We thank reviewer 3 for the constructive comments. Below we give the detailed comments to most of his/her points in italic.

GHM setup: When forced using ERAI, the GHM severely underestimates streamflow in the Rhine basin (Figure 5c), in particular in summer, when as much as half of the streamflow is missing. The authors recognise that the "GHM is too dry in the summer months", since ERAI precipitation estimates are "good". I raised this issue in a previous round of revisions, and the authors replied that they "have not performed an in-depth analysis on the performance of W3RA as this study focuses on the sensitivity to resolution." I regret to say that I do not agree with this argument. Before increasing the resolution, the authors should have made sure that the basic hydrological behaviour of the basin (its water balance) is adequately captured. In this specific case, the streamflow underestimation in summer indicates that the GHM is not adequately setup for the Rhine catchment.

In our analysis, we found comparable simulated discharge with ERAI for the Rhine as Photiadou et al (2011), who used a high resolution HBV96 calibrated hydrological model for the Rhine forced with ERAI. In contrast to their claims of underestimated precipitation, we found that the total precipitation over the Rhine catchment in ERAI is comparable with the observations. Due to resolution issues the distribution of the rainfall over the catchment might not be optimal, which could be important for snow build up in the Alps and hence could cause too low discharges in spring/summer. We found that the total actual evaporation over the Rhine basin is overestimated when compared to GLEAM. This could further contribute to the simulation results we obtained. However, the observed precipitation, and in particular evaporation contain large uncertainties (Bosilovich, 2008 ; Miralles, 2011). In our manuscript, we present a transparent and fair assessment of the mismatch between model and data. We could conduct a more in depth analysis of this issue. However, in our opinion, this will hardly contribute to the main aim of this work, which is testing the effect of resolution on simulating the hydrological cycle. We leave it up to the editor to decide if additional analysis is needed.

From a more general perspective, the authors mentioned in their reply that "The global hydrological model which we use is not calibrated for the specific catchments of this study. In general, calibration of global hydrological models is limited." I recognise that, when run at the global scale, GHMs are challenging to calibrate, but for this study, I consider that the authors should have taken the time to adjust model parameters in order to provide acceptable streamflow simulations for the two basins they selected, before increasing the GHM resolution.

This study focuses on the effect of increased resolution in simulating the hydrological cycle of river basins. Therefore, we use a global climate model and a global hydrological model. We examine our hypothesis by focussing on two large river basins, with different characteristics. To decrease the computing time we only run the global hydrological model for the two selected basins. However, we still do this from the viewpoint of 'global hydrology', i.e. using a global hydrological model not specifically designed for the two basins. If we want to obtain the best results in modelling the hydrological basins over the Rhine and Mississippi basin, we should have chosen specific regional models build for the specific basins, which is not the research question of the study. We intend to add an extra paragraph on the issue on global modelling and calibrated catchment scale modelling in response to reviewer 3.

Parameter estimation under higher resolution: In the revised manuscript, the authors mention that for the GHM, "to allow a fair comparison between the two model resolutions, we remapped these parameters from the 0.5_ to the 0.05_ resolution." The authors explained in a previous round of reviews that "we remap the parameters from the low to the high resolution using the resample statement in PCRaster (Karssenberg et al., 2010)". How the remapping exactly works is unfortunately not documented in the revised manuscript, but my impression is that a resampling routine alone is not sufficient to incorporate the new data and knowledge necessary to enable the model to resolve smaller-scale processes. In their revised abstract, the authors report that "Increasing the resolution of vegetation and orography in the high resolution GHM (from 0.5_ to 0.05_) shows no significant differences in discharge for both basins, likely because the hydrological processes depend highly on model parameter values that are not readily available at high resolution." All this indicates that although the GHM was run at higher resolution, it was probably run using parameters at an effective resolution close to the original (coarser) resolution, because model parameters are higher resolution were not "readily available". This likely explains why the benefits of increasing the GHM resolution were limited. Arguably, increasing the resolution of a model goes beyond decreasing its grid spacing, it should also involve the incorporation of new data and knowledge on process representation across scales. It is my impression that this part is essentially missing and this is preventing us from truly assessing the benefits of the increased resolution.

In every version of the manuscript we mentioned "to allow a fair comparison between the two model resolutions, we remapped these parameters from the 0.5 to the 0.05 degrees resolution." Apparently, this was not clear enough and we will adjust the text to make this procedure clearer.
The reviewer understood correctly that in this study we only tested the impact of increased resolution by increasing resolution and not by including more/different smaller-scale processes or involving new data. We decided to do so, since applying all suggested adaptions to the high-resolution version would imply a different model which is not directly comparable with the low resolution version. Therefore, we investigated if rescaling parameters, and only using known high-resolution parameters (topography and vegetation), at higher resolution would give better results. It was shown by Melsen et al (2016) that parameters can be transferred across the spatial scales, on regional scales from 1 km^2 to 100 km^2, and our work could be seen as a large-scale test of their work. We conclude our paper by stating exactly the view of the reviewer: quick gains are not straightforwardly achieved (line number 9-10-11 on page 27).

References

Bosilovich, M.G., Chen, J., Robertson, F.R. and Adler, R.F., 2008. Evaluation of global precipitation in reanalyses. *Journal of applied meteorology and climatology*, *47*(9), pp.2279-2299.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M., and Uijlenhoet, R.: Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin, Hydrol. Earth Syst. Sci., 20, 2207-2226, https://doi.org/10.5194/hess-20-2207-2016, 2016.

Miralles, D.G., De Jeu, R.A.M., Gash, J.H.C., Holmes, T.R.H. and Dolman, A.J., 2011. Magnitude and variability of land evaporation and its components at the global scale.

Photiadou, C. S.,Weerts, A. H., and van den Hurk, B. J. J. M.: Evaluation of two precipitation data sets for the Rhine River using streamflow simulations, Hydrology and Earth System Sciences, 15, 3355–3366, doi:10.5194/hess-15-3355-2011, 2011.