Interactive comment on "The critical role of uncertainty in projections of hydrological Extremes" *by* Hadush K. Meresa and Renata J. Romanowicz

Hadush K. Meresa and Renata J. Romanowicz

hadush@igf.edu.pl

Dear Editor and Reviewers,

Thank you very much for the constructive comments that will help to considerably improve and clarify the manuscript. The reviewers, put enormous effort into proof-reading our paper line by line and trying to clarify all the less-than-satisfactory statements and mistakes. We believe that the input will improve the manuscript significantly. All comments have been addressed point-by-point. Following the reviewers' feedback we will make the corresponding changes in the manuscript. Reviewers' comments are in italics.

Anonymous Referee #1

RC1. General comments

This paper is about the uncertainty of extreme flows with climate change. For that purpose, the authors use seven combinations of global climate models (GCMs) and regional climate models (RCMs) with one greenhouse gas concentration scenario to represent uncertainty in climate change. Furthermore, they use the GLUE method to represent hydrological parameter uncertainty and uncertainty in extreme value distribution parameters to represent the uncertainty in the statistical extreme value distribution. These three sources of uncertainty are investigated using the HBV hydrological model applied to a medium-sized Polish catchment. Although the topic is interesting and relevant for this journal, the paper is moderately written, lacks clarity in parts of the methodology and only briefly discusses results and insufficiently puts outcomes into perspective. For instance, the seemingly arbitrary choice to consider the three uncertainty sources is not justified. Are these three sources the most important ones or the easiest ones to quantify? Furthermore, the uncertainty due to the use of a particular extreme value distribution is not clearly and completely incorporated. A final example is the presentation and analysis of results, such as the analysis of annual maximum precipitation and temperature in relation with annual maximum flows and in particular annual minimum flows. In this case and several other cases it is not always clear which results are shown, why they are shown and what can be concluded from the results. Many other specific (and important) comments can be found below. Furthermore, the English writing style and grammar is moderate (including several typos); some examples can be found in the section 'technical corrections'.

AC1. General answer The paper lacks clarity in parts of the methodology and only briefly discusses results and insufficiently puts outcomes into perspective. Following the reviewer's general and specific comments, the clarity of the methodology will be

improved and the outcomes will be described in a wider perspective.

For instance, the seemingly arbitrary choice to consider the three uncertainty sources is not justified. Are these three sources the most important ones or the easiest ones to quantify

The choice of three particular sources of uncertainty, namely, a set of climate model ensembles, hydrological model parameter uncertainty and uncertainty in fitting extreme value distribution, was dictated by one of the aims of the research – i.e. an assessment of influence of hydrological model uncertainty on projections of low- and high-flow extremes and the relative contribution of that "predictive" uncertainty in the spread of extreme indices related to the climatic model spread and the distribution fitting error.

This choice followed from a discussion on all the sources of uncertainties and a review of research done so far on the assessment of uncertainty of projections of hydrological extremes. The "predictive" model uncertainty is the only one which can be decreased when conditioned on the observations. The other sources of uncertainty have an "epistemic" nature and cannot be decreased. Bearing in mind the aims of the study, we restricted the sources of epistemic uncertainty to those which have the largest impact – i.e. climate model spread, omitting the uncertainty related to bias correction or geography. The error related to the distribution fit was included as an essential part of the extreme index evaluation, which requires extrapolation of annual maximum or minimum flow distributions to higher order quantiles (e.g. 1-in-100 year, or 1-in-200 year). In this paper, the error related to the evaluation of maximum and minimum annual flow statistic was treated as epistemic, that means, not conditioned on real observations.

In addition, Osuch et al. (2016) presented the influence of emission scenario, climate model, bias correction method and geography on flow indices in a case study that included the same catchment, Biala Tarnowska. Therefore we wanted to avoid the repetition.

In this regard, our paper is an extension of the former paper, focusing on the influence

of hydrologic model uncertainty on annual maximum and minimum flow projections. In our opinion, including the other sources of uncertainty would obscure our aim.

The uncertainty due to the use of a particular extreme value distribution is not clearly and completely incorporated.

The choice of extreme value distribution followed the validation of suitability of this distribution to describe the projected annual maximum and minimum flows using the probability plots. The MATLAB- based GEV distribution fitting algorithm was applied to all the climate models and the a posteriori hydrological model parameter set. This algorithm provides the estimates of 0.95 confidence bands for the distribution parameters. These parameters were subsequently used to obtain upper and lower confidence bands of the distribution through the inverse GEV model. In order to simplify the procedure, instead of sampling from the GEV parameters within the parameter space common to all hydrologic and climate model simulations, we sampled from each set of parameters assuming a normal distribution with the variance specified by the parameter upper and lower 0.95 confidence value, and in addition, assuming the independence of the GEV model parameters. The obtained 0.95 GEV distribution confidence values were used to estimate the spread of results related to the distribution fit.

Bearing in mind the large number of simulations, it was not possible to choose the best distribution for each projected time series. Furthermore, the aim of this study is to assess the ranges of uncertainty of extreme indices rather than their exact values.

A final example is the presentation and analysis of results, such as the analysis of annual maximum precipitation and temperature in relation with annual maximum flows and in particular annual minimum flows. In this case and several other cases it is not always clear which results are shown, why they are shown and what can be concluded from the results.

We agree with the reviewer that the presentation should be much improved and clarified. The following explanation will be added.

In the following section, we present an analysis of the variability of maximum precipitation and temperature series on an annual basis to see the correlation between the projected hydrological extremes and the input climate extremes. In Fig. 2, raw annual maximum daily precipitation and temperature time series for the Biala Tarnowska catchment obtained from the seven GCM/RCM models under the RCP4.5 scenario are shown. The periods cover the whole length of historical and projected years (1971-2100). The upper panel of Fig. 2 presents annual minimum precipitation based on corrected precipitation projections (the upper panel), annual maximum precipitation based on raw projections (middle panel) and temperature mean projections for corrected data are presented in the lower panel of Fig. 2.

The results show a visible increase of the annual maximum temperature and an increase of temporal variability with time, in particular for the maximum precipitation values from 2016 onward.

The English writing style and grammar

The English and grammar was, and will be, checked by a native English speaker.

Specific comments

RC1. P1, L7-9: It is not clear what is meant with a 'multi-model approach' and which steps are followed. AC1. The 'multi-model approach' is an approach which considers multiple climate models and multiple hydrological parameter sets. To avoid possible confusion this wording will be changed: "The approach followed is based on …."

RC2. P2, L9-11: The first question probably is related to the magnitude of the uncertainty, since this is still largely unknown and not systematically investigated. AC2. The sentence will be changed to: The question arises as to how large the uncertainty is and if it is acceptable to the end-user in adaptations to climate change and flood and drought risk assessments.

RC 3. P2, L15-16: "….can never be accurately evaluated …." is a very strong state-

ment, please rephrase. AC3. The sentence will be rephrased to: "However, complex hydrological and climate models are difficult to be accurately evaluated, because of uncertainty in observations, parameters and model structure simplifications."

RC 4. P2, L24-P3, L2: The authors mainly consider hydrological model and parameter uncertainty in their review. It might be worthwhile to firstly give an overview of all uncertainties involved in this type of studies including a classification. One such classification could be input, (hydrological) model system and output, and the literature can be reviewed accordingly. Now, uncertainties in the input (scenarios, GCMs, RCMs, downscaling, initial conditions etc.) are hardly reviewed. A complete overview of the uncertainties will also enable a better justification of the uncertainty sources considered in this study (see also page 3, lines 4-5). AC 4. As already discussed, the influence of other sources of uncertainty, including the choice of emission scenario, climate models (GCM/RCMs), downscaling and catchment type was performed by Osuch et al (2016) using a case study that included the catchment used in this paper. This paper was focused on predictive hydrological uncertainty to show that different objective functions should be applied when high and low flow extremes are considered. Apart from hydrological model parameters, seven climate models were also used and the spread relating to extreme index distribution was taken into account. However, the reviewer made the important point that our aims were not clearly enough presented and that the review of different sources of uncertainty would help to improve the presentation of that aim considerably. This part of the paper will be changed to justify better the choice of those three sources of uncertainty.

RC 5. P3, L14-15: The question is whether you can determine the uncertainty due to the choice of the extreme value distribution ('distribution fit') using time series of different lengths. When assessing effects of time series with different lengths on the results you might get an estimate of the influence of data quantity on the uncertainty in the results, but not of the influence of the goodness-of-fit of the distribution on the uncertainty. Furthermore, it seems only part of the statistical uncertainty is assessed

in this way, since for instance the influence of different extreme value distributions and extrapolation uncertainty is not taken into account. AC 5. Thanks to the reviewer, it is good point. However, we did not use different lengths of time series in order to determine the uncertainty due to the choice of the extreme value distribution. The sentence was misunderstood. In order to make our presentation more clear the sentence should read as follows: The uncertainty related to the distribution fit is analysed in two stages, using, separately, two different lengths of flow record to derive the quantiles of maximum and minimum annual flows, the 30-year long and 130-year long time series of future flow projections. The popular method of a comparison of changes in flow quantiles between the reference period and future periods is based on relatively short (e.g. 30-year) periods. It is well known that an extrapolation of a distribution function based on 30-year long time series towards 1-in-100 year quantiles involves very large errors (Strupczewski et al. 2011). Even the estimates of 1-in-30 year quantiles based on the 30-year long data are biased with large errors. We compare these errors with those involved on 1-in-100 year estimates obtained using the 130 year long time series. The question we pose is whether the estimates of future trends of extreme indices and their relative changes can be useful at all in view of the uncertainties involved.

RC 6. P3, L29-30: How many precipitation stations have been used to assess the catchment average precipitation (assuming lumped hydrological modelling has been carried out)? Has any elevation (or other) correction been incorporated? AC 7. We used five gauging stations to derive aerial precipitation in the catchment using Thiessen polygons. We did not use any elevation correction in this paper. However it was applied to the same catchment by Benninga et al (2017) and showed that the increase in precipitation due to the elevation is about 3%. RC 7. P4, L11: An important uncertainty source in climate impact studies is the uncertainty due to greenhouse gas emission scenarios. Hence, a limitation of this study is the use of only one emission scenario (RCP4.5) while one would expect the use of at least two scenarios (which are available in EURO-CORDEX). At least the authors should explain the implications of this limitation for their results. AC 7. The RCP 4.5 was applied because it is a sta-

bilization scenario and thus assumes the imposition of emissions mitigation policies. The RCP 4.5 is derived from its own "reference", or "no-climate-policy", scenario. This reference scenario is unique to RCP 4.5 and differs from RCP 8.5, RCP 6.0 and RCP 2.6 (Smith and Wigley 2006; Clarke et al. 2007; Wise et al. 2009). The influence of the emission scenario on flood indices was studied by Osuch et al. (2016) whilst the low flows were analysed by Osuch et al (2017; submitted for publication). Both those studies indicated that emission scenario choice has a relatively small influence on the results. The implication of the choice of only one emission scenario will be explained in the revision.

RC 8. P5, L9: Why is QM applied in this study? The reasoning behind this choice is not completely clear from the preceding sentences. AC 8. Many popular existing bias correction methods have been reviewed and compared and quantile mapping (QM) was found to outperform other methods (Gudmundsson et al., 2012; Teutschbein and Seibert, 2013; Chen et al., 2013; Osuch et al., 2016). More recently, the standard non-parametric QM method has been adapted to more explicitly preserve the raw modelled climate change signals (Willems and Vrac, 2011; Sunyer et al., 2014; Cannon et al., 2015). This means, in the QM method, that a raw modelled value is always corrected by the same value of bias or error that is determined by its respective quantile in the reference period.

RC 9. P5, L18-19: Did Osuch et al. (2015) model the same catchment as in this study and therefore, can it be assumed that the same five parameters are sensitive? And are the same five parameters sensitive for low flows and for high flows? That would be remarkable. AC9. The HBV model was applied in different hydro-climatic condition in Poland by different researchers and they found the five most sensitive parameters for both high flow and low flow characteristics. The set of five parameters chosen in this study was dictated by the most common catchment conditions. Therefore it is not surprising that the same parameters are sensitive in both high and low flow conditions. However, in this study we used two objective functions to encapsulate the high and low

flow characteristics instead of selected best parameters only belonging to low flow and high flow.

RC10. P6, L15-16: How many Monte Carlo simulations have been executed and is this number sufficient (compare with literature)? AC 10. 20000 MC simulations were executed. Many research papers recommend above 10, 000 MC (e.g. Xiaoli Jin et al., 2010; Romanowicz et al., 2013; Houska T. et al., 2014).

RC 11. P6, L22: Is it common practice to determine the thresholds in an iterative way? The determination of the threshold based on the requirement that 95% of the observations should be in the 95% confidence interval seems to be reasonable. However, please refer to other studies employing the same approach. AC 11. To our knowledge, it is a common practice. The thresholds determine the variance of the predictions. Too high a threshold results in too narrow confidence bands. By iteration we meant the "trial and error approach" which does not involve any algorithm. We would be surprised if the iterative determination of threshold values has not yet been introduced, but we are not aware of any studies that have followed this approach. We will change the wording to avoid confusion.

RC 12. P7, L4-5: In general it is doubtful whether distributions with a 'large' number of parameters will model data in a more accurate way than distributions with a small number. This partly depends on the data quantity and quality and similarly as in hydrological modelling there will be a balance between the complexity of the distribution (i.e. number of parameters) and the amount of data (and quality). AC 12. We agree with the reviewer that there must be a balance achieved between the complexity of the distribution (i.e. the number of parameters) and the quality of data. We admit that this sentence can be deleted as it is a too large generalization.

RC 13. P7, L7-8: What does an 'overall good performance' mean? Compared to which other distributions? AC 13. A number of distributions was tested including a three-parameter lognormal and an inverse Gaussian. GEV was the only distribution

that performed well both for the high and low flow extremes. Although it was not necessary to use the same distribution for both extremes, it made our discussion more transparent.

RC 14. P7, L25-27: It is not completely clear why the analyses are performed for a period of 130 years. Since the manuscript is about impacts of climate change on hydrological extremes, you would expect a comparison between historic and future climate conditions. Furthermore, climate change automatically implies the existence of nonstationarity and as such, by considering a period of 130 years assuming stationarity by using the same extreme value distribution will result in serious flaws. AC 14. We do not think that the impact of climate change on hydrological extremes should be based on a comparison between historic and future climate conditions. What we propose here is to study the trend of projected indices instead of the "change". The Biala Tarnowska catchment does not show any non-stationarity in the extreme flow events (Meresa et al. 2017, submitted for publication). Therefore it is a suitable catchment to compare both approaches. We are aware that non-stationary flood frequency analysis has to be applied for non-stationary extreme events. We want to show here that taking 30-year long time-series to compare between reference and future periods involves large uncertainty even for 30-year return flows. The uncertainty ranges of 30-year return period flows obtained using 130-year long time series can be nearly four times smaller.

RC 15. P8, L7-12: The idea behind this section is not clear. Why is the trend in daily annual maximum precipitation and temperature analysed while the interest is in uncertainty in hydrological indices with climate change? Moreover, why is the daily annual maximum precipitation of interest and not for instance the two-day or three day precipitation (which might be stronger correlated to annual maximum discharge values)? Which temporal resolutions of precipitation are relevant for annual minimum flows? And what is the supposed role of daily annual maximum AC 15. The idea behind presenting the precipitation and temperature patterns was to show the variability of driving forces behind the changes in flow extreme index. However, the idea was not properly

explained and followed. For a catchment of that size, daily maximum and mean sums of precipitation are well correlated with the flow patterns. The temperature patterns, on the other hand, present the changes in the evaporation losses and possibly, indicate the changes in flood regime.

RC 16. P8, L14-20: How have the different criteria for high and low flows been applied in continuous hydrological modelling for periods of 30 years (or more)? When is the 'high flow' parameter set being used and when the 'low flow' one? What is the threshold for low flows and high flows; a specific discharge value or exceedance frequency? AC 16. As explained in section 3.3, we applied a stochastic formulation to the estimation of the HBV model parameters. That means that 20000 simulations of the HBV model were run for the 30-year long calibration period with parameters sampled randomly within the assumed parameter ranges. The calibration is performed using the observed precipitation, temperature and flow records. We applied logNSE criterion for low flow and NSE criterion for high flow index to all the simulated flow series. Then we evaluated thresholds for the criteria, called likelihood thresholds, based on the requirement of 95% of the observations should be in the 95% confidence interval separately for high and low flows (Table 3). In other words, we do not have one "high" or "low" parameter set but we have two multiple sets (each including thousands of parameter sets) representing "high" and "low" flow models.

RC 17. P9, L5: Which best parameter sets are meant here? When is the best low flow parameter set used and when the best high flow parameter set? AC 17. Results shown in Fig. 4 were obtained from the HBV model simulations fed by the precipitation and temperature projections obtained from the seven GCM/RCM models under the RCP4.5 scenario for the best parameter sets from the MC parameter samples, giving the highest weights derived from the NSE for the high flows and logNSE for the low flows, respectively. The raw hydro-meteorological projections were applied to study the high flow index whilst bias corrected precipitation and temperature data were used for the low-flow index studies.

RC 18. P9, L7-8: 'twice as large'; where do we see that? AC 18. Sentence will be corrected to: Obtained flow projections shown in Fig. 4, follow the rainfall projections shown in Fig. 2, with annual maximum flow values even four times larger than historical events occurring after 2016 for some GCM/RCM model projections.

RC 19. P9, L12-22: This evaluation is not clear to me. Why do the authors evaluate results at a monthly scale? How can you assess annual maximum flows for each month? What do the authors mean with 'range' of annual maximum flows? AC 19. Thank you, it is corrected in the main manuscript, "annual" was replaced by "monthly". Analysis of variability of monthly flows was presented to illustrate seasonal changes of extreme flows in the near future period and the uncertainty related exclusively to hydrological model uncertainty for each climate model projection. Some changes of seasonality are visible for high flows, but low flows do not show any distinctive differences between reference period and near future. We agree with the reviewer that changes in seasonality are not explored further and therefore this section can be removed, together with Fig. 5.

RC 20. P10, L9-10: The decrease in the spread of Q30 in the far future compared to the near future is strange. The authors should reflect on this. Is it related to the fact that only one RCP scenario is taken into account? AC 20. This smaller spread of the far-future projected changes was also observed in the other climate impact studies on the same catchment (Osuch et al., 2017) for both the RCP4.5 and RCP8.5 emission scenarios using the HBV model for hydrological simulations. Research is on-going to explain that phenomenon.

RC 21. P10, L20-22: Also this observation needs discussion. Why the spread is more evenly distributed for minimum flows compared to maximum flows? AC 21. This is related to the influence of the climate model spread on the simulations (Osuch et al., 2017). It is much bigger for high flows and not very big for the low flows. We also have to remember that low flow simulations used bias-corrected meteorological drivers whilst the high-flow simulations were driven by the raw data. Bias correction decreases

the variability of climate models.

RC 22. P11, L13-14: Are the relative differences for annual minimum flows also smaller? AC 22. Yes. The sentence should read: The relative differences obtained for the annual minimum flow Q30 estimates are smaller, suggesting that low flow quantiles are less susceptible to the errors related to the length of the evaluation period.

RC 23. P12, L7-9: This is an interesting topic, but has not been investigated in this study since only one catchment has been considered. AC 24. We agree with the reviewer, this sentence is out of context and should be deleted.

RC 24. P12, L11-14: This is an interesting result assuming that all methodological steps are logical and correctly carried out. What is the reason for the importance of uncertainty due to climate models for high flow and the important of hydrological model parameter uncertainty for low flows? This is very important and interesting to discuss. AC 24. The important role of hydrological model uncertainty in low flow predictions was already noticed in forecasting (Beninga et al., 2017). That effect can be explained by the ratio of the prediction noise (in this case described by the hydrological model uncertainty) to the input signal which is much higher for low flows.

RC 25. P12, L23-24: What do the authors mean with 'this allows the problem of nonstationarity of model parameters to be avoided'? AC 25. The sentence should read: (iii) conditioning of the hydrological model was performed using different criteria for low and high flows in order to ensure the best model fit for the extremes; this does not solve the problem of non-stationarity of model parameters but allows for focusing on parameter sets adequate for low or high flow regimes.

RC 26. P12, L29-31: This statement seems to be obvious; the larger the ratio of return period vs. data length the higher the uncertainty. However, this extrapolation uncertainty is not explicitly assessed in this manuscript. AC 26. Thank you for the comment. This statement should read "analysis of the influence of length of time series records on the uncertainty bands of the high flow quantile estimates and their changes

C13

suggests that the ranges of quantiles of return periods Q30 are up to four times smaller when the long-term flow projections are used (Table 4). The low flow Q30 quantiles are less influenced by the length of the record.

RC 27. P23, Table 2: The ranges defined by the lower and upper bounds frequently do not match with the optimal values (e.g. for ALFA, PERC, CLFUX). Can you explain this? Furthermore, some lower and upper bounds are exactly the same. Does this indicate that these parameters are deterministic? What about CFMAX (not mentioned as sensitive in section 3.3)? Finally, an upper bound of 2 for LP is impossible and an optimum value of 1 is remarkable at least (it would mean only potential evapotranspiration under fully saturated conditions). AC 27. We thank the reviewer for this comment. There was a mistake in Table 2. The HBV model was calibrated using GLUE and optimal parameter sets were derived in the form of multiple parameter sets, different for the high and low flows. Therefore this table should not include the "optimal" parameter values. The Table 2 will be corrected

Technical corrections

RC 1. P1, L11: What is the distribution fit? AC 1. 'distribution fit' will be changed into "theoretical distribution fit error"

RC 2. P1, L13: What kind of weighting do the authors mean? AC 2. "with a separate criterion for high and low flow extremes"

RC 3. P1, L16: What is the difference between climate model variability and climate projection ensemble spread? Please use a consistent terminology. AC 3. The meaning of "variability" is not the same as "spread". Here we meant "variability".

RC 4. P2, L3: What is inverse modelling in this respect? Is this term commonly used for calibration and validation purposes based on observed (historic) data? AC 4. "Conditioning" can be used here instead of "inverse modelling", if it is clearer. Inverse modelling refers to model parameter calibration based on historical data.

C14

RC 5. P2, L6: "weighting" instead of "weighing". AC5. Corrected: "weighting" instead of "weighing".

RC 6. P3, L8: What is the 'relevant variability' of extreme index estimates? AC 6. Changed into "a direct assessment of variability of extreme index estimates"

RC 7. P3, L19: The case study has already been mentioned. AC 7. Thank you, the sentence will be deleted.

RC 8. P3, L30: The maximum daily precipitation? During which period? AC 8. Thank you, corrected. 'Maximum precipitation was 68.3 mm d-1 and annual mean streamflow is 0.4 m3s-1 over the observation period.' Changed to 'The annual maximum precipitation, annual minimum streamflow and annual mean streamflow of the catchment were 68.3 mm, 0.4 m3s-1 and 5.43 m3s-1 respectively over the observation period (1971-2000)'.

RC 9. P3, L30-31: Which period for the streamflow) Isn't 0.4 m3/s a very low value for catchment area of about 1000 km2? AC 9. Thank you. It is the same as with the previous comment. It is a minimum streamflow.

RC 10. P4, L12-14: Why do the authors use these complex abbreviations for the GCM/RCM combinations? It is not clear what the meaning of all the numbers is. Try to be consistent with the descriptions in Table 1. AC 10. The Table will be corrected

RC 11. P5, L12: Do you have a reference for the Matlab version of HBV? AC 11. The MATLAB version of HBV used in this study was based on Lindstrom et al (1997). The original MATLAB code from Twente University NL, was further developed and adjusted for the purpose of climate impact studies in the Institute of Geophysics PAS.

RC 12. P5, L15-17: Only 12 out of 14 HBV parameters are mentioned. In which routines can we find CFLUX and PERC (see line 19)? AC 12. Thank you. It is changed to 'These routines are governed mainly by fourteen HBV parameters, of which, six (TT, TTI, CFMAX, DTTM, CFR, WHC), three (FC, LP, BETA, CFLUX), two (KF, ALPHA) and

C15

one (KS, PERC) parameters are representing each routine respectively.'

RC 13. P5, L17: 'routines' instead of 'routing stage'? AC 13. Thank you; corrected: 'routines' instead of 'routing stage'?

RC 14. P6, L24-P7, L3: This general description of the GEV distribution is not necessary here and can be found in many text books. AC 14. It will be deleted

RC 15. P7, L16-17: What do the authors mean with ": : : aggregated speared of flow quantile change : : :"? AC 15. By "aggregated speared of flow quantile change" we meant "integrated spread . . . ".

RC 16. P7, L19: 'squared' instead of 'squere'. AC 16. Thank you, it is corrected in the main manuscript.

RC 17. P7, L22: The title suggests that the results of this study will be described. Please rephrase the title. AC 17. "Description of the results" would be better?

RC 18. P7, L23: Different temporal resolutions? Shouldn't it be different lengths of data periods? AC 18. Agree: 'different temporal resolutions' will be changed to 'different lengths of data periods'

RC 19. P7, L18: The meaning of all variables should be explained in the text. AC 19. All the variables will be explained. Where: T_SSijk is total sum square error for the specific hydrological extreme indicator (e.g. relative change in Q30) for the ith parameter sets range, jth climate model and kth distribution parameter range and μ is the overall mean and ε_ijkdenotes the white Gaussian error.

RC 20. P8, L6: "Results and discussion"? AC 20 Thank you, it is corrected in the main manuscript. As "Discussion of the results"

RC 21. P9, L2: 'the 10-year moving average from the ensemble mean'? AC 21. Thank you, it is corrected in the main manuscript. Corrected as 'the 10-year moving average from the ensemble mean' changed to 'mean from the ensemble of seven climate

C16

models'

RC 22. P9, L15-16: Fig. 5a is mentioned twice. AC 22. Fig. 5 will be deleted

RC 23. P9, L29-30: Decreases in minimum flows and increases in maximum flows? Shouldn't it be the other way around (according to the caption of Fig. 6)? AC 23. Thank you, it is corrected in the main manuscript: 'Figure 6. Empirical flow quantiles of annual maximum flow (upper panels) and annual minimum flow (lower panels) under baseline and future climates (near and far future periods); the climate model spread is presented as a shaded area; green line denotes the mean value from all the GCM/RCM model realizations, red line denotes the averaged results obtained for the reference period.'

RC 24. P10, L6-7: Here, the annual minimum flows increase (see previous comment). AC 24. Thank you. It is corrected as previous comment

RC 25. P10, L9: What is Q30? Commonly, that is a discharge with a non-exceedance frequency of 30%. However, here it seems to be an annual maximum flow with a return period of 30 years? AC 25. Yes, it is annual maximum flow with a return period of 30 years. For annual maximum flow those two definitions have the same meaning. However, now to avoid confusions, we used as Qt30 in the main manuscript.

RC 26. P11, L7: 'Table 4' instead of 'Table 3'. AC 26. Thank you, it is corrected in the main manuscript. 'Table 4' instead of 'Table 3'.

RC 27. P11, L26-P12, L2: The first part of the conclusions can be omitted (can be part of introduction section). AC 27. Thank you, it is corrected in the main manuscript. Deleted

28. P12, L9: 'hydrological parameter uncertainties' instead of 'hydrological model uncertainties'? AC 28. Thank you, it is corrected in the main manuscript. Corrected to 'hydrological parameter uncertainties' instead of 'hydrological model uncertainties'

RC 29. P12, L24-27: This is a repetition of lines 11-14. AC 29. Thank you, it is corrected in the main manuscript. Deleted

RC 30. P13, L3: A paper in preparation should not be included in the reference list. AC 30. This paper has already been submitted: "Meresa, H., Romanowicz, R.J., Napiorkowski, J.J.: Trends of projections of hydrological extreme indices in the 21st century, submitted, 2017." RC 31. P13-17: The reference list and referencing contain many errors, typos and inconsistencies. This should be carefully and thoroughly double-checked. AC 31. The reference list will be corrected.

RC 32. P18, Fig. 1: What is the unit of the DEM map? AC 32. The unit is meter. Thank you, it is corrected in the main manuscript.

RC 33. P18, Fig. 2: The interquantile range of what? Of the seven GCM-RCM combinations? In that case it would be better to show the individual model results, i.e. one annual maximum for each combination so 7 points per year. AC 33. Figure 2. Annual maximum daily precipitation and mean air temperature time series for the BialaTarnowska catchment obtained from the seven GCM/RCM models under the RCP4.5 scenario; upper panel: projected raw annual maximum daily precipitation; lower panel: projected corrected annual mean daily temperature in the 1971-2100 period; boxplots show interquartile range of 75th and 25th; whiskers show 50th percentiles. We think that boxplots are better suited to present the precipitation and temperature variability than individual model results.

RC 34. P19, Fig. 3: In particular the scale of the upper panel looks strange. Flows in cubic mm? How accurate is your model? Please use the same (realistic) x-axis ranges. AC 34. We guess that the reviewer means Fig. 4. The y-axis units should be in cubic meters per second. The figure will be corrected.

RC 35. P19, Fig. 4: This figure (and also Fig. 2) is too small. What do we see here? AC 35. Fig. 4 presents projected annual maximum and minimum flow. The figures 2 and 4 will be enlarged.

RC 36. P20: The differences between historic and future periods cannot be clearly seen in these figures. AC 36. Following the first reviewer comments we decided to

remove Fig.5 together with the subsection 4.4.

RC 37. P21, Fig. 6: What are the different lines in these figures? And is baseline and reference period the same? AC 37. Thank you, it is corrected in the main manuscript. Changed to 'Fig. 6. Empirical flow quantiles of annual maximum flow (upper panels) and annual minimum flow (lower panels) for the reference period and future climates (near and far future); the climate model spread is presented as a shaded area; green line denotes the mean value from all the GCM/RCM model realizations in each period (near and far future period), red line denotes the averaged results obtained for the reference period. Each black line represents individual climate models'

RC 38. P21, Fig. 7: In the caption 'right hand panel' is mentioned twice. AC 38. The figure caption will be changed to: Fig. 7. Total uncertainty ranges of theoretical GEV-based annual maximum (left hand panels) and minimum (right hand panels) flow quantiles over 30 year periods; upper panels - the reference period 1971-2000; middle panels - near future 2021-2050; lower panels - far future 2071-2100. RC 39. P22, Fig. 8: Idem, annual minimum flow is mentioned twice. AC 39. Figure 8 captions changed to: Fig. 8. Total uncertainty ranges of flow quantiles over 130 years period; annual minimum flow as a function of return level (right hand panel) and annual maximum flow as a function of return level (left hand panel), based on a GEV distribution fitted to the projected annual flow (1971-2100).

RC 40. P23, Table 1: Which meteorological institute is connected to RACMO? AC 40. Netherlands.

RC 41. P23, Table 2: The caption is not clear. AC 41. Changed to 'Table 2. HBV parameter ranges: upper band (UB), lower band (LB) value' RC 42. P24, Table 4: What do the authors mean with 'change in width of . . .'? What compared to what? AC 42. Table 4. Change in the width of 0.95 confidence intervals for Q30 for annual maximum and minimum flows evaluated using different time periods of different lengths, first three columns corresponding to Fig. 7 (30-year period) and the fourth column corresponding

to Fig. 8 (130-year period).

References Benninga, J-H., Booij, M., Romanowicz, R.J., Rientjes, T.H.M., 2016, Performance of ensemble streamflow forecasts under varied hydrometeorological conditions, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-584, 2016. Blasone, R., Vrugt, J., Madsen, H., Rosbjerg, D., Robinson, B., & Zyvoloski, G. (2008). Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo Sampling. Advances in Water Resources(31), 630-648. Meresa, H., Romanowicz, R.J., Napiorkowski, J.J.: 2017, Trends of projections of hydrological extreme indices in the 21st century, Acta Geophysica, submitted, for publication Osuch, M., R.J. Romanowicz, W. Wong, 2017, Analysis of low flow indices under varying climatic conditions in Poland, Hydrology Research, submitted for publication