Interactive comment on “A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models” by S. N. Gosling et al.

S. N. Gosling et al.
s.n.gosling@reading.ac.uk

Received and published: 14 December 2010

This paper is extremely well written and contains some important research and useful discussion. The paper demonstrates future runoff projection comparison of global and various catchment modelling approaches over catchments with varying properties. Overall I think the paper is certainly publishable subject to some minor points detailed below:

Thank you for taking the time to review our research and for your positive comments and suggestions to improve the manuscript. We have addressed each comment in
turn. Our responses are provided below.

*****************************************************************************************

1. A missing step is perhaps to compare like with like – catchments using the same catchment model and groups of similar catchments. I imagine due to the size of this task (time/computer and person power) this was not undertaken – but some brief commentary about how a completely thorough test could be done would be useful. It would be helpful to give more information on the computational resource (nodes, speed etc.) required. I’m guessing a full uncertainty approach was impossible because of this problem. Could be further discussed.

We have added the following text to the Discussion section, and added some extra references to support the discussion, which are also shown below. We have also created a new figure, which highlights which uncertainties we have considered, and which we have not (the new figure is attached to this HESSD comment).

“A key conclusion is that climate model uncertainty dominates hydrological model uncertainty. However, it is acknowledged that this conclusion is based on the prior uncertainty assigned to both climate and hydrological models. Moreover, we have not sampled downscaling uncertainty, emissions uncertainty, and hydrological model parameter uncertainty (see Fig. 1). Therefore, we are likely underestimating the magnitude of climate and hydrological uncertainty in our analysis. Given the constraints of computational resources, we considered seven climate models and two hydrological models for each catchment. It can be argued that the application of seven climate models presents a reasonable representation of climate model structural uncertainty, given that previous climate change hydrological impact assessments have tended to apply a similar or lower number of climate models (Arnell et al. in review; Hayashi et al. 2010; Prudhomme et al. 2003). The prior uncertainty from climate model structural uncertainty could be reduced by comparing GCM simulations of baseline climate with observations. Such considerations have led to the calculation of performance metrics
for GCMs, such as ranking them according to a measure of relative error (Gleckler et al., 2008). Forming a single index of model performance, however, can be misleading in that it hides a more complex picture of the relative merits of different models. Furthermore, for one specific region, Chiew et al. (2009) concluded that there was no clear difference in rainfall projections between the ‘better’ and ‘poorer’ 23 GCMs included in the CMIP3 archive (7 of which we applied here) based on their abilities to reproduce observed historical rainfall. Therefore in their analysis, using only the better GCMs or weights to favour the better GCMs gave similar runoff impact assessment results as the use of all the 23 GCMs. Moreover, on a conceptual level, it has been argued that, because of deep and structural uncertainty, it is not appropriate to seek to estimate the relative weight of different GCMs, and to do so would lead to significant over-interpretation of model-based scenarios (Stainforth et al., 2007): all models are only partial representations of a complex world, and miss important processes. For these reasons, in the present analysis, we assumed that all the GCMs are equally credible, although they are not completely independent. The computational resources required to perform multiple GHM simulations are relatively small compared with those required to run multiple CHMs because in previous work ClimGen was integrated with the GHM and adapted to run by high throughput computing (HTC) on the University of Reading Campus Grid, which reduced simulation time by a factor of over 80 relative to running on a single compute node (see Gosling et al. 2010). A more thorough consideration of downscaling uncertainty would apply climate projections from regional climate models (RCMs), which have been dynamically downscaled, and/or a range of different statistical downscaling algorithms other than that included in ClimGen (e.g. see Maraun et al. 2010). However, this would effectively at least double the computing and time resources required from what was used in the present analysis. A more thorough consideration of hydrological model uncertainty would explore 1) hydrological model parameter perturbations, and 2) the application of several CHMs for each catchment. However, this would be demanding in terms of computational and human resources. For instance, to address the latter suggestion above, each CHM (SLURP, SWAT, etc.)
would need to be calibrated for each individual catchment (Liard, Mekong etc.) and would then involve performing 216 CHM simulations (6 CHMs x 6 catchments x 6 increases in global-mean air temperature) for a single GCM pattern. As such, a computer cluster with around 216 nodes would be ideal, but each CHM would need to be adapted for running by HTC. This is not straightforward; see Gosling et al. (2010) for a detailed discussion on the issues regarding adapting a hydrological model to run by HTC. To address the former suggestion, Multi-Method Global Sensitivity Analysis (MMGSA; Cloke et al., 2007) presents a method for systematically perturbing all model parameters systematically but again, the extensive computing resources required for this precluded such an analysis here. Moreover, each CHM and GHM will include different parameters, so a like-with-like comparison is not straightforward. Nevertheless, Arnell (this issue) demonstrates that the uncertainty associated with 100 CHM model parameter sets is vastly smaller than the uncertainty across 21 GCM climate projections, which supports our conclusion that climate model uncertainty dominates hydrological model uncertainty. Moreover, evidence from other climate change impact assessment sectors (e.g. agriculture; Challinor et al. 2009) suggests that climate model uncertainty is effectively damped once other non-climatic uncertainties, such as decision-making processes or socio-economic uncertainties are considered, in a wider decision-making framework.”


2. I cannot see any baseline comparison of the precipitation that is being fed into the hydrological models, especially on a catchment integrated basis. It is difficult to compare the outputs if you do not have a benchmark comparison of the inputs. Can you provide this as figures/table?

The other reviewer made a similar comment, and suggested that we replaced the maps (original Fig. 2 and Fig. 3) with catchment precipitation data. We are in agreement that this is more useful. Therefore we have replaced the maps with charts that show percentage change in average annual runoff for each catchment. They are attached with this HESSD comment.

Furthermore, we have edited the text in Section 3.1. to read: “Precipitation is the main driver of runoff (Chiew et al., 2009) so it is important to understand the magnitude by which it changes in each of the climate change scenarios we considered. Fig. 3 shows the percentage change from baseline in total-annual precipitation for UKMO HadCM3 prescribed warming of 1-6°C, for each catchment. The greatest changes in precipitation are observed for the Liard (around +33% with 6°C prescribed warming), Xiangxi (around +31% with 6°C prescribed warming) and Okavango (around -44% with 6°C prescribed warming). Harper’s Brook is associated with a small change in precipitation with 6°C prescribed warming (-7%). Analyses in Section 3.2. demonstrate how the simulated changes in precipitation from each prescribed increase in global-mean air
temperature are realised in changes in runoff. Fig. 4 shows the percentage change from baseline in total annual precipitation projected by seven GCMs for a prescribed increase in global-mean air temperature of 2°C, for each catchment. Whilst all GCMs simulate increases in precipitation with climate change for the Liard, there is not consensus in the sign of precipitation change across the seven GCMs for the remaining catchments. For instance, with the Mekong, four GCMs simulate increases in precipitation with climate change and three GCMs simulate decreases. It could be argued that this precludes a hydrological analysis using all seven GCMs. However, given the large dependence of runoff on precipitation (Chiew et al., 2009) and that complex non-linear interactions are common between climate forcing and runoff (Majone et al. 2010), it is important to demonstrate how the uncertainty in the projections of precipitation across GCMs translates into runoff projections. Moreover, the consequent uncertainty across runoff simulations could have important implications for water resources management. Analyses in Section 3.3. demonstrate how the simulated changes in precipitation from each GCM are realised in changes in runoff.”

3. The discharge results are not presented along with the known performance or bias with observed-baseline projections (as it would also be useful to do for precipitation). How well do the models perform compared to observed data? Obviously this does not guarantee good performance for future datasets and indeed calibrations may not be optimal in a future state. Some discussion on this would be useful as well as a description of bias/performance of models.

We have included in Table 1 now, a summary of Nash-Sutcliffe model efficiency coefficients that are calculated in each of the respective CHM papers in the Special Issue. The edited table is attached to this HESSD comment. Furthermore, we have edited the Methods section to include the following text: “All the CHMs had already been calibrated typically using local gauge networks. For each catchment, the CHM was recalibrated for use with gridded (0.5°x0.5°) climate data from the CRU TS 3.0 dataset.
(Mitchell and Jones, 2005) for the period 1961-90. This process is described in each of the individual papers in this issue, listed in Table 1. A summary of the Nash-Sutcliffe model efficiency coefficients (E) (Nash and SutcliiñAe, 1970) that were calculated in validation exercises presented by each paper is also presented in Table 1. According to the classification scheme of Henriksen et al. (2008), the CHMs generally performed “fair” to “excellent”, although for a very small number of gauging stations in the Okavango and Mekong, the performance was “poor” (see Hughes et al. (this issue) and Kingston et al. (this issue) for more details)."

The references list has been updated accordingly, with the following new references:


*****************************************************************************************

4. The time period of runoff projections is not clear throughout the text.

The time periods are arbitrary. The following text has been added to the Methods section to explain this: “ClimGen generates 30-year long monthly timeseries of forcing data for a given GCM and prescribed increase in global-mean temperature (e.g. UKMO HadCM3 2.0°C). This means that the 30-year long climate change scenarios for a given GCM are representative of a world that is warmer from baseline by a prescribed temperature, but they are not assigned a specific time period in years, which is arbitrary. Therefore the runoff simulations are also presented for arbitrary 30-year periods, representative of worlds where global-mean temperature is a prescribed amount warmer than baseline (1.0, 2.0, 3.0°C etc.).”

*****************************************************************************************
5. You have only used one type of GHM. Worth discussing. What are the uncertainties because of this? Do you think your conclusion that GHMs are equally feasible to apply as CHMs stands up for all cases? I think you undersell the GHM in your text. You have some good evidence that it can be used for this type of assessment. Cite it clearly and sell your conclusion that GHM is useful.

We have added the following text to the Discussion, to address this comment: “Our analysis demonstrates that the GHM is able to represent the broad climate change signal that is represented by the CHMs, for each catchment. Therefore where future climate change impacts assessments seek to quantify and assess the range of hydrological projections across an ensemble of GCMs, it may be as equally feasible to apply a GHM as it is to apply a CHM to explore catchment-scale changes in runoff with global warming. However, in the present analysis, we only considered only one GHM, Mac-PDM.09 (Gosling and Arnell, 2010). Recent work highlights that there is uncertainty across different GHMs in the simulation of runoff (Haddeland et al., in review), so it cannot be assumed that all GHMs will perform in the same way as the GHM presented here.”

Also, to highlight the value of the GHM, we have added the following text:

To the Abstract: “This implies that for studies that seek to quantify and assess the role of climate model uncertainty on catchment-scale runoff, it may be equally as feasible to apply a GHM (Mac-PDM.09 here) as it is to apply a CHM, especially when climate modelling uncertainty across the range of available GCMs is as large as it currently is.”

To the Conclusion: “Therefore, where future climate change impacts assessments seek to quantify and assess the range of hydrological projections across an ensemble of GCMs, it may be as equally feasible to apply a GHM (Mac-PDM.09 here) as it is to apply a CHM to explore catchment-scale changes in runoff with global warming.”

Furthermore, given that it appears we “undersell” the GHM, we have removed the second sentence from the text below, which appeared in the original version of the
manuscript: “The results imply that the GHM we applied here may be a useful and complimentary tool to the set of CHMs we applied for assessing catchment-scale changes in runoff where ensembles (instead of a single GCM) of GCMs are applied. However, this does by no means advocate the application of any GHM over any CHM for catchment-scale studies – the results merely suggest that the GHM we applied could be seen as complimentary to the CHMs we applied”

******************************************************************************

6. The sensitivity of the ClimGen/weather generator downscaling approach is not discussed. This is not the only approach to downscaling (e.g. Maraun et al 2010 Reviews of Geophysics doi:10.1029/2009RG000314) – and of course adds further uncertainty into your results. What might be the ‘true’ uncertainty bounds for your results? Can you discuss this? Surely they are bigger than those shown in figure 8. Can you be explicit about all of the different components that go to make up the total uncertainties in your results – what are they and how big are they relatively speaking?

We have created a new Figure, Fig. 1, which is attached with this HESSD comment. The figure summarises the main stages of a climate change hydrological impact assessment and the inherent uncertainties. The figure highlights which uncertainties we sample in our analysis. We have also added the following text to the manuscript, in a new sub-section (“1.3 Uncertainties in climate change hydrological impact assessment”), to support the new figure: “Climate change will affect the global terrestrial hydrological system (Kundzewicz et al., 2007) and there is evidence that it has already responded to the observed warming over recent decades (Bates et al., 2008). The most common method for assessing the magnitude of this impact is to run a hydrological model driven by various climate projections from general circulation models (GCMs, i.e. global-scale climate models) as input forcing data (e.g. Gosling et al., 2010). The simulations of key hydrological indicators, such as river runoff, can then be used to assess the potential impact of climate change and to inform policy- and decision-making. However, there are a number of uncertainties associated with making such projections.
Fig. 1 summarises the four main stages of performing a climate change hydrological impact assessment, which is broadly similar to other climate change impact sector assessments (Gosling et al. 2009). The first stage is to determine the greenhouse gas emissions scenarios with which the climate model (e.g. a GCM) will be driven with, in order to produce the climate change projections (the second stage). GCMs typically represent the atmosphere, ocean, land surface, cryosphere, and biogeochemical processes, and solve the equations governing their evolution on a geographical grid covering the globe. Some processes are represented explicitly within GCMs, large-scale circulations for instance, while others are represented by simplified parameterisations. The use of these parameterisations is sometimes due to processes taking place on scales smaller than the typical grid size of a GCM (a horizontal resolution of between 250 and 600 km) or sometimes to the current limited understanding of these processes. Different climate modelling institutions will use different plausible representations of the climate system, which is why climate projections for a single greenhouse gas emissions scenario will differ between modelling institutes. Two main methods can be used to sample this so called “climate model structural uncertainty”. The first is to use a range of climate projections from ensembles of plausible GCMs, to produce an ensemble of impact projections for comparison. Such multi-model datasets are often described as “ensembles of opportunity”, e.g. the World Climate Research Programme Third Coupled Model Intercomparison Project (WCRP CMIP3; Meehl et al. 2007). A second approach generates a “perturbed physics ensemble” (PPE) that introduces perturbations to the physical parameterisation schemes of a single climate model, leading to many plausible versions of the same underlying model. If sufficient computer power is available, then very large ensembles can be generated in this way. For example, Stainforth et al. (2005) ran an ensemble of 2,578 simulations that sampled combinations of low, intermediate, and high values of 6 parameters. As well as climate model structural uncertainty, climate models are sensitive to the initial conditions with which the models are initialised, which adds a further level of uncertainty. The third stage of a climate change hydrological impact assessment is to downscale the climate model
output to a finer resolution, suitable for application to a hydrological model. Two approaches are typically available, statistical downscaling and dynamical downscaling. The former uses statistical relationships to convert the large-scale projections from a GCM to fine scales. Different statistical methods can be applied for the downscaling, which introduces uncertainty. The latter approach uses a dynamic model similar to a GCM to cover a region. The dynamic model is then forced at its lateral boundaries using results from the coarse scale GCM. The dynamic method is typically more computationally expensive but does not rely on the central assumption of most statistical downscaling, that the downsampling relationship derived for the present day will also hold in the future. In the final stage, the downscaled climate data is applied to a hydrological model. Uncertainty at this stage can arise from the application of different hydrological models, e.g. CHMs and GHMs (similar in essence to the uncertainty that can be sampled from a GCM ensemble of opportunity), and from different parameters sets and perturbations within a given hydrological model, i.e. parameter uncertainty (similar in essence to the uncertainty that can be sampled from a GCM PPE). For six catchments, we compare the simulated runoff response of a GHM and CHM to projected future climate associated with (1) several prescribed increases in global-mean temperature from a single GCM to explore simulated response to different amounts of climate forcing, and (2) a prescribed increase in global-mean temperature of 2.0°C for seven GCMs to explore response to climate model structural uncertainty. The main sources of uncertainty sampled by this methodological framework are shaded in Fig. 1. Note that emissions uncertainty and downsampling uncertainty are not sampled, i.e. they are held constant, and nor do we consider GCM perturbed physics or hydrological model parameter uncertainty.

We have also added the following text to the Discussion, which references the Maraun et al. (2010) paper you cited: “A key conclusion is that climate model uncertainty dominates hydrological model uncertainty. However, it is acknowledged that this conclusion is based on the prior uncertainty assigned to both climate and hydrological models. Moreover, we have not sampled downsampling uncertainty, emissions uncertainty, and
hydrological model parameter uncertainty (see Fig. 1). Therefore, we are likely underestimating the magnitude of climate and hydrological uncertainty in our analysis. Given the constraints of computational resources, we considered seven climate models and two hydrological models for each catchment. It can be argued that the application of seven climate models presents a reasonable representation of climate model structural uncertainty, given that previous climate change hydrological impact assessments have tended to apply a similar or lower number of climate models (Arnell et al. in review; Hayashi et al. 2010; Prudhomme et al. 2003). The prior uncertainty from climate model structural uncertainty could be reduced by comparing GCM simulations of baseline climate with observations. Such considerations have led to the calculation of performance metrics for GCMs, such as ranking them according to a measure of relative error (Gleckler et al., 2008). Forming a single index of model performance, however, can be misleading in that it hides a more complex picture of the relative merits of different models. Furthermore, for one specific region, Chiew et al. (2009) concluded that there was no clear difference in rainfall projections between the ‘better’ and ‘poorer’ 23 GCMs included in the CMIP3 archive (7 of which we applied here) based on their abilities to reproduce observed historical rainfall. Therefore in their analysis, using only the better GCMs or weights to favour the better GCMs gave similar runoff impact assessment results as the use of all the 23 GCMs. Moreover, on a conceptual level, it has been argued that, because of deep and structural uncertainty, it is not appropriate to seek to estimate the relative weight of different GCMs, and to do so would lead to significant over-interpretation of model-based scenarios (Stainforth et al., 2007): all models are only partial representations of a complex world, and miss important processes. For these reasons, in the present analysis, we assumed that all the GCMs are equally credible, although they are not completely independent. The computational resources required to perform multiple GHM simulations are relatively small compared with those required to run multiple CHMs because in previous work ClimGen was integrated with the GHM and adapted to run by high throughput computing (HTC) on the University of Reading Campus Grid, which reduced simulation
time by a factor of over 80 relative to running on a single compute node (see Gosling et al. 2010). A more thorough consideration of downscaling uncertainty would apply climate projections from regional climate models (RCMs), which have been dynamically downscaled, and/or a range of different statistical downscaling algorithms other than that included in ClimGen (e.g. see Maraun et al. 2010). However, this would effectively at least double the computing and time resources required from what was used in the present analysis. A more thorough consideration of hydrological model uncertainty would explore 1) hydrological model parameter perturbations, and 2) the application of several CHMs for each catchment. However, this would be demanding in terms of computational and human resources. For instance, to address the latter suggestion above, each CHM (SLURP, SWAT, etc.) would need to be calibrated for each individual catchment (Liard, Mekong etc.) and would then involve performing 216 CHM simulations (6 CHMs x 6 catchments x 6 increases in global-mean air temperature) for a single GCM pattern. As such, a computer cluster with around 216 nodes would be ideal, but each CHM would need to be adapted for running by HTC. This is not straightforward; see Gosling et al. (2010) for a detailed discussion on the issues regarding adapting a hydrological model to run by HTC. To address the former suggestion, Multi-Method Global Sensitivity Analysis (MMGSA; Cloke et al., 2007) presents a method for systematically perturbing all model parameters systematically but again, the extensive computing resources required for this precluded such an analysis here. Moreover, each CHM and GHM will include different parameters, so a like-with-like comparison is not straightforward. Nevertheless, Arnell (this issue) demonstrates that the uncertainty associated with 100 CHM model parameter sets is vastly smaller than the uncertainty across 21 GCM climate projections, which supports our conclusion that climate model uncertainty dominates hydrological model uncertainty. Moreover, evidence from other climate change impact assessment sectors (e.g. agriculture; Challinor et al. 2009) suggests that climate model uncertainty is effectively damped once other non-climatic uncertainties, such as decision-making processes or socio-economic uncertainties are considered, in a wider decision-making framework.”
The results of these edits, is that we are more explicit about all of the different components that go to make up the total uncertainties in our results. However, we do not attempt to quantify how large the “total uncertainty range” could be. Without extensive, detailed and systematic analysis, this would be highly subjective.

7. For the discharge, I would have thought that also looking at the whole discharge distribution would be interesting at the catchment scale. Did you consider this at all?

We did consider this. However, the paper explores hydrological impacts for 6 different degrees of global-man warming, across 7 different GCM climate change projections, and explores mean annual runoff, mean monthly runoff, and Q5 and Q95 runoff. The inclusion of a distribution analysis would risk making the paper too lengthy. Moreover, we believe that a great deal of information can be gained by exploring the Q5/Q95 and mean runoff, which are inherently indicators of the distribution, generally. Also, we deliberately avoided exploring only mean runoff in this paper. For instance, the complex nature of the response of river discharge to climate change highlights under-recognised limitations in the common use of mean river discharge as a measure of (1) the response of hydrological systems to climate change and (2) freshwater availability (e.g. water stress index, relative water demand) (Taylor, 2009). As our paper shows, mean river discharge can mask considerably greater intra-annual (seasonal) variations which are of fundamental importance to water management and our understanding of freshwater availability. For example, reductions in low (Q95) flows can lead to acute water shortages as well as affect environmental ṭhĆow requirements and dry-season water allocations; changes in high (Q05) flows can impact flood risk and basin storage requirements.

8. End of section 3.1: If there was no catchment for which all 7 GCMs agreed on ppt change then what is the point of carrying out the hydrological analysis. Think that you need to make your story more concrete here and discuss this point. Are the signs of the runoff response shown in the results directly attributable to the catchment precipitation used as input? If there was a more thorough front end analysis of the precipitation (as suggested above) then this would probably be more obvious.

This is a good point. However, it could also be argued that perhaps it is even more important to carry out the hydrological analysis because of differences in the sign of projected precipitation change. This is because the climate-runoff relationship is often non-linear and moreover, it is important to demonstrate the potential uncertainty there is in runoff simulations that use climate projections from different climate models. This has important consequences for the decision-making process in water resources management, for instance. Therefore we have edited the text in Section 3.1. to reflect your point above, as well as the points we have just discussed. The text now reads: “Whilst all GCMs simulate increases in precipitation with climate change for the Liard, there is not consensus in the sign of precipitation change across the seven GCMs for the remaining catchments. For instance, with the Mekong, four GCMs simulate increases in precipitation with climate change and three GCMs simulate decreases. It could be argued that this precludes a hydrological analysis using all seven GCMs. However, given the large dependence of runoff on precipitation (Chiew et al., 2009) and that complex non-linear interactions are common between climate forcing and runoff (Majone et al. 2010), it is important to demonstrate how the uncertainty in the projections of precipitation across GCMs translates into runoff projections. Moreover, the consequent uncertainty across runoff simulations could have important implications for water resources management. Analyses in Section 3.3. demonstrate how the simulated changes in precipitation from each GCM are realised in changes in runoff.”

*****************************************************************************************

9. Page 7205 line 10 – extra ‘that’
10. How does each component of your results compare with previous climate impact runoff studies from around the globe?

Unfortunately, such a comparison is precluded by the application of different climate change scenarios, different climate models, different CHMs and GHMs, and different catchment locations and sizes, which are applied in other studies. This means the results presented here are not directly comparable with previous climate change hydrological impacts assessments. Indeed, neither are many other hydrological impacts assessments, for this very reason. The purpose of this analysis was to compare a GHM with several CHMs for six specific catchments, rather than to compare the results with other studies. Readers interested in multiple model-comparisons are referred to the EU Water and Global Change Project (WATCH) for more information: http://www.eu-watch.org.

However, Thorne (this issue) does make some inter-study comparisons based upon his CHM simulations, because he used a number of SRES emissions scenarios, which are comparable with some other studies. We used prescribed warming scenarios, which are less comparable, but were necessary in our case to aid inter-model comparisons. Also, Kingston et al. (this issue) notes that their results are comparable to those of previous studies of the Mekong.


Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/7/C4103/2010/hessd-7-C4103-2010-supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 7191, 2010.
Fig. 1. Figure 1. The four stages of a climate change hydrological impact assessment and the inherent uncertainties. The shaded areas denote the uncertainties we considered in this analysis.
Fig. 2. Figure 3. Change in total-annual precipitation relative to baseline (vertical axis; %) for UKMO HadCM3 prescribed warming of 1-6°C (horizontal axis), for each catchment.
Fig. 3. Figure 4. Change in total-annual precipitation relative to baseline (vertical axis; %) for the 7 GCMs under 2°C prescribed warming (horizontal axis), for each catchment.