1. General comments

This study presents a multi-model ensemble simulation approach (3 GCMs and 8 GHMs) to project the future climate change impact (time period 2071 - 2100) of the global water resource on the base of 2 emission scenarios (A2, B1).

Before using the GCM output in the global hydrological models (GHMs) a bias correction method for precipitation and temperature is applied.

The results show basically a large spread in projected changes often with an increase of the available water resources as well as partly even a severe decrease in specific regions.

The study states a large uncertainty associated with the choice of the GCM and underlines that the influence of the different GHMs cannot be neglected in many regions.

The authors suppose the reasons for the uncertainties in the GHMs in significant differences in – I would call it - the “model philosophy” (physical based or more conceptual) and their formula parameterization as well. Explicit mentioned are the effects due to models with or without a detailed land surface energy balance and the impact of different formulations of evapotranspiration in the GHMs.

Unfortunately only as an outlook and not as the core of the paper the authors state the requirement for tools and methods which allow quantification of uncertainties and an assessment of robustness of climate change impacts without generating to much data in combination with the absolute necessity to demonstrate the results in a fairly simple and equally reproducible way.

The paper is well written and addresses with the subject of available global water resources in the future definite scientific relevant questions. However forcing hydrological models with climate model output is in general not a new approach. The difficulties and challenges in coupling climate and hydrological models should be considerable and more critical discussed (e.g. bias correction, consistency, conservation aspects, feedback mechanisms in land–atmosphere coupling).

Novel but not unique are the 8 GHM and the purpose of determination the variability of the different model accesses. That could be better highlighted in the paper.

The specified and described variability is quite hard to evaluate. For me sometimes important aspects and information in the explanation are missing, skipped or too short cited.

The presented work is part of the quite huge European Union project WATCH. Many aspects of the wide and varied project results are already covered in numerous papers. It’s understandable, that not all aspects could always be covered in all details in all papers. Please try to provide the necessary information to any obvious questions adequate within the paper.

More focus on the uncertainty issue and the social and scientific need of evaluating them is from my point of view the more exciting and challenging aspect in this work.

More comprehensive presentation of the major results and a critical evaluation of the applied methodology are preferable. Many of the figures are too small sized and missing partly easy accessible information to comprehend the illustrated results in a suitable time.

2. Major Remarks

2.1. Impact evapotranspiration

One main statement seems to be: uncertainties of the future change is not only caused in the climate models (GCM) but also significant in the hydrological models (GHM).

In this context the variability of the calculated evapotranspiration (ET) is pointed out but only vague attributed to the different model formulation in the 8 GHM.

A more detailed look to the different input-parameter for the evapotranspiration calculations (partly output of the climate model) could give useful hints. First a table with the 8 hydrological models and their model concept and the input forcing data would be very helpful.

Temperature and precipitation are bias corrected (“fitted to the observation”), but what’s about humidity or radiation – e.g. not necessarily independent of temperature? Can we guarantee that within the bias correction consistency and conservation aspects are sufficiently regarded?

How are humidity and/or radiation involved in the ET calculations/formula? May the observed strong uncertainty in the ET calculation of the GHM are caused by systematic errors in climate model output which are not considered in the bias correction?

Influence in ET-calculation by also using bias corrected humidity, radiation and wind GCM output?

Haddeland et al. (2012) seems to cover these aspects, but is only cited in P1333 L3 with a statement like “bias corrections of other GCM variables have no influence on the projected relative changes”.

As I understand Haddeland et al. (2012) they pointed out that in addition to precipitation and temperature also radiation, humidity and wind play a significant role when simulating the terrestrial water balance especially in energy-limited areas.

These variables are all dynamically coupled by various feedback processes but with the assumption that the links and feedbacks between the meteorological states and fluxes (temperature, humidity, precipitations, evapotranspiration, etc.) are not of key importance?

I think the uncertainties statements in the paper are in the main interest of the audience. The paper would win with a more detailed and critical discussion to where the uncertainties come from.

2.2. Handling and document bias correction

The observed bias impairs the direct application of the GCM data for hydrological modeling. With a bias correction method you improve the model outputs towards the observations in a post processing step.

However bias correction has a large influence on the GCM/GHM output in absolute terms and likely also on climate change signals (i.e. the relative change between a control and future period).

The bases of BC are some crucial assumptions: i.e. stationarity of the method under non-stationary conditions, lack in the physical basis; not satisfying conservation laws etc.

I recommend a more transparent analysis and discussion to the obvious BC impact.

I absolutely appreciate the section “Comparison to direct GCM output”

P1329 L 25ff (referring Fig. 4) I would expect more discussion about the fact that the mean GCM ensemble with BC (blue bar) and original – “no BC” (yellow bar) are mostly significant different (“even weaker than B1” is not the point – change “no BC” partly 2-3 times less than with BC, from my point of view that’s the remarkable conclusion in Fig. 4).

Tracking the argumentation on P1330 with the references to Fig. 5 and Fig. 6 is quite challenging. One point seems to be the spread differences for the simulations with original GCM Output and the bias corrected output. If I take a look to Fig. 4 and compare the green (with BC) and the yellow spread (original – no BC) I come to the same findings?

Despite any original performance applying BC on GCM output (by definition) increases the agreement with observations and narrows the uncertainty range of any further examination.

2.3. Natural climate variability

It would be quite interesting to add the natural climate variability to your analysis. In ECHAM5 for example you have different realisations, they could represent the natural variability of climate. To consider the natural variability in the ensemble could be quite helpful to evaluate the uncertainties in the model chain and the projected changes.

2.4. Available water resource - a seasonal perspective stringent?

At least at the point of rating the global available water resources a presentation of the annual change in the unit [mm/day] is for any reader a quite hard to evaluate target (first calculate an annual value, than compare the change with the absolute values …). Maybe relative changes in percentage would be the better, more self-explanatory unit (see figures).

Additional in many regions the annual water resource values are less important. The water need is highly seasonal driven and anyhow connected to natural inter-annual variability.

In addition the inter-annual change could be quite significant despite the annual changes are only marginal.

I see these circumstances crucial for the significance of results and would therefore already expect information in the main part of the paper and not only in the outlook.

2.5. Essential information to model calibration and validation

Due to the main question of the model uncertainties in the paper i missed a section with essential information to GCM and GHM calibration and validation. This could be a table or figure with GCM and GHM simulation results compared with the observation period. For precipitation and temperature the results before bias correction (with successful BC, simulation and observation should be statistical comparable). In this association the comparison of the GCM values radiation, humidity and wind simulated and observed data for the period 1971 – 2000 would be enlightening (see ET remarks above).

Do the GHM simulations with GCM input match the observed hydrological data?

ET is somehow part of GCM and GHM: congruent, differences?

3. Minor Remarks

P1324 L17 bias correction factors are derived from simulated data? – please explain

P1324 L17-25 the cited “successfully applied bias correction” should not wipe the critical discussion and statements in Ehret et al (2012) – no doubt we have to deal with them but not forget a detailed assessment of the impact of bias correction....

P1325 L7 (section 2.2.) a figure with a schematic ensemble setup would be very helpful and could also easily introduce the many acronyms

P1325 L12: do all readers know what emission scenarios A2 and B1 means – short classification like high/low green gas emission may helpful?

P1325 L15 of each hydrological variable – witch?

P1326 L10 similar magnitude – tell the magnitude / P1340 fig 1a demonstrate the change in the unit [mm/day] that's quite hard to judge- percentage change e.g. could be quite better realized

P1326 L17 missing for what emission scenario these statements are made (e.g. legend fig 2 tells A2 scenario?). Make it always clear in text and figures.

P1327 L12: probably Fig 2g not "and h"?

P1327 L13: probably Fig 2h not "2e and f"?

P1327 L25: where is Fig. S1

P1327 L24+25 & P1328 L1+2: there is missing information to fulfill this argumentation

P1328 L16ff / Fig.1342 "projected change is considered robust if the change is larger than the maximum spread. Fig. 4 shows …". In Fig.4 it initially seems that all spreads are larger than any mean change? Where is the misunderstanding?

Improve and better explain Fig. 4 (6 line legend is hard to follow). Suggestion: short any text passages with detailed description for each catchment (give that information is the job of the figures)

P1329 L14-16: maybe note "without bias correction for P and T"

P1332 L20ff conclusion - missing a (critical) statement to the use of bias correction

P1333 L1 what means bias correction on an individual basis?

P1322 L10 "robust signal" = models agree on the sign of change – I'm wondering about that weak evaluation – an ensemble approach should allow some more statistical testing

P1340 Fig 2 should be improved toward easily readability – at least please larger

P1344 Fig 6 should be improved toward easily readability – 8 line legend is way too much – hard to get the message of the figure