Review of

Scale-dependent biases in Alpine sub-daily areal precipitation extremes: added value of convection permitting models

by Rashid Akbary, Eleonora Dallan, Paul C. Astagneau, Raul R. Wood, Francesco Marra, Manuela I. Brunner & Marco Borga

General Comments

The given paper studies how bias of simulated extreme precipitation, from both convection-permitting models (CPM) and their driving regional climate models, changes with spatial and temporal aggregation over complex terrain. It thereby constitutes an important addition to the current scope of literature, where CPM evaluation is commonly performed at the native resolution only. While added value at high resolution has been proven throughout a variety of studies, it remained unexplored how this improvement in model performance reflects at areal aggregation of up to several thousands of kilometers and at temporal aggregation of up to 24 h. Such scales are however highly relevant for hydrological modelling in large river basins and of long-duration flood drivers. Given the high computational demand of CPMs, knowledge on how and whether added value of CPMs reflects at higher areal and temporal scales is of high practical relevance. A better understanding of the performance of CPMs, such as potential changes in the sign and magnitude of model bias with aggregation, furthermore pinpoints to physical inconsistencies in the current climate models and thereby supports model development. In conclusion, the presented paper bridges an important research gap, is of direct practical relevance and advances our understanding of CPM performance. It is therefore greatly welcomed as a contribution to HESS.

The paper is well structured, presents the findings in a clear and concise manner and is of high scientific quality. The current version of the manuscript would benefit from a sharper contextualisation and a more in-depth discussion on the scales of hydrological processes and their importance in the generation of different flood types through a revision of its introduction, as outlined below. The methodology section leaves a few open questions, which are addressed in the following. The results and their discussion are of high quality, with some rather technical revisions needed before publication.

Specific Comments

P. 1/ L. 17: It would be relevant to mention over which study area you’re conducting the analyses, namely over Switzerland, as it gives the reader an idea of the physiography at hand.

P. 1/ L. 20: The information of the length of the used time series could be useful, as it is a proxy for the robustness of the results (20 years).

P. 1/ L. 23: With only the abstract, it is not clear what the exact coefficient is that is under-/ overestimated (it is the 20-y extreme precipitation return level).

P. 24/ L. 24: The statement is given that: „RCMs consistently underestimate precipitation extremes across all spatial scales“. The RCMs’ bias is not quantified in the abstract, yet at the end (L. 27) it is concluded that CPMs outperform the RCMs for short-duration extremes. It would be helpful to underline this statement with a number for the RCMs’ bias. This allows the reader to follow the conclusion and gain an estimate of the degree of added value by CPMs.

P. 1/ L. 28: This statement could be refined by specifying for which catchments and flood types in particular scale-dependent biases are of major importance.

P. 2 / L. 32: One important parameter for flood generation, modified by spatial averaging, is indeed extreme precipitation intensity. However, it is not the only one, with precipitation volume also being a defining factor. This aspect could be mentioned here.

P. 2/ L. 34 f.: The given statement does not only apply to complex terrain, but is valid universally. However, there is indeed great value in conducting the analyses over mountainous terrain, as you highlight in L. 82 ff. In the introduction, you could elaborate more deeply on what's special in complex terrain and why your works finds particular relevance here.

P: 2/ L. 52 f. & L. 58: A more in-depth discussion on the scales of hydrological processes and their importance in the generation of different flood types, in conjunction with the meteorological scales would be needed. CPMs are known to offer added value for the representation of small-scale heavy summer storms, with Hortonian overland flow being a major driver of resulting flash floods in fast-responding catchments. Over longer time scales and larger catchments with longer response times, long-duration flood drivers take the lead, with saturation-excess overland flow being a major player. A discussion on how flood drivers change with scale and where CPMs are expected to offer added value for hydrological impact modelling, would be very valuable. Conclusionary, the types of catchments and floods for which the work presented here has primary relevance can be refined (c.f. comment to P. 1/ L. 28).

P. 2/ L. 60: Without having read the study, it is ambiguous whether the „spatial and temporal scales" mentioned refer to scales of modelling or aggregation. I would hence suggest rewriting this sentence.

P. 3/ L. 66: A specification which types of biases are quantified would make it clearer for the reader. Furthermore, the term of „event duration" might be misleading, since it refers to the temporal aggregation window, not however to the duration of the studied rainfall events.

P. 5/ L. 106: Is there any quantification of the uncertainties of CombiPrecip available? Such work has e.g. been done for the radar-based, gauge-adjusted quantitative precipitation estimates over neighbouring Germany (RADOLAN; Kreklow et al., 2020). If the uncertainties of CombiPrecip are known, this would be valuable information to be added.

P. 6/ L. 152: Here you could add a note that 20 years is too short for a robust climatological analysis.

P. 8/ L. 178: Areal mean precipitation time series are not only constructed for the climate models but also for the observations, right?

P. 8/ L. 179: This sentence only becomes clear later on, as here it is ambiguous whether e.g. a window size of 4 is a window with 4 pixels or a window with 4 x 4 pixels. I suggest rephrasing it.

P. 9/ Eq. 3: The given formula suggests that in the computation of the bias, g is the same number for the RCMs and the observations. With the RCMs at 12 km resolution and the observations at 3 km resolution, as an example $g = 3$ would mean an effective area of 1296 km² for the RCMs, but of 81 km² for the observations. In keeping, it is not entirely clear, how the metrics are computed e.g. in

Fig. 2 f). The question arises whether the RCMs median at 1296 km² is compared to the observations at 1089 km² or at 81 km² (see preceding comment). A reformulation and some more information regarding the computation of the bias would help clarify this.

P. 10/ L. 236 f.: At what spatial resolution is the partitioning into elevation groups being performed?

P. 11/ Fig. 2 & P. 12/ Fig. 3: You have chosen to juxtapose the observations and CPMs with an 11 x 11 window (1089 km²) to the RCMs with a 3 x 3 window (1296 km²). What favored this choice against an 12 x 12 window for the observations and CPMs (which would have given an equal effective area of 1296 km² for all sets)?

P. 15/ L. 305: I suggest changing the word „strong“, as it suggests higher magnitudes of bias, which are not visible from the plot. It only becomes clear that the heavy underestimation is more widespread, but not necessarily more pronounced.

P. 16/ L. 322: A specification that it is the CPM ensemble median which shows near-zero bias would be recommended, given the large member spread.

P. 16/ L. 341: The statement is given that mid elevations show mild overestimation of 20-y return levels by CPMs at both 1-h and 24-h durations. While this is true across all scales at 24-h durations, however at 1-h durations, the bias only grows beyond 1 for areas above 1000 km². This could be clarified.

P. 20/ L. 400: Please present Fig. S8 in the results section as well. In the current version of the manuscript, if is referred to for the first time in the discussion.

P. 22/ L. 476: A few studies have included analyses on the scale dependence of convection-parameterizing RCM bias, e.g. Prein et al. (2016) and Fantini et al. (2018). The study presented here does so for the first time for CPMs. I suggest making this distinction.

Technical Corrections

P. 2/ L. 44: The abbreviation CPMs has already been given in L. 42.

P. 3/ L. 85: Based on the map given in Fig. 1, Ticino would be situated in the Southern Alps, rather than in the south of the climatological region of the Alps. If this is correct, I would suggest capitalising the S (i.e. „the Southern Alps“, rather than „the southern Alps“).

P. 7/ Table 1: „WRF3.8.1DA (12km)“ and „HadGEM3 (25km)“ both need a space between the number and the unit.

P. 9/ L. 233: mean Bias → mean bias

P. 13/ L. 271: The observations at the smallest spatial scale are shown in Fig. 3a not in Fig. 3c.

P. 15/ L. 301 & Appendix/ Fig. S4: The native resolution RCM resolution is of approximately 144 km², not of 114 km².

P. 15/ L. 302: An area of 1296 km² equals to 3x3 RCM grids, not 6x6 grids.

P. 15/ L. 310: The abbreviation mAM is not used in the plots, rather „mean AM“ is used.

P. 16/ L. 328: In the manuscript text, the reference to Fig. S7 comes before any reference to Fig. S5 and Fig. S6.

P. 19/ Fig. 8 & Appendix/ Fig. S1, S2, S3, S4, S7: I suggest including the abbreviation (AM) into the respective figure captions or into the main text.

P. 22/ L. 462: There is a spare point here.

References

Fantini, A., Raffaele, F., Torma, C., Bacer, S., Coppola, E., Giorgi, F., Ahrens, B., Dubois, C., Sanchez, E., and Verdecchia, M.: Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-CORDEX and EURO-CORDEX experiments against high resolution observations, Clim Dyn, 51, 877–900, https://doi.org/10.1007/s00382-016-3453-4, 2018.

Kreklow, J., Tetzlaff, B., Burkhard, B., and Kuhnt, G.: Radar-Based Precipitation Climatology in Germany—Developments, Uncertainties and Potentials, Atmosphere, 11, 217, https://doi.org/10.3390/atmos11020217, 2020.

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., van Meijgaard, E., Déqué, M., and Nikulin, G.: Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits?, Clim Dyn, 46, 383–412, https://doi.org/10.1007/s00382-015-2589-y, 2016.