Answer to comments of Anonymous Referee #1

The original comments of Referee #1 are in black color and indicated by “R1:”. Replies by the authors (“A”) are colored in green. Actions are introduced by “Action:”, changes in the manuscript are in italics.

R1: General Remarks:
This manuscript uses the global water resources assessment model WaterGAP to conduct a series of experiments by changing meteorological forcing and land use data, model physics, parameters, and processes involved with human water use. Results from these experiments are compared with those from the standard version of the model to study the sensitivity of simulated water fluxes and stores to the variations made in the experiments. The study concludes that the basin-specific calibration leads to the largest change among different variants considered, with the changes in climate forcing to be the second dominant factor. While the results are interesting and suitable for HESS, I find that there are important shortcomings in the presentation of this study which should be addressed before acceptance for publication. Some of the issues are related to limited discussion of the possible reasons behind the differences in hydrologic components simulated by different model variants. The introduction and several other sections can be significantly refined with additional efforts. I summarize my specific comments in the followings.
A: We thank the reviewer for the detailed and constructive comments. We will answer the specific comments below.

Specific Comments:
R1: P1589,L11: “Parameter uncertainty is neglected”. What is its significance here? Please clarify.
A: We would like to answer this comment also considering general comment 2 of anonymous referee #2. In short, we want to state that we do not consider parameter uncertainty throughout the study, but only the other main sources of uncertainty, model structure and spatially distributed input data. We deleted the sentence and added at the end of the paragraph the following sentence for clarification: “*The impact of parameter uncertainty is not investigated in this study as previous studies have already identified its strong impact on WaterGAP output (e.g. Kaspar 2003, Werth and Güntner, 2010).*”

R1: P1591,L19: Wouldn’t the sudden change in input forcing cause abrupt changes in storages and fluxes in the model? I assume that the model would show certain spurious trends in states and fluxes until it reaches a new equilibrium in the new climate. Please discuss this issue here.
A: This is correct. As we stated in the results section P 1596, L20-22, AET increases when using WFDEI starting in 1979. Also, the changes in freshwater storages (Table 3, col. STANDARD comp. to CLIMATE) indicate a decrease of some storages due to the AET increase. To our knowledge, there is currently no long term daily climate forcing data set available which includes radiation information (as WFD or WFDEI). WFD ends 2001 which reduces the potential to use newer observed discharge data for calibration. Therefore we combined both datasets. As we stated in P1609, L9 it would be of great value for large scale modelling to have a consistent dataset available. Action: We cannot provide a thorough discussion in the methods section, but we will include an additional sentence and the reference to the results section. P1591, L22: “*The effect of switching the climate input data-set in 1979 is described in Sect. 3.1*”.

R1:P1594,L8: "…entirely satisfied from surface water resources…". How is the demand satisfied if surface water supplies are not enough?

A: We thank the reviewer for pointing to this unclear expression. We wanted to state that, in STRUCTURE, water is only abstracted from surface water storage and not from groundwater. As in all other model variants, it is possible that water supplies are not sufficient to satisfy the demand. See also Table 2, column STRUCTURE, where the demand (NA_s) is 1082 $km^3 yr^{-1}$ and the satisfied demand is 983 $km^3 yr^{-1}$ due to limited surface water resources. Action: we will modify this bullet point. P1594, L8: "*Water for human water use is abstracted only from surface water bodies, i.e. there are no groundwater abstractions as introduced by Döll et al. (2012)*".

R1:P1595,L11: Why is ET less when water use is considered? I would assume that ET increases due to irrigation. Also, soil water storage should increase if water is taken from rivers and is used for irrigation.

A: We agree with the reviewer. What was lacking in the manuscript was the explanation that what is termed AET does not include evapotranspiration caused by irrigation and other human water use, but that this evapotranspiration is called "actual water consumption". Action: We added as third sentence of the first paragraph of section 3.1 (after "… and selected climate forcing"): "*Please note that actual evapotranspiration AET in Table 2 and Figure 3 does not include evapotranspiration caused by irrigation and other human water use. This part of evapotranspiration is called actual water consumption WC_a.* In addition, we rephrased the P1595, L11 as follows: *"When human water use is not taken into account (NoUse), AET increases by 131 km^3 yr^{-1} because evaporation from open water bodies increases as they are not depleted by water uses."* In Table 2 we added a footnote to row 2: "*AET does not include evapotranspiration caused by human water use, i.e. actual water consumption WC_a*". In Table 5, where we summed up WC_a and AET for STANDARD, we added a footnote after the value for STANDARD *"sum of AET and WC_a"*.

Regarding the question why soil water storage is not affected: Within WGHM, addition of irrigation water to soil, evapotranspiration of part of the added water and return flows of the rest are not modeled. Instead, net abstractions from surface water and groundwater are computed in a pre-processing step by the sub-model GWSWUSE (as also noted in P1596, L 14-17). Action: We feel that no action is required.

R1: P1595 L17: Is it 1031 or it is because of roundup? You may note in Table captions that the numbers don't sum up due to rounding.

A: Thanks! Changes in TWS should be 142 km^3 yr^{-1} instead 143 km^3 yr^{-1} after checking the numbers again (Table 2), hence the overall sum is 1031 km^3 yr^{-1}. Action: We corrected the number in the text.

R1: P1596, L16: "without passing the soil compartment": Where does the groundwater go if it is used for irrigation? Isn't it added to the soil?

A: As we answered two comments above, GWSWUSE calculates net abstractions from groundwater resources and surface water resources (e.g. based on the irrigation water use model). Hence, water goes directly into evaporation or back to groundwater / surface water as return flow without passing the soil compartment. Action: no action is required.

R1: P1596, L20: How do you attribute the difference in ET from WFD and WFDEI? High ET could be related to higher temperatures but how do precipitation differ in two datasets?

A: The precipitation totals in both datasets differ only slightly (1979-2001 global average value for WFD: 110309 $km^3 yr^{-1}$; for WFDEI: 110812 $km^3 yr^{-1}$). The same holds true for temperature (WFD: 13.73 °C, WFDEI: 13.80 °C) and longwave downward radiation (WFD: 325.71 W m^{-2}, WFDEI: 325.01 W m^{-2}) while shortwave downward radiation differs largely (WFD: 172.74 W m^{-2}, WFDEI: 189.21 W m^{-2}) (all values excluding Greenland, Antarctica and inland sinks to be consistent with Table 2). Hence, the shortwave radiation bias is the reason for higher AET with WFDEI. We stated this on P1609, L4. Action: We added the reason for the increase already on P1596, L22: "*With WFDEI that is based on ERA-Interim, AET is around 70 000 $km^3 yr^{-1}$, compared to 65 000 $km^3 yr^{-1}$ in case of WFD. This is caused by differences in the shortwave downward radiation.*"

R1: P1597 and Fig 3: Significant spatial differences can be seen in Figure 3. More detailed explanation on the reasons would be appreciated. For example, why is ET so high around the great lakes in STRUCTURE-STANDARD? Likewise, what causes the huge blue blob in eastern China in NoCal-STANDARD?

A: There are two separate reasons for the spatial differences the reviewer pointed to. 1. The difference of the Great Lakes between STRUCTURE and STANDARD: With STANDARD, the correction factors CFA have to be applied in order to reduce runoff from land surfaces (and thus increase evaporation) to achieve a good fit of simulated and observed river discharge in the river basin. Simply, in STANDARD, AET is too low in this basin. This is caused by evaporation reduction factor (P1615, L21-25), which reduces evaporation from the Great Lakes to between ¼ and ½ from potential evaporation due to low modeled lake storages. Within STRUCTURE, no evaporation reduction factor is considered. Hence, the great lakes evaporate with PET within STRUCTURE, resulting in a much larger AET (and even more lower lake storages) of the 4 grid cells within STRUCTURE. Those grid cell values are unfortunately not visible in Fig. 3 as the outflow grid cells (where the whole water balance is modeled) are overlaid by the boundary of the lakes. As in STRUCTURE the lakes themselves evaporate much more than in STANDARD, AET in the surrounding land in the basin is much smaller in STRUCTURE than in STANDARD, hence the red blob occurs in Fig. 3. Consequently, RWR is higher (blue colors in Fig. 4) around the Great Lakes. Action: We added some sentences after P1597,L26: *"In addition, more complex effects occur. The Great Lakes, for example, evaporate with PET in STRUCTURE, even when the lake storage is relatively low. This results in a relatively low modeled discharge which fits well to the observed ones. Hence, neither the correction factors CFA nor CFS are required in the Great Lakes basin. However, in STANDARD the reduction factor reduces evaporation by up to ¾ of PET. The resulting higher modeled discharge has to be reduced by an increased AET in STANDARD (and in the other model variants) on the land around the lakes as compared to STRUCTURE (red areas around Great Lakes in Fig. 3)."*

2. The huge blue blob in eastern China in NoCal-STANDARD: The blue colors are mainly within the Yangtze River which is used for calibration. The calibration parameter γ is between 0.1 and 0.5 which means that a high runoff from soil is modeled (P1618, L10). Within NoCal, values for γ are globally set to 1.0. Thus, less runoff and more AET is modeled with NoCal compared to STANDARD. Action: We added the following sentences before P1597,L27: "*Differences between NoCal and STANDARD are resulting due to the calibration parameter γ which differs from 1.0 (NoCal) in most cases in STANDARD. For example, there are blue patterns in China and South America. In both regions, γ is less than 1.0 in*

STANDARD which results in higher runoff and less modeled AET. In many other regions (red areas), γ is greater than 1.0 in STANDARD."

R1: Figure 4: Why is NoUse missing in Figure 4? Adding it would make Figure 3 and 4 identical and that makes it easy to follow.
A: all model variants for Figure 4 are run without considering water use as we defined renewable water resources in this way on P1598, L21. However, it is erroneous to label Fig. 4 (a) as STANDARD while it is NoUse. Action: adapt within the Figure (NoUse instead STANDARD) and change Figure caption to "*Renewable water resources (mean annual runoff from each cell if water use is neglected) calculated by WaterGAP 2.2 NoUse variant (a) and absolute differences to other variants (variants here run without considering water use) (b-e).*". Within the text (Sect. 3.3), we feel that no additional action is required.

R1: Figure 4: Why is there a red blob in northwest India in the difference between in landcover-standard (c)? Specifically, why does land cover affect so hugely in some regions?
A: We thank the reviewer for pointing that. We are sorry that we have missed to give some more explanation of the reason of those "blobs". Those are the artifacts which we briefly mention on P1603, L18. In this study, we integrated for the first time the station correction factor CFS into the outputs of AET and RWR. We did this as follows: Firstly, the runoff from grid cells CR_{basin} is summed up for the river basin. Secondly, discharge at basin outlet that would occur without CFS is calculated by using the model output and CFS $Q_{simnoCFS}$ (by dividing the simulated discharge Q_{sim} including CFS by CFS). Thirdly, the inflows from upstream basins are summed up (Q_{inflow}). A correction factor $f_{crunoff}$ is calculated as (Q_{sim} - Q_{inflow})/($Q_{simnoCFS}$ - Q_{inflow}) and applied to each grid cell of the specific basin. This method enables the consistency of the global maps with the values given in Table 2, as CFS is integrated there. For the Indus basin and LANDCOVER, $f_{crunoff}$ is 2 meaning that all runoff values from the grid cell (which can be negative in those cases, when more water inflowing from the upstream grid cell is evaporated in (global) lakes and wetlands than is generated from precipitation in the grid cell) are multiplied. Within the red blob in northwest India (it is the Indus basin), the following happens: Figure 4 shows the runoff from grid cells which occur in model runs without water use. Q_{sim} is nearly doubled and Q_{inflow} is increased by around 30%, CR_{basin} is negative. When applying the correction method for taking into account CFS, $f_{crunoff}$ gets a value of -60 due to inconsistent input data. Hence, all runoff values for the grid cells are multiplied by this factor. At the Indus, positive and negative runoff values occur which gets now unrealistic high. Again, the correction method attempts for a consistent visualization of RWR (and slightly adapted also AET) taken the CFS into account.

Further analysis of this attempt to consistently integrate CFS into grid cell values of AET and RWR since submission of the manuscript lead us to the conclusion that it is advisable to abstain from this approach because physically implausible AET and RWR values may result in some cases probably due to inconsistencies between precipitation input and observed discharge. For example, in grid cells without global lakes or wetlands, we obtained negative RWR due to the correction which is physically impossible. Action:

1) We now show in Figs 3 and 4 with model output that was not modified by the CFS correction as described above.
2) We deleted the last sentence of section 3.2 "Moreover, in calibration basins, AET is adjusted in such a way that it is consistent with precipitation and simulated discharge and affected by correction factors CFA and CFS (calibration details see Appendix B)."

3) As fourth and fifth sentence of section 3.1, we added: "*For computing global values of AET and renewable water resources RWR, AET and RWR in calibration basins were adjusted taking into account station correction factor CFS such that a closed global water balance is achieved (for calibration details see Appendix B). Grid cell values of AET and RWR (Figs. 3 and 4), however, do not reflect CFS, to avoid physically implausible values that likely result from inconsistencies between precipitation data and observed river discharge.*"

R1: P1599,L21: It would be interesting to see NoCal in Fig. 5 if the limits are not way too off.
A: Thanks, we added NoCal to Fig. 5 and modified figure header. In some basins, limits are way too off and sometimes only single months are shown. Action: We deleted at P1599, L21 "(NoCal is not shown as the Y-axis would have a very large spread)."

R1: Figure 5: I wonder why STRUCTURE does a very good job in Lena. Conversely, why does it show mediocre performance in Parana? Please discuss in more details.
A: The STRUCTURE and the STANDARD variant differ concerning the representation of several processes. In the Parana, the poor performance of STRUCTURE (compared to all other variants) results from the missing reservoir management algorithm.
In the Lena, the good performance of STRUCTURE is related to the flow velocity algorithm. The algorithm apparently underestimates flow velocity in the lower reaches (where bed slopes are very low) which leads to a shift in the mean monthly discharges. We will add this explanation to Sect 3.4.1.

R1: Figure 5: Please make the legends bigger.
A: Thanks! Action: we increased the font size and line thickness of the legend.

R1: P1600, L1-L11: Please add more discussion on the varied response to different factors in different regions.
A: We were concerned about the length of the paper of our initial submission, but we are happy to include more explanation. Action: We modified the text starting at P1599, L21 to *"Where seasonality of climate is high, like in the monsoon-dominated Mekong basin, only marginal differences occur due to land cover and model structure. Structural model refinements have also important effects on discharge seasonality. For example, the constant flow velocity of STRUCTURE (in contrast to variable flow velocity in the other variants) leads to a higher peak in the Danube, Volga, Mississippi, Amazonas, Lena and Yangtze. Here, the variable flow velocity algorithm underestimates flow velocity in the lower reaches where bed slopes are very low. For the Lena, this leads to a strong underestimation of peak flow (which explains the improved seasonality of STRUCTURE compared to observed discharge). The reservoir algorithm which is not enabled in STRUCTURE has impacts at the Yangtze, Rio Parana, Mississippi and the Volga in terms of smoothing the discharge. For the Rio Parana, this is the main influence in the STRUCTURE variant. The representation of snow in STANDARD leads to a more heterogeneous snow coverage as compared to the STRUCTURE variant. The strongest impact occurs for the Rhine, where the snow algorithm is the dominant reason for the differences to STRUCTURE. In STRUCTURE, the snow water storage of the Rhine headwater (Alps) is generally lower. In particular between May and October (the Alps are modeled as snow-free between June and September), this leads to a decrease of discharge as snowmelt cannot contribute any longer as it does in STANDARD. The importance of the climate forcing can be seen in the Mississippi and the Rhine where CLIMATE results in*

overestimated peak seasonal discharge. In the Danube, WFD/WFDEI climate input (in STANDARD) is particularly beneficial, as the fit to observed seasonality is much better than with CRU TS 3.2/ GPCC v6 climate (in CLIMATE).
For the Mackenzie River, all model variants are close to each other but far away from observations. Here, freezing and thawing of the river are not reproduced as none of the model variants represents these processes. Interestingly, the Lena river basin is also frozen during winter time but here, low flows are simulated quite well. In Amazon, the model variants underestimate the delay of peak discharge which might be explained by the lack of modeling dynamic floodplain inundation.

The impact of alternative land cover is only slightly influencing discharge seasonality. Most effects occur at the Rhine, where CORINE-based land cover (variant LANDCOVER) consists dominantly of cropland. Many grid cells in the other model variants consists of mixed forest or cropland / natural vegetation mosaic which both have a lower albedo, resulting in more evaporation and less discharge especial in the summer months. Additional effects occur due to deeper roots at mixed forest class. Only for the Mackenzie, Lena and Yangtze, mean monthly river discharges of NoCal within the range of all other variants in some months. The NoCal values for the Orange river are so high that throughout the year, they are higher than the highest observed value (and the values of the other variants) (Fig. 5). This supports the use of a calibrated model for discharge analyses."

R1: P1601,Figure 7: It is not clear what the variable shown in 7(a) is? Is it the difference between the minima and maxima of the seasonal cycle of TWS? Please explain in the text.
A: We indeed missed to explain how the seasonal variation was calculated. Action: modify sentence on P1601, L11: "*The dominant seasonal changes of TWS can be characterized by the difference between the minimum and the maximum value of mean monthly TWS (1971-2000).*"

R1: P1602, L17-L26: Please consider revising some sentences to improve readability.
A: We have revised the section to: "*Discharge estimates differ due to the applied estimation method and precipitation data set. Mueller et al. (2013) do not consider precipitation undercatch correction and assume a global precipitation of ~ 99 000 km^3 yr^{-1} which is low compared to recent estimates of Schneider et al. (2014) (117 000 km^3 yr^{-1}) or the values used in this study (Table 2). Regarding WaterGAP estimates of global discharge, model refinements have led to an increase of discharge. The value for STANDARD is approx. 450 km^3 yr^{-1} higher than for STRUCTURE (Table 2), and previous estimates (Döll et al., 2003) are even lower as precipitation undercatch was not taken into account.*"

R1: P1604, L25: Why changes in ET and runoff compensate each other, given that the total input precipitation could be different? Does it mean that storage change is huge?
A: The precipitation data which are used (CLIMATE: GPCCv6, WFD: GPCCv4, WFDEI: GPCCv5/v6) do not differ too much in the regions which are described in brackets (South East Asia, Australia, Saudi Arabia), but radiation differs. Thus, AET is higher and to keep the water balance, RWR is lower by approximately the same amount. Action: we have adapted the text, P. 1604, L25: "*In those regions with similar precipitation amounts but different radiation, RWR decreases by the same amount as AET increased (e.g. South East Asia, Australia, Saudi Arabia). In other regions, no clear effect on RWR is detectable (e.g. North America).*"

R1: P1605,L17: In figure 5 why don't we see any improvements in the Mackenzie as in Danube caused by the difference in snow melt timing? In general, please provide more detailed discussions.
A: The Mackenzie River system is different to the Danube River in that the Mackenzie is freezing in the winter months, which is not considered in all model variants (see also P1600, L4). Action: we already have added some discussion at Page 1600.

R1: P1606,L20: Again, as I also noted earlier why does ET reduce when water use is considered?
A: As we answered earlier, AET in Fig. 3 does not include the (additional) evaporation from WC_a. Action: we rephrased the sentences P1606, L18: "*In regions with intense water use, in particular from surface water bodies (e.g. in Pakistan), AET without considering additional evaporation from WC_a (Table 2) is reduced due to human water use (Fig. 3e). This effect occurs because human water uses decrease surface water storages and thus the reduction factor (Appendix A5) decreases evaporation from surface water bodies.*" In addition, we deleted the subsequent sentence.

R1: P1606, L22: It is not clear why ET from irrigated crops in not considered in AET? Please explain this clearly in the manuscript here and elsewhere.
A: We have already answered this above.

R1: Table 2: P for CLIMATE: Is it possible to show the averages for WFD and WFDEI in the footnote.
A: We added a footnote to Table 2: "*mean annual P (1979-2001) is 110309 km^3 yr^{-1} in WFD, and 110812 km^3 yr^{-1} in WFDEI.*"

R1: Table 3: soil for STANDARD and NoUse: Why are they same? It is not clear where the irrigation water goes as it never affects soil water. Please clarify.
A: As we have already stated, the soil compartment is not affected by water uses. Action: we extended to footnote a: "*In WaterGAP, increase of soil water storage by irrigation is not taken into account such that storage values for STANDARD and NoUse variants are the same.*"

R1: Table 2, Column 1 (STANDARD): Is water consumption limited to the availability from surface water only? What is the -72 groundwater use? Is it return to groundwater from surface water use? Please clarify.
A: Yes, only net abstraction from surface water is limited by availability in WaterGAP while net abstractions from groundwater are unlimited as groundwater can be depleted. The negative global net groundwater abstractions (-72 km^3 yr^1) results from return flows from surface water irrigation to groundwater being larger than the groundwater abstractions for the global sum. Concerning spatial patterns of net groundwater and net surface water abstractions, please refer to Fig. 4 in Döll et al. 2012 (modeled with a different WaterGAP version but the general pattern is the same). Action: We modified footnote d to: "*negative values indicate that return flows from irrigation with surface water exceed groundwater abstractions*".

R1: Editorial Issues (I list some editorial issues below but the list may not be exhaustive. Please carefully proofread the manuscript before submission.)
A: We have carefully revised the manuscript according to the comments listed below.

P1586:
R1: L4: add “to” after “due”

R1: L7: change “sum” to “total”

R1: L15 onwards: The first paragraph of introduction looks a little patchy. Please consider re-writing/re-structuring the whole section.
R1: L20: delete “a”
R1: L25: change ‘it’ to “they”
R1: L26: It is not clear what it means by “estimation of groundwater recharge is equivalent to ….”. Please clarify.
R1: Page 1587:
R1: L5: “strategies” to “methods/ways”
R1: L6: “on the other hand” looks abrupt and awkward. Please re-phrase
R1: L13: change “equations” to “schemes” or something alike
R1: L20: change “yr” to “years”
R1: L27: I suggest re-phrasing
R1: P1588:
R1: L10: “,while…” is not clear
R1: L13: please revise this sentence
R1: L13: change “of” to “in”
R1: L14: “simulations of” should be “simulated by”
A: Thank you for the detailed comments. We will re-write and re-structure the whole introduction and include all your comments above.

R1: P1590,L18: change “done” to “made”

R1: P1591,L27: “(“ is missing.

R1: P1596, L3: it should be “(RWR)”
A: We also changed this to (AET) on P1597, L2.

R1: P1598, L21: “RWR” is already defined.
A: we have deleted “Renewable water resources”, thanks.

R1: P1602,L20: “currently available”

R1: P1603, section 4.2: You may want to use “advantages” or alike instead of “benefits”.

R1: P1603,L16: Re-phrasing required. It should be “ET is largely reduced in one half of the basin ….”
A: We modified the sentence to “*ET is largely reduced in one half of the basin (and vice versa) at the river basin Yenisey at station Igarka (western Siberian Plain) when using alternative climate forcing*”, thanks.

R1: P1605, L13: This sentence is difficult to comprehend

A: Thanks. We adapted the sentences to: P 1605, L3: *"For example, the different elevations of the 100 subgrids used for the improved snow modeling (Schulze and Döll, 2004) lead to different temperatures (see Appendix A2) and thus to more differentiated snow melting within one 0.5° grid cell in STANDARD as compared to STRUCTURE where snow within the whole cell either melted or not on any day."*

R1: P1608,L14: delete "well"

A: deleted, thanks.

R1: P1609,L19: "but also" please revise this sentence.

A: We modified this sentence: "*The improved representation of hydrological processes of WaterGAP within the last decade led to a more complex model structure. In most cases, those modifications resulted in a better fit to observed river discharge*."

R1: P1609, L27: please replace "weak" by a better term.

A: we replaced "weak" by "*poor*", thanks.

R1: P1615,L5: please add "," before "which".

R1: Table 1: "Like STANDARD": change to "Similar to STANDARD"

R1: Table 1: change "land cover of the…" to "land cover for the…"

References

Döll, P., Kaspar, F. and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270(1-2), 105–134, doi:10.1016/S0022-1694(02)00283-4, 2003.

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G. and Scanlon, B. R.: Impact of water withdrawals from groundwater and surface water on continental water storage variations, J. Geodyn., 59-60, 143–156, doi:10.1016/j.jog.2011.05.001, 2012.

Kaspar, F.: Entwicklung und Unsicherheitsanalyse eines globalen hydrologischen Modells, University of Kassel., 2003.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y. and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, Hydrol. Earth Syst. Sci., 17(10), 3707–3720, doi:10.5194/hess-17-3707-2013, 2013.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M. and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, Theor. Appl. Climatol., 115(1-2), 15–40, doi:10.1007/s00704-013-0860-x, 2014.

Schulze, K. and Döll, P.: Neue Ansätze zur Modellierung von Schneeakkumulation und -schmelze im globalen Wassermodell WaterGAP, in Tagungsband zum 7. Workshop zur großskaligen Modellierung in der Hydrologie, edited by R. Ludwig, D. Reichert, and W. Mauser, p. 11, Kassel University Press, Kassel., 2004.

Werth, S. and Güntner, A.: Calibration analysis for water storage variability of the global hydrological model WGHM, Hydrol. Earth Syst. Sci., 14(1), 59–78, doi:10.5194/hess-14-59-2010, 2010.

New Figure 3, without considering CFS.

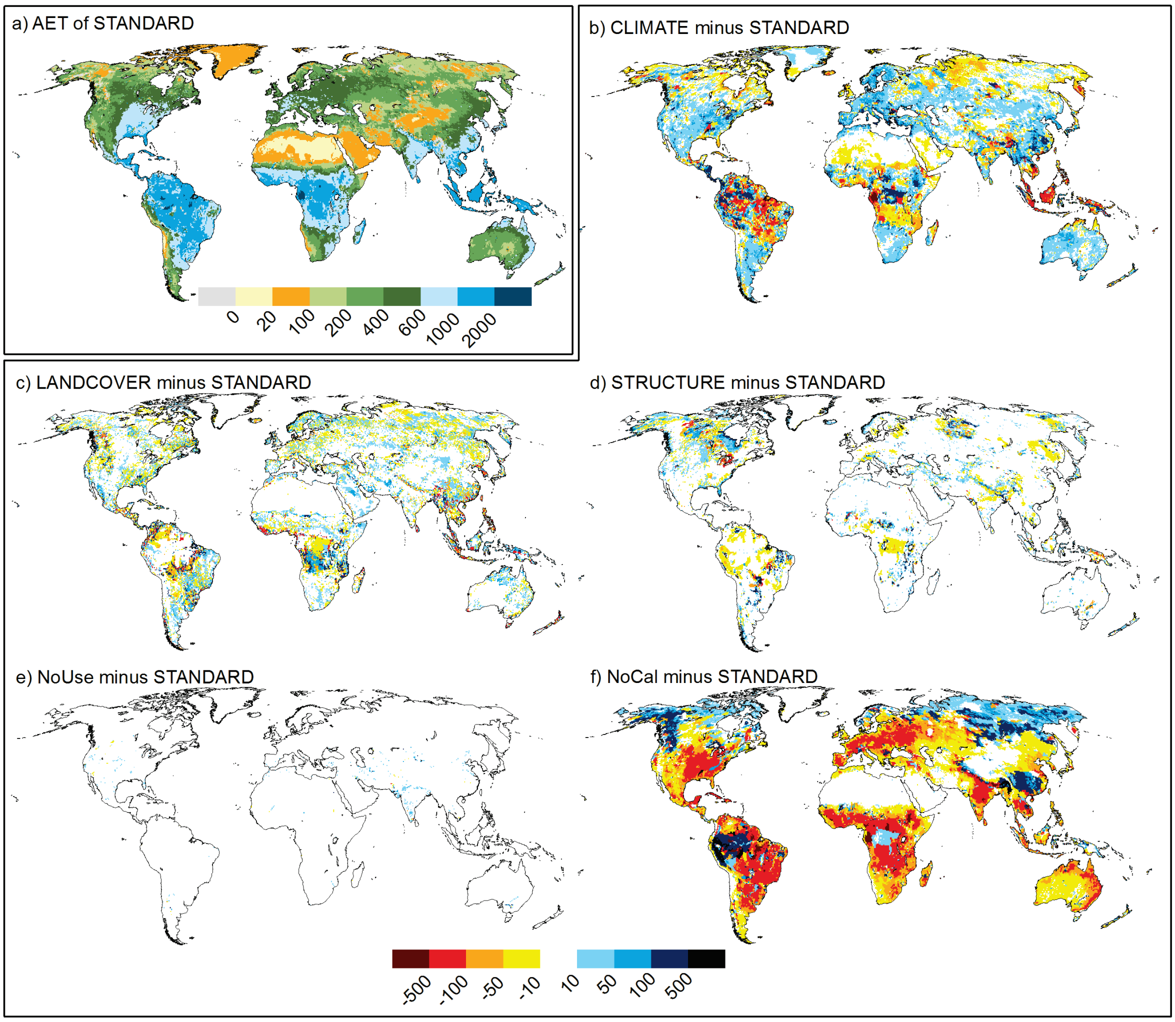

Figure 1. Actual evapotranspiration AET for STANDARD (mean value 1971-2000, in mm yr^{-1}) (a) and differences between the model variants and STANDARD in mm yr^{-1} (b-f).

New Figure 4, without considering CFS.

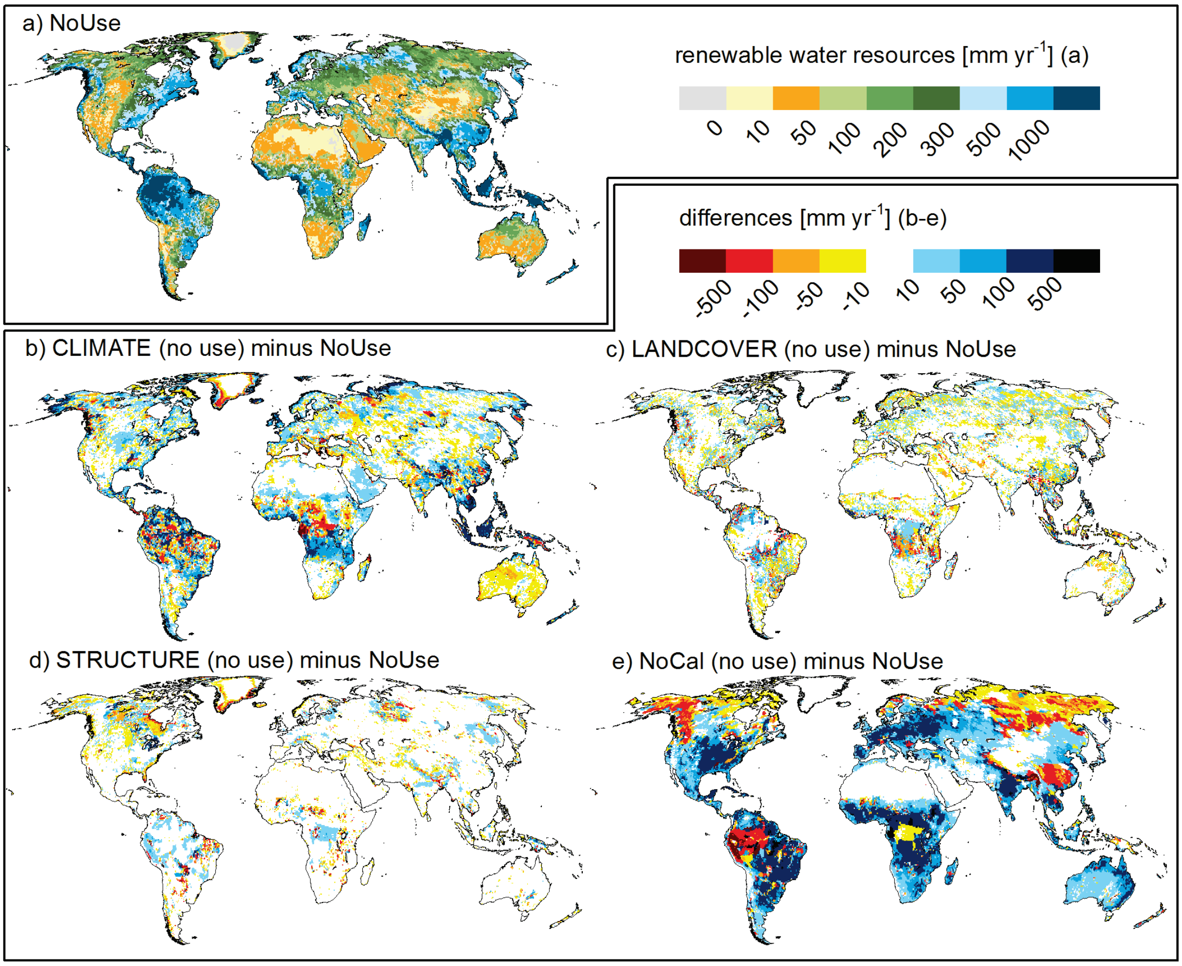

Figure 2. Renewable water resources (mean annual runoff from each cell if water use is neglected) calculated by WaterGAP 2.2 NoUse variant (a) and differences to other variants (variants here run without considering water use) (b-e).

New Figure 5, including now NoCal (purple line) and a bigger legend.

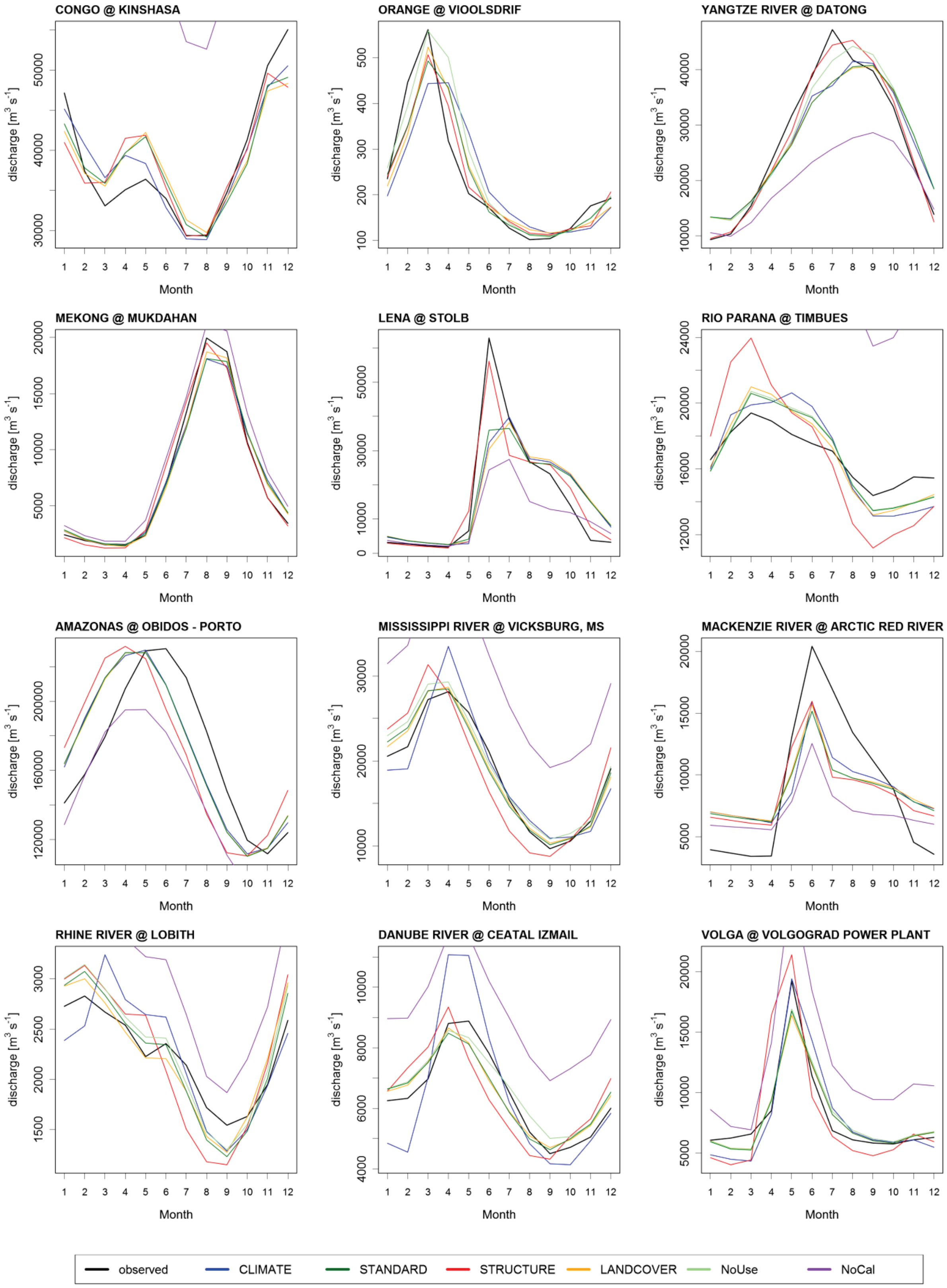

Figure 3. Discharge seasonality for selected basins and the calibrated model variants. Values for NoCal are only visible if they are in the range of calibrated model variants.