Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-77-RC2, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Towards the Development of a Pan-European Stochastic Precipitation Dataset" *by* Lisa-Ann Kautz et al.

Peter Stucki (Referee)

peter.stucki@giub.unibe.ch

Received and published: 2 April 2019

Review of the manuscript Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-77

General comments

In this manuscript, the authors present the configuration and results from dynamical downscaling and subsequent bias correction for a number of heavy-precipitation events in five large river catchments in Central Europe since 1979. They make use of the ERA-Interim and ERA20C reanalyses for initializing the COSMO-CLM regional climate model with horizontal grid sizes of 25 km. Calibration and validation of the obtained new datasets are done using the gridded E-OBS daily precipitation dataset. With this

Printer-friendly version

Discussion paper

approach, the authors claim to provide a proof of concept for producing a consistent pan-european precipitation dataset, which I think would be a relevant scientific contribution within the scope of this journal.

In general, the structure, language and style of the article are on a good level, and I have only minor comments on this. Also, I acknowledge the amount of conducted work, provided material and analyzed data. In contrast, I do have a number of major concerns about the concepts, datasets and methods used in the study. These need to be addressed before I can recommend accepting the manuscript for publication.

Major comments

My first concern is about the input data used and the downscaling resolution. In this study, the ERA-Interim dataset is downscaled from an 80-km grid to a 25-km grid. With the advent of the ERA5 at a 31-km grid, such a downscaling procedure results in hardly any gain in resolution, while it has substantial computational costs. The authors argue that a 25-km grid is adequate for driving hydrological models. Although I am not a specialist in this field, I would argue that convection-permitting model resolution is needed for many applications, and the authors also mention this in the manuscript (P16 L15). I can clearly see that research must be based on data availability at the time of the analyses. Nevertheless, I assume that the development of ERA5 was already known when this study was conceptualized, and also that future applications would rather switch to using the better resolved dataset. In addition, the study downscales from the long-term ERA-20C reanalyses for comparisons of performance. However, the current study does not exploit the long-term character of the reanalyses; early events that would be covered by ERA-20C only are not included. To conclude, I am reluctant to insist on redoing the experiments using the currently available datasets because of the massive additional work load, although it seems the way to go to achieve a solid contribution in the sense of a proof of concept for future hydrological applications.

My second concern is about the application of bias correction. In the first place, I have problems to follow the argumentation that the chosen methods result in a 'consistent' precipitation dataset. I take from the experimental setup that after downscaling from 'consistent' reanalyses (I agree with the authors, P2 L6), the bias correction based on E-OBS actually re-introduces inhomogeneities, because the E-OBS dataset is only quality checked, but by far not homogeneous (see also P16 L20). In addition, I see that a number of studies conduct calibration and validation of bias correction using the same data. However, in a clean experiment, the training dataset must be independent from the test dataset. Therefore, I suggest dividing the available data into these sets, or apply cross-validation or similar techniques. Moreover, there are also disadvantages of applying bias correction, especially EQM, which should be addressed in a proof of concept. These are (i) the assumption of climatological stability of the derived transfer function (ii) the potential physical distortion of the dataset by statistical correction including (iii) potential physical disconnection from other variables like temperature, for instance, and (iv) the problem of how to deal with unobserved extreme values when applying EQM to events outside the calibration period. I appreciate that the study tests a range of options for bias correction before selecting EQM, such as recommended by Maraun et al. (2010) and others. In turn, this means that finding EQM the best-performing method is not a novel contribution of the study. The authors discuss potential improvements, and there are even more options, e.g. applying a combination of change factor and quantile mapping, using detrended or region-aware quantile mapping, applying multivariate bias-correction or bias-correction based on synoptic weather situations. Such a comparison might be a valuable and novel step.

My third major concern is about uncertainties along the downscaling – bias-correction – validation path. On a number of occasions, I miss information about potential or quantified uncertainties along the process of downscaling, bias correction and validation. An important feature of current observational and reanalyses products like more recent versions of E-OBS, ERA5 and CERA-20C (which cover the same periods as the ones used in the manuscript) is that they provide an ensemble to assess uncertainties or sensitivities. I do not insist on a thorough probabilistic analysis here, but I tend to think that such ensemble information should be exploited in the current study in one way or the other. Furthermore, no information about uncertainties in the configuration of the regional model are given, just a statement that it is ‘suitable’.

My fourth concern is about the achieved enhancements and the chosen metrics and variables. I agree with my fellow reviewer that the absolute differences between corrected model and observation-based dataset are still enormous at places. This contrasts the argument of the authors that the study provides a proof of concept for hydrological modeling. Then, the added value often comes from the second or third decimal place, or even slight worsening occurs regarding some metrics (e.g. LS and LOCI in Table 1). This leaves me uneasy, and I wonder if such differences are still (statistically) significant, can be called enhancement and worth all the involved efforts for practical applications. At one point, the authors mention that such small differences rather deserve the adjective ‘similar’ (P7 L14), or that missing the dynamics is a much larger problem (e.g. Figure S1). Furthermore, the distinctions between the ‘captured’ categories are comprehensible, but involve a lot of subjective judgement. For better comparability, I suggest using standard ETCCDI precipitation indices like annual precipitation maxima in terms of Rx5day, R20mm, R95pTOT or R99pTOT. For instance, I wonder how many and which heavy-precipitation events are missed or captured in terms of three- or five-day rainfall totals. Finally, please check if non-parametric measures should replace parametric measures (Pearson correlation, standard deviations, also involved in Taylor diagram) for precipitation.

Please refer to the attached supplement for the specific comments.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-77/hess-2019-77-RC2-supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-

77, 2019.