in places where there is both a large rainfall change predicted by the ensemble mean, and that change is fairly uncertain.

Fig 4. Put unit of measures as in Fig. 2

Thanks. I have added the units now.

Line 367. “8%” would be “7%”

Thanks.

Line 443. “.... Root mean squared size ...” please explain better

I have rewritten this section to explain how this is calculated.

Line 454. “Fig. 2” should be “Fig. 3”

Thanks for noticing that.

Fig 9c. Adjust the y-axis label

Thanks for noticing that.

Best regards,

Steve Jewson