General comments:

Hanasaki et al present their effort to enhance the H08 global hydrology model with schemes to attribute water abstraction to different water sources. They detail the functionality of every scheme, explain its impacts in different regions and finally discuss some sources of uncertainty that should be kept in mind. The study is clearly structured and (even though it is quite long) easy to read and follow. The combination of this number of water sources in one model definitely merits the publication of their work. However, there are some points I like to be discussed before the final paper should be accepted:

Thank you very much for taking the time to review this paper. We are grateful for your detailed and helpful comments. We have responded to all comments and made revisions as indicated below.

• [R1-M1] The authors validate their model version by comparing its results to TWS anomalies measured with GRACE. They utilize a simulation with naturalized setup (e.g. no human impacts) with a simulation with all human impacts. However, the latter does not only include their improvements but also the river regulation and dam management scheme implemented in H08 earlier. I would claim this two aspects are already explain most of the improvement in the TWS anomaly. If the author do want to demonstrate an improvement due to their recent chances, they instead need to compare to a simulation with the original H08 and its human impacts enabled.

As you have pointed out, we compared naturalized and human-impacted simulations, not the new and old versions of H08. We have rephrased the text in the revised manuscript to make this clear. Following your suggestion, we have added a new simulation mode to run the new H08 model using a configuration similar to that of the original model (we termed this the ORIG simulation). We have described the method and the difference in simulation performance between ALL and ORIG in Supplemental Text S4 as follows:

"The H08 model has been enhanced by six new schemes. We disabled some of its components, such that it works similarly to the original H08 (hereafter the ORIG simulation mode). To disable the groundwater scheme, we set the groundwater recharge factors (i.e., f_r, f_t, f_h, and f_{pg} in Eq. 1) to zero globally. Thus, groundwater recharge is disabled, and groundwater fluxes and storage become constant at zero.

To disable the groundwater abstraction scheme, we set the fraction of the water requirement assigned to groundwater (f_{gw} in Eqs.4 and 5) to zero globally. This setting assigns the entire water requirement to surface water, preventing water abstraction from non-renewable groundwater. To disable aqueduct water transfer and seawater desalination, we set empty maps of implicit and explicit aqueducts, and the area utilizing seawater desalination. To disable return flow and delivery loss, we set the ratio of consumption to withdrawal (e in Eq. 10) and the proportion lost during delivery (l in Eq. 11) to unity and zero globally, respectively. We then fed the consumption-based (not withdrawal-based, as in the main text) water requirement into the H08 model. Finally, to reconfigure the original local reservoirs, we set the catchment area of a local reservoir (A_{lres} in Eq. 7) to unity globally, and then fed the original global and local reservoir distributions into the model.

We compared performance metrics of ORIG with ALL (H08 with new schemes) for the heavily human-affected basins described in Table S4. Regarding TWSA, ALL outperformed ORIG in five of six basins in terms of NSE and CC. The good performance of ALL in the TWS anomaly is attributable primarily to the inclusion of the groundwater recharge scheme, which provides greater amplitude and a delayed peak in the TWS anomaly, agreeing well with observations. Other factors, e.g., the inclusion of return flow and aqueduct water transfer, showed marginal effects because they have little effect on monthly-scale water storage in the basins. Regarding river discharge, we observed considerable improvement in NSE in four of six basins. This result is attributed to the inclusion of groundwater, which supplies stable baseflow throughout the year."

• [R1-M2] At several points the authors point out that the water balance is strictly closed. Technically this will be true as, no doubt, the models tracks all water storages and fluxes and no water is generated or vanished which the authors are not aware of. However, the do use unlimited water sources to satisfy the water requirements and, thus, the water balance is actually violated. Please reformulate such statements to avoid misunderstandings.

To make our intentions clearer, we have rephrased this as "water source is traceable" throughout the text.

[R1-M3] The authors present numbers about how much water for a given sector is extracted from which source. However, such number seem to rely on very arbitrary decisions about the order of water extractions (see specific comment P18L4). As the numbers are presented as important parts of their results, I would want to see a justification why this order of water abstraction (and therefore this numbers) is more valid than any other order. Are there economic reasons for this prioritisation? What about allowing the different sectors to share a commonly used source according to their relative water demand fraction. This would much better reflect the simultaneous use of a source by different sectors.

> As you have pointed out, the order of water abstractions affects the results. We added the rationale for our priority assessments to Section 2.1.7., as follows: "The order of water withdrawal reflects the distinct differences in water use intensity on the general premise that priority should be given to high value–added products in resource allocation. Municipal, industrial, and agricultural water use intensities per value added (service, manufacturing and power generation, and agricultural sectors) are estimated to be 0.012, 0.063, and 2.2 $10^6 m^3\ 10^6 USD^{-1}$, respectively." The idea of allocating water by the volume of water required is interesting, and perhaps this approach is in practice in some regions. However, this method would not likely yield practical results if it were adopted as a globally uniform algorithm. For example, in Asian countries, irrigation water accounts for a considerable fraction of water used; hence, if the algorithm were adopted, only a small volume of water would be assigned to municipal and industrial sectors during the dry season, which is unrealistic.

• [R1-M4] In section 3.4.3, the authors discuss different reasons about why the available water is significantly less than the required amount of water. Here, I would ask them to reflect about some maybe related points: why do you actually chose to satisfy the water requirements instead of just diagnosing the missing water. In this way you would also avoid the water balance violation via return flows.

> Thank you for this interesting question. We had logical and technical reasons for making this choice. To explain the logical reason, we added the following text to Sect 2.1.7: "A precondition of this study was that water withdrawal estimates were based on values reported in the AQUASTAT database. As shown in Section 2.2.1 and Appendix A1, municipal and industrial water requirements were taken from the database, and simulated irrigation water was carefully compared with these data (Supplemental Text S1). We presumed that AQUASTAT reported the volume of water that was actually abstracted in each nation. Option 1 strictly followed this condition and compensated with unspecified surface water in cases where surface

water data were not available.” Regarding the technical reason, we added the following explanation: “Note that adding water from imaginary sources to the H08 model is nearly the only way to quantitatively estimate the volume of missing source water, particularly for irrigation. As shown in Appendix A1, the irrigation water requirement was determined by the soil moisture deficit, which shows highly nonlinear behavior and interacts with other components. An imaginary source of water fills the deficit at every calculation interval, such that the accumulation of water equals the volume of missing source water.”

[R1-M5] based on the information from the appendix I understand that the water withdrawn for industrial and municipal sector is consumed. However, in reality both sectors produce a large amount of waste water which, after treatment, goes back into the water cycle. Is this somehow accounted for or does your data source explicitly include the consumed water?

Yes, this water is accounted for. Fractions of the water withdrawn as municipal and industrial water (15% and 10%, respectively) are consumed and removed from the system. The remaining fraction is drained to rivers as return flow. See the latter part of Section 2.1.6.

[R1-M6] If not this might be part of the missing water. what about the possibility that the water is not actually missing. You derive the surface water / groundwater water withdrawal ratio from quite large scale data. Thus, the real ratio at grid cell level might be extremely different. For grid cells with either a large groundwater or surface water storage the use of this large scale average might cause a depletion in the surface water (groundwater) storage even though there would be enough water in the groundwater (surface water) storage. To me this seems to be a much larger source of uncertainty than e.g. the model resolution itself.

Yes, a substantial volume of unspecified surface water could be attributed to the fixed surface and groundwater fraction. We added this point to Section 3.4.3 (Potential sources of uncertainty) as “Water source separation into surface water and groundwater was determined by a single factor, termed the fraction of water requirement allocated to groundwater. Due to a lack of available data, the same factor was applied for vast areas, ignoring local heterogeneity, which is also a source of uncertainty.” Additionally, as mentioned in our response to a comment from Reviewer 3, we conducted a new simulation (SWT) that allows additional abstraction from renewable groundwater in cases where unspecified surface water is used. The results and discussion of this simulation are shown in Supplemental Text S5. In short, the option reduced the volume of unspecified surface water by approximately 200 km^3yr^{-1} (30%), but increased the total groundwater use far in excess of the reported estimation range; hence, this option is less likely to improve the overall simulation performance.

[R1-M7] Considering these points, I'd ask the author to either justify the robustness of their existing results or adapt some of my proposed changes where possible. Of course, some points (like missing surface water / groundwater abstraction ratio on grid scale) cannot be changed but should be discussed in the uncertainty part. Alternatively, the authors could consider publishing their research in a journal like GMD (http://www.geoscientificmodel-development.net/) where the focus is rather on the development of new model components and, thus, less changes in the manuscript would be needed.

As stated above, we tried to incorporate your valuable comments as frequently as possible. We believe that these additions further enhance the robustness of this paper.

Specific comments

• [R1-S1] P1L24: Do these numbers refer to the simulation or to the GRACE data?

These numbers refer to the simulation. We have added "simulated" in the text to improve clarity.

• [R1-S2] P3L14: I am confused about the local reservoirs. So the local reservoirs were already in the model? It does seem strange to write like ...six things were added, but one not/was already there... please clarify.

We have rephrased this section as "Six schemes or additional components were developed and implemented in the H08 model (Hanasaki et al. 2008a, b, 2010, 2013a, b): groundwater recharge, groundwater abstraction, aqueduct water transfer, local reservoirs, seawater desalination, and return flow and delivery loss schemes. Note that the local reservoir scheme was replaced with that of the original H08 model, whereas the other five schemes were new additions."

• [R1-S3] P5L11: How do you know the total water requirement? Is this computed by your model (if so, how) or based on external data (if so, which dataset)? I see it is explained in the appendix, so just add a link here.

> Thank you for the suggestion; we have added the link to Appendix A.

• [R1-S4] P6L3: While I agree with your decision to use the country with the larger population, I'd like to know what uncertainty is introduced due to it. Would a different sampling affect your fractions distinctively? How important do you consider the (not represented) spatial variation of this fraction within the national borders?

> We have added the following discussion: "As the groundwater use fraction varies considerably among countries (and among regions within countries), this assumption propagates notable uncertainties in the results." We speculate that considerable spatial variation must exist within each nation, but making a more specific statement on this subject is difficult due to a lack of data.

• [R1-S5] P6L11: So you fulfil the groundwater abstraction requirements by take water from an unlimited reservoir. Considering that the extracted water will partly end up as irrigation water, some of it will enter the soil and eventually the renewable groundwater storage. How does this agree with your statement in the abstract, that your water balance is closed at any time. For me it sounds like you (at least potentially) add water to the system and therefore effectively violate the water balance (although you probably technically close it by accounting for this violation).

> Your point is well taken. As mentioned above, we rephrased this text in the Abstract, from "the water balance was always strictly closed" to "all water fluxes and storage were strictly traceable." We hope that this adjustment resolves the conflict between the model concept and the wording of the text in the abstract.

• [R1-S6] P7L14: Why is the water transport via aqueducts considered to be a withdrawal. I'd assume you just move water in the river network from one cell to another. Please rephrase.

> The primary objective of this paper was to specify water sources for human use in a model simulation. The quantity and location of the movement of "water in the river network from one cell to another" are important in meeting this objective. We have

added the caveat, “Note that water withdrawal via aqueducts is generally not distinguished from water withdrawal from a river in reality, and is seldom recorded independently. This point is revisited in Section 3.1.3.”

• [R1-S7] P8L11: What is a storage area of a grid cell?

Thank you for pointing out this oversight. We have rephrased this text as “The catchment area of a local reservoir was equal to that of the largest within a grid cell, unless the area did not exceed the area of the grid cell.”

• [R1-S8] P8L22: Considering you remark (P3L14) I am now confused about whether this is the old local reservoir scheme of original H08 or the new one that was not implemented...

In this section, we described the new local reservoir scheme, which was replaced with the old scheme from the original H08 model. The treatment differs substantially from the original, which is shown in Appendix A. Because we have already rephrased the earlier remark (P3L14), this part should now follow logically.

• [R1-S9] P9L12: Does this mean you (simply) define seawater desalination to be equal to the water requirements from municipal and industrial sector? Could you please add an equation as you did for the other sources.

Yes, you are correct. As the condition is complex (i.e., seawater is available for municipal and industrial water use only where the three geographical conditions are met), the mathematical equation is so complex as to be unhelpful for the reader. We would prefer to leave this equation out of the manuscript.

• [R1-S10] P9L31: What would water lost through percolation be in you model? I thought you only have one soil layer? What is the storage water percolates from?

Our original meaning was that “water consumption” due to leakage and similar effects was added to the return flow. As this section appears to be confusing, we have removed this sentence from the revised manuscript.

• [R1-S11] P10L16: I assume you mean you take the water from the origin of an aqueduct that ends in the actual grid cell, right?

Yes, you are correct. Although this sentence is long, we believe that it conveys our intention to readers.

• [R1-S12] P10L24: Again, using such an unlimited source is a water balance violation.

We have rephrased "maintaining the water balance" as "strictly tracking all water fluxes and storage; therefore, the model contains no unexplained water imbalance."

• [R1-S13] P10L26: What do you mean by statistically based OR well validated?

We have rephrased this section as follows: "statistically based (i.e., the national annual water withdrawal volume for municipal and industrial use was derived from the AQUASTAT database; Appendix A) or well validated (i.e., the national annual simulated irrigation water withdrawal volume agrees well with AQUASTAT data; Appendix A, Figure S2)."

• [R1-S14] P11L20: Do you need all of the 8 forcing variables, or just a subset?

We used all eight forcing variables, which are needed to solve the surface energy balance. We added the following sentence: "All variables are indispensable in the H08 model to solve the land surface water and energy balance."

• [R1-S15] P11L26: From what I read so far, I disagree with this statement. I think you mean that you track all fluxes, sources and sinks and therefore have no unexplained water imbalance in the model, but you can never have a closed water balance while assuming unlimited water reservoirs. Please either reformulated these remarks concerning the water balance or convince me otherwise.

We have replaced "strictly maintain the water balance" with "strictly tracking all water fluxes and storage; therefore, the model contains no unexplained water imbalance."

• [R1-S16] P11L29: What does it mean with respect to the global reservoirs which are already part of the original H08? Where they active in the NAT simulation as well? Is the difference between NAT and ALL just the use of the new sub-models (thus you can clearly show their effect on the simulation) or is it naturalized vs humanimpacted (in which case you would not know whether a given effect comes from the human-impact related processes already being part of the original H08 or from your new processes)? Furthermore, it would be important to know, to what extent unlimited reservoirs contribute to the results.

> To avoid confusion, we have added two sentences: "As mentioned above, the original H08 model consists of six sub-models (land surface hydrology, river routing, reservoir operation, water abstraction, crop growth, and environmental flow requirement). We developed six new schemes. Two simulations with different combinations of sub-models and schemes were conducted in this study." To respond to your questions, global reservoirs were excluded in NAT, which is a naturalized simulation, but included in ALL, which incorporates human impact. Additionally, as mentioned earlier, we added the ORIG simulation (Supplemental Text S4) to reproduce the H08 model simulation with the original settings. Unlimited water sources can appear in the ALL and ORIG simulations when the water abstraction sub-model is enabled.

• [R1-S17] P13L17: Or does it rather demonstrate the validity of your irrigation water requirement computation (and not the full model)? Because (as you said yourself) the requirements for other sectors as well as the separation into surface water / groundwater abstraction comes from data.

> To avoid confusion, we have replaced "implies the validity of our model" with "demonstrates the validity of the irrigation water requirement computation."

• [R1-S18] P15L22: So also the global reservoirs are only active in the ALL simulation? Which means it is hard to clearly separate the simulation improvement coming from the reservoir operations already being part of H08 and the new source schemes.

> Global reservoirs are active in the ALL simulation, but not in the NAT simulation. We also performed the ORIG simulation, which reproduces the function and configuration of the original H08 model. Differences between ALL and ORIG show the effects of the new schemes, as described in Supplemental Text S4. We added the following sentence to the end of introductory paragraph of Section 3.2: "In this subsection, we compare the NAT and ALL simulations to investigate their performance in representing human activity in the enhanced H08 model. A direct

comparison between the original and enhanced H08 models is shown in Supplemental Text S4."

• [R1-S19] P17L18: Your renewable groundwater storage is rather stable in all six river basins, but the unlimited storage shows a clear trend towards depletion. Looking at equation 4 and 5, I wonder how this can be because I'd expect that first the renewable storage has to run dry (Eq 4) before the unlimited storage is used (Eq 5). Is this an effect of the monthly and/or spatial averaging? Please explain.

> Thank you for bringing this to our attention. We have added the following description: "Although non-renewable groundwater storage shows a negative trend, basin-average renewable groundwater storage (Figure 10) was not necessarily depleted because abstraction from the non-renewable part took place only in a limited number of grid cells."

• [R1-S20] P18L4: Why is this only partly true? If you would withdraw water for irrigation first, probably all water for industrial and municipal sectors would have to come from unsustainable storages. Thus, different numbers for the different sectors do not appear to be results, but rather reflect your computation choice or the amount of water available. As I understand it, the only robust value in those numbers would be the percentage of unsustainable water use (accumulated over all sectors). Please comment on this.

> I agree with your assessment that the fraction of non-renewable groundwater was determined by the sectoral order of abstraction (highest priority given to the municipality, followed by industry and irrigation). We have removed "partly" from the text. As we stated above, however, we retained the original order because we believe that it represents general water use priority in the real world.

• [R1-S21] P19L9: Why 'introducing USW OR taking option 1'? As I understand it, option 1 means introducing the USW.

> Thank you for pointing out our error. The text now reads, "by introducing USW (see description of Option 1 in Section 2.1.7)."

• [R1-S22] P20L2: You mean of volume of the extracted non-renewable groundwater, not the volume of the aquifer itself, right? Please be concise here, because the paragraph sound like the latter.

You are correct. We have carefully edited this section to avoid potential confusion.

• [R1-S23] P21L17: Is this improvement really due to the six new schemes or rather due to the already existing global reservoirs and dam operations?

As stated above, the comparison between ALL and ORIG is shown in Supplemental Text S4. We found that ALL (the H08 model with the new schemes) outperformed ORIG (the original H08 model).

• [R1-S24] P22L20: I don't see how the economy (maybe apart from desalination part) and environmental aspects are accounted for in H08. From this paragraph I would expect that as a result of H08 simulations you could come up with kind of a cost-benefit analysis for different sources. Thus, this statement seems a bit strong for me. Please be concise about what exactly you can do with this model version.

In this section, we intended to emphasize the meaning of specifying water sources and their possible usage in further studies. Indeed, the acquisition of detailed global information about water sources must be the first step toward more advanced studies, including those investigating economic and environmental aspects.

Technical comments

• Tab S1: Please repeat the header for every table page

We will consult the editorial staff about this matter.

• Tab A1: It seems this data could be easily displayed in portrait format. Please only use sideways tables if really necessary.

We wished to avoid frequent page rotation because this often causes editorial problems. Most of the tables fit better in landscape orientation. We apologize for the inconvenience.

• All figures: You seem to prefer to use landscape format. However, it makes reading the paper more difficult especially in digital format and is not necessary for all of you figures (e.g. 1,4,7,8 and others). Please use portrait whenever possible.

Again, we apologize for the inconvenience.

• P8L12: You describe the general mechanics of your scheme. Better use present tense for such paragraphs.

Thank you for the suggestion; we have changed this section to the present tense.

• P8L23: Please revise this paragraph with regard to duplicates (estimation of extent areas) and unnecessary information (implementation difficult, still we implement edit). Your paper is quite long anyway, thus it should be shortened wherever possible.

We have removed the last sentence of this paragraph, per your suggestion.

• Fig 6: In the enlarged figure there is not much to see thanks to the text. As you refer only to a few selected basins in your results section, please remove the labels from most points and add them only to those you actually discuss.

Thank you for this suggestion; we have removed as many labels as possible from the enlarged figures.

• Fig 8: Consider shading the ocean area in a light grey for an easier overview. Also both regions are so close together that you could show them in one map.

We have now shaded the ocean area in gray.

• Fig 11: This is an awesome figure! You may think about changing the color of the region borders to avoid low visibility for patches where the background color matches the border color. Also it might be worthwhile to add small lines from the region to the circle to avoid any confusion about what belonging to what.

Thank you for this comment. We have changed the color of the regional frame to blue, per your suggestion. In consideration of your comments and those of Reviewer 2, we have labeled each region to avoid potential confusion.

• P18L25: Typo CAN→CNA (same typo in table S2)

Thank you; this error has been corrected.

• P19L8: This is not an inconsistency but just the difference

We prefer the term “inconsistency” because water supply and demand are always matched in reality (as in economic theory). From this perspective, water availability and requirement should be consistent.