Response to Referee #2 comments

by J. Jarsjö et al., 2011, regarding HESSD paper “Hydrological responses to climate change conditioned by historic alterations of land-use and water-use”

- The study seems to rely very much on earlier efforts of modelling the Aral sea basin. I support the desire to be concise and avoid repetition of work that is already published. However, enough information should be provided to allow interpretation of the results and scrutinization of the authors’ claims, which in my opinion is currently not the case. For instance, one of the conclusions of the paper is that their offline modelling is an improvement over the online implementation of land-surface processes in GCMs, but this is not clear. At least some information should be given on how land-use and landcover information is used in the model, how irrigation is dealt with (e.g., efficiency, separation of recharge and evapotranspiration fluxes), what, if any, observation data were used to calibrate the model. Also, more details on the model performance should be given, if only to be able to compare the model’s uncertainty with the uncertainty of the GCM projections.

→ We have now considerably extended the information on the modelling steps and theoretical background (sections 3.1 and 3.2), addressing the above comments in the following way:

(i) Regarding the adopted offline modelling, a main advantage is that hydrological processes can be handled with the same high spatial resolution for all GCMs in a reasonably time-efficient way, regardless of the varying coarseness of available GCM grids, see answer to specific comment 7609/11-14 (part II)

(ii) Regarding the treatment of irrigation in the model, and more generally how various flows (e.g. evapotranspiration, surface water, groundwater) are modelled, we now explain that the model independently quantifies the three principal outflow components of lake drainage basins, namely the discharges of the principal rivers into the lake, the diffuse flows along the shoreline of the lake (from groundwater and small, transient streams) and evapotranspiration over the land and water surfaces of the lake drainage basin. We also explain in more detail how the irrigation water is redistributed over the irrigated fields as extra precipitation. For instance, in response to several of the detailed questions below, we enhance sections 3.1 and 3.2 by clarifying that <u>water application</u> constitutes an input to the model (taken from independent data), whereas the (consumptive) <u>water use</u> is an output of the model, determined by (the modelled) amount of irrigation water that remains in the basin after losses to the atmosphere through evapotranspiration.

(iii) The model was found to match river flow (Q) observations at pre-irrigation conditions (before the large-scale irrigation development during the1950’s) and post-irrigatrion conditions (after the 1950’s), without need of calibration; the modelled ET differed by only 2% from the ET derived from water balance closure. We explain that the model match with the pre-irrigation conditions shows that it yields consistent results under relatively undisturbed conditions. Furthermore, the correspondingly good model match with the post-irrigation conditions shows that it also reproduces the additional ET losses caused by the inefficient/ excessive application and spreading of irrigation water on the agricultural fields. Quantitative expressions for model performance are now given in the methods section, and used in subsequent sections.

(iv) Land-use information is used to delineate the spatial extent and location of the areas where the land-surface is irrigated and therefore kept at a wetter state than given by the precipitation data from CRU. Furthermore, in the revised version of the manuscript, we use land cover information to investigate the possible influence of open water evaporation from the river on the presented results (see also our response to reviewer #1, comment 7602/8).

- The paper concludes that it is better to use the model ensemble mean rather than results based on one single GCM model, based on the fact that the ensemble mean tends to be closer to the CRU reanalysis data than individual models. Although this is a sensible approach, it should be noted that:
* the CRU reanalysis data may not be an independent source of information, especially for the AR4 models. Given that the CRU data were available during the implementation of the models for AR4, climate modellers may have aimed at improving agreements between the reanalysis data and the simulations. This may also be a reason behind convergence of the AR4 ensemble compared to TAR.
→ There are uncertainties in the openly available CRU dataset. However, the discrepancy between the CRU observation data and the historical output of individual GCMs is too large (up to a factor two for precipitation data) to be explained by uncertainties in reported CRU observations alone. Two independent facts converge on supporting this conclusion. 1. CRU data is unlikely to be systematically biased to such an extent, when averaged over such large areas as the Aral Sea Drainage Basin (1.3% of the earth's land surface). 2. Previous site-specific results of Shibuo et al. 2007 [*Hydrological responses to climate change and irrigation in the Aral Sea Drainage Basin, Geophysical Research Letters, vol. 44, L21406*] show that the precipitation input from CRU is consistent with water balance closure (in which ET is modelled independently, and resulting river runoff has been shown be consistent with observations). This is now explained in the revised manuscript.

* It may be erroneous to extrapolate agreement between historic observations and simulations towards predictive capacity in the future, especially when anomalies are used (see e.g., D. A. Stainforth, T. E. Downing, R. Washington, A. Lopez, and M. New. Issues in the interpretation of climate model ensembles to inform decisions. Philosophical Transactions of the Royal Society A, 365:2163–2177, 2007).
→ The model results are compared with long-term (multi-decadal) changes, not anomalies. We explain explicitly in the manuscript (7597/28-7598/12) that it is such slow changes that are relevant for the present modelling (rather than short-term changes, or anomalies). This is because initial system conditions, such as current groundwater levels and lake areas, can influence system characteristics during a short-term perturbation (anomaly). However, if the forcing of the perturbed state instead persists for long times, system conditions will have time to respond to the changed forcing, which can fundamentally change system characteristics. Calibration of models to system characteristics observed during short-term events may therefore introduce errors in predictions of long-term, boundary-condition driven future changes – we have reworded the text to make clear that the used method is not associated with such problems; see specific response to comment 7597/28 - 7598/12 for details.

* An advantage of implementing a model ensemble is not only the possibility of calculating a mean (or median) but also the spread of the ensemble as a measure of uncertainty. This is currently only superficially discussed in the paper.
→ We now explicitly show all the model results outside of the standard deviation as points in Figure 2 (see also comment 7606/1). Direct comparisons of the full T, P, ET and R ranges have also been added to the text. In addition, implications of these relatively large ranges are mentioned and discussed in several paragraphs (7604/17-20; 7605/1-6; 7606/1-16; 7610/4-6;16-25), which taken together thoroughly addresses the ensemble spread and its implications. Please note that the error bars did not cover the full standard deviation in the previous figure,

due to a now corrected mistyping in the graph-program; the now corrected illustration therefore more clearly supports the description found in the main text

- Although the paper claims to deal with interactions between climate change and water use, in fact both processes are dealt with separately. All water use is held constant for the future, and the impact of climate change without water use is compared to that without water use. This is a useful first approximation because of the non-linearities in the relation between water extraction and evapotranspiration, but it is a very coarse approximation. Water use is very likely to change in the future, either because of a higher demand, or because of limitations in supply. The fact that some scenarios result in a negative runoff (Fig. 3), which I interpret as groundwater depletion, highlights that water shortages will be important in the near future. This should be discussed in more detail.

→ The statements that interactions between climate change and water use are not dealt with, and that all water use is held constant in the future, are based on misunderstandings, since the considered future "irrigation" and "no-irrigation" scenarios (see p.7603, lines 17-20) imply very large differences in future water use. We consider the cases of zero water application (the no-irrigation scenario, p.7603, line 20) and a water application of 50 km^3/year (the irrigation scenario, p.7603, line 19). The difference in water application between the considered cases is hence 50 km^3/year. Furthermore, even for the subset of model cases that have the same water application, the modelled consumptive water use (equaling the amount of applied irrigation water that is lost through ET, see the first response paragraph), will differ for the varying future climatic conditions given by the 20 considered GCMs.

Regarding water use and future water shortages, some countries (e.g., Turkmensitan) plan to increase the water diversions and extend the canal-system. If the plans were to be realized, this could dramatically increase water shortage in the downstream-most part of the ASDB. There are also plans of upstream countries like Tajikistan to increase the use of hydropower. Water would then be increasingly be discharged during cold periods, which would imply more upstream water storage and less downstream water availability during the growing season. In the downstream part of the basin, water shortages and salinization prevent any further regional expansion of irrigation. On the other hand, the prevailing irrigation system of ASDB counts as one of the world´s most inefficient, which means that considerable amounts of water could be saved by changing irrigation techniques, even if the water demanding cotton and rice production is maintained. The water application could be reduced even more if alternative, less water demanding crops were introduced. We now extended the discussion on these issues in the end of the manuscript.

specific comments

7596/18: define R here (which is where it is used for the first time). However, it is preferable to use the full word throughout the paper for readability.

→ OK - R defined.

7598/1: what are "long-term change performances"?

→ We here refer to how well the model reproduces the long-term (hydrological) changes: this is now clarified in the manuscript.

7597/28 - 7598/12: What is the relevance of this discussion about short-term versus long-term modelling?

→ The discussion justifies the relevance of the adopted method. In order to further emphasize the point, we add that "…we here condition the hydro-climatic modelling to observational data on responses to long-term changes in forcing. This is because the system then will have had time to respond to the changed forcing, hence revealing system characteristics that are independent of the initial conditions, which is relevant for slow, boundary-condition driven changes at focus in this study".

7598/24: different change pressures -> different types of change
→ OK, changed the text accordingly

7598/13 - 29: this paragraph is vague and little to the point. What you want to argue is basically, that if one is to determine the impact of changes in land cover and water use on river flow, it is important to close the water balance?
→ The paragraph does not focus on "impact of changes in land cover and water use on river flow". The discussed problem importantly includes the boundary condition of a changing climate. We think this is clearly stated by the first sentence in the paragraph: "Regionally, the impacts on water resources from changes in climate overlap with the impacts from land-use and water-use changes". The rest of the paragraph logically focuses on this issue.

7598/27: runoff: Are you able to separate river flow from groundwater flows. Does the model deal with this?
→ Yes. As explained in the general comments section, we now clarify how the model independently quantifies the three principal outflow components of lake drainage basins, namely the river discharges at the outlets of the principal rivers, the diffuse flows occurring at all other locations along the shoreline of the lake (from groundwater and small, transient streams) and evapotranspiration over the land and water surfaces of the entire lake drainage basin.

7599/15: Do you really resolve interaction? Interaction would mean either impact of land-use change on climate, or impact on climate change on water extraction and/or land-use. It seem that you treat both separately (i.e. keeping one constant while varying the other) and compare the relative changes.
→ We do resolve interaction, since we do not – as the reviewer suggests – keep one factor constant while varying the other. As mentioned earlier, results on the magnitude of interaction effects are obtained by model quantification of:

(1) ET for the historic period with zero water application
(2) ET for the historic period with a water application of 50 km3/year
(3) ET for the future period with zero water application
(4) ET for the future period with a water application of 50 km3/year

Both water application and climate differ between the considered cases (1) and (4). Interaction effects between climate and water application are therefore resolved. For example: Without interaction effects, the ET change in response to a changed climate would be the same with irrigation [given by (4) minus (2)] as without [given by (3) minus (1)]. However, as figure 5 shows, there is a clear difference in climate-driven ET change for the different water application practices. Specifically, our scope, as reflected in the title of the paper, and stated in more detail on p.7599, lines 14-16, is to resolve effects of interactions between climate change and water re-distributions on future water fluxes and water resource availability. Both reviewers found this scope to be relevant and of high interest. It is hence not true that interaction would mean either the studying the impact of land-use change on climate, or impact of climate change on water extraction / land-use, which are very different issues (targeting atmospheric responses or human responses), and inconsistent with the hydrological response focus of the study.

7599/22: main uncertainties: basically only one type of uncertainty is dealt with, which is the uncertainty in the climate projections. There is no uncertainty analysis of the hydrological model at all.

→ We disagree. Two main uncertainties of quite different nature are studied. The first regards the influence of the spread of different GCM projections on the hydrological modelling results (error bars in Fig. 2). The second regards effects of different approaches to hydro-climatic model coupling, and in particular how sensitive the hydrological modelling step is to observed biases in GCM output for the historical period. The latter issue is addressed by use of two different hydrological modelling approaches: with and without a bias correcting step (blue and red bars of Fig. 2). Furthermore, in the revised manuscript, we address two additional sources of uncertainty in more detail. 1. In new hydrological simulations, we complement the previous results from the ET model of Langbein by applying the ET model of Thorntwaite, which uses monthly data and hence can account for effects of seasonal variations. 2. We test the influence of refined routines for free water evaporation from the rivers.

7600/4 & 5: the text would be more fluent if you use precipitation and runoff instad of P and R throughout the text.

→ We would like to keep a notation that makes it easy to distinguish between e.g. ET and ΔET, P and ΔP etc, since the logics of paragraphs like 7604/1-14 would be hard to follow if the frequently used symbols would have to be replaced by more lengthy (combinations) of words. Also, the used abbreviations P (precipitation), R (runoff), ET (evapotranspiration) etc are well-established.

7601/7-11: Here a case is made for the direct use of GCM outputs. The size of the Aral sea basin may indeed reduce the bias in the total precipitation flux over the basin as represented by the GCMs, but much depends on how the water balance model deals with the spatial distribution of precipitation, as well as precipitation intensity and duration. Each of these characteristics may have an important impact on model behaviour, for instance whether precipitation evaporates or runs of into the drainage network.

→ The spatial distributions of precipitation, temperature, evapotranspiration etc are well resolved in the modelling, since e.g. the grid is sufficiently fine to account for the spatial variation in climate data from CRU. Regarding possible effects of seasonal changes in precipitation and other climate parameters, the issue is addressed by incorporating new results from modelling at monthly resolution, see our response to the general comments of reviewer #1.

7601/ section 3.1: some equations of the main algorithms would be useful here.

→ OK, changed the text accordingly

7602/1-15: Even with the presented references, a more detailed description of the hydrological model is necessary. Is it correct that that the implemented model does not do any time stepping, but only calculate the water balance over a 30 year period? If so, land-use is necessarily held constant over the entire time period. Is this realistic for the past? What land-use discretisation was used? If the Langbein formula was used, I presume that no vegetation properties incorporated? How was irrigation accounted for?

→ We have now considerably extended the direct information on the modelling steps and theoretical background. The quantifications are based on input of long-term average conditions. The considered average land-use conditions represent the period 1961-1990, i.e., after the considerable irrigation expansion of the 1950´s. We clarify that land-use information (with resolution of about 500m×500m) is used to delineate the spatial extent and location of the areas where the land-surface is irrigated and therefore kept at a wetter state than given by the precipitation data from CRU. Furthermore, in the revised version of the manuscript, we use land cover information to investigate the possible influence of open water evaporation from rivers and reservoirs on the presented results (see also our response to reviewer #1, comment 7602/8).

7602/21: fully consistent with effects of historical, multi-decadal land-use and wateruse driven changes: this is a bold statement. Even though the model performance is probably fully presented the given reference, it would be good to provide more details, such as the goodness of fit and the number of observations that the model was compared (calibrated?) to.

→ We perform a more detailed assessment, including analysis of the bias-factor as defined Jarsjö et al 2008 [*Spatial distribution of unmonitored inland water discharges to the sea, Journal of Hydrology, 348,59-72*]:

$$X = \frac{Q_{obs}}{Q_{mod}} + \left(1 - \frac{Q_{obs}}{Q_{mod}}\right) \cdot \frac{P_{obs}}{ET_{mod}}$$

in which Q_{mod} and Q_{obs} are the modelled and observed total river discharge at the basin outlet, P_{obs} is the observed total precipitation over the basin and ET_{mod} is the modelled ET. The value of X represents a bias-factor by which the modelled total ET would need to be scaled, in order to obtain a calibrated model that reproduces the observed river discharge Q_{obs}. The factor X equals unity if there is no bias, i.e. if the model independently can reproduce Q_{obs}, without any scaling of ET_{mod}.

For the historical, pre-irrigation period in ASDB, Shibuo et al. 2007 [*Hydrological responses to climate change and irrigation in the Aral Sea drainage basin, Geophysical Research Letters, 34, L21406*] reports P_{obs}=467 $km^3 \cdot yr^{-1}$ Q_{obs}=71 $km^3 \cdot yr^{-1}$, Q_{mod}=77 $km^3 \cdot yr^{-1}$, and ET_{mod}=391 $km^3 \cdot yr^{-1}$, which yields a bias factor X of 1.02 for the modelled ET implying that it would need to be just 2% higher to exactly obtain Q_{mod} =Q_{obs}. This shows that the model yield consistent results under relatively undisturbed conditions. For their considered period 1983-2002, during which 50 $km^3 \cdot yr^{-1}$ were re-routed to irrigated fields, Shibuo et al. 2007 reports P_{obs}=487 $km^3 \cdot yr^{-1}$ Q_{obs}=12 $km^3 \cdot yr^{-1}$, Q_{mod}=16 $km^3 \cdot yr^{-1}$, and ET_{mod}=458 $km^3 \cdot yr^{-1}$, which yields a bias factor similarly close to unity (X=1.02) as for the pre-irrigation period. This shows that the model results are also consistent with the observed effects of the additional ET losses caused by the water re-routings to irrigated fields.

7603/1-8: Both Langbein and Thronthwaite are known to be prone to large bias. Have any studies been done to compare these methods with more recent methods such as Penman Monteith? At least for the historical period (were data about humidity, radiation and others are available) this should be possible.

→ We use a tougher test of the Langbein (and Thorntwaite) models than comparison with the Penman Monteith model, since we close the water balance and compare with actual runoff observations (under different ambient conditions, see above). For instance, the study of Kite and Droogers 2000 [*Comparing evapotranspiration estimates from satellites, hydrological models and field data, Journal of Hydrology, 229,3-18*] shows that the Penman-Monteith type of ET models can yield errors of 30% - 50% unless calibrated to runoff observations. This is considerably larger than errors of 10%-15% that are involved in ET estimation from water balance closure (see Asokan et al. 2010 [*Vapor flux by evapotranspiration: Effects of changes in climate, land use and water use, Journal of geophysical research, 114, D24102*], and references therein.

7603/15: keeping irrigated area stable may be consistent with the impossibility to expand irrigation, but it may be inconsistent a forced decrease of irrigation because of a

lack of water under future drying scenarios.
→ The basin may experience a combination of expansion in its upper parts, and forced decrease in its lower parts, see also our response to the last general comment above.

7603/29 - 7604/2: the use of calibrated and uncalibrated is wrong here. No calibration was done. It would be more correct to distinguish between the two methods as biascorrected and not bias corrected. Also, where absolute or relative anomalies used?
→ Yes, we agree that the suggested terminology is better. The manuscript has been changed accordingly. Regarding "anomalies", we *added* the GCM change projections to (i) the GCM output for the reference period, and (ii) the CRU observational data for the reference period (7603/26-28); hence, in both cases the hydrological modelling was performed using absolute values.

7604/21: the corresponding ensemble mean value: is this for TAR or AR4 or both?
→ This is for AR4: we changed the text accordingly.

7604/25: what is the TAR model ensemble average P?
→ This is shown in table 2. This is now clarified in the main text, and information on how to conveniently convert units are included in the table heading (see also follow-up question below).

7605/13 and further: The results would be much easier to interpret if they were expressed as mm/year instead of km3/year
→ This is a matter of taste. We think that water balances, which are a key focus of the study, are much easier to understand if expressed in volumes per year. We however appreciate that other units also can be useful and therefore provide information on how to conveniently convert units (see response to previous question).

7605/25-28: I disagree with this conclusion, especially since the reanalysis data and the model simulations may not be independent (see higher). It is much more useful to look at the ensemble spread.
→ The conclusion follows directly from comparing our own results on the two alternative assumptions of hydrological model coupling, and cannot be disregarded on the basis of a possible dependence between reanalysis data and GCM simulation results. The conclusion specifically regards the hydrological model sensitivity to observed biases in GCM output for the historical period, which differs from elsewhere discussed effects of GCM ensemble spread (see also 7599/22). This point is now clarified by the change of notation from "calibrated/ uncalibrated GCM results" to "GCM results with/ without bias correction" (see our response to 7603/29 - 7604/2). The existing GCM biases need attention in hydrologic modelling, regardless of the extent to which climate modelers have improved agreements between their models and reanalysis data. The presented results support the conclusion that systematic biases in GCM output (note the changed wording for increased clarity) do not have large effects on the trend of runoff change, as seen in Figure 2. This means that the bias-correcting step in the hydrological modelling (previously called the calibrated approach to model coupling) may not be critical for the modelled change trends. Note that (other) results regarding the considerable effects of GCM ensemble spread on runoff change are discussed elsewhere (7604/17-20, 7605/1-6; 7606/1-16;7610/4-6;16-25).

7606/1: Given that the model structures are not independent, the standard deviation may not be the best indicator of spread. I suggest you use either the total prediction envelope, or add the models outside the standard deviation as points in Figure 2 (see also Mote, P., L. Brekke, P. B. Duffy, and E. Maurer (2011), Guidelines for constructing

climate scenarios, Eos Trans. AGU, 92(31), doi:10.1029/2011EO310001.)
→ We added the models outside the standard deviation as points in Figure 2.

7606/13-16: this discussion seems to start from the premise that the ensemble mean is the correct prediction which is not true. If some models do predict a positive change in runoff, then there is a non-negligible chance that future runoff will indeed increase.
→ We do not assume, or claim, that the ensemble mean is correct. We simply explain that we avoid errors associated with single GCMs by using the ensemble mean.

7606/20-7607/3: This discussion assumes that the CRU data set is correct which is unlikely to be the case.
→ We do not assume that the CRU data is correct. Such an assumption is not needed either, since different facts converge on supporting the reasoning of this section. See further our response to the reviewer's general comments on the CRU data interpretation.

7609/11 – 14 (part I): This is not true. In fact, most hydrological models incorporate landsurface schemes that are much more complex than the model implemented in this study.
→ This statement does not refer to incorporation of land surface schemes in hydrological models, it refers to the adopted water balance approach, and the incorporation of hydrological routines that resolve basin-scale water balances in GCMs (7609/ 11-14). Such routines do not exist in the presently considered 20 GCMs, and it would indeed be a huge task to implement them, not least due to issues of scale. Resolving water balances in river basins around the globe would require much finer GCM grids than those available in AR4, for instance.

7609/11 – 14 (part II): The model itself is perhaps implemented at higher resolution, but it is unclear whether this yields any improvement, because (1) precipitation is not disaggregated, and (2) it is unclear how the model takes benefit from higher resolution land-cover maps. For instance, it seems that the evapotranspiration routine does not use information about land cover. One advantage of the offline approach would be the ability to incorporate interactions with water extraction for irrigation. However, in the current setup, such interaction is not dealt with. Rather, both processes (water use impacts and climate change impacts) are modelled independently.
→ As stated earlier, this study does resolve effects of interactions between climate change and water re-distributions on future water fluxes: the views of the reviewer that it does not, and that the processes are modeled one at a time, are misunderstandings (see further response 7599/15, and our last response paragraph in the "general comments" section). Furthermore, the reviewer assumption (1) above, that the precipitation is not disaggregated in the model, is not correct. As stated upfront in the modelling section 3.1 (7601/13) and explained in the following text (7602/1-5), we use distributed modelling at high resolution that accounts for the spatial variation in climate data from CRU. With the current revisions, including an extended description of the model and modelling method, we think these points are made clear, as well as issue (2) above. As stated earlier, we also investigate possible effects of seasonal changes in precipitation and other climate parameters, by incorporating new results from modelling at monthly resolution (see also our response to the general comments of reviewer #1).

Fig. 3: On what kind of observations are the observed runoff data based?
→ On data from two independent runoff data bases: the Global Runoff Data Centre (Koblenz, Germany, available at http://grdc.bafg.de), and from a scientific synthesis report of regionally available measurement data, by Mamatov 2003 (INTAS project 1014, EU-INTAS Aral Sea Basin Call 2000, Group CR5 status report of April 30, 2003). References are now included in the text.

Fig. 3: please also indicate the spread of the GCM model ensemble for the projections
→ The spread has already been shown in Figure 2 (which has been enhanced according to the reviewer comments, including both standard deviations and models outside of the standard deviations as points). As previously mentioned, the implications of this relatively large spread (e.g. that even the projected direction of runoff change can differ depending on choice of individual GCMs) are mentioned and discussed in several paragraphs (7604/17-20; 7605/1-6; 7606/1-16;7610/4-6;16-25), which taken together thoroughly addresses the ensemble spread and its implications. A main new point that we want to illustrate with Figure 3 is that the different ensemble averages converge on showing runoff decreases; this point is best made without repeating the already discussed issue on ensemble spread.

Fig. 3: are the groundwater water depletions projected by the uncalibrated models realistic? Would it not be more likely that irrigation will decrease because of a lack of water? Again expressing the runoff in mm/year rather than km3/year would make this easier to interpret.
→ See response to the last general comment, and specific comment 7605/13.

Fig. 4: the sketches on the right-hand side are unclear. What do the bended shapes represent? Do the different colours have any meaning? It would probably be much clearer to show the values as bar charts.
→ We changed the shapes to more closely resemble catchment areas and removed the colour difference.