**Authors responses (in blue and italics) to referees comments (in black)**

*We thank the two reviewers for their constructive comments and suggestions and believe our proposed revisions in response to their comments, outlined below, will significantly improve the manuscript.*

**Reviewer #1**

**Major remarks**

The authors consider the performance of 22 GCMs from the CMIP3 exercise. They compare and rank the GCMs using various statistical measures for precipitation (P) and temperature (T), and they compare their results to ranking from literature based on different performance indices. In this way they identified the five best performing GCMs with regard to their plans that they want to use output from these best performing models to simulate runoff in a future study that shall be published in another paper as part 2 of the study.

*Noted. The revised paper is retitled so that it is independent of the complementary paper presented in the earlier submission.*

Even though I found it interesting to see how ranking based on P and T statistics only compares to more sophisticated multi-variable performance indices, I don't see what is really novel in the present manuscript.

*Our prime objective in this paper was to select a sub-set of GCMs that require the least bias correction as input to a hydrologic model for estimating runoff in a second paper. Although this objective still stands, the paper has been revised to be read as a contribution to GCM assessment but from a hydrologist's point of view. The hydrologist's requirements are very different to that of the climate modeller.*

*Hydrologists and water managers are specifically interested in the performance of GCMs in reproducing hydrologically relevant metrics. It is appropriate, therefore, that the metrics used to assess GCM performance be relevant to hydrologists and engineers (for example, related to bias and absolute RMSE) rather than those generally reported (for example, spatial correlations, normalised RMSE, Taylor plots, PDF and Bayesian approaches, skill scores, etc). Specifically, mean and variance of precipitation depths and mean temperature values are the key GCM variables for hydrologic simulation (which will lead to runoff). As noted from Table 2a in the revised paper, only one paper (Raisanen, 2007) provides numerical metrics (mean or bias and root mean square error of annual precipitation, variance of annual precipitation, mean or bias and root mean square error of mean annual temperature) of direct interest to hydrologists. Our revised paper provides a missing link in the large number of performance assessments of CMIP3 GCMs. This belief is based on 35 papers published from 1992 to 2014 dealing with GCM assessments. Only seven studies report metrics of bias of simulated precipitation and/or temperature of interest to the hydrologic community. We are not suggesting the approach in this paper is novel but rather pragmatic and offers a different approach to assessing GCMs. We note the reviewer "... found it interesting to see how ranking based on P and T statistics only compares to more sophisticated multi-variable performance indices ...". The revised paper incorporates a replacement Table 2 in which numerical and ranking procedures to assess CMIP3 GCM performance are listed and discussed in the manuscript.*

I also wonder why the authors do not look additionally at characteristics of annual runoff simulated by the GCMs, and compare these values to observed river discharge. At annual time scales or coarser, the lateral routing does not play a role, and such intercomparison may give insights on how the GCMs generally simulate the terrestrial part of the hydrological cycle (as P is also considered).

*There are two reasons why we did not look at GCM generated runoff in this study.*

*(1) Arora (2001) demonstrated that the quality of GCM generated runoff is dependent mainly on the quality of GCM precipitation and any bias in GCM precipitation is amplified in the resulting GCM runoff. In our paper we assess GCM bias in reproducing observed precipitation conditions and find substantial biases for all GCMs, thus we would expect significant bias in runoff generated by a GCM Land Surface Scheme.*

*(2) Sperna Weiland et al. (2012) assessed the performance of GCM runoff for two GCMs where the Land Surface Schemes had been tuned to reproduce realistic runoff estimates. They found that runoff estimates from an external hydrologic model generally outperformed the GCM runoff estimates.*

*Therefore in our companion study we used an external hydrologic model for generating runoff with bias corrected inputs. This is consistent with most hydrologic climate change impact assessments that use an offline hydrologic model rather than GCM runoff.*

*We have added a comment to the revised companion paper (originally part 2 of this paper) indicating the reasons why we are not using GCM generated runoff.*

While a future paper plan is sufficient to motivate why the authors look at temperature and precipitation measures, it does not really justify larger subsections on "estimating streamflow" and "potential evapotranspiration" as both subsections do not contribute to the main objective of the manuscript. They certainly would contribute to the follow up paper, but this is not part of the present paper. In this respect the complete Appendix is not necessary. The content of the appendix is interesting and probably appropriate for the second paper, but does not contribute to the understanding of the main results of this paper.

*Agreed. In the revised paper we have substantially reduced the "estimating streamflow" and "potential evapotranspiration" sections and embedded them into the Introductory section and refocused the context on a brief justification of the variables (precipitation and temperature) to be investigated. The original appendix provides justification for adopting temperature as a surrogate for potential evapotranspiration and a discussion of the errors entailed in that decision. This material is more appropriate in Part 1 than Part 2 and has been moved to Supplementary Material of this paper.*

By the way if it should be justified why only precipitation and temperature are necessary to simulate runoff based on literature, then this part clearly lacks to take into account recent research that addresses global multi-model studies with several GCMs and global hydrology models conducted within the EU project WATCH (see, e.g. Haddeland et al. 2011 Multi- Model Estimate of the Global Terrestrial Water Balance: Setup and First Results. J. Hydrometeor. 12, 10.1175/2011JHM1324.1, 869-884.) and the ISIMIP exercise (see http://www.isi-mip.org/). In this respect, the authors stated in the beginning of p. 4535.: "…precipitation and temperature, which are sufficient to estimate the mean and variability of annual runoff from a traditional monthly rainfall–runoff model (Chiew and McMahon, 2002) and from a top-down annual rainfall–runoff model (McMahon et al., 2011)." While I agree that they are sufficient to estimate current climate runoff characteristics, I doubt their validity under global warming conditions. Recent research has shown that temperature-only based estimates of

potential evapotranspiration tend to fail under global warming conditions (e.g. Hagemann et al. 2011 Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J. Hydrometeor. 12, 10.1175/2011JHM1336.1: 556-578.).

*We agree with the reviewer's comment that temperature-only PET models, like Thornthwaite, can be problematic under future projected conditions. In the former Appendix (now included as Supplementary Material to this paper) we provide a comprehensive discussion of recent research about the pros and cons of PET models in hydrologic modelling of present and future conditions, particularly in energy limited environments. In that discussion we note that conceptual hydrologic models calibrated for a specific catchment are flexible enough to extract the PET information they need from whichever PET data (temperature based or more complex) are used (see Chapman, 2003). The situation of using a calibrated conceptual model on a single catchment is different to the approach described in Haddeland et al. (2011, 2012) and Hagemann et al. (2011) in which the models are not calibrated for a specific catchment.*

*The references to Haddeland et al. (2011; 2012) and Hagemann et al. (2011) provide interesting case studies in the issue of using a simple, or complex, estimate of PET in hydrologic modelling. However, the resulting mean annual runoff and evapotranspiration estimates for present conditions from the models reported in Haddeland et al. and Hagemeann et al. exhibit large bias. For example, in Haddeland et al. (2012, Table 2) mean annual runoff is estimated from four hydrologic models driven by three climatological datasets (WFD, bias and non-bias corrected ECHAM Cntrl and IPSL Cntrl) for the period 1971-2000. The hydrologic models adopt a range of evaporation formulations: LPJmL (Priestley-Taylor); VIC (Penman-Monteith); MPI-MH (Thornthwaite); and Water GAP (Priestley-Taylor). The average mean annual runoff estimate for each model, driven by WFD and bias corrected ECHAM and IPSL, is 63,242 $km^3$ $year^{-1}$ (LPJmL), 45,343 $km^3$ $year^{-1}$ (MPI-MH), 54,617 $km^3$ $year^{-1}$ (VIC), and 55,214 $km^3$ $year^{-1}$ (Water GAP) respectively. These values are in stark contrast to the observation-based mean annual runoff value of 37,288 $km^3$ $year^{-1}$ from Dai and Trenberth (2002, page 670; Trenberth et al., 2007) for the period 1979-2000. The large difference in runoff between these models and the observed value of Dai and Trenberth, in the order of 50% for all models except MPI-HM, is concerning. Although Haddleland et al.(2011, page 873) imply this overestimation may be due to inadequate definition of the terrestrial land area, the methodology of Dai and Trenberth's (2002) and Trenberth et al.'s (2007) was based on runoff volumes rather than depth $\times$ area and therefore is largely independent of the global terrestrial land surface. For evapotranspiration, Hagemann et al. (2011, Figure 6) present bias estimates for 12 catchments over the period 1961-1990 and find the LPJmL (Priestley-Taylor PET model) driven with bias corrected inputs results in an average negative bias of approximately -18%, whereas, the MPI-HM (Thornthwaite PET model) has an average positive bias of approximately 1% for the same catchments.*

*In summary, if the references suggested by the reviewer had presented runoff and evaporation results that were consistent with observations, then the argument of the reviewer that "Recent research has shown that temperature-only based estimates of potential evapotranspiration tend to fail under global warming conditions (e.g. Hagemann et al. 2011)" would have provided a stronger justification for adopting a more complex evaporation model. Although we agree that temperature-only PET estimates can be problematic under future conditions, we do not agree that comparison of biased runoff and evapotranspiration estimates presented in these references can be used to support this argument. In fact the results presented reflect our concerns outlined in the former Appendix, that the input data for complex PET models need to be reliable, which supports our methodology of driving our hydrologic model with temperature as a surrogate for PET while being aware of likely errors in energy limited environments.*

*We note that in our companion paper that although the PERM hydrologic model uses T as an input, it does so through a transfer function (PERM parameter a) that allows the model to estimate actual evaporation. This transfer function is calibrated at each catchment and effectively represents the evaporative environment of each catchment. The evaporation model in PERM is not a temperature-only model like Thornthwaite, which estimates only the 12 monthly average values, but is a time-series broadly representing the radiation and aerodynamic evaporation processes experienced by the catchment during the calibration period. A necessary assumption of any hydrologic modelling under climate change is that the calibrated parameters apply into the future.*

*In view of our comments above and as we have revised the paper to be independent of the original Part 2, we see no need to change the revised manuscript.*

I suggest a) either to clearly focus on the ranking and the comparison to other skill scores, i.e. removing all parts of the present manuscript that do not contribute to this issue. But here question arise whether such a study would qualify to be a full HESS paper? A paper has to qualify by itself, and not because there shall be a second part that is more hydrology oriented. Thus, it might also be appropriate to b) merge this manuscript with the second paper part to have a complete study. But then the limitations addressing the use of temperature-only based PET estimates need to be thoroughly taken into account.

*We note the reviewer presents two options. We have given serious consideration to option (b), but have decided not to merge the two papers for the following two reasons. (1) We believe the original paper Part 1 does qualify as a full paper (see second comment above) following the proposed revisions outlined in response to both reviewers. (2) A merged paper would significantly dilute the significant material presented in Part 2 and be far too long. We have been invited to provide a revised paper of Part 2 to HESS which we will do.*

Please note that I also wrote comments to paper parts that may be removed if option a) is chosen. In summary, I suggest major revisions to be conducted before the paper may be accepted for publication.

*Noted.*

**Minor Comments**

In the following suggestions for editorial corrections are marked in Italic.

p. 4534 – line 1

… to characterize the …

*Sentence has been deleted as part of revision.*

p. 4543 – line 21-24

Even though the citation is from a published paper, I disagree with this. While I agree that it is very important that a GCM captures the current climate reasonably well, this does not mean that it will be closer to the future greenhouse response of the real world. Here, it is important that the GCM has the right climate sensitivity and that it adequately captures certain feedbacks that play an important role in this response. The latter, e.g., is currently being (and will be) investigated in research focusing on emergent constraints.

*We agree this was an inappropriate quote. We have replaced it with one from IPCC AR4 [Randall, R.A. and Wood, R.A. (Coordinating lead authors): Climate models and their evaluation. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change AR4, Chapter 8, 589-662, 2007].*

p. 4545 – line 8

… data set is presented …

*Agreed, corrected in the revised manuscript.*

p. 4547 – Sect. 5.1

The findings in this section are really interesting. They should be more highlighted and are worth to be explored further.

*We agree these results are interesting, but further exploration is beyond the scope of the paper. We have supported our observation with reference and comment to a recent paper by Masson and Knutti (2011).*

p. 4550 – Sect. 5.5

You shouldn't talk about results that are not considered in this study. E.g. it is mentioned that "…for a range of catchment scales world-wide."

On one hand one would ask, Which scales? On the other hand, it is written: "We have not reported the results of this catchment comparison here because many catchments in our data set are smaller than a GCM grid cell and, therefore, the comparison is not strictly appropriate."

Thus, it is sufficient to talk about the results of large catchments, as it doesn't really make sense (as also realized by the authors) to look at small catchments if GCM results are considered.

*Agreed, we have removed this section and Figure 7 from the revised manuscript.*

p. 4553 – line 23-25

This conclusion does not apply for global climate change studies as relevant climate changes are ongoing exactly in those areas that are strongly energy and water limited, especially a sit is projected that many dry areas will become even drier, i.e. more areas become stronger water limited.

*We note that our conclusion was for areas that 'are not strongly water or energy limited', which the reviewer has misread as 'strongly energy and water limited'.*

On the other hand the largest warming signal is observed in high latitudes that are strongly energy limited. Thus, I cannot follow that argument that complex PET formulations are not necessary. Especially if there are indications that the very simple ones based on temperature only fail under climate change conditions (see above).

*We make two points here. Firstly as noted above we reject the argument based on the references listed by the reviewer. These in fact strongly supported our argument rather than those of the reviewer. Secondly, in the former Appendix A (now incorporated as Supplementary Material), we reviewed the literature dealing with this issue. Based on the studies of Chapman (2003), Andréassian et al. (2004) and Oudin et al. (2005a, b) we concluded that for estimating PET for use in calibrated*

*hydrologic models a method based on temperature alone is satisfactory for the purposes of our study. In the Appendix we acknowledged that using temperature-only information to drive a calibrated hydrologic model could result in errors in runoff estimation in energy limited environments, as noted by the reviewer. However, we also argued that errors in runoff estimation would be small in water limited regions and regions that are not strongly water or energy limited as the key hydrologic driver is precipitation and soil moisture, not energy input. For example, in a hot dry region a change in PET (either increase or decrease) will have negligible impact on runoff as the limiting variable is precipitation.*

*We do not believe incorporating the additional climate variables required to estimate potential evaporation by the Penman-based formulation of wind speed (at 2 m height), net radiation (requiring both short and long wave, incoming and outgoing, to be estimated) and relative humidity estimated at 2 m is justified given the difficulty in obtaining GCM data that are credible. Johnson and Sharma (2009) have shown that in terms of their Variable Convergence Score (VCS, scaled between 0 and 100, where 100 is perfect convergence between GCMs) the predictions of the surface wind and specific humidity have VCS scores of approximately 40, net long wave radiation about 20 compared with surface temperature and net short wave radiation of about 70 and precipitation at 10.*

*Our revised manuscript and Supplementary Material reflect this argument.*

p. 4554 – line 17-21

But less confidence should not outweigh wrong physical behaviour in the projections. I.e., just because there is less confidence in some variables, their effect on water and energy cycles at the land surface shouldn't be neglected.

*We understand, but do not accept this argument. There are two issues here: (1) the physical representation of PET; and (2) the quality of the input data to the PET equation. For hydrologic climate change impact assessments the practitioner is faced with a trade-off between these two issues. Using a Penman-based formulation will include all the PET relevant physical processes, but the results may not be useful if the input data are not of adequate quality (see Kay and Davies, 2008; Kingston et al., 2009). Furthermore, it should be noted that if the Penman equation were used for example, a decision will be required as to which wind function to adopt which can affect the resultant evaporation estimate by up to ~19% (McMahon et al., 2013, Supplementary Material page 28), so there is some uncertainty even if perfect data were available. Whereas, a simpler method based only on reliable input data may not include all PET physical processes, but there will be more confidence in the processes modelled. We adopted the latter approach in our study.*

p. 4555 – line 7

It is written: "GCM projections of those process variables, other than temperature, may be unrealistic."

But if Temperature-only based parameterizations behave unrealistic in certain regions, too, then all climate change impact maybe unrealistic. That's actually some uncertainty one has to live with and pay regard to it.

*We agree with this comment. See previous response, no change in manuscript.*

p. 4555 – line 18

It is written:

"This error in PET trend is unlikely to be important for hydrologic modelling of water limited catchments, where changes in precipitation are the main driver of changes in runoff."

As mentioned above: While this may be true for the present climate it has been shown that errors in PET behaviour can become important under future climate conditions.

*In water limited environments now, and in the future, changes in PET will have negligible impact on runoff as precipitation (hence soil moisture) is the dominant driver of runoff. In this environment, errors in PET have little to no impact on estimates of runoff from a calibrated rainfall-runoff model. As discussed in the Appendix, the case for energy limited environments is different.*

*We see no need to amend the manuscript nor change the old appendix which is now supplementary material.*

## Reviewer #2

This manuscript and its companion paper attempts to quantify the impact of internal climate variability (as simulated by GCMs) on runoff risk assessments. This (the first paper) focus on selecting a subset of GCMs to focus on for further analysis in the companion paper.

*Noted.*

I have three comments on Part 1 that the authors may want to address.

1) Typically one would see uncertainty in climate change simulation to arise from three types of uncertainty: a) emission scenario uncertainty, b) model structural uncertainty (between GCMs) and c) internal climate variability (as represented by running an ensemble using one GCM changing only the initial conditions). In this paper the authors only discuss the latter two sources, though it could well be the case that emission scenario is the one with the largest uncertainty (particularly in the far future and for variables with strong direct relationship to global warming, i.e. temperature). It would be useful for the authors to provide some details on all sources of uncertainty and maybe comment on what sources are more dominant at what time horizons (see e.g. papers cited at http://climate.ncas.ac.uk/research/uncertainty/ as a starting point).

*We agree that scenario uncertainty is important, particularly in the far future, and will modify the text accordingly in the companion paper.*

*In this study we focused on two periods, 1965-1994 and 2015-2044, where within-GCM and between-GCM uncertainties dominate over scenario uncertainty. This view is supported by Wilby and Harris (2006) who assessed the uncertainty in future low flows for the River Thames and ordered the uncertainties from largest to smallest as follows: GCM > (empirical) downscaling method > structure of hydrologic model > hydrological model parameters > emission scenario. In a paper on understanding uncertainty in regional temperature prediction, Hawkins and Sutton (2009) estimated the contributions to the total prediction uncertainty from internal variability, model uncertainty and scenario uncertainty where scenario uncertainty is least important. Hawkins and Sutton carried out a similar analysis for regional precipitation again concluding that internal variability contributes most of the uncertainty and scenario uncertainty is generally small (Hawkins and Sutton, 2011).*

2) In Part 1, the focus lies on selecting a sub-set of GCMs from the CMIP3 archive. However, given that the methods that the researchers are using are statistical, would it not be possible to conduct experiments on all possible GCMs bar those that are outright bad (e.g. use metrics to remove poor models rather than selecting a few good ones?). The reason for keeping more models in would be that a) you could well imagine that some models perform better in some regions than others but overall isn't one of the high scorers b) others have shown that models that do well in current climate do not necessarily agree on a future pathway, thus would you not want to be inclusive rather than exclusive? At the very least it would be good to see where the models you selected fall within the spectra of CMIP3 models, do they simulate an overall wet future/dry future/ relative to the CMIP3 ensemble?

*We remind the reviewer that we do analyse, assess and report on the performance of precipitation and temperature for all GCMs and associated runs in the paper. In the revised paper it is not until Section 5.3 (which is the penultimate sub-section before the Conclusions) that we discuss which models we assess as better performing ones.*

*For this project our resources were limited so that the adoption of five GCMs to be used in our subsequent analyses and reported in the complementary paper was a trade-off between more analyses and interpretation as against less resources being devoted to runoff uncertainty and reservoir yield analysis. We believe we have achieved a reasonable compromise. This decision does not affect the approach taken nor our conclusions presented in this revised paper.*

*As requested by the reviewer, we have included a new Figure 7 in which we compare the ratio of mean annual precipitation from 2015-2034 to 1965-1994 against the mean annual temperature difference between 2015-2034 and 1965-1994 for the global land surface. Figure 7 shows the relative positions of the selected five GCM runs compared with the 23 CMIP3 GCMs including all 44 ensemble members. The five selected GCM runs are well distributed amongst the 44 GCM ensemble members, which indicate that they are reasonably representative of the range of future GCM projections. We observe that most GCM runs are clustered around the median response, except for the seven CCSM runs in the top right quadrant with a precipitation ratio > ~1.04.*

*In the revised manuscript we include a sub-section (Section 5.4) that addresses this issue and discusses Figure 7.*

3) The metrics is based on absolute data for rainfall, could this be an issue? All models have a bias but are we note more interested in the relative change? Since you bias correct before you use the data for the impact work, should bias really be equally important in selecting models as the other metrics used?

*No, we don't believe this is an issue. To assess GCM performance for hydrologic simulation requires absolute, rather than relative, values and our aim was to identify a sub-set of GCMs that required the least bias correction in precipitation and temperature that would be used in our later simulations to be presented in the companion paper.*

*No change to manuscript.*

Finally just on a structural comment, the content in Appendix A doesn't really sit comfortably within this paper. It would probably sit better in Part 2.

*Agreed. The Appendix has been moved to Supplementary Material (see earlier comment in response to Reviewer 1).*

## References

Andréassian, V., Perrin, C., and Michel, C.: Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models, J. Hydrol., 286, 19–35, 2004.

Arora, V.: Streamflow simulations for continental-scale river basins in a global atmospheric general circulation model. Advances in Water Resources, 24, 775-791, 2001.

Chapman, T. G.: Estimation of evaporation in rainfall–runoff models, MODSIM 2003 International Congress on Modelling and Simulation, vol. 1, 14–17 July 2003, Townsville, Australia, 148–153, 2003.

Dai, A. and Trenberth, K. E.: Estimates of freshwater discharge from continents: Latitudinal and seasonal variations, Journal of Hydrometeorology 3, 660-687, 2002.

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. and Yeh, P.: Multimodel estimate of the global terrestrial water balance: Setup and first results, J. Hydrometeor, 12, 869-884, 2011.

Haddeland, I., Heinke, J., Voß, F., Eisner, S., Chen, C., Hagemann, S. and Ludwig, F.: Effects of climate model radiation, humidity and wind estimates on hydrological simulations. Hydrol. Earth Syst. Sci., 16, 305-318, 2012.

Hagemann, S., Chen, C., Haerter, J.O., Heinke, J., Gerten, D. and Piani, C.: Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J. Hydrometeor. 12, 556-578, 2011.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions. Bulletin of the American Meteorological Society DOI:10.1175/2009BAMS2607.1, 1095-1107, 2009.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, Climate Dynamics, 37, 407-418, 2011.

Johnson, F. M. and Sharma, A.: Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments, Journal of Climate, 22, 4373-4382, 2009.

Kay, A.L., Davies, H.N.: Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. Journal of Hydrology 358: 221-239, 2008.

Kingston, D.G, Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W.: Uncertainty in the estimation of potential evapotranspiration under climate change. Geophysical Research Letters 36, L20403, 2009.

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R.: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, Hydrol. Earth Syst. Sci., 17, 1331-1363, 2013.

Masson, D. and Knutti, R.: Spatial-scale dependence of climate model performance in the CMIP3 ensemble, J. Clim. 24, 2680-2692, 2011.

Oudin, L., Michel, C., and Anctil, F.: Which potential evapotranspiration input for a lumped rainfall–runoff model?, Part 1 – Can rainfall–runoff models effectively handle detailed potential evapotranspiration inputs?, J. Hydrol., 303, 275–289, 2005a.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling, J. Hydrol., 303, 290–306, 2005b.

Raisanen, J.: How reliable are climate models? Tellus, 59A, 2-29, 2007.

Sperna Weiland, F. C., van Beek, L. P. H., Kwadijk, J. C. J. and Bierkens, M. F. P.: On the suitability of GCM runoff fields for river discharge modelling: A case study using model output from HadGEM2 and ECHAM5, Journal; of Hydrometeorology, 13, 140-154, 2012.

Trenberth K. E., Smith, L., Qian, T., Dai, A. and Fasullo, J.: Estimates of the global water budget and its annual cycle using observational and model data, Journal of Hydrometeorology, 8, 758-769, 2007.

Wilby, R. L. and Harris, I.: A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, Water resources Research 42, W02419, doi:10.1029/2005WR004065, 2006.