Interactive comment on “Statistical downscaling with the downscaleR package: Contribution to the VALUE intercomparison experiment”

https://doi.org/10.5194/gmd-2019-224-RC1

Response to reviewer #1

We thank the referee for the time devoted to review our manuscript, and the positive feedback provided. Along the next lines, the different comments posed by the reviewer are reviewed point by point. The referee’s comments are indicated in black, and the author responses in blue fonts.

A step commonly carried out when assessing the ’quality’ or ’value’ of climate data is the comparison with observed data, normally applying a downscaling step. This paper presents a reproducible R-based workflow in the context of the COST action VALUE. The paper presents a workflow (also shared as R Markdown notebook) which start with data loading to the visualisation of the results. In this workflow the authors compare different downscaling techniques.

I have a few comments here that I think would improve the submitted paper:

1. In the Section 4.1 the authors might add some numbers to Figure 6 (even a separate table) showing average (possibly also std or quantiles) values of RMSE, Correlation and variance ratio. Comparing M1, M6 and their -L version graphically is not easy.

We have included the numbers corresponding to the validation of the pan-european experiment in the new Table 3 of the revised manuscript version.

2. Again in Figure 6 I don't understand the meaning of 'A factor of 0.1 has been applied to RMSE for better comparability of results.', why not leaving the original values?

We decided to apply a scaling factor of 0.1 to the RMSE values in order to make their magnitude comparable to that of the other validation measures, so they can be visually compared in the same plot. We have replace the caption indicating that "[...] The colour bar indicates the mean value of each measure. A factor of 0.1 has been applied to RMSE in order to attain the same order of magnitude in the Y-axis for all the validation measures", hoping that it is now more clear.

3. The authors should say something on the computation time needed for the experiments described in the Figure.

We have included a new section in Appendix 1 devoted to a more detailed analysis of computing times. Please find attached some figures to be included in the new revised version of the manuscript addressing the efficiency, in terms of computing (user) times, of the different downscaling methods. As an example, Fig. 1 shows the computing times required to accomplish each of the methods used in the Iberian Peninsula experiment. A more detailed discussion of these results and additional figures/tables are included in the revised version of the manuscript.

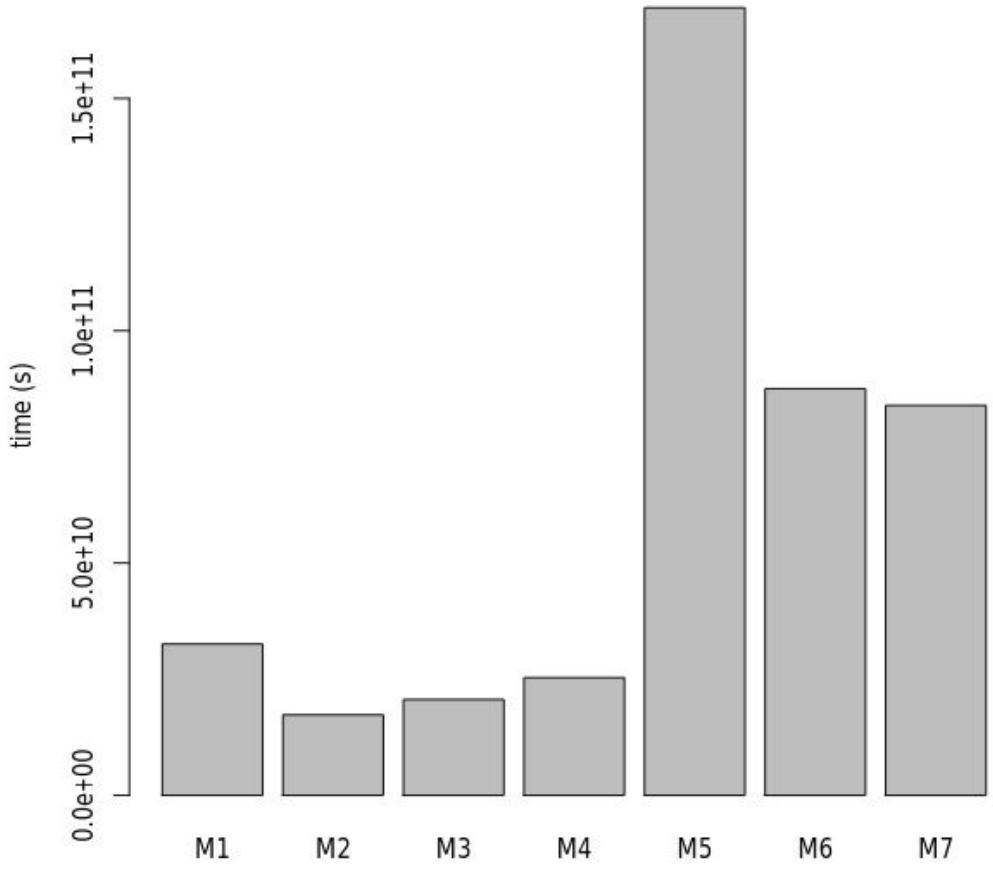

Figure 1. Cross-validation times required for the downscaling models developed in the Iberian experiment. The computational times of the generalized linear models configurations (see Table A1) includes both the downscaling of the occurrence and amount of precipitation, whereas for the analogs both aspects are downscaled simultaneously. More information about the configurations can be found in Tables A1 and 2, or in the companion paper notebook

4. How the developed package is able to deal with large datasets (10-100-500GB)? Is there any support to larger-than-memory computing (e.g.Python Dask)?

Current on-going work is being done in order to handle larger matrices using the bigmemory package. Also, we are considering future developments in order to be able to run scalable applications in the climate4R Hub (a

cloud-based facility allowing to remotely running climate4R applications). In the meantime, some large tasks can be conveniently sliced using the helper function downscaleChunk. For brevity, we have not included further details on these new developments in this paper. However, there is a related article currently under interactive discussion in this journal in which some of these features are presented. The application of deep learning in downscaling applications is presented and some features to handle large datasets are here presented: https://www.geosci-model-dev-discuss.net/gmd-2019-278/

5. Can the authors say something about the importance of choosing the right domain to compute the EOF? Sometimes the results can be very sensitive to the choice of the domain.

As the referee points out, the domain selection is an important part of model building, being an important decision affecting model performance. In this paper, we show how domain selection can be very easily accomplished with just changing simple parameters (lonLim and latLim) either on the predictor dataset loading (function loadGridData) or by recursively subsetting the already loaded predictor set (using the function subsetGrid). This allows for a flexible configuration of experiments in which different alternative domains can be easily tested. However, in this paper we stick to the domains already well tested in previous studies over the Iberia Peninsula (Gutiérrez et al. 2013) and over Europe, using to this aim the standard experimental settings of the VALUE experiment, to which downscaleR has contributed the methods analysed. Even though domain screening is out of the scope of this article (focused on the presentation of the downscaling tool), we indicate how these type of experiments can be easily undertaken. As an interesting alternative to this time-consuming task, we show that a local predictor approach, based on the use of local predictors close the predictand location can be used without significant changes in the future deltas obtained. To better illustrate this finding, we include further details on the future climate deltas in the new revised section 4.2, and the new Fig. 8 of the revised manuscript shows how the local predictor approach does not significantly alter the deltas obtained.