Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-318-AC2, 2018

Hydrology and Earth System Sciences
Discussions
Open Access

Interactive comment on "Quantifying the impacts of human water use and climate variations on recent drying of Lake Urmia basin: the value of different sets of spaceborne and in-situ data for calibrating a hydrological model" *by* Seyed-Mohammad Hosseini-Moghari et al.

Seyed-Mohammad Hosseini-Moghari et al.

hosseini_sm@ut.ac.ir

Received and published: 17 December 2018

Referee #1: This manuscript uses WaterGAP Global Hydrology model to quantify the effects of human water use on inflow to Lake Urmia, lake water volume, and groundwater. The model was manually calibrated 4 for time using different observation data sets (remote sensing of irrigated area, monthly total water storage anomaly, in-situ observations of stream flow, and groundwater levels from 284 wells). Strengths of the work

Printer-friendly version

Discussion paper

include a focus on the pressing problem or Lake Urmia decline and identification of the effects on groundwater. With these strengths, there are also several issues that I feel need to be addressed to accept this manuscript for publication.

Response: We would like to thank you for the thorough consideration and critical comments that helped us improving the manuscript. We will try to do our best to consider all your recommendations in the revised version. Below, we have provided a point-by-point response to your comments.

Referee #1: Is the finding that humans affected lake decline new? There have been several recent studies that report this finding (Alborzi et al. 2018; Chaudhari et al. 2018; Shadkam et al. 2016) and some of these studied used the same model inputs as this work and also report groundwater changes. What is new in this work?

Response: We agree that different studies have been conducted on the Lake Urmia for quantifying the anthropogenic and climatic impacts on shrinking of the lake. But most of these studies focused on drought events, inflow into the lake or lake levels. None of these studies considered the effect of human water use on TWSA and groundwater storage (quantitatively), even distinguishing surface water from groundwater use. Also none of previous studies used such a diverse observation data that allow understanding of the different storage compartments of the basin. Furthermore, it should be noted the study does not only aim at quantifying the different impacts of climate and human water use but equally at understanding, for the example of the investigated basin, the value of different observational data types for calibrating a hydrological model and thus understanding dynamics and flows in a basin (see page 5, lines 30ff). Our study shows that for a comprehensive hydrological modeling that captures correctly the main dynamics in a basin, a multi-objective calibration is needed, and then previous studies missed some part of a comprehensive hydrological process modeling (e.g. the specific impact of groundwater pumping as compared to surface water use). Below please find the shortcomings of the studies you mentioned.

Printer-friendly version

Discussion paper

Alborzi et al. (2018):

1) Did not use a hydrological model.

2) Did not have any discussion on groundwater, lake storage, total water storage under natural condition.

Chaudhari et al. (2018):

1) Used global hydrological model (HiGW-MAT) without calibration for Lake Urmia basin.

2) In hydrological modeling the focus is only on streamflow and there is no discussion on groundwater, lake storage, total water storage under natural condition.

3) Did not take into account impacts of domestic and industrial water use.

4) Did not use in-situ data for irrigation water requirement and used FAO Penman Monteith method for estimating irrigation water requirement.

5) Did not use in-situ climatic data for estimating irrigation water requirement (used 6 h atmospheric reanalysis data provided by Japanese Meteorological Agency (JMA) Climate Data Assimilation System (JCDAS) as meteorological data)

Shadkam et al. (2016):

1) Only considered a single objective calibration of VIC model based on the inflow into the lake.

2) In hydrological modeling the focus is only on streamflow and there is no discussion on groundwater, lake storage, total water storage under natural condition.

3) All irrigation requirements need to be fulfilled by available surface water.

4) Did not take into account the impact of domestic and industrial water use.

As mentioned above, none of previous study provide a comprehensive modeling such

Printer-friendly version

Discussion paper

as our manuscript.

Referee #1: Is a global hydrologic model appropriate for a basin level analysis?

Response: Even though global hydrological models have a coarse spatial resolution, they contain a lot of different data (climate, physiographic, water use etc.) for which local information may not be available or of better quality. The study has shown that for the relatively large Lake Urmia basin, the global WGHM model can be informative after multi-observation calibration as after calibration the fit to those different types of observations is rather good.

Referee #1: The description of how the model simulates relevant processes is scant.

Response: Due to the fact that the length of manuscript already is long and also there are some other literatures that have described the model simulates relevant processes, we explained the model in a summarized form. In the revised version we will add the following sentence to section 2.1 (line 27):

"Lake water storage is simulated as the difference of precipitation on the lake, evapotranspiration, inflows and outflows. For end lakes like Lake Urmia, outflow is zero. The temporal variation of lake area, affecting precipitation and evapotranspiration from the lake, is simulated as a non-linear function of lake water storage. For more information on data and model algorithms used in WaterGAP please refer to Müller Schmied et al. (2014) and Döll et al. (2014)."

Referee #1: Given resonance times, how appropriate is the temporal spacing (daily) relevant to the spatial grid size?

Response: In our opinion, a daily time step fits well to simulation of water flows and storage at the 0.5° grid scale and is the usual time step in global hydrological modeling. Land surface models that also simulate energy flow require a smaller time step.

Referee #1: Is it computationally efficient to run a global model for 15 or 20 grid cells of interest? There needs to be a much stronger justification for why the modeling and calibration methods are the correct approaches to use to answer the motivating questions.

Response: It should be noted there is no need to run the model for whole globe. The model can be run for a specific basin, in the case of Lake Urmia for just 22 grid cells. Therefore, simulations are highly efficient. We will add this explanation to section 2.1. Using WaterGAP to answer the motivating questions is also efficient in that the model for the Lake Urmia basin was already set up at the beginning of the study as WaterGAP as a global model is ready for simulating any (large enough) basin around the world. We do not think that it is necessary to further explain in the text why the modeling and calibration approach is suitable, as e.g. section 2.1 describes that WaterGAP allows to consider a complex hydrological system (including surface water, groundwater, lake, human water use, etc).

Referee #1: There is a lot of focus in the text on the multiple calibration variants run with different input data sets. What was learned from this activity? How do those result effect Lake Urmia management?

Response: We have described the lesson we learned from this activity in details in section “3.2 What we learn from the calibration?”. We learned that there is no guarantee that a single objective calibration improves the model performance with respect to the simulation of other components of hydrological system. As a result, for defining the Lake restoration plans a comprehensive modeling framework like our study is needed. Specifically, quantification of return flow from irrigation is paramount for managing irrigation in the basin. We will add the following paragraph to the conclusion, as the last one:

“Our study has shown that management of the Lake Urmia basin should be based on a comprehensive assessment of all water storages and flows in the basin, including human water uses of groundwater and surface water. We recommend refining the estimated net abstractions from surface water and groundwater by a basin-wide spatially explicit quantification not only of water abstractions but also return flows to groundwater and surface water."

Referee #1: Also, what could one potentially learn from 4 model calibrations that use different calibration data and yield four different models?

Response: It is very common to use a single or two objective calibrations for calibrating a hydrological model. In this study, we have tried to understand which level of data can reveal that our modelling is holistic. Based on the results, for a holistic modeling, at least remote sensing, discharge and groundwater levels data are required. In addition, we have investigated with adding each in-situ data in calibration process how the model performance improved in simulating different water resources components.

Referee #1: What are the limitations of this study? The discussion of uncertainty in the results needs to go much deeper and be more specific. This uncertainty is real and likely plays a large role in the interpretation of the results.

Response: In the revised manuscript, we will add, as section 4, a short discussion.

"4. Discussion

Even after multi-objective calibration of a state-of-the-art comprehensive hydrological model, there remain many uncertainties that affect the accuracy of the model results. Like the results of all hydrological models, our results are affected by uncertainties in model input, model parameters and model structure. Model parameter uncertainty was reduced by the comprehensive multi-observation calibration, albeit conditioned on just one climate input data set and using just one model (instead of the state-of-the-art multi-model ensemble approach, compare www.isimip.org, last accessed: 14 Dec. 2018). Given the low spatial model resolution ($0.5^\circ \times 0.5^\circ$), the model results are only valid for the basin as a whole and results for individual grid cells are very uncertain. Also due to a lack of data at the basin scale, the hydrogeology of the basin was not taken into account in the model. Information on irrigated area in each grid cell was taken from a global data set of areas equipped for irrigation from groundwater and surface water (Siebert et al., 2010), which was adapted in this study by scaling it by basin-wide correction factors to better capture the temporal development of irrigation. Calibrated modeling results are also affected by uncertainties of the observation data. GRACE TWSA data are more reliable for larger (100,000 km2 (Landerer and Swenson, 2012)) areas than the basin area. Estimation of groundwater storage changes based on water level data for unevenly distributed wells is rather uncertain due to the unknown heterogeneities in the subsurface. Validation results, here the good fit of simulated to "observed" lake water volume decline, are be affected by a likely underestimation of the actual decline by the "observed" value derived from remote sensing of lake water level elevation and lake water area by Tourian et al. (2015) assuming a constant bathymetry. However, there was an increase in the elevation of the lake bottom due to sedimentation and salt precipitation (Shadkam et al., 2016) so that the "observed" water volume decline was likely lower than the actual one, and our model would underestimate the lake storage decline, too."

Referee #1: I found the writing difficult to follow in numerous places, particularly the results section. There are lots of acronyms, run-on sentences, and text that digresses from the section headers or topic sentences of paragraphs. The writing here made it difficult for me to see the main results and findings of the work.

Response: We will revise whole of manuscript and reformulate the results section for the revised version.

Referee #1: pp. 2-4. The first three figures recount results from prior work. I would much prefer to see figures and tables focus on new insights gained from the work. For example, new figures that show uncertainties.

Response: In our opinion these figures are needed to inform the reader about the story that happened at Lake Urmia basin. Also, only Figure 2 is taken from previous studies. Due to the length of manuscript, we prefer not to add new figures.

Referee #1: p. 3, line 5. I think "somewhat recovered" is overstated. Hard to tell from Figure 3. Maybe stabilized.

Response: We agree with you regarding the extent of lake stabilized but if you consider the south part of the lake a minor increase in the lake storage is seen. So, we will revise the sentence as follow:

"After 2015, lake extent has stabilized (Fig. 3) and lake storage has slightly recovered".

Referee #1: p. 3, line 18. Is the value -11.2 mm/yr correct? It seems incredibly small. In The Hashemite Kingdom of Jordan, drawdowns are 1+ m/yr, in numerous wells. In the U.S., we talk about drawdowns of ft/year.

Response: The value taken from the study of Forootan et al. (2014) refers to groundwater storage, not to a drawdown of the groundwater table. And it is the average loss over the whole basin and not a drawdown in an extraction well which of course can be much higher than an average drawdown over the whole basin. Change in groundwater storage can be calculated by multiplying change in groundwater level with the specific yield.

Referee #1: p. 6, line 5. Only the anomalies? Or at all time periods? If the former, please explain what is meant by "anomaly", how determined, and why anomaly is the appropriate frame to discuss. I would want to calibrate a model across a range of conditions some of which might include anomalies.

Response: The model has no information about the real bathymetry and initial value of lake storage; it can only simulate changes as compared to some initial condition. As a result, we can compare only lake storage change or lake storage anomalies. Lake storage anomaly at a given time is equal to the lake storage at that time minus the long-term average of lake storage. In this study, as mentioned in p. 20 line 24, anomalies were calculated with respect to the mean lake storage during 2004-2009 (baseline period used for the provided GRACE data).

Referee #1: p. 6, lines 18-28. I'm not familiar with WaterGAP. How does this model actually work? Explain.

Response: While the WaterGAP description in section 2.1 is brief, the reader is referred to publications that provide more information on the model. However, as written above, we plan to extend the WaterGAP description slightly, in particular with respect of lake modeling.

Referee #1: p. 7. What is total water storage anomaly (TWSA)? This term seems rather central to the paper. Please explain.

Response: Total water storage (TWS) is amount of water which is stored in different components of the continents, e.g. as follows (Scanlon et al., 2018):

TWS= SnWS+CWS+SWS+SMS+GWS

where SnWS is snow water storage, CWS is canopy water storage, SWS is surface water storage, SMS is soil moisture storage, and GWS is groundwater storage. Neither hydrological models nor GRACE can compute the total amount of stored water. They can only compute varisations according to a temporal average. Therefore, TWS anomalies (TWSA) are evaluated, defined as TWS(t) – mean (TWS).

We will insert the following sentence as the second sentence of the GRACE section (p. 7, line 9):

"TWSA describes the total amount of water stored on the continents, including water storage in surface water bodies, groundwater and soil, as compared to the mean value of total water storage over a reference period."

Referee #1: p. 8, lines 13-18. This method of applying (1- return flow multipliers) to the abstractions to estimate consumptive use assumes that water is used by only one water user. Is this a realistic assumption? If the return flow is used by another agricultural user and then again by a 3rd or 4th user, the basin-wide consumptive use fraction will be much different than the values reported. The large grid size magnifies this error. Table 1. How sensitive are study results to the values in this table?

Response: In our opinion due to the fact that most return flow returns to groundwater, there is no concern in this regard. On the other hand, in arid area e.g. Urmia basin the return flow to surface water in each irrigation is not too much that can be used by another user. Anyway, we do not know exactly how the authors of the values determined them. However, the study results are not sensitive to the independent return flow estimates that were used only in one variant (RS_Q_GW_NA). We write in on p. 19, line 13:

"Consideration of regional estimates of human water withdrawals in a specific year as well as regional estimates of return flow fractions in variant RS_Q_GW_NA does not improve the fit to observations significantly and only leads to slight parameter adjustments. This indicates a reasonable simulation of per hectare water consumption for irrigation by the WaterGAP model."

Referee #1: p. 10. Lines 5-15. So the correction factors are needed because WaterGAP does not get the underlying physical hydrology correct? The correction is linear? Is the process causing the error also linear?

Response: Irrigated area in the standard version of WaterGAP is constant during the period of investigation. The correction factor based on MODIS remote sensing time series adjusts the WaterGAP value in each year homogeneously across the basin. Standard WaterGAP irrigation area is multiplied by the correction factor.

Referee #1: p. 11, line 1. Which parameters were varied to calibrate this model?

Response: These parameters are presented in Table 3.

Referee #1: p. 11, line 4. What is meant by optimal fit?

Response: Optimal fit means the best possible match between observed and simulated time series. Performance of the fitted model is quantified by three different performance indicators in Table 4.

Referee #1: p. 15, line 2 and Table 3. Shouldn't these parameter values be the same across all the model variants? What is physically changing in the system that these parameter values would change across the model variants?

Response: These parameters should not be the same because model calibration means change in model parameters to achieve the best fit to observations. As different observations are used in each variant, the values of calibration parameters are different, too. There is no physical change in the basin but different parameter values indicate that model parameterization cannot be uniquely determined. Calibration just to streamflow observations, as is usually done in hydrological modeling, does not assure a correct simulation of water storage changes, for example. In our study, the parameters optimized by just using remote sensing information for model calibration (variant RS) lead to an unsatisfactory simulation of inflow into the lake (see Table 4).

Referee #1: p. 16, lines 2-3. There could be a net groundwater abstraction but still areas where there is recharge. Is this an issue of coarse spatial resolution?

Response: Yes, there could be areas of recharge due to irrigation with surface water. However, the study just analyzes the mean behavior over the whole basin, based on the results in 22 grid cells.

Referee #1: p. 18, lines 1-10. The discussion of uncertainty here is missing a fundamental point. Calibration cannot help if the model structure is uncertain (or in error) or the temporal or spatial scaling of the model is mismatched to the modeled parameters of interest. This discussion also heads in a different direction than "what do we learn from the calibrations?" The text never explains what was learned. What was learned? Please discuss.

Response: Regarding uncertainty of model structure, we plan to refer to multi-model ensembles in the new discussion section (see above). In this section we do explain what we learned from our calibration exercise or rather from adding more types of observations. It is outside the scope of the paper to discuss the impact of temporal or spatial scales on model results as we did not investigate this (or e.g. uncertainty due to the applied climate forcing). Thus, when we discuss different calibration variants we discuss how uncertainty can be reduced by additional observational data types.

We believe that in the “What we learn from the calibration?” section, the key findings were reported. The most important of them reveals that a single objective/observation calibration cannot capture hydrological dynamics and there is no guarantee a well-simulated model based on a tuned variable can properly simulate other components of the model. As a result, for a general statement about water resources in a given region a multi-objective calibration is required.

Referee #1: p. 18, lines 11-20. These statements are better placed in the introduction to justify the use of global hydrologic models. Still, why is a global model the appropriate choice when the domain of study is limited to one hydrologic basin (Urmia)?

Response: We prefer to leave this paragraph in the results section to clarify the context of the calibration exercise that is on the one hand done to efficiently analyze the specific situation in the Lake Urmia basin but also to evaluate the value of calibrating global hydrological models against multiple observation types. The reasons for using a global model have been given above.

Referee #1: p. 19, line 22. What beta?

Response: At the beginning of the sentence, there is the reference to Eq. 2. Beta adjusts the net abstraction from groundwater.

Referee #1: p. 19, line 30. “much less overestimated” means what?

Response: Its means that although the model variant still overestimates inflow to the lake, the degree of overestimation has decreased strongly.

Referee #1: p. 19, line 31-32. This doesn’t make sense to me. How come it is ok to change the parameter in one model variant but not others?

Response: It is another thing we can learned from the calibration. As shown in Figure 6, in each variant the model is calibrated against different observations. For example, in the first variant the WGHM evaluated based on TWSA and in the second one WGHM evaluated based on TWSA and inflow into the lake. In the first variant we did not need to change the parameters which have no effect on simulated TWSA. But in the second variant we have to calibrate model against both TWSA and inflow into the lake. So, in this variant the parameters which have an effect on inflow into the lake also should be adjusted. It means that when we have a single objective calibration it is possible we reach to appropriate results via different combination of parameters values. When we add more and more objectives or observational data types in the calibration process, the number of parameters which should be changed.

Referee #1: p. 20, lines 1-5. I would expect to see better calibration with more observational data (i.e., stream flows and lake levels).

Response: We agree with you. Our results agree with your expectations, and this is what we express in the sentence starting in line 4: "Still, calibration to both observational data types leads to the best simulation of both annual lake inflow and lake volume anomalies."

Referee #1: p. 20, line 22. I'm confused. The scenario "with reservoirs but without human water use" does not fit either of the two scenarios described in the prior sentence.

Response: WHGM has the capability to assess the effect of dam building on water resources without considering human water use from reservoirs. Our results showed that the results with and without reservoirs has only 2% difference which indicated insignificant effect of reservoirs on water resources over the basin. So the run with or without reservoirs effect can be considered the same. As a result, we can consider the run without human water use as WGHM run under natural condition.

Referee #1: p. 20, line 24. What is meant by anomalies? This term has still not been defined.

Response: Anomalies is the difference from an average, or baseline for example in this study GRACE data are total water storage anomalies based average of its observation between 2004-2009. Will be added to the text when introducing GRACE (see response above).

Referee #1: p. 21, line 17, "The lower lake water loss..." What are the loss terms besides evaporation? How are these other loss terms smaller when inflow is larger? Explain.

Response: Evaporation is the only loss term of the lake. There was less decrease in lake storage due to more inflow into the lake. We will revise the sentence as:

"The smaller decreasing trend for lake water volume under naturalized conditions is clearly caused by more inflow into the lake, even though lake evaporation is somewhat higher under naturalized inflow conditions due to the larger lake extent."

Referee #1: p. 21, lines 20-23. I don't follow this explanation. There are too many NAs in this sentence. What causes the difference between the naturalized and anthropogenic scenarios?

Response: In naturalized runs, abstraction from the surface water or groundwater are assumed to be zero, while and water flows and storages vary only due to climate variations while in the anthropogenic scenario we simulate both the effect of water abstraction by human and of climate variations. We will delete the sentence in lines 20-22.

Referee #1: p. 21, lines 25-28. I don't follow. What is the connection between the first part of the sentence and the second part?

Response: We wanted to state that inflow was less than the minimum environmental water requirements (3,085Âů106 m3) and that therefore a loss of lake water volume is expected. So, the first part of the sentence indicated under anthropogenic situation the inflow into the lake since 2008 never has reached to minimum environmental water requirements. The second part is related to naturalized condition which showed that under naturalized condition only in 2008 and 2009 the inflow into the lake were less than minimum environmental water requirements. We intend to reformulate the first sentence as follows:

“Since 2008, inflow has never reached 3,085Âů106 m3/yr, the value estimated to be the minimum environmental water requirements that compensates the amount of annual evaporation from of the lake surface (Abbaspour and Nazaridoust, 2007). Therefore, a decrease of lake water storage can be expected for the best estimate of WaterGAP of 2,639Âů106 m3/yr.

Referee #1: p. 21, lines 28-32. Is a run-on sentence.

Response: We agree with you. We will revise it.

Referee #1: p. 21, lines 32 – 21. Put these ratios in context. What is desirable? Undesirable? What has implications for lake health? What values are acceptable?

Response: We think that it does not make sense to talk about acceptable values. From the perspective of lake health, a higher value may be more desirable, or rather that the values under anthropogenic conditions become closer to the values under naturalized conditions. However, we do not think it is necessary to state this explicitly in this scientific publication, it will be clear to the reader even without stating this explicitly.

Referee #1: p. 23, line 15. How can water storage be negative?

Response: Total groundwater storage is not computed in WaterGAP, only storage relative to a storage that occurs when the heads in surface water and groundwater are the same. Negative values of groundwater storage computed by WaterGAP indicate that net abstractions from groundwater are larger than natural groundwater recharge, while baseflow is zero. Groundwater levels can be assumed to have dropped below the surface water heads. In WaterGAP, groundwater recharge from rivers is not taken into account.

Referee #1: p. 23, line 18. This is an interesting result. It needs to be much more strongly emphasized. These cells are the locations where groundwater declines and there could be problems.

Response: As WaterGAP does not include reliable high-resolution information on irrigated areas and groundwater use infrastructure, the computed cell-specific net abstractions from groundwater are highly uncertain. This is why our study focuses on basin averages, and cell-specific results are less prominently shown.

Referee #1: p. 23, lines 21-23. This qualification and limitation seems rather important. Why should the model results be trusted or used if the model does not get groundwater storage correct?

Response: No model is perfect, and calibration as done here is a way to compensate for a lack of process accuracy. Due to calibration, we do get groundwater storage (more or less) right.

Referee #1: p. 23, lines 23-25. Run-on sentence. What is meant by the clause with maximum?

Response: As mentioned in p. 23 line 22, WaterGAP cannot simulate a possible drop of the groundwater table below the surface water level in the absence of groundwater abstractions, and groundwater storage might in reality have been lower before the start of groundwater abstraction than simulated in the naturalized run. Thus, contribution of human water use to groundwater storage decline might therefore be overestimated.

Referee #1: p. 24, lines 18-19. Is this result surprising? More calibration data means a better fit model. How does this result improve understanding of the Lake Urmia system?

Response: There are a few hydrological modelling studies on Lake Urmia basin but all of them (except Chaudhari et al. (2018) who has not implemented calibration at all) have only a singly observational data type/objective for calibration. So as first multi objective calibration study of Lake Urmia basin we show how a single objective calibration does not have the proper capability for a comprehensive modeling in the basin. As a result, this study provides a unique information for understanding the Lake Urmia basin system not only the Lake Urmia which does not reported in any previous studies. In addition, if the model structure or the input data are wrong, it may not be possible to improve the simulation of an increasing number of observational data type that are used for calibration. Some trade-off may occur; for example, total water storage anomaly simulation may decrease but streamflow performance may increase if streamflow is added as a second calibration data type, in addition to total water storage anomaly.

Referee #1: p. 24, lines 25-28. Is this finding new? If so how? I feel the Urmia Lake Recovery Program has been working under the assumption that agricultural water use was a large contributor to lake decline and that they have been taking steps in recent years to address.

Response: Regarding lake inflow, Shadkam et al. (2016) reported similar results as our study. However, results regarding TWSA, lake storage and groundwater storage are the new findings.

Yes, Urmia Lake Restoration Program (ULPR) is working under the assumption that agricultural water use was a large contributor to lake decline. ULPR cannot manage climate variations or changes, so they certainly need to focus on management water use over the basin, if according to our study, human impact is significant.

Referee #1: p. 24, line 29. 90% of what?

Response: 90% of groundwater storage losses. We will reformulate the sentence as follows:

“90% of groundwater storage loss is estimated to be caused by human water use but this value may be somewhat overestimated by WGHM because climate-driven loss under naturalized conditions may be underestimated due to the simplified representation of groundwater-surface water exchanges in the model."

Referee #1: p. 24, lines 19-24. I disagree. There are lots of other similar systems in the world – Great Salt Lake, Owens Lake, Dead Sea, Ural Sea, etc. each satisfy the first two criteria listed. What of these results is generalizable?

Response: Thank you for pointing this out. We do not know what is generalizable as we have not done this calibration exercise in other basins. We intend to add one sentence at the end of the paragraph that provides a recommendation based on our study.

"In basins with large lakes, and in particular with end lakes, remotely sensed time series on lake area and the elevation of the lake water table should be used to estimate time series of lake water storage as these observational data can be expected to be of high value for understanding the freshwater system by hydrological model calibration. Groundwater storage cannot be observed from space but relies on in-situ observations on groundwater heads in wells but, as in the case of Lake Urmia basin, but such data may be crucial for a correct understanding of the freshwater system."

Referee #1: p. 24, lines 31-34. How do the model results inform the 2014-2017 trends? Also, how can climate change be constrained in this basin? Explain.

Response: Unfortunately, almost all the input datasets are not available from 2014 onwards. Thus we did not have the model results for recent years. With "constraining climate change" we meant the global reduction of greenhouse gas emissions, nothing at the basin scale. We will replace the term "constrained" by "globally mitigated".

Referee #1: Figure 9. What is being shown in panels A, B, and D? The y-axis labels were mentioned in the text but never explained.

Response: All panels have been explained in following pages and lines: Figure 9a on p. 21 lines 1-5. Figure 9b on p. 21 lines 15-16. Figure 9d on p. 22 lines 1-14.

Referee #1: Figure A1a. The color scheme makes it difficult to differentiate grid cells.

Use only three colors to differentiate the 3 types of storage. How can storage volume be negative?

Response: We will change the colors according to your suggestion. Total groundwater storage is not computed in WaterGAP, only storage relative to a storage that occurs when the heads in surface water and groundwater are the same. Without NAg, storage can therefore never become zero or less than zero. If NAg becomes larger than groundwater recharge, storage can obtain negative values.

Referee #1: Data availability. I don't follow. If the authors do not have permission to share the data, then how can they share by author request? The HydroSat site underwritten by the University of Stuttgart is neat. What is the original source data for Urmia? Also, there is no water storage anomaly data for Urmia.

Response: We have no permission to share data except by personal request for research purposes. We have indicated the source of all data in the manuscript. The data for Lake Urmia was obtained from various sources. Regarding HydroSat site, website you can download original data for lake level and extend with related reference which is Tourian et al. (2015) after registration. Also about water storage anomaly data for Urmia, for the lake water storage anomaly, it should be noted that anomalies can be calculated easily by subtracting the mean lake water storage in baseline (2004-2009) from the lake water storage time series.

The section on data availability will be reformulated as follows:

"In-situ data from "Iran Water Resources Management Company" including groundwater levels, precipitation and temperature publicly are available upon request from the corresponding author. GRACE data is available through http://www2.csr.utexas.edu/grace/RL05_mascons.html (last accessed: 17 Jul. 2018). Lake water surface extents and water levels are available at http://hydrosat.gis.uni-stuttgart.de/php/index.php (last accessed: 17 Jul. 2018). All simulation results are available in the supplement."

****

Thank you very much again for your time and for providing valuable comments.

****

References

Alborzi, et al.: Climate-informed environmental inflows to revive a drying lake facing meteorological and anthropogenic droughts, Environmental Research Letters, 13(8), 084010, doi:10.1088/1748-9326/aad246 ,2018.

Chaudhari, S., Felfelani, F., Shin, S., and Pokhrel, Y.: Climate and anthropogenic contributions to the desiccation of the second largest saline lake in the twentieth century, Journal of Hydrology, 560, 342-353, doi:10.1016/j.jhydrol.2018.03.034, 2018.

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T., and Eicker, A.: GlobalâĂŘscale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites, Water Resources Research, 50, 5698-5720, doi:10.1002/2014WR015595, 2014.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M., and Döll, P.: Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, Hydrology and Earth System Sciences, 18, 3511-3538, doi:10.5194/hess-18-3511-2014, 2014.

Siebert, S., Burke, J., Faures, J.-M., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T.: Groundwater use for irrigation-a global inventory, Hydrology and Earth System Sciences, 14, 1863-1880, doi:10.5194/hess-14-1863-2010, 2010.

Landerer, F. W., and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water resources research, 48, doi:10.1029/2011WR011453, 2012.

Scanlon, B. R. et al.: Global models underestimate large decadal declining and rising

Printer-friendly version

Discussion paper

water storage trends relative to GRACE satellite data, Proceedings of the National Academy of Sciences, 115(6), 1080-1089, doi:10.1073/pnas.1704665115, 2018.

Shadkam, S., Ludwig, F., van Oel, P., Kirmit, Ç., and Kabat, P.: Impacts of climate change and water resources development on the declining inflow into Iran's Urmia Lake, Journal of Great Lakes Research, 42, 942-952, doi:10.1016/j.jglr.2016.07.033, 2016.

Tourian, M. J., Elmi, O., Chen, Q., Devaraju, B., Roohi, S., and Sneeuw, N.: A space-borne multisensor approach to monitor the desiccation of Lake Urmia in Iran, Remote Sensing of Environment, 156, 349-360, doi:10.1016/j.rse.2014.10.006, 2015.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-318, 2018.