Reply to the comments of reviewer 2

We thank the reviewer for the thorough reading of the manuscript and the valuable remarks that helped us to improve the manuscript. In the following, the original reviewer comments are given in italic and all line numbers and figure numbers refer to the original submitted version that was reviewed if not mentioned otherwise.

Reply to review of reviewer 2

Detailed river runoff data are needed for the coastal ocean models. Unfortunately, the observational data are often with too coarse resolution and/or of insufficient accuracy. Data coverage may be improved by hydrological modelling, but it creates specific errors. As a result of freshwater input bias, modelled ocean state may drift away from the real salinities and related variables, especially in climate-related long-term studies. Therefore, the approach of the present MS to elaborate and test bias correction for simulated river runoff is highly needed and scientifically interesting.

Obtained results are convincing. In particular, the model system gives too large peak discharges in the Elbe, Rhine, Weser and Odra rivers, while the modelled low discharges are close to the observed values. This range-dependent bias is effectively corrected by the applied method.

The MS has a good quality and should be published. When reading I found some problems, resolving of which in the revised MS would help the readers.

Thank you for the positive evaluation of our manuscript.

A) *One problem lies within the title, that brings forward "three-part". It is not clear what does it tell scientifically; what is the difference from "two-part" or "four-part" bias correction? I would prefer a scientific name for the developed bias corrected method. Is it the "quantile mapping bias correction"? Later it has been explained that the method has been specifically applied to "different correction factors for low, medium and high percentile ranges of river runoff over Europe". Note that in the literature there are numerous uses of "three-step" bias correction that differ from the basic principles as well as from applications. Another possible interpretation of "three-part" is: (1) quantile mapping and correction at measurement sites, (2) transfer of bias correction to the river mouths, (3) interpolation of bias to the unsampled coastal sections. In short, the MS would benefit from having a well-defined title, not leaving space for ambiguity.*

We modified the title to pay regard to your concerns:

New title**: A three-quantile bias correction with spatial transfer for the correction of simulated European river runoff to force ocean models**

Within the text, we will mainly replace 'three-part' by 'three-quantile' bias correction.

B) *Quantile mapping is a well-known approach for bias correction in climate studies. The MS would benefit, if some basic statistical concepts and references are introduced in the introduction. Presently, the introductory paragraph in lines 53-66 is rather fragmentary and does not provide sufficient background information for the study. Specific to the river discharge, this paragraph is too much listing specific case studies and is lacking the references to the well-known studies, except for Madadgar et al (2014).*

We modified several parts of our introduction to make the aim of our study clearer and why we chose the bias correction method that is presented here.

We modified the text in Line 7-8:

> "… Baltic Sea. To achieve low biases in riverine freshwater inflow in large-scale climate applications, a bias correction is required that can be applied in periods where runoff observations are not available and that allows spatial transferability of its correction factors. In order to meet these requirements, we have …"

We modified the text in Line 54-55:

> "The bias correction of river runoff is an approach that has been used particularly for short term hydrological forecasts and ensemble predictions of up to six months. However, these approaches (see, e.g., those listed in Kim et al. 2021 and Madadgar et al. (2014)) are often specifically trimmed to flood forecasts. Therefore, they often require the existence of observed values from previous time steps so that that are not applicable in climate change studies, such as autoregression models (Kim et al., 2021) or components of a Bayesian forecasting system (Krzysztofowicz and Maranzano, 2004). Others like non-parametric methods based on Bayesian approaches as proposed by Brown and Seo (2010; 2012) need a large number of ensemble members (Madadgar et al. 2014).
>
> Recently, bias correction …"

We added new text in line 60:

> "… (USA). A criticism of using quantile-mapping in flood forecasting is that it does not maintain the pairing of corresponding simulated and observed flows (Madadgar et al. 2014). Madadgar et al. (2014) also noted …"

We added new text after line 78:

> "Therefore, we decided to not apply methods that employ detailed modifications of the discharge curves for specific rivers such as those methods that use complex matrix arithmetic of observed and simulated discharge time series (e.g. Zhao et al., 2011), or the common quantile-mapping approaches, The latter are conducted using a lot of bins, so that the bias in the discharge curve of a specific river can be strongly reduced. However, these detailed correction factors for every bin may likely not be transferred to other locations. It may work for the same river if station and river mouth are relatively close to each other, but certainly may not be valid for the transfer to neighbouring catchments."

C) *Introductory paragraph in lines 37-52 could be made more informative, backing the statements with appropriate references. In particular, the principle that "the river runoff is consistent with the atmospheric forcing" could be more elaborated and referenced. How bias correction changes this consistency? (Later this question is briefly discussed in lines 71-78).*

We modified the text in line 41:

> "… atmospheric forcing (e.g. Vinayachandran et al., 2015; Hagemann and Stacke, 2022), i.e. that the impact of weather events and trends in the atmospheric forcing is transferred via the river runoff into the ocean."

In order to clarify that we specifically consider temporal consistency, we modified the text and added some analysis on how our bias correction affects this consistency. Hence, we modified the text in:

Line 71: "…high degree of temporal consistency …"

Line 73: "…, a bias correction with as little fitting or modification of the daily sequence of runoff as possible is desired."

In addition, we added the supplementary figure S1 and the following text after line 320.

> "In order to analyse how much the bias correction affects the daily sequence of river runoff at the station locations, we calculated the correlation between the simulated discharges and the observations. Supplementary Figure S1 shows that the correlation patterns of HDW and HDW-BC with observed discharges are quite similar. For rivers where differences can be identified, the correlation mostly increases for HDW-BC. The correlation between HDW and HDW-BC is generally higher than 0.95, and only a very few rivers show correlations lower than 0.9. These rivers are usually rivers that are heavily influenced by human activities, such as the Volga and the Luleaelven."

We also modified the text in line 281-282:

> "… mean bias, the Pearson correlation coefficient and the Kling-Gupta efficiency (KGE; Gupta et al., 2009; Kling et al., 2012). All metrics …"

D) *The sub-section "2.4 Bias correction of river runoff" lacks references to the basic principles. It should have clear description of already known procedures (with references) and new/specific approaches. If the method is completely novel, this could be spelled out; then also more justification should be presented.*

We modified the text in lines 147-148:

> "We have developed a bias correction method for river runoff that uses correction factors for three quantiles and includes a spatial transfer of these factors. We note that our three-quantile bias correction is similar to a very coarse quantile mapping. The latter has been introduced in climate change impact research to correct for significant biases

in data produced by global and regional climate models. Quantile mapping is a distribution mapping in which the distribution function of climate values is corrected to match the observed distribution function. Details of such mapping applied to precipitation and surface air temperature can be found, for example, Piani et al. (2010) and Teutschbein et al. (2011). Our bias correction method involves several steps. First, different correction factors for low, medium and high percentiles are calculated at the station locations and then applied at the respective river mouths."

E) Lines 181-183 say that "bias correction can lead to spurious daily jumps in discharge when the percentile boundary is crossed and the bias correction factors differ between the percentile ranges". This is suppressed by applied smoothing. It would be interesting to know if these jumps could be avoided (not suppressed) by some elaboration of the methods, for example by introducing continuous correction factors instead of stepwise correction.

Using continuous correction factors might improve the performance of the bias correction at the station locations but may have a degrading effect on the transfer to river mouth locations and neighboring catchments. On purpose, we did not generate more than three percentile ranges as we think that the higher the number of correction factors is, the less spatially transferable they are. As the spatial transfer of bias correction factors for river runoff is a rather new approach (This, we point our more clearly in our response to the comment of reviewer 1 on line 174-180.), we cannot undermine this statement by other literature. However, our evaluation of inflow into sea basins is supporting the validity of our approach.

Note also that the smoothing work around is actually a linear interpolation of the correction factors between the two neighboring percentile ranges (low & middle, middle & high).

Some technical remarks.

1. Principles and approaches of the HydroPy model and HD model (is it a unique name?) should be shortly outlined in the beginning of sections 2.2 and 2.3.

Yes, HD model is a unique name that has not been duplicated since the publication of its first version by Hagemann and Dumenil (1998).

We added in line 121:

"HydroPy (Stacke and Hagemann, 2021) is a state-of-the-art global hydrology model for which no model calibration was performed for its setup. Within global hydrological modelling, the usage of uncalibrated models is rather common (see, e.g., Haddeland et al., 2011), even though some models exist that are calibrated for global studies. In the present study, HydroPy was driven …"

We added in line 134:

"The HD model (Hagemann et al., 2020) is a well-established river routing model that is implemented in a range of global and regional model systems. As noted in Hagemann et al. (2020), no river specific parameter adjustments were conducted in the HD model to enable its applicability for climate change studies and over catchments, where no daily discharges are available at a downstream station. To simulate discharge with the HD model, we used …"

Please note that we originally planned to add more details in the sections 2.2 and 2.3. However, the editor correctly pointed out that these details comprised a lot of doubling of text with the information provided in Hagemann et al. (2020) and Hagemann and Stacke (2022). Consequently, we removed these details and referred to the respective publications.

2. Smoothing formulae in lines 185-188 have different notations for q and Q than in lines 153-155.

In lines 185-188, we corrected the notation by using Q_p and Q_{100-p}.

3. Titles of Table 5 and Figure 12 contain "fractional area coverage" that is not defined.

We modified the titles of Table 5 and the caption of Figure 12:

"… fractional catchment coverage (see Table 1) is …"

4. Legends of Figures 5, 7-10 contain "HD5.2-HydroPy-WFDE5" that are not defined.

According to the comments of reviewer 1, we replaced HD5-WFDE5 by HDW, HD5-GSWP3 by HDG, and HD5-Bias C. by HD-BC. In this respect, HDW-BC and HDG-BC refer to the bias corrected data of HDW and HDG, respectively. In the respective figures, HD5.2-HydroPy-WFDE5 is replaced by HDW and HD5.2-HydroPy-WFDE5 by HDG.

5. I counted 37 abbreviations; alphabetically from "20CR" to "WFDE5". Some of the abbreviations are well-known like ECMWF, ESA, HELCOM, NASA, OSPAR and they do not complicate the reading. At the same time, there are abbreviations that occur only once (AHOI, GCOAST, RCSM, they appear only in the conclusions, BSH, RSME – they are not defined, GSM) or a few times (DB, EMS, GRDC, ISIMIP …). Reader would benefit from less abbreviations.

We reduced the number of abbreviations:

We removed 20CR, GCOAST, AHOI, ESA; GRDC, GSM, NASA; RCSM, SMAP; SMOS; WFD, BSH, DKRZ, WDCC and used only their respective full names. We will check the manuscript for further obsolete abbreviations.

We added the definition of RSME in Sect. 2.7

Added references

Brown, J. D., and Seo, D. J.: A nonparametric postprocessor for bias correction of hydrometeorological and hydrologic ensemble forecasts, J. Hydrometeorol., 11, 642-665, https://doi.org/10.1175/2009jhm1188.1, 2010.

Brown, J. D., and Seo, D. J.: Evaluation of a nonparametric post-processor for bias correction and uncertainty estimation of hydrologic predictions, Hydrol. Process., 27, 83-105, https://doi.org/10.1002/hyp.9263, 2012.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel estimate of the global terrestrial water balance: setup and first results, J. Hydrometeorol., 12, 869-884, https://doi.org/10.1175/2011jhm1324.1, 2011.

Krzysztofowicz, R., and Maranzano, C. J.: Hydrologic uncertainty processor for probabilistic stage transition forecasting, J Hydrol, 293, 57-73, https://doi.org/10.1016/j.jhydrol.2004.01.003, 2004.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, J Hydrol, 395, 199-215, https://doi.org/10.1016/j.jhydrol.2010.10.024, 2010.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram., J Geophys Res-Atmos, 106, 7183-7192, https://doi.org/Doi 10.1029/2000jd900719, 2001.

Teutschbein, C., Wetterhall, F., and Seibert, J.: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale, Clim. Dyn., 37, 2087-2105, https://doi.org/10.1007/s00382-010-0979-8, 2011.

Vinayachandran, P. N., Jahfer, S., and Nanjundiah, R. S.: Impact of river runoff into the ocean on Indian summer monsoon, Environmental Research Letters, 10, https://doi.org/10.1088/1748-9326/10/5/054008, 2015.

Zhao, L., Duan, Q., Schaake, J., Ye, A., and Xia, J.: A hydrologic post-processor for ensemble streamflow predictions, Advances in Geosciences, 29, 51-59, https://doi.org/10.5194/adgeo-29-51-2011, 2011.