Dear Editor and Reviewers,

We have revised the manuscript according to the insightful comments provided by the editor and reviewers. All recommendations have been addressed in the revised manuscript. We would like to thank you for the thorough consideration and critical comments that helped us to improve our manuscript. In the following, we have provided a point-by-point response (in blue) to the comments.

Editor Comments

Editor Comment: Both referees agree the paper is improved a lot and methods - concepts are well explained. However, both referees are also critical on the interpretation of your own results and especially on the discussions currently in literature on the Urmia lake drying causes. I do agree with the referees that interpretations and discussion should be extensive and complete and take into account various views.

Please respond in detail to the referees' reports and submit a revised version of the paper including a track changes version. I will review the paper and your responses before considering publication.

Response: Thank you very much for giving us this opportunity to revise the manuscript. We have revised the manuscript based on your and referee's comments and added a more extensive discussion. We hope the changes have made the manuscript suitable for publication, and we look forward to your response.

A point-by-point response to the Referee #2

Referee #2: As I asked in the previous version, several revisions are made. I think the revised article is much better and a lot of time and effort are given to the study. As I underlined before, the drawn conclusions are over-exaggerated. For instance, the reality that your model can not be calibrated using groundwater data can never be interpreted as there is no interaction between groundwater and the lake. In addition, in P 18, Line -5-10, it is better to be aware that, no modeling was done by ULRP (2015c) and, Danesh-Yazdi and Ataie-Ashtiani (2019) but they emphasized that "it is not likely" to have a relation between. Also based on the studies of Ashraf et al. (2017), and Rodell et al. (2018) there is an undeniable relationship between groundwater and lake water in Lake Urmia basin. So I recommend adding this publication as well to your discussion as the cases which concluded that there is a relationship between groundwater and lake.

Response*:* Firstly, thank you very much for your time to review our manuscript. We are not sure whether we understood your point about over-exaggerated correctly. However, it seems that it related to the contribution of climatic variations on lake shrinkage. It should be noted that all studies that conducted hydrological modelling over lake Urmia support our results (except Chaudhari et al. (2019) who used uncalibrated model for this purpose). Shadkam et al. (2016), who calibrated the VIC model for Urmia Basin with a longer period (1960-2010), stated: "Our results show that annual inflow to Urmia Lake has dropped by 48% over the study period. About three-fifths of this change was caused by climate change, and about two-fifths was caused by water resources development." Also, Farokhnia et al. (2018) who developed and calibrated SWAT model for separating anthropogenic factor and climate impact on Lake Urmia during 22 years period ending in 2009 stated: "the cumulative effect of climatic and human-induced changes in reducing the inflow into Urmia Lake was almost equal, but due to the intensification of climate variability in the second half of this period, the effect of climatic factors was dominant in the rapid negative trend of lake water level." These statements show that there are some other research that emphasize the contribution of the climatic factors. Therefore, we believe that although all models suffer from uncertainties, these uncertainties cannot change the overall results.

About the interaction between groundwater and the lake, you are right that there is no modelling in ULRP (2015c), but field tests by the Japan International Cooperation Agency (JICA) Japanese company confirmed their assumption. We could not find any statement in Rodell et al. (2018) about the relationship between groundwater and lake across the Lake Urmia basin. Of course, there are some sentences about groundwater and lake in Urumqi (in China), not Urmia!

We have revised the part related to the relationship between groundwater and lake by considering your comment as follows:

"It is worth mentioning that WGHM as a hydrological model that does not include a gradient-based groundwater model has some limitations for studying groundwater-lake water flows. We attempted to calibrate WGHM under the assumption that there are direct water flows between lake and groundwater. Under this assumption, the seasonality of the groundwater storage was strongly misrepresented. Therefore, as accepted by ULRP (2015c), we assumed there is no direct flow between the lake and groundwater. This is consistent with Danesh-Yazdi and Ataie-Ashtiani (2019) who stated that a significant water exchange between the lake and groundwater is unlikely. Also, Amiri et al. (2016) based on isotope and chemical tracer analyses rejected any significant relationship between lake. However, some studies, e.g. Ashraf et al. (2017) and Vaheddoost and Aksoy (2018), stated the opposite. In conclusion, the results of this study support the idea that there are no significant direct interactions between lake and groundwater in the Lake Urmia basin."

Referee #2: It is better to use "water level" instead of "Water table" for the case of lake water level.

Response: The "lake water table" was replaced with the "lake water level" in the whole manuscript.

Referee #2: The statistics obtained for your models are not satisfactory especially when Genetic Programming or optimization is addressed.

Response*:* We agree with you if, for example, we have done a single calibration against Q when all observed data be used (i.e., variant RS_Q_QW_NA), the KGE would be larger than the current KGE. However, for ensemble outputs from a multi-objective calibration the KGE value of 0.86, 0.82, 0.85, and 0.89 for simulating TWSA, Q, GWSA, and LV, respectively, are quite good as Pechlivanidis et al. (2014) stated that a KGE value of greater than 0.75 is acceptable for the calibration of a hydrological model.

Referee #3: First and foremost, I would like to sincerely apologize for my late reviews – due to family reasons. I have now reviewed the new manuscripts and authors' responses to previous reviews. By and large, the authors have addressed most of the comments, updated the modelling, condensed the manuscript and improved the use of language, and improved the discussion of results (particularly adding a new section on the limitations of the study). So, I would like to thank authors for their efforts in this round of the review.

Response: Thank you very much for your time to review the revised version of our manuscript. It is our pleasure that the revised version satisfied you. We have considered all your points in the new revised version. Below, we have provided a response to your comments in detail.

Referee #3: I would like to also thank authors for dedicating a section on limitations (section 4.3). Not only it is good practice for scientific studies, but also it directs readers to design future studies to address past limitations. That said, a few major points are missing in the discussion of limitations:

• The study period is limited and do not include some of the most important years of the lake (and the basin) old/recent history and evolution, e.g. recent changes in the lake since 2013 (beyond the study period of this manuscript). Compared to most hydro-climatology studies of the lake, this study is based on a very short period (2003-2013), and hence generalizing its results beyond this domain is difficult, particularly due to the non-stationarity of the lake system. I expanded on this in section 3 of this review.

• Role of dam constructions, groundwater withdrawals and its hydrological connectivity to the lake, and seasonal variations of the flow overlooked in the model calibration.

Response: We respond to this comment in later parts where you describe each limitation separately.

Referee #3: Moreover, section 4.2 "Comparison to human vs. climatic contribution as determined in previous studies" comes short of providing an adequate and accurate characterization of the ongoing debate within the literature:

• I acknowledge that the Lake Urmia desiccation has been an ongoing contested debate, i.e. whether the main driver of drying is management-related and human activities or climatic. This very question is indeed the crux of the matter, and hence it is in the best interest of both authors and readers to be more rigorous on this discussion. As opposed to several previous studies, this study puts more weight on the climatic drivers of the lake drying. However, the authors (in section 4.2) misrepresented or overlooked some of those studies which argued for human activities over climatic drivers. Regardless of my personal position in this debate, authors are unduly framing the results and merits of the previous studies to justify their own side of the argument. Further, they have failed to discuss a few important studies on the lake. I demonstrate this in section 2 of this review. To help the authors, I discussed several points in details. I have also suggested few additional references and edits throughout. Therefore, in my opinion section 4.2 of the manuscript is inadequate and must be improved.

Response: Thank you very much for your detail points. We have expanded the discussion to apply your comments. The changes that have been made based on your points are presented in your following comments separately.

Referee #3: Section 4.2 “human vs. climatic contribution”

The discussion is not elaborative, lacks a clear discussion line, and previous studies are not appropriately represented/discussed. Here I discuss a few points in this section as an example, to help authors better discuss this important point.

Response: Thank you very much for the help to improve the discussion part, we have revised the manuscript accordingly.

Referee #3: Page 18 line 17-21: “Chaudhari et al. (2018) concluded that human-induced changes accounted for 86% of the lake volume decline during 1995-2010, while we determined values of 39-43% for 2003- 2013. According to our study, human water use was the reason for 39-45% inflow reduction into the lake during 2003-2013 which is very similar to the values of Shadkam et al. (2016) for the years 2003-2009 (comp. their Fig. 8). Discrepancies are likely due to different analysis methods but different analysis periods and conceptualizations make a direct comparison of the estimated contributions difficult.” Re the underlined part: it is a general statement and not good enough to simply overlook the details leading to these differences. It is essential to discuss in more detail what are the main differences between these studies e.g. in terms of data type, analysis approaches, fundamental assumptions, etc. For instance, Chaudhari et al. (2018) studied a considerably longer period. They also studied the land use changes in detail: over 1987-2016 showed ~98% and ~180% increase in agricultural lands and urban areas, respectively. They accounted for human impact during 1995-2010 (based on simulation of streamflow into the lake). Various studies identified two distinct periods of pre- and post-change in the lake dynamics, e.g. Khazaei et al. (2019) identified year 2000 as the change point and Fazel et al. (2017) identified year 2001. Given that, studies such as Chaudhari et al. (2018) take into account a wider range of the non-stationarity of the lake than the present study where only a part of the post-change period is investigated. It is plausible to expect that if your model was successfully calibrated over a longer period including years prior to 2000, it would have lead to different results.

Response: Firstly, it should be noted that some other studies like Shadkam et al. (2016) who calibrated the VIC model for Urmia Basin with a longer period (1960-2010) stated: “Our results show that annual inflow to Urmia Lake has dropped by 48% over the study period. About three-fifths of this change was caused by climate change, and about two-fifths was caused by water resources development.”. Also, Farokhnia et al. (2018) who developed and calibrated SWAT model for separating anthropogenic factor and climate impact on Lake Urmia during 22 years period ending in 2009 stated: “the cumulative effect of climatic and human-induced changes in reducing the inflow into Urmia Lake was almost equal, but due to the intensification of climate variability in the second half of this period, the effect of climatic factors was dominant in the rapid negative trend of lake water level.". These studies that include the change point year (2000 or 2001) support our results. Therefore, we believe our results are valid for the studied period without considering the change point year(s).

As we mentioned in the manuscript, Chaudhari et al. (2018) used an uncalibrated global hydrological model. Although, the model performance was assessed against the inflow into the lake (inflow data from Hassanzade et al. (2012)). The used data for assessing the model is quite far from observations. For instance, simulated annual inflow into the lake was estimated to be $3{,}700 \cdot 10^6$ m^3 in 2003 (their Fig. 8), while observed inflow was much higher, $5{,}835 \cdot 10^6$ m^3. In 2009, observed inflow, with $1{,}036 \cdot 10^6$ m^3, was only half of the simulated one. Therefore, the source of the main differences is using the uncalibrated model in Chaudhari et al. (2018). We believe that considering a longer period and land-use change based on an uncalibrated model can not challenge our results while Farokhnia et al. (2018), who also considered the land-use change support our findings. We modified this part as follows:

"Discrepancies are likely due to different analysis methods, but different analysis periods and conceptualizations make a direct comparison of the estimated contributions difficult. Chaudhari et al. (2018) performed a comprehensive hydrological modelling of Lake Urmia basin. They also studied the land use changes in detail over 1987-2016 and determined a ~98% and ~180% increase in agricultural lands and urban areas, respectively. However, their uncalibrated global hydrological model that represented the basin by 5-6 cells only was not able to simulate well the flows and storages in the basin. For example, simulated annual inflow into the lake was estimated to be $3{,}700 \cdot 10^6$ m^3 in 2003 (their Fig. 8) while observed inflow was much higher, $5{,}835 \cdot 10^6$ m^3. In 2009, observed inflow, with $1{,}036 \cdot 10^6$ m^3, was only half of the simulated one. Therefore, the very high human contribution to the lake volume decline of 86% determined by Chaudhari et al. (2018) may arise from the poor performance of the uncalibrated model. In addition, Chaudhari et al. (2018) studied a considerably longer period, i.e., 1995-2010, that includes the change point of lake dynamic (the year 2000 based on Khazaei et al. (2019) and 2001 based on Fazel et al. (2017)). Although including years prior to 2000 might be lead to different results, some other studies like Shadkam et al. (2016) and Farokhnia et al. (2018), who their modelling included years 2000 and 2001, support the results of the current study. Shadkam et al. (2016) stated that climate change was responsible for three-fifths of inflow reduction into the lake, and the rest was caused by water resources development between 1995-2010. Also, Farokhnia et al. (2018) showed that during a 22 years period ending in 2009, the effect of anthropogenic and climatic factors in reducing the inflow into Lake Urmia was almost equal."

Referee #3: "While Ghale et al. (2018) seem to support the results of Chaudhari et al. (2018) as they state that 80% of drying of Lake Urmia is due to anthropogenic impacts during 1998-2010, their statistical analysis assumes that lake inflow from rivers can be considered to reflect "anthropogenic impacts" while precipitation and evaporation reflect climatic variation. <u>However, inflow is in reality also affected by climatic variations</u>." Your argument here is incomplete, as the impact of climate vs. human activities on river networks is different for headwaters and lower river reaches. Fazel et al. (2017) investigated this in detail, analyzing the flow regime changes across the lake basin (57 flow gauging stations) over the period 1949-2013 (perhaps the longest record of the basin flow studied so far). Their study showed that while "flow regime in river headwaters appeared to be dominated by natural forces", "the reduction in river flow magnitude increased from headwaters to downstream reaches for all rivers" due to dam river regulations and dam constructions. They further argued that "Changes in river flow in the period 1965–2013 cannot be explained by climate change, the effects of which occur much more slowly than those of land use change in the region". They concluded that "The results showed that irrigation was by far the main driving force for river flow regime changes in the lake basin. All stations close to the lake and on adjacent plains showed significantly higher impacts of land use change than headwaters. As headwaters are relatively unaffected by agriculture, the non-significant changes observed in headwater flow regimes indicate a minor effect of climate change on river flows in the region".

Response*:* Fistly, please note that our paper does not make any claims about the impact of climate change, but abouth the impacts of climatic variations. Our results do not deny the impact of human water use. We considered human water use as a significant factor, while the climatic factors are also significant. Therefore, although our result is not entirely consistent with Fazel et al. (2017), it is not in conflict with it either. Contrary to your statement, the study period of Fazel et al. (2017) is filled with missing data (please see Figure 7 in their study), and the period with completed data is considerably shorter than 1949-2013. They considered the post-change period, i.e., 2001-2013, compared with 1965–1977 as the pre-change period for downstream stations, while for headwater stations, the pre-change period was 1989–2001 or 1965–1977 due to availability of data (please see Figures 4 and 5 in their study). Therefore their analysis suffers from some errors caused by the different comparison periods. However, among the 6 stations that they plotted in Figure 5 (for headwater), three of them also have a reduction in inflow. We believe that comparing a period of 12 years under this condition cannot be a useful approach for quantifying the effect of climate change over the basin. However, more flow reduction in downstream stations due to agricultural development in upstream expected in any basin. Our study also confirmed that human activities have significantly reduced the inflow in the downstream of the basin, i.e., inflow into the lake. We modified this part as follows:

"While Ghale et al. (2018) seem to support the results of Chaudhari et al. (2018) as they state that 80% of drying of Lake Urmia is due to anthropogenic impacts during 1998-2010, their statistical analysis assumes that lake inflow from rivers can be considered to reflect "anthropogenic impacts" while precipitation and evaporation reflect climatic variation. However, although inflow into the lake is surely affected by human water use in upstream, also affected by climate variations over the basin."

Referee #3: "Using a statistical change point analysis and without modelling, Khazaei et al. (2019) stated that given the stable conditions of precipitation and temperature, climatic changes could not explain the dramatic decline of the lake level; however, they did not use in-situ data (except lake water level data) for their analysis" Study by Khazaei et al. (2019) is more than a simple statistical change point analysis, they estimated the land use change (particularly vegetation dynamics and its associated hydrological loss in terms of evapotranspiration) and trends of various hydro-climatic variables across various time scales. While their study surely has its own limitations, lack of modeling and use of insitu data are not the major limitations – let alone this is too generic for a scientific criticism. One of the major limitations of their work, for instance, is that they did not account for the role of groundwater dynamics in their analysis.

Response: We agree with you that the use of global datasets is not a limitation for a study. However, using them for a small region like the Urmia basin to investigate the effect of climate variability is always questionable when their validity is not assessed over the region. Therefore, we believe that it can be a limitation and may lead to misleading results if the accuracy of the data is not reasonable. They used monthly GPCP precipitation data for assessing the trend of precipitation over the basin. However, the proportion of shared variance between GPCP and in-situ data over the basin is about 0.75 on a monthly scale (see Table 2 in Jalili et al. 2012). Therefore, their analysis includes some errors caused by the questionable quality of data. Additionally, their analyses were done on a monthly scale that neglected changes in the daily precipitation pattern. We modified this part as follows:

"Using a statistical change point analysis and without modelling, Khazaei et al. (2019) stated that given the stable conditions of precipitation and temperature, climatic variations could not explain the dramatic decline of the lake level. They also estimated the change of vegetation dynamics and its associated hydrological loss in terms of evapotranspiration. They used monthly GPCP precipitation data for assessing the trend of precipitation over the basin. However, the proportion of shared variance between GPCP and in-situ data over the basin is about 0.75 on a monthly scale (see Table 2 in Jalili et al. 2012). Therefore, their analysis suffers from the poor quality of precipitation data. Moreover, their analysis was done on a monthly scale that cannot capture the sub-monthly variability of climatic variables. Also, they did not account for the role of groundwater dynamics in their analysis."

Referee #3: "For quantifying human and climatic contributions to observed hydrological changes, a comprehensive modeling approach that takes into account, for example, the impacts of changing temperatures on runoff and thus river inflow and on evapotranspiration of the lake itself is preferable." Preferable to what exactly? I tend to disagree that modeling is preferable to comprehensive analysis of historical data. Modeling introduce various sources of new uncertainty to a problem (such as model structural uncertainty, parameter uncertainty, over-parameterization, parameter transferability across time and space, etc.), which are not preferable to the simplifying assumptions underlying statistical analyses (such as trend, correlation, or linear regressions). In general, I believe, both approaches of modeling and data analysis can inform us in some ways, while each has its own shortcomings in other ways.

Response: We generally agree with you, but statistical analyses such as trend and correlation sometimes cannot give us useful information about what would happen. For example, when there is no trend in precipitation, but having a significant trend in river flow, we conclude that human activities are the leading factor for river flow reduction (most of the trend studies over lake Urmia suffer from this issue). While, in hydrological modelling the depth of precipitation in each event, the interval between rainfall events (represented in soil moisture) and other involved elements to generate runoff would be considered. Meanwhile, we agree with you that the models have their own disadvantages, i.e., more uncertainty. In our study discussed that the main point of disagreement among researchers on the contribution of each factor to the lake shrinkage comes from this issue. All modelling studies (except Chaudhari et al. (2018) that used uncalibrated model), i.e., shadkam et al. (2016), Farokhnia et al. (2018), and also our study highlighted the impact of climate factor could not be ignored over the basin. However, trend and correlation analysis studies like Khazaei et al. (2019) and Ghale et al. (2018) stated the climate contribution is negligible compared to anthropogenic impacts. Trend analysis of daily precipitation for different classes, i.e., tiny, moderate, and heavy like Bavil et al. (2018) can be helpful if we would like to use a trend analysis approach. Bavil et al., (2018) showed that there is a significant increase in the frequency of daily precipitation of less than 5 mm and a significant decrease in the frequency of daily precipitation of 10-15 mm, suggesting a runoff reduction even in case of constant annual precipitation. We modified this part as follows:

"For quantifying human and climatic contributions to observed hydrological changes, a comprehensive modelling approach that takes into account, for example, the impacts of changing temperatures and land use change (e.g., urbanization and cropland expansion) on runoff generation and thus river inflow and on evaporation of the lake itself is preferable to statistical analyses such as trend and correlation analysis. Such statistical analyses may be misleading about reasons for certain temporal changes. For example, when there is no trend in precipitation but a significant trend in streamflow, it may be concluded that human activities are the dominant case of streamflow reduction; most of the trend studies for Lake Urmia suffer from such a hasty conclusion. In hydrological modelling, more detailed information such as the depth of precipitation in each event, the interval between rainfall events (represented in soil moisture) and other involved elements to generate runoff are considered. All modelling studies (except Chaudhari et al. (2018) who used an uncalibrated model), i.e., Shadkam et al. (2016), Farokhnia et al. (2018) and our study, found that the impact of climatic variations could not be ignored over the basin, while, trend and correlation analysis studies such as Khazaei et al. (2019) and Ghale et al. (2018) stated the climate contribution is negligible compared to anthropogenic impacts. We suggest to do trend analysis of daily precipitation distinguishing different intensity classes (e.g. Bavil et al. 2018)."

Referee #3: "Chaudhari et al. (2018) but their uncalibrated global hydrological model that represented the basin by 5-6 cells only was not able to simulate well the flows and storages in the basin." This is a mischaracterization of Chaudhari et al. (2018), undermining their extensive modeling setup and evaluation. Although Chaudhari did not explicitly discuss the model setup and calibration, they demonstrated the adequacy of their model by evaluating various model outputs against available knowledge and data of the LU basin. For instance, they compared their simulation inflow to the lake with the observed inflow record (previously gathered by Hassanzadeh et al., 2012). As the figure shows it is in good agreement. I agree with the authors' intent to critically review previous studies to elaborate their shortcomings, however this must be done rigorously and accurately.

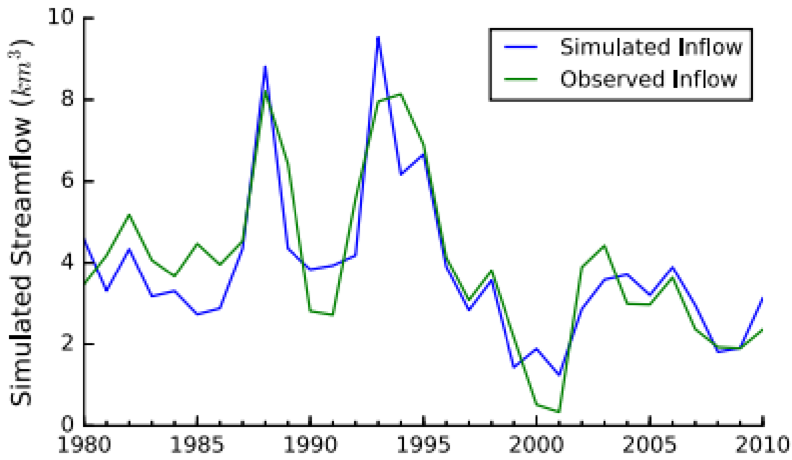


Fig. 7. Comparison of simulated river inflow to Urmia Lake from HiGW-MAT model and the inflow data from Hassanzadeh et al., (2012).

Response*:* We would not undermine their extensive modelling setup and evaluation. We reported only their shortcomings. From Chaudhari et al. (2018): "The simulations we use in this study are based on Pokhrel et al. (2015) and Felfelani et al. (2017) conducted at the global scale.". Therefore, no calibration was performed specifically for Lake Urmia. We could not find inflow data in Hassanzadeh et al. (2012). However, although their model can be simulated inflow into the lake relatively well, the problem is that this inflow data is not correct and even far from observations. As we mentioned in the manuscript, for example, simulated annual inflow into the lake was estimated to be $3{,}700 \cdot 10^6$ m^3 in 2003 (their Fig. 8) while observed inflow was much higher, $5{,}835 \cdot 10^6$ m^3. In 2009, observed inflow, with $1{,}036 \cdot 10^6$ m^3, was only half of the simulated one. We believe under this situation, their results are not reliable and at least are not comparable with other modelling studies over the basin; however, we remain at your disposal for any further information and modification.

Referee #3: Page 19 lines 2-5: "Hosseini-Moghari et al. (2018) showed that an increasing frequency of days with less than 5 mm precipitation in combination with decreasing monthly precipitation has led to the observed reduced inflow into two dams in the Lake Urmia basin that are located 5 downstream of areas with insignificant human water use." This study is not available online and it is not possible to confirm whether it is peer-reviewed or not.

Response*:* This paper was presented in the 8th Global FRIEND-Water Conference, Beijing, China, 6-9 November 2018. Please check the paper here http://m.iahr.org.cn/Resource/abstractinfo/3955.

Referee #3: "We conclude that analyses should be done on a daily time scale or smaller." What type of analyses exactly? Such a generic statement. Needless to mention that the very present study of the authors is not done on a daily time scale either. The time scale of a study depends on its objective.

Response: All types of analysis would be done for climate change or variability impact. Because climate change affects not only the climate variable amount but also their pattern. It should be noted that our modelling was done on a daily scale and presented on a monthly or annual scale. Daily assessment does not mean to provide results on a daily scale; it means to assess the effect of variation in the daily time scale on the results. We modified this part as follows: "We conclude that for assessing the effect of climatic variability on hydroclimatic variables, the analyses should be done on a daily time scale or shorter to consider the change in amount and patterns of variables."

Referee #3: "we examined the ratio of annual inflow into the lake (based on the ensemble mean) over annual precipitation during the study period. This ratio reached maximum values in 2003 (0.29 and 0.41 for the anthropogenic and naturalized conditions, respectively) and minimum values in 2009 (0.07 and 0.15). Averaged over the period 2009-2013, these ratios are, with 0.11 (ANT) and 0.22 (NAT), much smaller than the values for 2003-2007, 0.20 and 0.32. Thus, the drought year 2008 as well as the relatively small ratio of inflow into the lake over precipitation in the last five years of the study period play a significant role in the decline of inflow and lake water storage" There are various issues with this argument. First, the period 2009-2013 is a very short period to build a hydro-climatic analysis on, particularly for LU with remarkable non-stationarity. So, the naturalized scenario based on this period is not reliable. Second, the considerable extraction of groundwater resources has been an additional source of water for irrigation and consequently hydrological loss in this basin. The impact of groundwater withdrawal (and its consequent hydrological loss) would have had a direct impact on the lake and possibly on streamflow generation in the basin as well (e.g. as the land coverage of the basin has changed). Urbanization in this basin (discussed by Chaudhari et al. (2018)) together with the expansion of agricultural and irrigated areas would have an impact on streamflow generation (both magnitude and generation mechanisms).

Response: Firstly, we did not do any specific hydro-climatic analysis. We mentioned the ratio of annual inflow into the lake over annual precipitation for the sub-periods. We did not state any general statement from a statistical point of view that requires much data; therefore, non-stationarity for sub-period of 5 years does not matter here. Regarding the effects of the extraction of groundwater, we agree with you. Although comprehensive modelling of the relationship between groundwater withdrawal and streamflow generation is not a simple simulation, the WGHM model can model this relationship with promising accuracy; hence there is no serious concern in this regard. Under the naturalized conditions, there is no groundwater withdrawal and land-use change, so under this condition, your mentioned point does not matter. Finally, our results support your statement, and the difference between the ratio in anthropogenic and naturalized conditions comes from the expansion of agricultural and irrigated areas.

Referee #3: "For quantifying human and climatic contributions to observed hydrological changes, a comprehensive modeling approach that takes into account, for example, the impacts of changing temperatures on runoff and thus river inflow and on evapotranspiration of the lake itself is preferable." Also, estimating the impact of land use change (e.g. urbanization and cropland expansion) on runoff generation in the basin.

Response: We have revised it as follows:

"For quantifying human and climatic contributions to observed hydrological changes, a comprehensive modelling approach that takes into account, for example, the impacts of changing temperatures and land use change (e.g., urbanization and cropland expansion) on runoff generation and thus river inflow and on evaporation of the lake itself is preferable to statistical analyses such as trend and correlation analysis."

Referee #3: 2.1. On the role of atmospheric drought

Figure 2 and the last paragraph of page 2: This figure and its associated text provide an incomplete overview of the lake dynamics. The decline of the lake water level started around the year 2000, which is way more abrupt than 2003 onwards.

Response: Unfortunately, the lake level data that was provided by Tourian et al. (2015) started in the middle of 2002. Therefore, we could not extend that back before 2000. We attempted to show the lake dynamic within our study period. However, we modified the text as follows:

"In the 1970s and 80s, the water level of Lake Urmia was approximately at 1,276 m above sea level and then increased to more than 1,278 m in 1995 due to a few wet years (Shadkam et al., 2016). Khazaei et al. (2019) identified the year 2000 as the change point of lake dynamics. The water level dropped to 1,274 m in 2003 because of the severe drought in 1999-2001 exacerbated by human water use (Shadkam et al., 2016). From 2003 to 2014, lake extent was approximately halved, and water level declined by another 3 m, while seasonal variability of lake water extent increased (Tourian et al., 2015)."

Referee #3: Page 1 line 30: "The study shows that even without human water use Lake Urmia would not have recovered from the significant loss of lake water volume caused by the drought year 2008." First, you have not provided any evidence that the drought year 2008 caused a significant loss in lake volume, this causal link is nonexistent in your study. The authors are trying to over-emphasize the role of atmospheric droughts, specially the 2008 one. There has been stronger atmospheric droughts in previous years than year 2008. Here is a figure from Alborzi et al. (2018). The historic droughts during 80s and early 90s are more severe than the 2008 drought, yet the lake has survived (AghaKouchak et al., 2015).

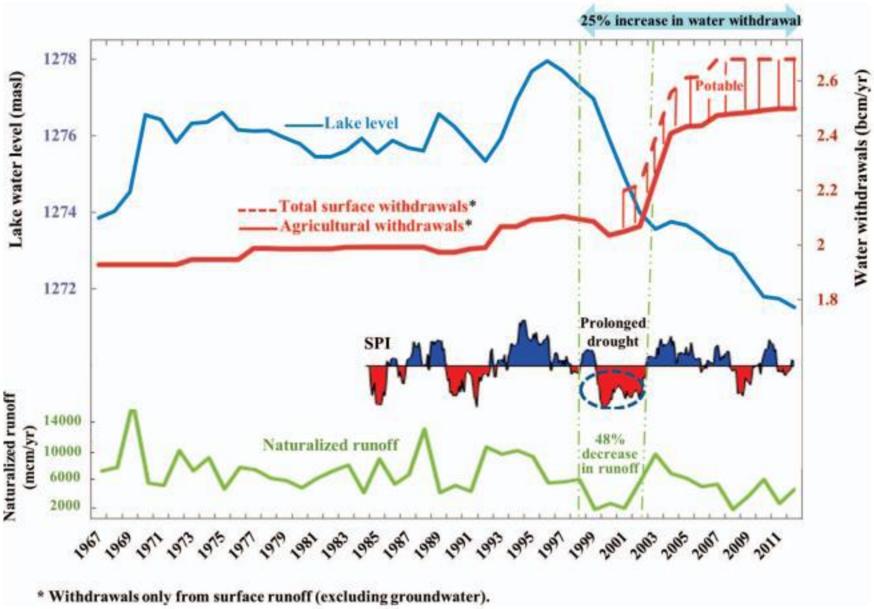


Figure 2. Key attributes of the lake-basin system prior to restoration program in 2013, including observed lake level, standardized precipitation index (SPI), basin-scale naturalized runoff, and surface water withdrawal. The basin's recent wet (blue) and dry (red) periods are illustrated in SPI and naturalized runoff curves. Post-1998 drop in lake level corresponds to a substantial increase (~25%) in surface water withdrawals during the prolonged drought of 1998–2002.

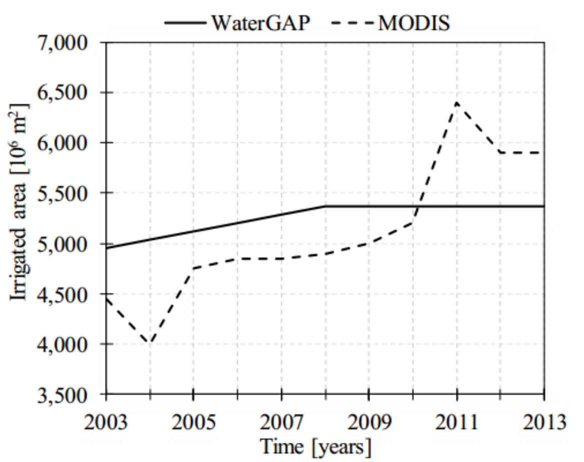
Response*:* Firstly, it should be noted that the significant loss in 2008 (in comparison with other years of the study period) was shown in Figure 8. About the rest of your comment, we see some issues as follows:

- According to the figure you provided, the droughts during the 80s and early 90s also reduced the lake level, but the following wet period not only compensated the drought impact but also reached the lake level to the maximum level during the past 50 years.
- As you mentioned in previous comments, there is a sharp reduction in lake level after 2000. Based on the figure you provided, the drought during 1999-2002 could be the main reason for this reduction. It should be noted that it takes time to lake loss a $30 \cdot 10^9$ m^3 water that had in 1995. Therefore, we should not expect during the previous drought, and in a few years, the lake level falls below the ecological level.
- The past three droughts are not comparable due to the lake had a significant volume in two first droughts, so we believe that it is hardly acceptable to conclude that because the lake survived in the past droughts, then the lake should also be resistant to other droughts.

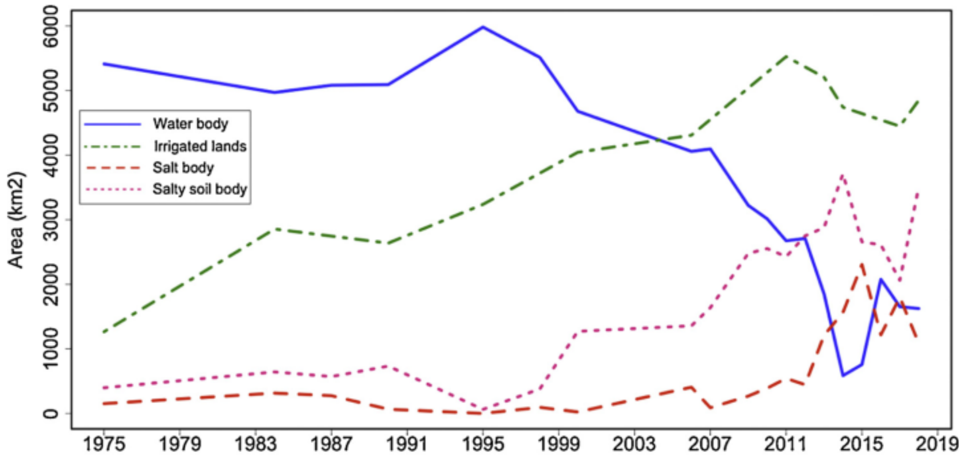
Based on the aforementioned discussion, let us disagree with your statement that due to this fact that lake survived from the 80s and early 90s so it could survive from the 2008 drought. Because the conditions of the lake in 2008 were completely different from previous droughts. Moreover, the conditions of the basin after 2008 were different, i.e., reducing the ratios of annual inflow into the lake over annual precipitation, as we discussed above.

Referee #3: 2.2. Irrigated area

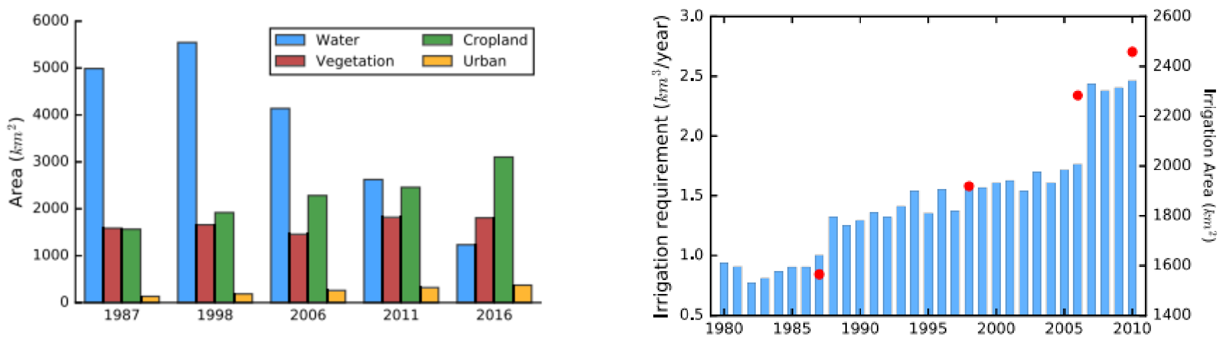
One aspect that has not been discussed in section 4.2 is the irrigated area: how it is differently estimated by different studies and its implications. Below I have extracted figures corresponding to the estimated irrigated areas by 3 different studies. Supplement page 3 line 5: "Considering that water management in the basin aims at preventing any increase of irrigated areas, it is assumed that the irrigated area in 2013 remained at the 2012 value (Fig. S3)". This assumption is questionable. For instance, Alizade Govarchin Ghale et al. (2019) estimated the irrigated lands to decrease by ~12% from 2012, but again increase in 2018-2019. Further, their estimated irrigated area is very different from the present study: year 2003 is different by ~500 km2 (~12%). The trend is also different, e.g. the increase during 2007-2011, or during 2003-2005. While both this study and Alizade Govarchin Ghale et al. (2019) estimated the irrigated area based on the overall vegetation coverage, Chaudhari et al. (2018) made a distinction between natural vegetation and cropland. They showed (see the figures below) that the while the natural vegetation has oscillated throughout the years (1998 to 2006 → decreased, 2006 to 2011 → increased, and 2011 to 2016 → decreased), the cropland has continuously increased. Moreover, they estimated the annual net irrigation requirement (NIWR) during 1980-2010 based on the crop evapotranspiration (FAO Penman Monteith approach), independent of the global hydrological model they used, and compared it with estimations based on Landsat classification (see Figure 9 in their study). So, their estimated irrigation is independent of how well or poorly their model was calibrated, and arguably more comprehensive than your study. While the present study demerited Chaudhari et al. (2018) (page 19 of the manuscript) and entirely overlooked Alizade Govarchin Ghale et al. (2019) and Fazel et al. (2017), the authors failed to acknowledge that these studies delved deep into land use changes, irrigation water requirement, and flow regime changes.



From the present study



From (Alizade Govarchin Ghale et al., 2019)



From Chaudhari et al. (2018)

Response: Firstly, we should note that our study did not demerit Chaudhari et al. (2018); we stated that the results of an un-calibrated model were evaluated based on wrong in-situ data would not be reliable. Now we found another issue in Chaudhari et al. (2018). They stated, "we use the Landsat imagery only for the month of September when the crops are completely matured in the study region." Despite their statement, some crops were harvested that month. There are two peaks for crop area in the basin 1) May for rainfed agriculture and 2) July and August for the irrigated area (Kamali and Youneszadeh Jalili 2015). As a result, their estimation of the irrigated area is significantly less than the actual one. For example, Kamali and Youneszadeh Jalili (2015) estimated the irrigated area for 2003 about 4,500 km^2 and based on FAO (WaterGAP dataset), it was about 5,000 km^2 while Chaudhari et al. (2018)'s estimation was about 2,000 km^2 which is less than half of actual one.

As you mentioned, the irrigated were used in Alizade Govarchin Ghale et al. (2019), and our study both were estimated based on the overall vegetation coverage. However, Alizade Govarchin Ghale et al. (2019) used April and August to the estimated irrigated area while Kamali and Youneszadeh Jalili (2015) used July and August that lead to some differences. Also, month April that was used by Alizade Govarchin Ghale et al. (2019) includes both irrigated and rainfed farms, the distinction between irrigated and rainfed cultivation may also make some differences. However, due to the fact, Kamali and Youneszadeh Jalili

(2015)'s report was approved by the ULRP, we believe that the use of the official report from ULRP would be more reliable than other sources.

About our assumption regarding the irrigated area in 2013, it should be noted all the works you mentioned here published after our manuscript submission. Therefore, we had no more reliable sources to use its data for 2013. The overall result may not change significantly by considering the small difference in the estimated irrigated areas in 2013.

We have added Fazel et al. (2017) to another part of our discussion based on your previous comments and added the following explanation to section 4.2:

"As a final word, the irrigated area used in this study obtained from the official report of ULRP (Kamali and Youneszadeh Jalili 2015). However, Chaudhari et al. (2018) estimated the irrigated area significantly less than the irrigated area used in the current study (Figure S3 compared to Figure 9 in their study). They used September for estimation of the irrigated area while the crops are completely matured in July and August in the basin. As a result, some crops are harvested in September. Therefore, it could be the main reason for such a significant underestimation of irrigated areas in the basin by Chaudhari et al. (2018). Also, Alizade Govarchin Ghale et al. (2019) estimated the irrigated area in the basin. Although their result is much closer to Kamali and Youneszadeh Jalili (2015) relative to Chaudhari et al. (2018), they used April and August to the estimated irrigated area, while Kamali and Youneszadeh Jalili (2015) used July and August that lead to some differences. Also, month April that was used by Alizade Govarchin Ghale et al. (2019) includes both irrigated and rainfed farms, the distinction between irrigated and rainfed cultivation may also make some differences. However, due to the fact, Kamali and Youneszadeh Jalili (2015)'s report was approved by the ULRP; we believe that the use of the official report from ULRP would be more reliable than other sources. However, the data reported by Kamali and Youneszadeh Jalili (2015) surly suffer some uncertainties that are inevitable."

Referee #3: 3 Comments on "Section 4.3 Limitations"

3.1. Study period

Reviewer comment: the time period 2003-2013 is inadequate for modeling the lake dynamics. Before 2000 the lake was not as heavily impacted by over-regulation of the river flows, and also between 2000-2003 there is significant variation in the lake level and annual inflows to the lake. Therefore, it is essential to include these years, for as many variable as possible. Otherwise, the model is biased and not representative of the lake dynamics.

Authors' response: We have considered this period due to the fact that the observed data was available for this period. We completely agree with you; it was better to consider a longer period for calibration. However, we don't prefer to reconstruct data, that is error-prone. The GRACE data and irrigated areas are not available for the period 2000-2003. Further, we don't want to use the model for out of calibration period, therefore we believe that for using the model in the calibration period there is no concern about the bias.

Reviewer's response: the point I argued is not simply about the length of data and model calibration. There are major aspects of the lake dynamics (and the basin evolution) that falls outside the 2003- 2013 period. While you evaluated your model within the calibration period using an independent variable, you tend to generalize your findings about the lake beyond the limited period of 2003-2013. To study the drivers of the lake drying, it is not adequate to build up your entire argument based on a limited time period that does not include the non-stationarities of the lake system: various studies identified two periods of pre- and post-change for the lake, e.g. Khazaei et al. (2019) identified year 2000 and Fazel et al. (2017) identified year 2001 as the change point. Given that, your study does not cover the pre-change period, and both anthropogenic and natural scenarios are defined based on only a sub-period (2003-2013) of the post-change period (2000 to date), which biases the scenario analysis. Further, most recent changes in the lake system is also not discussed. The lake has experienced considerable changes since 2013, e.g. see the extensive study by (Alizade Govarchin Ghale et al., 2019) on the land use changes within the lake basin. The figure below (extracted from Alizade Govarchin Ghale et al. (2019) shows the historic surface area as well as its increase since 2013– evidence of remarkable non-stationarity. To what extent your modelling assumptions and results are compatible with this non-stationarity, particularly the most recent changes of the lake?

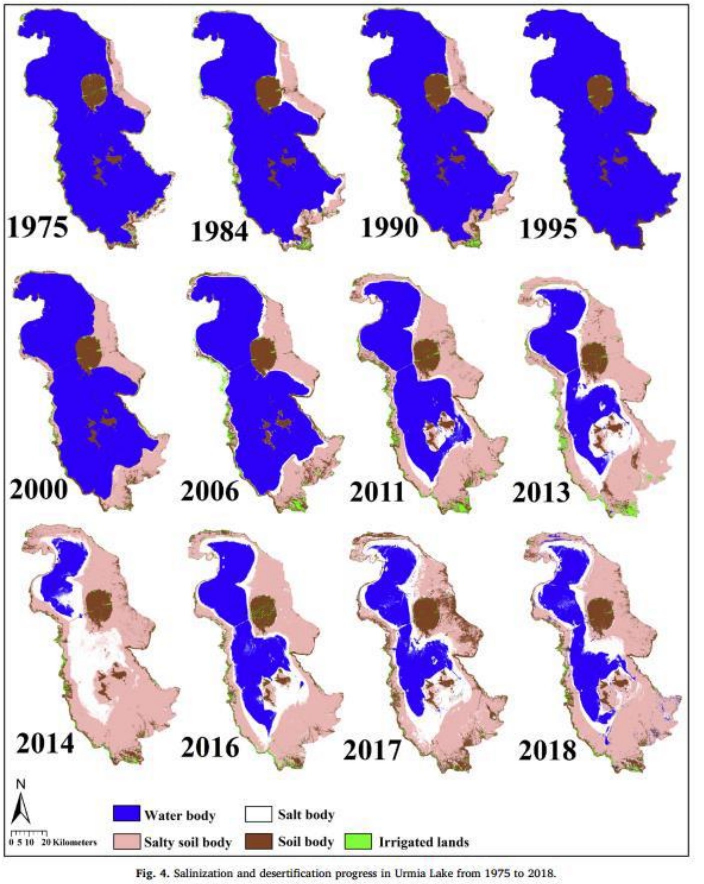


Fig. 4. Salinization and desertification progress in Urmia Lake from 1975 to 2018.

Response: We did not generalize our findings for the lake, beyond the period of 2003-2013 nowhere in the manuscript. As we mentioned, the results of two other modelling studies, i.e., Farokhnia et al. (2018) and Shadkam et al. (2016) that covered the pre-change period, support our results. However, we did all analyses for the calibration period and no extending for after 2013 or before 2003. We also discussed the most recent condition of water storage in the basin as: "GRACE TWSA data indicate an increasing trend in water storage in the basin during 2014-2017 due to both less water use due to water management (ULRP, 2015b) and the wet years 2015/2016. This trend is about half as strong as the decreasing trend during 2003-2013."

It should be noted, WGHM as a distributed hydrological model is less dependent on the behavior of data in the calibration period than the data-driven models. We believe that as WGHM simulated the non-stationarity behavior of the lake during the calibration period, if there are high-quality input data for running the model out of calibration period, it can simulate the water storage components within the basin with reasonable accuracy; at least for the most recent changes of the lake that the changes fall in bounds of calibration period has to be reliable. To address your comment, we have added the following sentences to the limitations section:

"Finally, the study period 2003-2013 does not include some of the years with significant changes in the dynamics of the lake and the basin (i.e., years 2000 and 2001 that identified as the change point of the lake by Khazaei et al. (2019) and Fazel et al. (2017), respectively) due to data availability. Therefore, our results cannot be generalized to previous decades."

Referee #3: 3.2. Other limitations and suggestions

Role of dams: Another aspect of the lake system that you did not accounted for explicitly is the dam construction within the lake basin over the past decades (24 dams were constructed during 1970- 2000, and 32 during 2000-2014), which studies such as Fazel et al. (2017) and Alizade Govarchin Ghale et al. (2018) accounted for explicitly.

Response: We have added the following sentences to the limitations section:

"We determined that the results of the naturalized run with and without reservoirs for annual inflow into the lake differ by less than 2%, whereas Fazel et al. (2017) and Ghale et al. (2018) stated that dams have a significant impact on the lake shrinkage. However, Shadkam et al. (2016) showed the role of dams in the reduction of inflow into the lake did not exceed 5% due to evaporation from reservoirs. Moreover, in this study, the inflow into the lake was assessed on an annual scale, and there is no correlation between the dams' operation and annual inflow in the basin (Fathian et al. 2014). Therefore, the error from this source to our result should be negligible."

Referee #3: **Seasonal flow and model calibration:** While all variables are calibrated/evaluated on a monthly basis, the streamflow is calibrated on annual scale. I suspect that it is due to the fact that the model could not adequately represent the seasonal variations in streamflow, which are significant for this basin (Alizade Govarchin Ghale et al., 2019; Fazel et al., 2017). The seasonal variations of the flow have direct implications on irrigation estimations and the lake dynamics.

Response: The reason was that the reliable monthly data for inflow into the lake was not available, and the ULRP's annual data was calculated based on monthly data in stations around the lake with filling several gaps. However, to consider your concern about the seasonality of inflow into the lake, we have plotted the seasonality of simulated inflow here. As you can see, the seasonality is almost similar to the seasonality presented by Fazel et al. (2017).

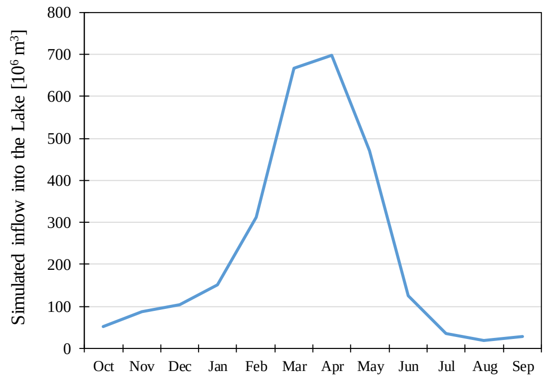
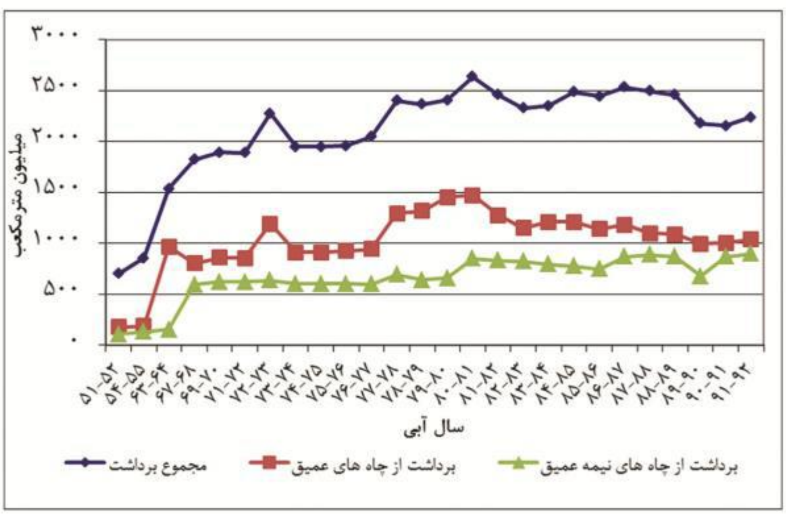


Figure R: Seasonality of simulated inflow into the lake

Referee #3: Groundwater withdrawal and its hydrological connectivity to the lake: The groundwater withdrawal is under-estimated in the model setup (a point that has been raised by reviewers before). While authors stated that "Observed decline of groundwater storage was $1.8 \cdot 10^9$ m3, i.e. 18% of the observed total water storage loss in the basin" (page 17 line 20), the groundwater withdrawal (including both shallow and deep wells, see the figure) shows at least 2.1 MCM withdrawal in the past 2 decades.



From LURP report (in Persian, attached to this review). Red is withdrawal from deep wells, and green is withdrawal from partial deep wells, and blue is the total extraction.

Response: Our sentence is about groundwater storage loss, not groundwater withdrawal. It should be noted some parts of withdrawal from surface water and groundwater would retune to the groundwater. The loss is the withdrawal amount minus returned amount. Therefore, it is reasonable that the withdrawal volume is more than the loss. Hence, it sounds not only there is no problem here, but also your provided reference approved the validity of our results.

Referee #3: Also, as discussed by Danesh-Yazdi and Ataie-Ashtiani (2019) the hydrologic connectivity between the lake and groundwater remains an under-studied aspect of the lake dynamics – which is a general limitation of most studies including the present one.

Response: To address your comment, we added the following sentences to the limitations section:
"Also, in this study, it is assumed that there is no significant direct relationship between the lake and groundwater. However, the hydrologic connectivity between the lake and groundwater remains an under-studied aspect of the lake dynamics (Danesh-Yazdi and Ataie-Ashtiani 2019)."

Referee #3: Page 20 lines 1-3: re lake bathymetry please also cite the below studies:

• Sima, S., & Tajrishy, M. (2013). Using satellite data to extract volume–area–elevation relationships for Urmia Lake, Iran. *Journal of Great Lakes Research*, *39*(1), 90-99.

• Karimi, N., Bagheri, M. H., Hooshyaripor, F., Farokhnia, A., & Sheshangosht, S. (2016). Deriving and evaluating bathymetry maps and stage curves for shallow lakes using remote sensing data. *Water Resources Management*, *30*(14), 5003-5020.

Response: Both references were added.

Referee #3: 4 Minor comments

Page 3 line 10: "Studies on various aspects of the Lake Urmia disaster abound. With decreasing lake water volume, salt concentration has increased". Please cite the recent study on salt concentration as a dust source:

• Boroughani, M., Hashemi, H., Hosseini, S. H., Pourhashemi, S., & Berndtsson, R. (2019). Desiccating Lake Urmia: A New Dust Source of Regional Importance. IEEE Geoscience and Remote Sensing Letters.

Response: The reference was added.

Referee #3: Page 3 line 11: "Precipitation reduction, temperature increase, agricultural development including construction of man-made dams and building a causeway across the lake have been identified as the reasons for the degradation of Lake Urmia (Abbaspour and Nazaridoust, 2007; Zeinoddini et al., 2009; Delju et al., 2012; Jalili et al., 2012; Sima and Tajrishy, 2013; Fathian et al., 2014; Farajzadeh et al., 2014; Banihabib et al., 2015; AghaKouchak et al., 2015; Azarnivand and Banihabib 2017; Alizadeh- Choobari et al., 2016; Ghale et al., 2018; Khazaei et al., 2019)". Please separate out the references and cite relevant references for each factor (underlined phrases) individually. It helps the readers to track back.

Response: It has been modified as follows:

"Precipitation reduction and temperature increase (Delju et al., 2012; Fathian et al., 2014; shadkam et al. 2016; Farokhnia et al. 2018), agricultural development including construction of man-made dams (Farajzadeh et al., 2014; Banihabib et al., 2015; Azarnivand and Banihabib 2017; AghaKouchak et al., 2015; Ghale et al., 2018; Khazaei et al., 2019) and building a causeway across the lake (Zeinoddini et al., 2009) have been identified as the reasons for the degradation of Lake Urmia."

Referee #3: Page 4 line 25: "a good fit of simulated and observed streamflow may not necessarily lead to an appropriate simulation of other flows and storages (Beven and Freer, 2001). Therefore, additional types of observations have to be added to avoid equifinality (Beven and Freer, 2001; Döll et al., 2016)." The second sentence does not follow the first sentence, and using "therefore" does not make sense here. Also, by adding further data types, one will not "avoid" equifinality, because equifinality is a general property of open complex systems (e.g. hydrological models) and cannot be avoided. The goal is to "reduce" equifinality when possible. Please also cite the following recent studies on equifinality which are directly relevant to the discussion:

• Kelleher, C., McGlynn, B., & Wagener, T. (2017). Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding. Hydrology and Earth System Sciences, 21(7), 3325.

• Khatami, S., Peel, M. C., Peterson, T. J., & Western, A. W. (2019). Equifinality and flux mapping: A new approach to model evaluation and process representation under uncertainty. Water Resources Research, 55, 8922– 8941.

Response*:* Both references were added. The mentioned part has been modified as follows:

"a good fit of simulated and observed streamflow may not necessarily lead to an appropriate simulation of other flows and storages (Beven and Freer, 2001). Moreover, additional types of observations have to be added to reduce the possibility of equifinality (Döll et al., 2016; Kelleher et al. 2017; Khatami et al. 2019)."

Referee #3: Page 12 line 17 "We determined that the results of the naturalized run differ by less than 2% from a run with reservoirs but without human water use". First, it is not clear 2% of what is discussed here exactly. Second, such a small difference between the two scenarios is clearly a red flag, indicating that the model setup and/or scenarios are problematic. Most of the recent studies concluded that the lake condition is heavily impacted by human water use.

Response*:* In this sentence, we are speaking about two scenarios of the naturalized run, not anthropogenic one. Therefore, there is no concern about the model setup. We stated that the difference between results annual inflow into the lake is less than 2% in naturalized run with and without reservoirs (there is no human water use in both scenarios). This value for Q and for a period longer than our study period estimated less than 5% by Shadkam et al. (2016). The mentioned part has been modified as follows:

"We determined that the results of the naturalized run for annual inflow into the lake differ by less than 2% from a run with reservoirs but without human water use."

Referee #3: Page 14 line 20 "In this way, efficient simulation of regional water flows and storages can be achieved, possibly as an alternative to a costlier setup of a regional model". I'm not sure if I understood this part. What is costly about a regional model that is discouraging? What do you exactly mean by "setup a regional model", do you mean to develop a model from scratch?

Response: We stated in its previous sentence, "a global hydrological model includes all data for simulating water flows and storages in specific regions of interest everywhere on the globe, and model calibration against multiple (regional) observations is a means for improving the performance of the global model regionally." Hence, a global model with a regional calibration is much preferable than setup a regional model. Set up a regional model takes more time and needs more labor cost compared with the use of an available model. "setup a regional model" could be "develop a model and its dataset from scratch" or "develop a dataset for an available hydrological model".

Referee #3: Page 18, reword the title of the subsection 4.2 "Comparison to human vs. climatic contribution as determined in previous studies", it does not read well.

Response: It has been modified as follows:

"Distinguishing the contributions of human water use and climate variability to lake shrinkage"

Referee #3: Page 19 line 8: "*respectivly*" → respectively

Response: It has been corrected.

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

References

Alizade Govarchin Ghale, Y., Altunkaynak, A., and Unal, A.: Investigation Anthropogenic Impacts and Climate Factors on Drying up of Urmia Lake using Water Budget and Drought Analysis, Water Resources Management, 32(1), 325-337, doi:10.1007/s11269-017-1812-5, 2018.

Alizade Govarchin Ghale, Y., Baykara, M., and Unal, A.: Investigating the interaction between agricultural lands and Urmia Lake ecosystem using remote sensing techniques and hydro-climatic data analysis, Agricultural Water Management, 221, 566-579, doi:1016/j.agwat.2019.05.028, 2019.

Amiri, V., Nakhaei, M., Lak, R., and Kholghi, M.: Geophysical, isotopic, and hydrogeochemical tools to identify potential impacts on coastal groundwater resources from Urmia hypersaline Lake, NW Iran, Environmental Science and Pollution Research, 23(16), 16738-16760, doi: 10.1007/s11356-016-6859-y, 2016.

Ashraf, B., AghaKouchak, A., Alizadeh, A., Baygi, M. M., Moftakhari, H. R., Mirchi, A., and Madani, K.: Quantifying anthropogenic stress on groundwater resources, Scientific reports, 7(1), 12910, doi: 10.1038/s41598-017-12877-4, 2017.

Bavil, S. S., Zeinalzadeh, K., and Hessari, B.: The changes in the frequency of daily precipitation in Urmia Lake basin, Iran, Theoretical and Applied Climatology, 133(1-2), 205-214, doi:10.1007/s00704-017-2177-7, 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.

Danesh-Yazdi, M., and Ataie-Ashtiani, B.: Lake Urmia Crisis and Restoration Plan: Planning without Appropriate Data and Model Is Gambling, Journal of Hydrology, 576, 639-651, doi: 10.1016/j.jhydrol.2019.06.068, 2019.

Farokhnia, A., Morid, S., Abbaspour, K., and Delavar, M.: Development of SWAT-LU model for simulation of urmia lake water level decrease and assessment of the proposed actions for its restoration; (role of anthropogenic and climatic factors on hydrological change of the basin and lake), Iranian Journal of Irrigation and Drainage, 12, 1041–1058 (in Persian with English abstract), 2018.

Fathian, F., Morid, S., and Kahya, E.: Identification of trends in hydrological and climatic variables in Urmia Lake basin, Iran, Theoretical and Applied Climatology, 119, 443-464, doi:10.1007/s00704-014-1120-4, 2014.

Fazel, N., Torabi Haghighi, A., and Kløve, B.: Analysis of land use and climate change impacts by comparing river flow records for headwaters and lowland reaches, Global and Planetary Change, 158, 47-56, doi:10.1016/j.gloplacha.2017.09.014, 2017.

Ghale, Y. A. G., Altunkaynak, A., and Unal, A.: Investigation Anthropogenic Impacts and Climate Factors on Drying up of Urmia Lake using Water Budget and Drought Analysis, Water Resources Management, 32, 325-337, doi:10.1007/s11269-017-1812-5, 2018.

Hassanzadeh, E., Zarghami, M., and Hassanzadeh, Y.: Determining the main factors in declining the urmia lake level by using system dynamics modeling, Water Resources Management, 26, 129-145, doi:10.1007/s11269-011-9909-8, 2012.

Hosseini-Moghari, S.M., Araghinejad, S. and Ebrahimi, K.: Monthly Precipitation Assessment: a misleading tool for understanding the effects of climate change, 8th Global FRIEND-Water Conference, Beijing, China, 6-9 November 2018.

Jalili, S., Kirchner, I., Livingstone, D. M., and Morid, S.: The influence of large-scale atmospheric circulation weather types on variations in the water level of Lake Urmia, Iran, International Journal of Climatology, 32, 1990-1996, doi:10.1002/joc.2422, 2012.

Kamali, M., and Youneszadeh Jalili, S.: Investigation of landuse changes in Lake Urmia Basin using remotely sensed images, Report of Urmia Lake Restoration Program (ULRP) (In Persian), 2015.

Khazaei, B., Khatami, S., Alemohammad, S. H., Rashidi, L., Wu, C., Madani, K., and Aghakouchak, A.: Climatic or regionally induced by humans? Tracing hydro-climatic and land-use changes to better understand the Lake Urmia tragedy, Journal of Hydrology, 569, 203-217, doi:10.1016/j.jhydrol.2018.12.004, 2019.

Pechlivanidis, I. G., Jackson, B., McMillan, H., and Gupta, H.: Use of an entropy-based metric in multiobjective calibration to improve model performance, Water Resource Research, 50, 8066–8083, doi: 10.1002/2013WR014537, 2014.

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.

ULRP: Urmia Lake - Causes of shrinkage and potential threats. Available at: http://ulrp.sharif.ir/fa (In Persian), 2015c.

Vaheddoost, B., and Aksoy, H.: Interaction of groundwater with Lake Urmia in Iran, Hydrological processes, 32(21), 3283-3295, doi:10.1002/hyp.13263, 2018.

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

Seyed-Mohammad Hosseini-Moghari[1,2], Shahab Araghinejad[2], Mohammad J. Tourian[3], Kumars Ebrahimi[2], Petra Döll[4,5]

[1]Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China.
[2]Department of Irrigation and Reclamation Engineering, University of Tehran, Karaj, Iran
[3]University of Stuttgart, Institute of Geodesy (GIS), Stuttgart, Germany
[4]Institute of Physical Geography, Goethe University Frankfurt, Frankfurt am Main, Germany
[5]Senckenberg Leibniz Biodiversity and Climate Research Centre (SBiK-F), Frankfurt am Main, Germany

Correspondence to: Seyed-Mohammad Hosseini-Moghari (Hosseini_sm@igsnrr.ac.cn)

Abstract. During the last decades, the endorheic Lake Urmia basin in northwestern Iran has suffered from declining groundwater tables and a very strong reduction in the volume as well as recently in the extent of Lake Urmia. For the case of Lake Urmia basin, this study explores the value of different locally and globally available observation data for adjusting a global hydrological model such that it can be used for distinguishing the impacts of human water use and climate variations. The WaterGAP Global Hydrology Model (WGHM) was for the first time calibrated against multiple in-situ and spaceborne data to analyse the decreasing lake water volume, lake river inflow, loss of groundwater, and total water storage in the entire basin during 2003-2013. The calibration process was done using an automated approach including a genetic algorithm (GA) and Non-dominated sorting genetic algorithm II (NSGA-II). Then the best-performing calibrated models were run with and without considering water use to quantify the impact of human water use. Observations encompass remote-sensing based time series of annual irrigated areas in the basin from MODIS, monthly total water storage anomaly (TWSA) from GRACE satellites, and monthly lake volume anomalies. In-situ observations include time series of annual inflow into the lake and basin averages of groundwater level variations based on 284 wells. In addition, local estimates of sectoral water withdrawals in 2009 and return flow fractions were utilized. Calibration against MODIS and GRACE data alone improved simulated inflow into Lake Urmia but inflow and lake volume loss were still overestimated, while groundwater loss was underestimated and seasonality of groundwater storage was shifted as compared to observations. Lake and groundwater dynamics could only be simulated well if calibration against groundwater levels led to an adjustment of the fractions of human water use from groundwater and surface water. Thus, in some basins, globally available space-born observations may not suffice for improving the simulation of human water use. According to WGHM simulations with 18 optimal parameter sets, human water use was the reason for 52-57% of the total basin water loss of about 10 km^3 during 2003-2013, for 39-43% of the Lake Urmia water loss of about 8 km^3 and for up to 87-90% of the groundwater loss. Lake inflow was 39-45% less than it would have been without human water use. The study shows that even without human water use Lake Urmia would not have recovered from

the significant loss of lake water volume caused by the drought year 2008. These findings can support water management in the basin and more specifically Lake Urmia restoration plans.

1 Introduction

Iran is a country with an arid and semi-arid climate where population growth and the government's aim of food self-sufficiency has led to increasing irrigated crop production and exploitation of surface water and groundwater resources. Climate change has resulted in increased temperatures and, in particular the northwest of the country, in decreased precipitation (Tabari and Talaee, 2011a, b) and thus decreased renewable water resources. In the last decades, numerous wetlands and lakes in Iran have dried up, and groundwater levels have strongly declined in most areas (Madani et al., 2016). The most serious disaster has occurred in the Lake Urmia basin, an interior basin in the northwest of Iran located in the three provinces West Azarbaijan, East Azarbaijan, and Kurdistan that covers an area of 52,000 km^2 (Fig. 1). At the downstream of the basin, 17 permanent rivers and 12 seasonal rivers discharge into the largest natural water body in Iran, Lake Urmia. Over the past two decades, climate variations and human activities (Hassanzadeh et al., 2012) have decreased inflow into the lake. Precipitation in the basin shows a decreasing trend over the period 1951-2013, with particularly low values after 1995, and evaporation has increased (Alizadeh-Choobari et al., 2016). Lake water volume is now approximately $30 \cdot 10^9$ m^3 below its historical maximum (ULRP, 2015a).

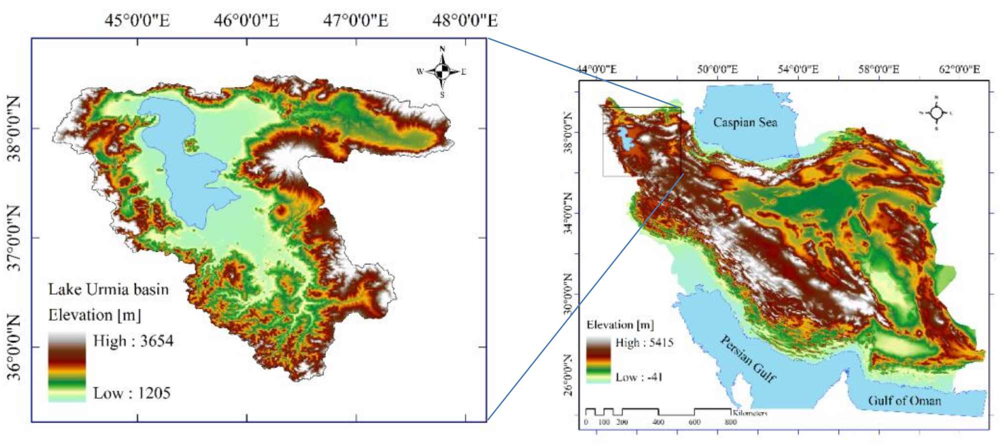


Figure 1: Location of Lake Urmia basin.

Lake Urmia is one of the largest hypersaline lakes in the world, which due to its ecological and natural features is a National Park, a Ramsar Site and a UNESCO Biosphere Reserve (Eimanifar and Mohebbi, 2007). It is a terminal lake that loses water only by evaporation (Hassanzadeh et al., 2012). Abbaspour and Nazaridoust (2007) estimated that inflows of at least $3 \cdot 10^9$ m^3/yr are needed to compensate for lake evaporation, while Alborzi et al. (2018) estimated values between $2.9 \cdot 10^9$ to $5.4 \cdot 10^9$ m^3/yr depending on climatic conditions. According to Alborzi et al. (2018), recovery of the lake could range from 3 to 16 years depending on climatic conditions, water use reductions, and environmental releases. Inflow from groundwater to the lake was estimated to be less than 3% of total inflow from precipitation, rivers, and groundwater (Hasemi, 2011). In the 1970s and 80s, the water ~~table~~level of Lake Urmia was approximately at 1,276 m above sea level and then increased to more than 1,278 m in 1995 due to a few wet years (Shadkam et al., 2016). ~~Afterward, the water table~~Khazaei et al. (2019) identified the year 2000 as the change point of lake dynamics. The water level dropped to 1,274 m in 2003 because of the severe drought in 1999-2001 exacerbated by human water use (Shadkam et al., 2016). From 2003 to 2014, lake extent was approximately halved, and water level declined by another 3 m, while seasonal variability of lake water extent increased (Tourian et al., 2015) (Fig. 2).

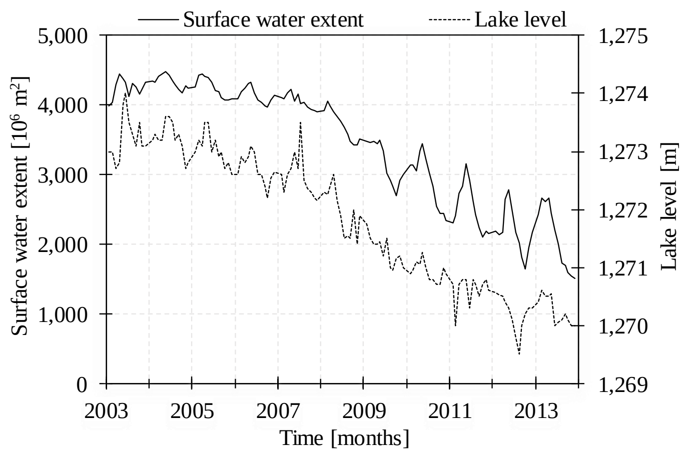


Figure 2: Time series of surface water extent and water ~~table~~level elevation of Lake Urmia (data from Tourian et al., 2015).

Studies on various aspects of the Lake Urmia disaster abound. With decreasing lake water volume, salt concentration has increased~~,~~ (Boroughani et al. 2019), endangering the aquatic biota feeing birds; exposed salt layers may lead to salt storms (Pengra, 2012). ~~Precipitation reduction, temperature increase, agricultural development including construction of man-made dams and building a causeway across the lake have been identified as the reasons for the degradation of Lake Urmia (Abbaspour and Nazaridoust, 2007; Zeinoddini et al., 2009; Delju et al., 2012; Jalili et al., 2012; Sima and Tajrishy, 2013; Fathian et al., 2014; Farajzadeh et al., 2014; Banihabib et al., 2015; AghaKouchak et al., 2015; Azarnivand and Banihabib 2017; Alizadeh-Choobari et al., 2016; Ghale et al., 2018; Khazaei et al., 2019).~~ Precipitation reduction and temperature increase (Delju et al., 2012; Fathian et al., 2014; shadkam et al. 2016; Farokhnia et al. 2018), agricultural development including construction of man-made dams (Farajzadeh et al., 2014; Banihabib et al., 2015; Azarnivand and Banihabib 2017; AghaKouchak et al., 2015; Ghale et al., 2018; Khazaei et al., 2019) and building a causeway across the lake (Zeinoddini et al., 2009) have been identified as the reasons for the degradation of Lake Urmia. By using Gravity Recovery And Climate Experiment (GRACE) satellite observations, altimetry data for Lake Urmia and outputs of the Global Land Data Assimilation System (GLDAS), Forootan et al. (2014) estimated the trend of groundwater storage changes in the Lake Urmia basin as -11.2 mm/yr between the years of 2005 to 2011, the largest decrease of the six investigated Iranian basins. Ahmadzadeh et al. (2016) investigated the effect of irrigation system changes in the basin from the surface to pressurized systems; they found that such changes would increase water productivity but would have no effect on lake inflow and would reduce groundwater levels by 20%.

Three hydrological modelling studies for Lake Urmia basin focused on quantifying the contributions of various factors on lake water volume (Hassanzadeh et al., 2012), lake inflow (Shadkam et al., 2016) or both (Farokhnia et al. 2018; Chaudhari et al., 2018). Using a lumped system dynamics modelling approach and observed time series of lake water volume for model calibration, Hassanzadeh et al. (2012), determined that about 65% of lake level decline between 1997 and 2006 was due to reduced river inflow, while four major man-made reservoirs contributed 25% and diminished precipitation on the lake surface 10%. Shadkam et al. (2016) evaluated the impact of climate, irrigation with surface water and reservoirs on inflow into the lake for the period 1960-2010 using a modified version of the macro-scale gridded hydrological model Variable Infiltration Capacity (VIC) model, which was calibrated against time series of river discharge at six observation station at the downstream end of six sub-basins draining into Lake Urmia. While the model was driven by global gridded WFDEI climate data set with a spatial resolution of 0.5°, basin-specific information on 41 reservoirs and on the temporal development of irrigated areas were taken into account. The study found that reservoirs had a very small impact on annual inflows and that climate variations accounted for 60% of lake inflow decrease of 48% over the 50-year period. In the model, all irrigation requirements need to be fulfilled by available surface water. Therefore, reduced availability of surface water during the 2000s due to low precipitation and high temperature resulted in unfulfilled irrigated water demand and a cap on the effect of human water use in the model while in reality, groundwater abstractions occurred and even increased (Delju et al., 2012; Hesami and Amini, 2016). In addition, the modelling study of Shadkam et al. (2016) did not consider the impact of domestic and industrial water use in the basin which can be expected to have increased during the last decades, given a population increase from 4.8 to 5.9 million from 2002 to 2010 (http://ulrp.sharif.ir/en/page/about-urmia-lake-basin, last accessed: 28 Apr. 2018). Chaudhari et al. (2018) used the output of the global HiGW-MAT model, with 1°×1° grid cell size of approx. 10,000 km^2, to distinguish climatic and anthropogenic contributions to the shrinkage of Urmia Lake. By running the model with and without human impacts (surface and groundwater use as well as reservoirs), they estimated that the human-induced river flow decline between 1995-2010 to account for 86% of the observed decrease of lake volume. However, a comparison with GRACE TWSA showed that the model overestimates the decrease in TWSA in the basin between 2003 and 2010. The HiGW-MAT model was not calibrated for the Lake Urmia basin but net irrigation requirements were simulated specifically for this study based on Landsat satellite images for 5 years between 1987 and 2016. The lake water balance is not simulated by the model such that no comparison with observed lake water levels was possible. A comparison with river discharge or groundwater observations was not done either. Farokhnia et al. (2018) developed a Soil & Water Assessment Tool (SWAT) model for quantifying the role of anthropogenic and climatic factors on hydrological change of the basin and lake during the 22-year period ending in 2009. By running the SWAT model under anthropogenic and natural conditions, they estimated the role of anthropogenic and climatic factors on the shrinkage of Urmia Lake. They concluded that the contribution of human activities and climate variability is almost equal to decreasing inflow into the lake and lake volume loss. They illustrated that in the second half of their study period, the climatic factors are responsible for 58% of the lake volume loss. However, they did not provide any results about the effects of human water use and climate change on groundwater across the basin. Besides, domestic and industrial water use was not considered in their study.

In previous hydrological modeling studies of Lake Urmia basin, there either no model calibration or calibration was only done using a single observation type, in particular surface water inflow into the lake. Although streamflow observations are very informative for hydrological modelling as they integrate over processes in the whole upstream basin, a good fit of simulated and observed streamflow may not necessarily lead to an appropriate simulation of other flows and storages (Beven and Freer, 2001). ThereforeMoreover, additional types of observations have to be added to avoidreduce the possibility of equifinality (Beven and Freer, 2001; Döll et al., 2016; Kelleher et al. 2017; Khatami et al. 2019). In this study, a multi-observation calibration approach was used to calibrate a hydrological model which was then applied to quantify the contributions of climate variations and human activities to the decrease of Lake Urmia water volume and river inflows. In addition, using Lake Urmia basin as a test case, we wanted to explore the value of different types of observation data for adjusting a global hydrological model by multi-observation calibration. Currently, global hydrological models are mostly uncalibrated but globally available space-born observations have increased the opportunity for model calibration at the global scale (Döll et al., 2016). For this purpose, the WaterGAP global hydrology model (WGHM) was calibrated by means of genetic algorithm (GA) and Non-dominated sorting genetic algorithm II (NSGA-II) for the Lake Urmia basins. Descriptions of the used data and the simulation setup are presented in section 2. The results of the different calibration variants and the impacts of human water use are shown in section 3. Section 4 discusses multi-observation calibration and the analysis of human impact as well as the limitations of the study. Finally, conclusions are drawn.

2 Methods and data

We analyzed the 11-year period from the beginning of 2003 until the end of 2013, as both GRACE data and global climate data to drive WaterGAP were available for this period. In the following sections, WaterGAP, its input data and the observational data used for calibration as well as the calibration approach are described.

2.1 WaterGAP

WaterGAP is a global hydrological model for assessing water resources under the influence of humans (Döll et al., 2003; Müller Schmied et al., 2014). With a spatial resolution of 0.5°×0.5°, it simulates water abstractions and consumptive water use (so-called net abstractions, i.e. the amount of water that evapotranspirates during use and does not flow to surface water bodies and groundwater afterwards) in five sectors (irrigation, livestock, domestic, manufacturing and cooling of thermal power plants); then net abstractions from either groundwater (NAg) or surface water bodies (NAs) are computed (Müller Schmied et al., 2014; Döll et al., 2012). Time series of NAg and NAs in each grid cells are then input to the WaterGAP Global Hydrology model WGHM that simulates their effect on water flows and storages. In its standard version, WaterGAP is calibrated against observed mean annual river discharge at 1319 stations worldwide by adjusting 1-3 model parameters related to runoff generation and streamflow (Müller Schmied et al., 2014), but due to lack of data not for any station in Lake Urmia basin. A previous WaterGAP version was calibrated, for 22 large basins, against streamflow and total water storage anomalies by adjusting 6-8 parameters (Werth and Güntner, 2010). WGHM can be run globally or for a specific basin. In this study, it was run only for the 22 0.5° grid cells that represent the Lake Urmia basin in WGHM (Fig. 3). A more detailed description of WGHM can be found in the supplement.

2.2 Data

We used the following observations for calibrating WGHM: (1) Remote sensing data including irrigated area in Lake Urmia basin and GRACE TWSA, (2) inflow into Lake Urmia Q, (3) groundwater levels from well observations, which were converted into groundwater storage anomalies GWSA (see section S2) and (4) statistical information on water withdrawals and consumptive uses in the basin. In addition, time series of lake volume based on remote sensing was used for validation. The 0.5° gridded EWEMBI data set was used as climate forcing. Irrigated area and Q are at the annual time scale, TWSA, GWSA and lake volume on the monthly scale and the climate forcing is on a daily scale. All data cover the period 2003-2013 (see section S2 for details).

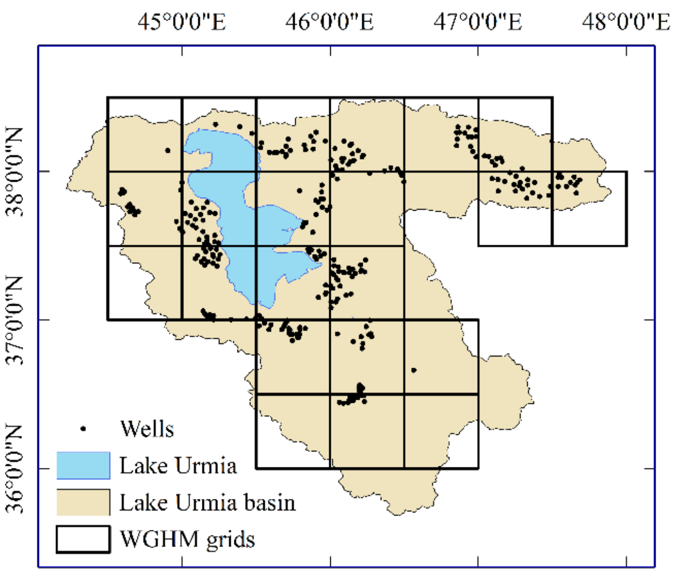


Figure 3: Grid cells in WGHM corresponding to Lake Urmia basin along with the locations of groundwater wells across the basin.

2.3 Calibration approach

Two calibration variants were applied. In the RS variant, only the remote sensing information was used for calibration, including irrigated area from MODIS and GRACE TWSA. In the variant RS_Q_GW_NA, ground-based information was used in addition to the remote sensing observations. This included inflow into the lake, groundwater data and statistical information regarding water use. Calibration was done using the genetic algorithm (GA) for variant RS, with just one calibration objective, and the non-dominated sorting genetic algorithm II (NSGA-II), a multi-objective version of GA, for the variant RS_Q_GW_NA. To integrate optimization algorithm with WGHM, we scripted the codes in shell and R environments by modifying *'GA'* (Scrucca, 2013), and *'nsga2R'* (Tsou, 2013) Packages in R. GA and NSGA-II are the most common evolutionary optimization algorithms in hydrological model calibration (e.g. Azarnivand et al. 2020). Both algorithms start with a random population (here WGHM parameters) and after evaluating the objective function(s) (here KGE) the better parameter sets are selected based on the value of the objective function (in GA) and non-domination and crowding distance (in NSGA-II). Then, the crossover and mutation operators are applied and the process will be continued until one stopping criteria met. The details of GA and NSGA-II can be found in Mirjalili (2019) and Deb et al. (2002), respectively. Because of the use of the random generators in GA and NSGA-II, we did five runs for each algorithm to achieve more reliable results. The selected parameters for each algorithm are presented in the supplement (Table S3). Fig. 4 shows the flowchart of these algorithms along with a schematic of the calibration process for the two calibration variants. In short, calibration included the modification time series of irrigated areas, of NAg and NAs, with different multipliers for individual years, as well as the modification of seven temporally constant model parameters or, in case of spatially heterogeneous parameters, multipliers (see Table 1). Modifications were done homogeneously for the whole basin. Months with assumed irrigation in Lake Urmia basin according to WaterGAP correspond to the actual irrigation months (Apr. and Oct.) in the basin according to Saemian et al. (2015). Thus no correction of seasonality was needed in the calibration process. More details are provided in the supplement. During calibration, seven model parameters (Table 1) were adjusted that are known to have an impact on TWSA, Q and GWSA. We used a modified version of the Kling Gupta efficiency (KGE) as the objective function, where the trend of the time series was added as a fourth component to the KGE (see Eq. 5 below).

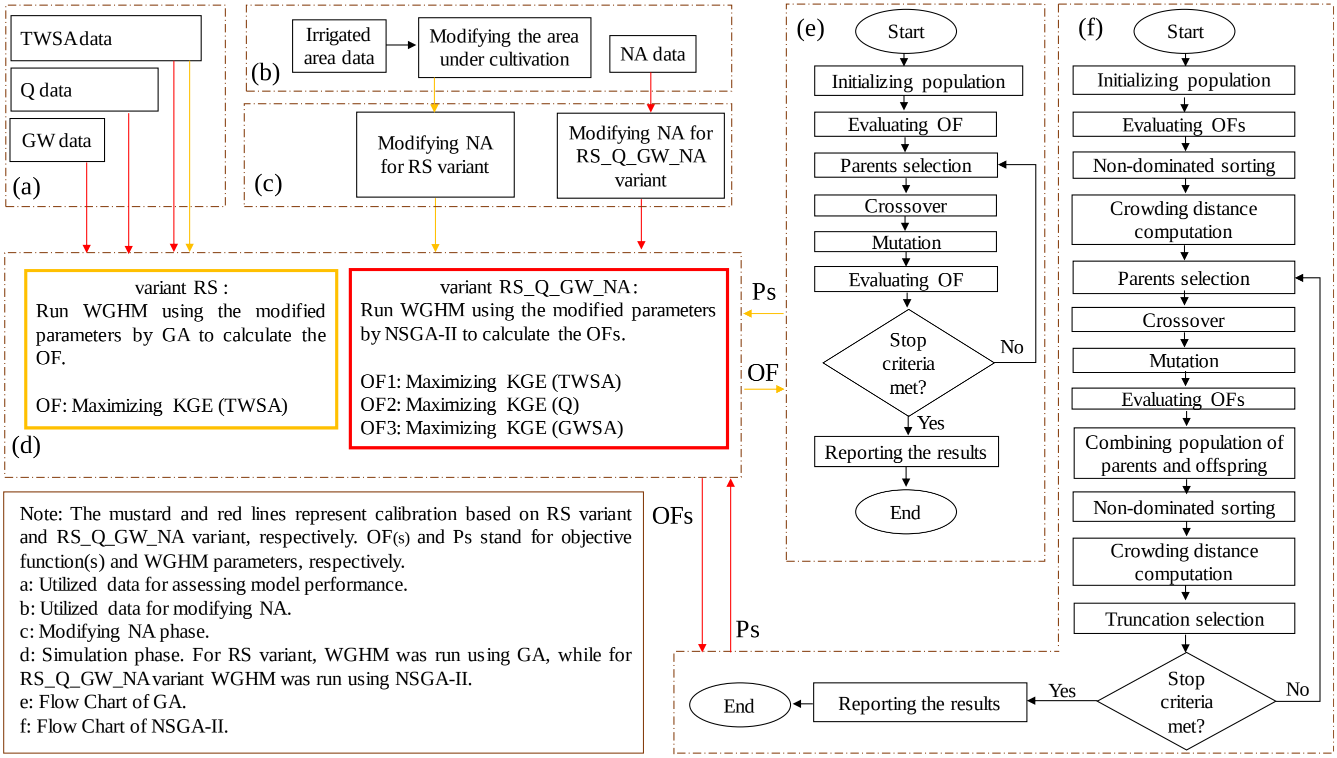


Figure 4: Flowchart of the WGHM calibration approach.

Table 1: WGHM parameters with the most effect on TWSA, inflow into the lake, groundwater storage.

Parameter	Value		
	Default	Minimum	Maximum
P1: Rooting depth multiplier	1	0.5	3
P2: Maximum active lake depth [m]	5	2	12
P3: Runoff coefficient multiplier	1	0.5	1.5
P4: Multiplier for the fraction of total runoff that becomes groundwater recharge	1	0.5	5
P5: Maximum amount of groundwater recharge per day multiplier	1	0.5	5
P6: Minimum amount of daily precipitation necessary in arid/semi-arid areas to get groundwater recharge [mm]	12.5	5	15
P7: Maximum canopy storage [mm]	0.3	0.1	1.4

2.4 Performance indicators

Performance of the WGHM was evaluated using the correlation coefficient (CC), Nash-Sutcliffe efficiency (NSE),, root mean square error (RMSE), relative absolute error (RAE), and a modified version of the Kling Gupta efficiency (KGE) with

$$CC = \frac{Cov\,(Obs, Sim)}{\sigma_{obs} \times \sigma_{Sim}} \quad (1)$$

$$NSE = 1 - \frac{\sum_{t=1}^{T}(Sim_{(t)} - Obs_{(t)})^2}{\sum_{t=1}^{T}(Obs_{(t)} - \overline{Obs})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Obs_{(t)} - Sim_{(t)}\right)^2} \quad (3)$$

$$RAE = \frac{\sum_{t=1}^{T}|Obs_{(t)} - Sim_{(t)}|}{\sum_{t=1}^{T}|Obs_{(t)} - \overline{Obs}|} \quad (4)$$

$$KGE = 1 - \sqrt{(CC-1)^2 + \left(\frac{\sigma_{Sim}}{\sigma_{obs}} - 1\right)^2 + (\frac{\overline{Sim}}{\overline{Obs}} - 1)^2 + (\frac{Trend_{Sim}}{Trend_{Obs}} - 1)^2} \quad (5)$$

where *Cov* is covariance function, σ refers to standard division, *Trend* indicates the linear trend of the time series, *Obs* is observed value, *Sim* is simulated value, *t* refers to time counter and *T* is the period length. Optimum values of CC, NSE and KGE are 1, and of RMSE and RE are 0. Trends and overall behaviour of the time series were also analysed.

3 Results

3.1 Model calibration

First, NA was adjusted based on either MODIS data only (variant RS) or MODIS data and information of basin water use (variant RS_Q_GW_NA) (section S3). Then, optimal model parameters were identified using GA and NSGA-II for both variants. Figure 5a shows the calibration history of WGHM based on the best performance of GA among five runs for the variant RS. GA started from a KGE value with respect to TWSA near 0.60 and reached to 0.87 after about 5,000 functional evaluations (WGHM runs). Figure 5b illustrates the final Pareto fronts obtained by five runs of NSGA-II for the variant RS_Q_GW_NA. For the variant RS_Q_GW_NA after about 12,000 functional evaluations (for each NSGA-II run), NSGA-II found 18 optimal parameter sets. Figure 6 shows the parameter ranges (5 and 18 values for each parameter for variants RS and RS_Q_GW_NA) obtained by five different runs of GA and NSGA-II in RS and RS_Q_GW_NA variants. Then, an ensemble of WGHM simulations was generated for the variants RS and RS_Q_GW_NA which comprises the model runs with the optimal parameter sets.

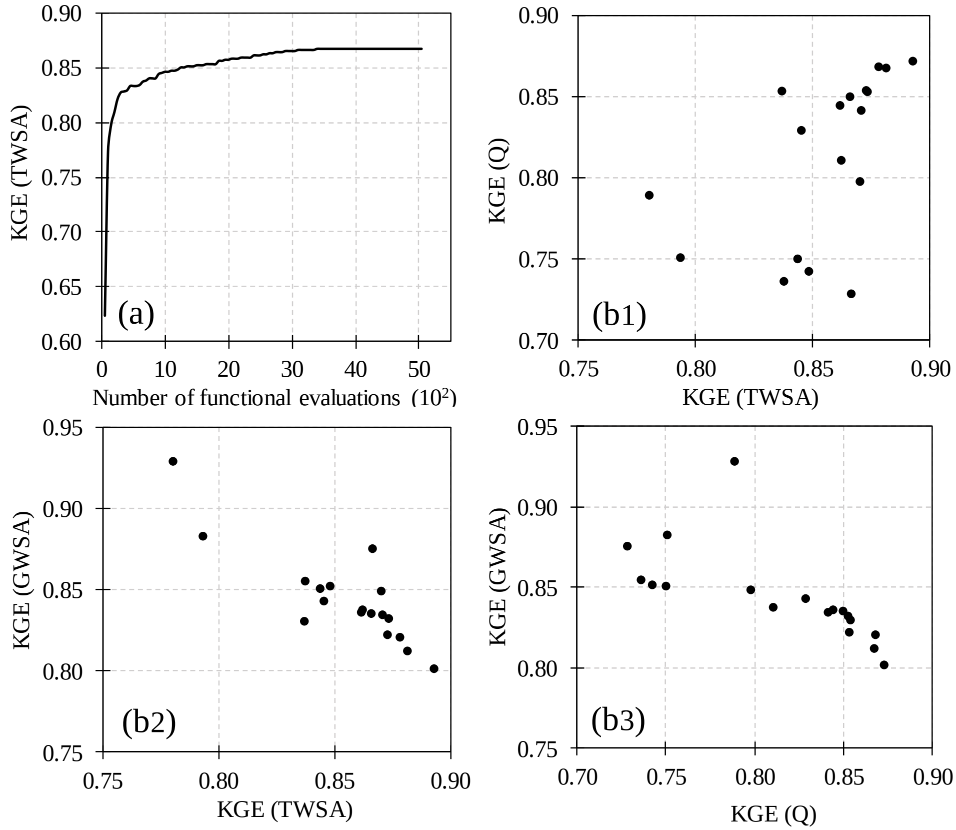


Figure 5: Best convergence history of GA in calibrating WGHM for the variant RS (a) and Pareto fronts for the multi-objective calibrations generated by NSGA-II for the variant RS_Q_GW_NA (b1), (b2) and (b3).

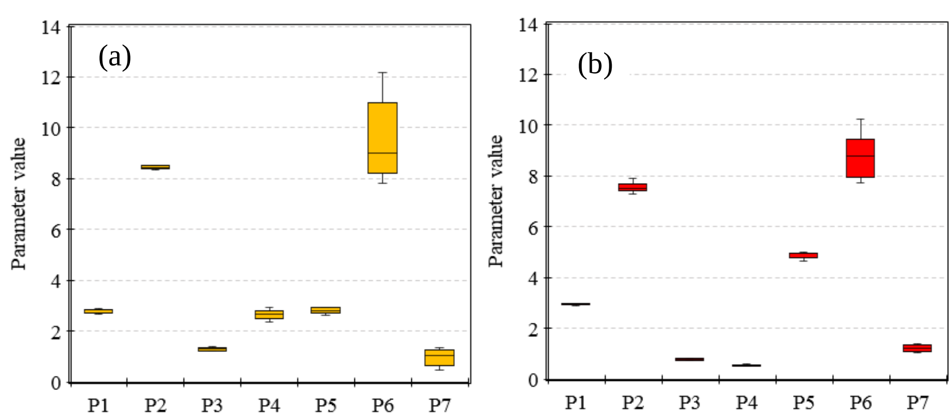


Figure 6: Adjusted WGHM parameter values for variant RS (a) and RS_Q_GW_NA (b).

Figure 7 compares the output of the calibrated model ensembles (variants RS and RS_Q_GW_NA) with observations and the output of the standard version of WGHM. The minimum and maximum value of each variable in each time period are shown as uncertainty bound of the results in each variant. Standard WGHM underestimates total water storage decline in the Lake Urmia basin between 2003 and 2013as compared to GRACE observations. A good fit to GRACE results in calibration variant RS, due to 1) a stronger increase of human water abstractions over time as indicated by MODIS (Fig. S4), 2) an almost tripling of rooting depth and thus soil water capacity (P1), 3) an increased fraction of runoff that recharges the groundwater (P4-P6) and a 4) a higher maximum canopy storage everywhere in the basin (P7) and 5) an increase of maximum active lake depth of Lake Urmia from 5 m to more than 8 m in variant RS (P2) (Figs. 6a and 7a). With the larger soil and canopy water storage capacities, runoff and thus inflow into Lake Urmia decrease as compared to standard WGHM (Fig. 7b). Still, simulated inflows into Lake Urmia computed in variant RS are still much higher than the observed values (Fig. 7b) and seasonality of groundwater storage is totally misrepresented (Fig. 7c). The required reduction of computed lake inflow (Q) can be achieved in variant RS_Q_GW_NA by adjustment of the runoff coefficient and a slight further increase in maximum soil and canopy storage (Fig. 6), while the fit to GRACE TWSA remains good (Fig. 7a). However, the seasonality of groundwater storage could only be achieved by adjusting the sources of total net abstractions in variant RS_Q_GW_NA (Fig. 7c). NAg in the standard and RS variants is negative, which means that there is an artificial groundwater recharge due to irrigation by surface water during the summer irrigation months, leading to an increase in groundwater storage. Groundwater storage observations, however, show a decrease during this period, indicating that irrigation causes a net abstraction from groundwater. Therefore, annual values of NAg as computed by WGHM were multiplied, in variant RS_Q_GW_NA, by a negative correction factors (Table S2).

Performance indicators CC, NSE, RMSE, RAE, and KGE with respect to monthly TWSA (Fig. 7a), annual Q (inflow to Lake Urmia, Fig. 7b) and monthly GWSA (Fig. 7c) are presented in Table 2 for the standard version and the ensemble means of the two calibration variants. Regarding the fit to TWSA observations, NSE increased from 0.48 in the standard version to 0.86 in the RS variant for which TWSA was the only observation considered, and increased slightly to 0.88 when groundwater observations were taken into account in RS_Q_GW_NA variant. This performance improvement is also reflected by CC, RMSE, RAE, and KGE. Although the performance of WGHM with respect to the observed lake inflow was improved in the RS variant, the variant does not yet provide reliable simulations of lake inflow. The calibration against inflow observations in variant RS_Q_GW_NA strongly improves inflow simulation, with NSE and KGE jumping from negative values for the standard variant to values 0.93 and 0.82, respectively. The good performance shown by CC for all model variants indicates that all model variants identify correctly high and low flow years. In case of GWSA, all performance indicators show that consideration of remote sensing data only does not lead to an acceptable simulation of groundwater storage. Only the variant for which groundwater observations were taken into account lead to satisfactory performance.

For model performance evaluation, we compared the lake volume simulated by WGHM with the observed lake volume of Tourian et al. (2015) (Fig.7d and Table 2). The standard model underestimates the decline in both lake water and TWSA, both calibrated variants simulate the TWSA trend correctly, but variant RS, overestimate the decline of lake water storage, thus compensating for not decreasing sufficiently groundwater storage (Fig. 7c) due to assuming a net groundwater recharge due to surface water irrigation. Only variant RS_Q_GW_NA simulates not only the groundwater dynamics but also the decline of lake water volume correctly. KGE for the monthly lake volume anomaly is 0.52 for the standard WGHM and improves to 0.75 for RS. Including groundwater level data further improved the fit to observed lake volume, leading to a very high KGE of 0.89 (Table 2). We conclude that the calibration of WGHM against diverse observations (that do not include lake volume observations) leads to improved simulation of lake volume dynamics.

Table 2: Performance of standard and calibrated WGHM variants with respect to observations of TWSA, inflow to lake, GWSA and lake volume anomaly.

Phase	Variables	Criteria	Standard	RS	RS_Q_GW_NA
Calibration	Monthly TWSA	CC	0.84	0.93	0.94
		NSE	0.48	0.86	0.88
		RMSE [mm]	77	40	37
		RAE	0.72	0.39	0.36
		KGE	-0.36	0.85	0.86
	Annual Q	CC	0.94	0.97	0.97
		NSE	-8.51	-0.75	0.93
		RMSE [10^6 m^3/yr]	4121	1767	356
		RAE	3.92	1.67	0.33
		KGE	-0.61	0.29	0.82
	Monthly GWSA	CC	0.03	0.16	0.95
		NSE	-0.31	-0.28	0.89
		RMSE [mm]	21	20	6
		RAE	1.07	1.04	0.30
		KGE	-0.87	-0.83	0.85
Evaluation	Monthly lake volume anomaly	CC	0.82	0.98	0.98
		NSE	0.68	0.92	0.96
		RMSE [10^6 m^3]	1922	928	656
		RAE	0.51	0.25	0.18
		KGE	0.52	0.75	0.89

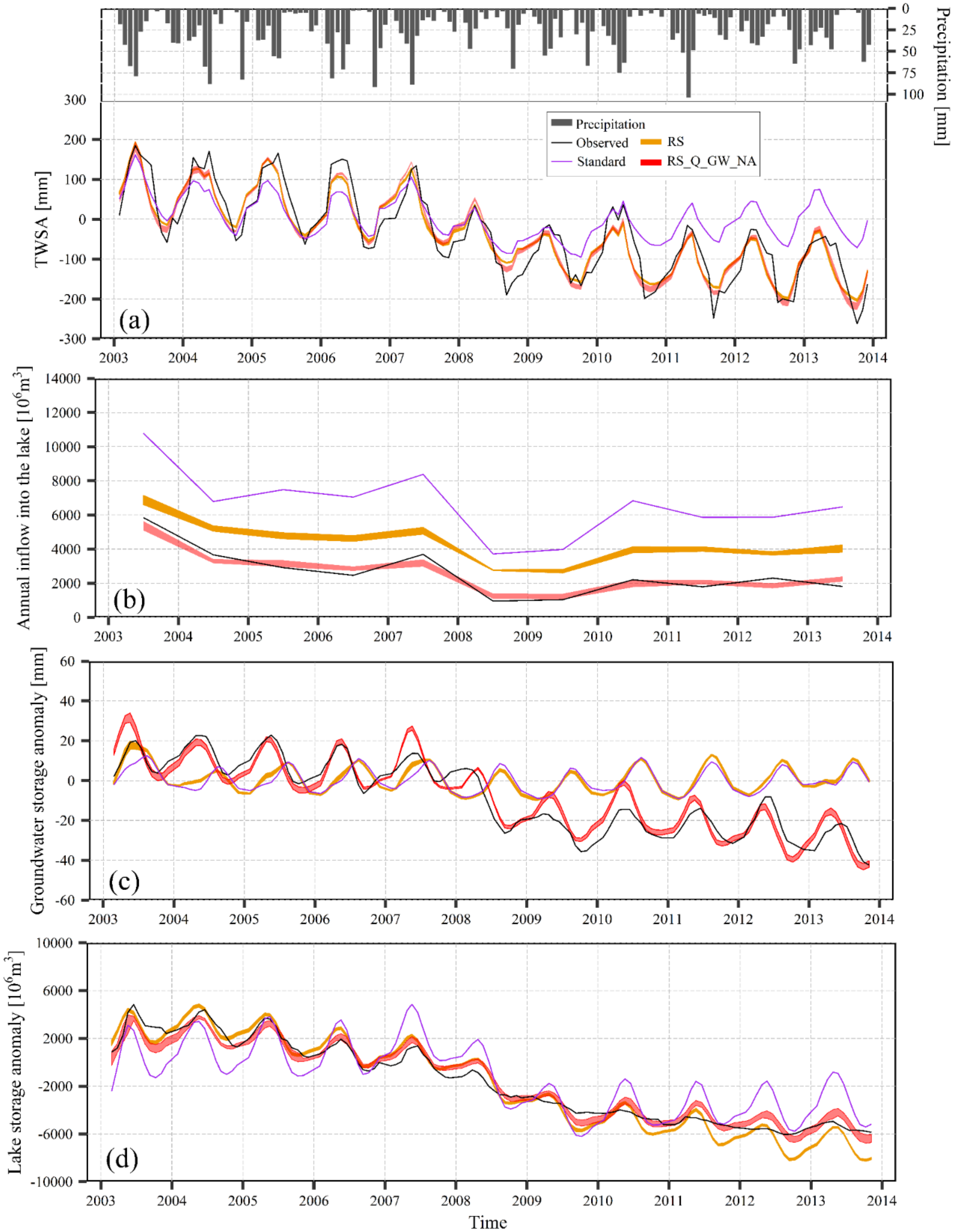


Figure 7: Time series of monthly total water storage anomaly TWSA (a), annual lake inflow Q (b), monthly groundwater storage anomaly GWSA (c) and monthly lake volume anomaly (d), from observations, standard WGHM and the two calibration variants RS and RS_Q_GW_NA.

3.2 Differential impacts of human water use and climate variation on Lake Urmia basin

The impact of human water use and man-made reservoirs on water flows and storages was quantified by comparing the output of WGHM in which human water use and man-made reservoirs are considered (this is normally done, now called WGHM-ANT) with the output of a model run for naturalized conditions, where it is assumed that there are no reservoirs and no human water use (WGHM-NAT). We determined that the results of the naturalized run for annual inflow into the lake differ by less than 2% from a run with reservoirs but without human water use. Therefore, differences between WGHM-ANT and WGHM-NAT outputs can be considered to be caused by human water use. It should be mentioned that all simulated and observed storages (total, groundwater, lake) are not absolute values but anomalies with respect to the mean water storage during 2004-2009 (baseline period used for the provided GRACE data). Moreover, to quantify the uncertainty in the model calibrations, WGHM-ANT and WGHM-NAT were run based on all 18 optimal parameter sets were obtained from Pareto front for variant RS_Q_GW_NA. All results were presented by min-max ranges.

When comparing TWSA under anthropogenic and naturalized conditions in Fig. 8a, remember that TWSA in Lake Urmia basin is dominated by water storage in Lake Urmia. Seasonal TWSA variation of WGHM-ANT and WGHM-NAT do not differ much. Starting after the heavy rain in April 2007 and strongly caused by the lack of spring precipitation in 2008, both WGHM-ANT and WGHM-NAT (as well as GRACE TWSA) show a decreasing trend that is only somewhat more pronounced in WGHM-ANT (Fig. 8a). Thus, this decrease is mainly due to dry climate conditions during the well-known severe drought of 2008, with annual precipitation of only 241 mm, i.e. 74% of the mean value for 2003-2013. Also in the absence of human water use, total water storage would not have recovered after 2009 but would have stayed 50-100 mm below the values occurring before 2008. However, while in WGHM-NAT the minimum storage in late summer, i.e. the period with high irrigation, remains almost at a constant level after 2009, it decreases each year in WGHM-ANT due to consumptive increasing irrigation water use (see Fig. S4). The linear trends of WGHM-ANT and WGHM-NAT TWSA time series for the period 2003-2013 are between -23.6 and -25.1 mm/yr (GRACE: -24.4 mm/yr) and between -10.1 and -11.9, respectively. The TWSA trend for two sub-periods before and after 2008, 2003-2007 and 2009-2013 [-11.7, -18.5] and [-10.6,-16.3] mm/yr, respectively, for WGHM-ANT and only [-1.8,3.3] and [-2.9,-0.6] mm/yr, respectively, for WGHM-NAT. The last-mentioned trends are not significant at the 5% confidence level based on Mann-Kendall's test. According to WGHM, the basin lost, on average during 2003-2013, between $1{,}226 \cdot 10^6$ and $1{,}305 \cdot 10^6$ m^3 water/yr, while in the absence of human water use, it would have lost between $524 \cdot 10^6$ and $618 \cdot 10^6$ m^3 water/yr, i.e. 52-57% less. Of this total water volume between $914 \cdot 10^6$ and $975 \cdot 10^6$ m^3/yr of lake water was lost, while only $523 \cdot 10^6$ and $598 \cdot 10^6$ m^3/yr would have been lost without human water use (Fig. 8b).

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. While mean inflow during 2003-2013 is computed to be between $4{,}323 \cdot 10^6$ and $4{,}685 \cdot 10^6$ m^3/yr under naturalized conditions, it decreases by 39-45% reached to between $2{,}463 \cdot 10^6$ and $2{,}742 \cdot 10^6$ m^3/yr under anthropogenically altered conditions (Fig. 8c). The difference is only 50% of NA as only a fraction of (potential) net abstractions from surface water NAs (required to allow optimal irrigation) could be made 1) due to a lack of water availability in the surface water bodies and 2) because a fraction oft of NAg is provided a decrease in groundwater storage. Since 2008 the inflow into the lake has never reached $3{,}085 \cdot 10^6$ m^3/yr. This is the value estimated to be the minimum environmental water requirement that compensates the amount of annual evaporation from the lake surface (Abbaspour and Nazaridoust, 2007). Therefore, a decrease in lake water storage can be expected for the estimated inflow by WaterGAP between $2{,}463 \cdot 10^6$ and $2{,}742 \cdot 10^6$ m^3/yr during 2003-2013. In WGHM-NAT, the inflow was lower than $3{,}085 \cdot 10^6$ m^3 only in 2008 and 2009. Still, the average inflow into the lake from 2009-2013 of between $3{,}528 \cdot 10^6$ and $3{,}840 \cdot 10^6$ m^3/yr would have been only enough to keep the lake from further losing volume (needed to compensate for lake evaporation). Thus even in the WGHM-NAT, inflow into the lake would not have been enough for a recovery to conditions between 2003 and 2007 (Fig. 8b).

Groundwater storage is estimated to decline by between $239 \cdot 10^6$ and $267 \cdot 10^6$ m^3/yr during 2003-2013 in WGHM-ANT, the decline is only between $24 \cdot 10^6$ and $35 \cdot 10^6$ m^3/yr in WGHM-NAT (Fig. 8d). Different from lake water storage, groundwater storage would have recovered after 2008/2009 if there had been no (increasing) net groundwater abstractions (Fig. 8d, compare Fig. S4b), even though mean groundwater recharge were between 2,340·106 and 3,103·106 m3/yr during 2009-2013 as compared to between $3{,}091 \cdot 10^6$ and $4{,}179 \cdot 10^6$ m^3/yr during 2003-2007. To summarize, human water use was the reason for 52-57% of the total water loss in the basin, for a maximum of 87-90% of the groundwater loss and for 39-43% of the Lake Urmia water loss during 2003-2013, and lake inflow was 39-45% less than it would have been without human water use.

4 Discussion

4.1 Model calibration

Global hydrological models suffer from a high uncertainty, in particular as model inputs are uncertain. For example, climate input data are based on low-density climate observations and information on water use is often very scarce and outdated. For modelling at the global scale, it is generally not possible to obtain, the same detailed data for a specific region compared to the case that modelling this region only. Still, a global hydrological model includes all data for simulating water flows and storages in specific regions of interest everywhere on the globe, and model calibration against multiple (regional) observations is a means for improving the performance of the global model regionally. In this way, efficient simulation of regional water flows and storages can be achieved, possibly as an alternative to a costlier setup of a regional model. More importantly, the regional-scale multi-observation calibration done in this study can serve to inform efforts for global-scale but region-specific multi-observation calibration of global hydrological models that would allow to strongly improve the performance of global hydrological models at the scale that they are made for (Döll et al., 2016).

Remote sensing data are the most accessible data for calibration of global hydrological models, including TWSA from GRACE. Therefore, the model variant RS only used globally available RS data, MODIS and GRACE data products.

However, MODIS data can only be used to determine the temporally variable extent of irrigated areas in dry regions of the globe such that the important adjustment of temporal dynamics of statistics-based irrigated areas is not possible everywhere. GRACE TWSA quantify the anomalies and changes of water storage aggregated over all land water storage compartments such as snow, soil, groundwater, lakes, wetlands, and rivers. Considering GRACE TWSA improved the simulation of the important water storage compartment Lake Urmia. However, the unsatisfactory simulation of inflow into Lake Urmia and of groundwater dynamics clearly shows that a good fit to observed TWSA does not guarantee a good simulation of river flows or groundwater storage. Still, calibration against TWSA did, even if only very slightly, improve model performance also with respect to lake inflow and groundwater dynamics.

To assess the value of using inflow into the lake (Q), groundwater observations (GW) and observed lake volume (LV) time series in model calibration, WGHM was calibrated manually based on some other variants i.e. RS_Q, RS_LV, RS_Q_LV and RS_Q_GW in a step-wise fashion (not shown). Based on the results, by adding discharge data (RS_Q variant), the model was able to simulate TWSA and Q accurately without changing the inputs of the model and only based on modifying the parameters. Groundwater level data were found (variants RS_Q_GW and RS_Q_GW_NA) to be necessary to identify that different from what is estimated by the standard version of WaterGAP, there is more irrigation with groundwater and less with surface water such that a net abstraction of groundwater and not artificial groundwater recharge occurs due to irrigation. Information on groundwater level dynamics with a suitable spatial density is not readily available for most regions of the globe. To simulate groundwater dynamics properly, it was not sufficient to adjust parameters of the hydrological model (in particular two groundwater recharge related model parameters (Fig. 6b), but it was necessary to alter the fractions of net water abstractions that come from groundwater and surface water bodies. Only then, groundwater storage decline by net groundwater abstraction was simulated, and lake water storage decline could be correctly simulated instead of being overestimated when only TWSA and lake inflow data are used for calibration. As in the case of adding lake inflow as calibration data type, no trade-off between the fits to the different data types occurred.

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 compared to variant RS_W_GW significantly and only leads to slight parameter adjustments. This indicates a reasonable simulation of per hectare water consumption for irrigation by the WaterGAP model. To summarize, consideration of more and more observations and other independent data results with improved fits to three types of observations, TWSA, lake inflow, and groundwater dynamics, while at the same time more parameters need to be adjusted (Tables 1 and 2 and Fig. 6). No trade-offs between the fits to the three observational data types occurred in the case of the Lake Urmia basin.

While the introduction of annually varying corrections for NAg and NAs (Table S2) for variant RS_Q_GW_NA leads to the best fit to multiple observation types, it may be preferable to have instead of 11 free parameters just 1, i.e. a temporally constant β.With a temporally constant β of -0.5, the fit to TWSA and inflow to the lake does not change at all, and groundwater storage is only slightly increased in the dry years 2008 and 2009. Thus, given the uncertainty of observed groundwater storage variations, a temporally constant NAg correction factor is sufficient for achieving a good fit for all observations.

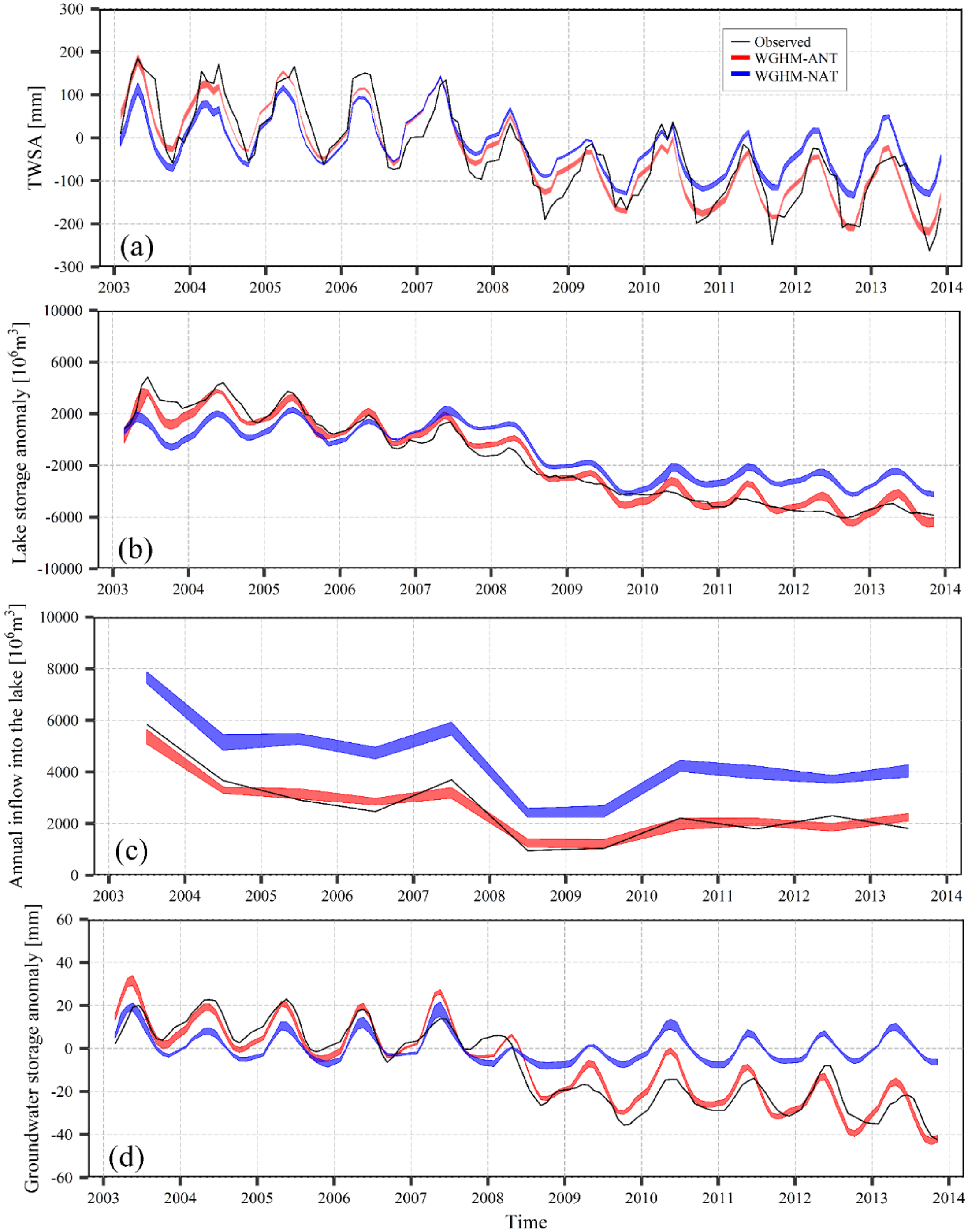


Figure 8: Time series of simulated (variant RS_Q_GW_NA) and observed monthly TWSA (a), lake water storage anomaly (b), annual inflow into the lake Q (c), and monthly groundwater storage anomaly GWSA (d), under anthropogenic (WGHM-ANT) and naturalized (WGHM-NAT) conditions.

In the RS_LV variant, simulation of TWSA and GWSA did not change appreciably but both simulated lake volume anomaly and lake inflow greatly improved as compared to the RS variant. NSE for monthly lake volume anomaly and annual lake inflow reaches 0.95 and 0.44, respectively. Inflow into the lake is much less overestimated than in variant RS. To achieve these fits, the variant RS parameters were adjusted the rooting depth multiplier to 2.5 and setting the potential evaporation multiplier to 2. Adding lake volume observations on top of lake inflow observations in RS_Q_LV variant leads to an improved fit to lake volume observations, with NSE increasing from 0.81 to 0.95, but the fit of observed inflow into the lake slightly worsens from 0.88 in RS_Q to 0.85 in RS_Q_LV. In this variant, the RS_Q variant parameters were used, except the maximum active lake depth was set to 9 m and the potential evaporation multiplier to 2. We conclude that in the case of the end lake, Lake Urmia, calibration against time series of lake volume anomalies could, in the absence of inflow data, help to improve the simulation of inflow, while calibration against time series of inflow could, in the absence of lake volume observation, improve the simulation of lake volume anomalies. Still, calibration to both observational data types leads to the best simulation of both annual lake inflow and lake volume anomalies. However, the groundwater storage dynamics could not be improved without calibration against groundwater level dynamics.

In many hydrological model calibrations, trends are not used as performance criterion. We found that model variants obtained by calibration without a trend criterion, and which have a very similar performance criterion, do not necessarily lead to similar estimates of total and compartmental water losses over the whole time period 2003-2013. For example, using variants RS_LV and RS_Q with similar NSE with respect to monthly time series of TWS, TWS loss between 2003 and 2013 is simulated to be $7.86 \cdot 10^9$ m^3 and $12.20 \cdot 10^9$ m^3, respectively (Table 3). TWS loss according to variant RS_Q_GW_NA (based on ensemble mean) is $9.84 \cdot 10^9$ m^3, even though NSE is only 0.04 higher, while modified KGE (Eq. 5) for RS_LV, RS_Q, RS_Q_GW_NA is 0.68, 0.71, and 0.86 respectively. We conclude that in the case of relevant trends, the calibration criteria should include the minimization of the difference between observed and simulated trends.

Table 3: Water loss in the storage compartments of Lake Urmia basin between 2003 and 2013 as observed and simulated by the WGHM variants that were calibrated using different observation variables.

	Water loss between 2003 and 2013 [10^9 m^3] (mean annual storage in 2003 minus mean annual storage in 2013)							
Storage compartment	Observed	Standard	RS	RS_LV	RS_Q	RS_Q_LV	RS_Q_GW	RS_Q_GW_NA
TWS Storage	9.90	3.62	10.30	7.86	12.20	8.24	9.78	9.84
GW Storage	1.80	0.17	0.33	0.06	0.02	0.03	2.68	2.26
Soil Water Storage	N.A.	0.15	0.26	0.20	0.29	0.24	0.25	0.25
Lake Storage	8.00	3.16	9.53	7.37	11.83	7.78	6.62	7.24

Based on spaceborne TWSA and lake level observations, total water storage in Lake Urmia basin declined by $9.9 \cdot 10^9$ m^3 from its annual average in 2003 to its annual average in 2013 and about 80% was due to the loss of lake water (Tourian et al. 2015). Observed decline of groundwater storage was $1.8 \cdot 10^9$ m^3, i.e. 18% of the observed total water storage loss in the basin. WGHM overestimates observed loss from groundwater in both calibrations variants that take into account groundwater observations. In WGHM simulations, groundwater decline and depletion below the level of surface water storages occur in only 7 out of the 22 0.5° grid cells within the basin (Fig. S5a). In 5 of these 7 grid cells, groundwater levels were stable during 2003-2007 and only declined from 2008-2013, caused by increased NAg and decreased groundwater recharge in the latter part of the study period. It is these 7 cells that cause the basin groundwater decline under the anthropogenic conditions shown in Fig. 8d. For naturalized conditions, peak seasonal water storages decrease somewhat but minimum water storages cannot drop appreciably given the already very low minimum seasonal storage values during the relatively wet five first years of the investigate period (Fig. S5b) because WaterGAP cannot simulate a possible drop of the groundwater table below the surface water level in the absence of groundwater abstractions. Thus, the contribution of human water use to groundwater storage decline might be overestimated as 1) groundwater storage decline under the impact of human water use is overestimated (Table 3, variant RS_Q_GW_NA as compared to observations and 2) groundwater storage decline under naturalized conditions without human water use may be underestimated.

It is worth mentioning that WGHM as a hydrological model that does not include a gradient-based groundwater model has some limitations for studying groundwater-lake water flows. We attempted to calibrate WGHM under the assumption that there are direct water flows between lake and groundwater. Under this assumption, the seasonality of the groundwater storage was strongly misrepresented. Therefore, as accepted by ULRP (2015c), we assumed there is no direct flow between the lake and groundwater. This is consistent with Danesh-Yazdi and Ataie-Ashtiani (2019) who stated that a significant water exchange between the lake and groundwater is unlikely. Also, Amiri et al. (2016) based on isotope and chemical tracer analyses rejected any significant relationship between lake. However, some studies, e.g. Ashraf et al. (2017) and Vaheddoost and Aksoy (2018), stated the opposite. In conclusion, the results of this study support the idea that there are no significant direct interactions between lake and groundwater in the Lake Urmia basin. ~~While Vaheddoost and Aksoy (2018) using traditional hydrograph separation methods claimed that there is a significant relationship between the lake and groundwater, Danesh-Yazdi and Ataie-Ashtiani (2019) rejected their claim. Equally, some studies that applied isotope and chemical tracer analyses (e.g. Amiri et al. 2016) rejected any significant relationship between lake and groundwater. In conclusion, the results of this study support the idea that there are no significant direct interactions between lake and groundwater within the basin.~~

~~4.2 Comparison to human vs. climatic contribution as determined in previous studies~~

4.2 Distinguishing the contributions of human water use and climate variability to lake shrinkage

In order to design the Lake Urmia restoration program, it is vital to know which factors contribute how much to the shrinkage of the lake. All previous studies (e.g. Hassanzadeh et al., 2012; AghaKouchak et al., 2015; Ghale et al., 2018; Chaudhari et al., 2018; Farokhnia et al. 2018) agreed that shrinkage is caused by both climate variations and human activities, but there is no consensus about the relative contributions. For example, Chaudhari et al. (2018) concluded that human-induced changes accounted for 86% of the lake volume decline during 1995-2010, while we determined values of 39-43% for 2003-2013. In line with our results, Farokhnia et al. (2018) showed that due to high climate variability in Lake Urmia basin during 1999-2009, the climate was the dominant factor of lake volume loss, causing 58% of observed loss. According to our study, human water use was the reason for 39-45% inflow reduction into the lake during 2003-2013 which is very similar to the values of Shadkam et al. (2016) for the years 2003-2009 (comp. their Fig. 8). Discrepancies are likely due to different analysis methods, but different analysis periods and conceptualizations make a direct comparison of the estimated contributions difficult. Chaudhari et al. (2018) performed a comprehensive hydrological modelling of Lake Urmia basin. They also studied the land use changes in detail over 1987-2016 and determined a ~98% and ~180% increase in agricultural lands and urban areas, respectively. However, their uncalibrated global hydrological model that represented the basin by 5-6 cells only was not able to simulate well the flows and storages in the basin. For example, simulated annual inflow into the lake was estimated to be $3{,}700 \cdot 10^6$ m^3 in 2003 (their Fig. 8) while observed inflow was much higher, $5{,}835 \cdot 10^6$ m^3. In 2009, observed inflow, with $1{,}036 \cdot 10^6$ m^3, was only half of the simulated one. Therefore, the very high human contribution to the lake volume decline of 86% determined by Chaudhari et al. (2018) may arise from the poor performance of the uncalibrated model. In addition, Chaudhari et al. (2018) studied a considerably longer period, i.e., 1995-2010, that includes the change point of lake dynamic (the year 2000 based on Khazaei et al. (2019) and 2001 based on Fazel et al. (2017)). Although including years prior to 2000 might be lead to different results, some other studies like Shadkam et al. (2016) and Farokhnia et al. (2018), who their modelling included years 2000 and 2001, support the results of the current study. Shadkam et al. (2016) stated that climate change was responsible for three-fifths of inflow reduction into the lake, and the rest was caused by water resources development between 1995-2010. Also, Farokhnia et al. (2018) showed that during a 22 years period ending in 2009, the effect of anthropogenic and climatic factors in reducing the inflow into Lake Urmia was almost equal.

While Ghale et al. (2018) seem to support the results of Chaudhari et al. (2018) as they state that 80% of drying of Lake Urmia is due to anthropogenic impacts during 1998-2010, their statistical analysis assumes that lake inflow from rivers can be considered to reflect "anthropogenic impacts" while precipitation and evaporation reflect climatic variation. However, inflow into the lake is in reality also affected by climatic variations. over the basin. However, although inflow into the lake is surely affected by human water use in upstream, also affected by climatic variations over the basin. Using a statistical change point analysis and without modelling, Khazaei et al. (2019) stated that given the stable conditions of precipitation and temperature, climatic changes could not explain the dramatic decline of the lake level; however, they did not use in-situ data (except lake water level data) for their analysis.variations could not explain the dramatic decline of the lake level. They also estimated the change of vegetation dynamics and its associated hydrological loss in terms of evapotranspiration. They used monthly GPCP precipitation data for assessing the trend of precipitation over the basin. However, the proportion of shared variance between GPCP and in-situ data over the basin is about 0.75 on a monthly scale (see Table 2 in Jalili et al. 2012). Therefore, their analysis suffers from the poor quality of precipitation data. Moreover, their analysis was done on a monthly scale that cannot capture the sub-monthly variability of climatic variables. Also, they did not account for the role of groundwater dynamics in their analysis. Based on an analysis of the Standardized Precipitation Index (SPI), a drought index, AghaKouchak et al. (2015) reported there was no significant trend in droughts over the basin during the past three decades and concluded from this that human activities and not climatic variations were the main reason for lake shrinkage. Different from our study and the modelling studies of Shadkam et al. (2016), Farokhnia et al. (2018), and Chaudhari et al. (2018), these

~~three~~the studies ~~consider~~by Ghale et al. (2018), Khazaei et al. (2019) and AghaKouchak et al. (2015) considered only the dynamics of monthly and annual precipitation and neglect changes in the variability of daily precipitation. During the last three decades, there was a significant increase the frequency of daily precipitation of less than 5 mm and a significant decrease in the frequency of daily precipitation of 10-15 mm, suggesting a runoff reduction even in case of constant annual precipitation (Fig. 2 in Bavil et al., 2018). Hosseini-Moghari et al. (2018) showed that an increasing frequency of days with less than 5 mm precipitation in combination with decreasing monthly precipitation has led to the observed reduced inflow into two dams in the Lake Urmia basin that are located downstream of areas with insignificant human water use. We conclude that for assessing the effect of climatic variability on hydroclimatic variables, the analyses should be done on a daily time scale or ~~smaller~~shorter to consider the change in amount and patterns of variables. Moreover, we examined the ratio of annual inflow into the lake (based on the ensemble mean) over annual precipitation during the study period. This ratio reached maximum values in 2003 (0.29 and 0.41 for the anthropogenic and naturalized conditions, ~~respectivly~~respectively) and minimum values in 2009 (0.07 and 0.15). Averaged over the period 2009-2013, these ratios are, with 0.11 (ANT) and 0.22 (NAT), much smaller than the values for 2003-2007, 0.20 and 0.32. Thus, the drought year 2008 as well as the relatively small ratio of inflow into the lake over precipitation in the last five years of the study period play a significant role in the decline of inflow and lake water storage.

For quantifying human and climatic contributions to observed hydrological changes, a comprehensive modelling approach that takes into account, for example, the impacts of changing temperatures and land use change (e.g., urbanization and cropland expansion) on runoff generation and thus river inflow and on ~~evapotranspiration~~evaporation of the lake itself is preferable~~.~~ to statistical analyses such as trend and correlation analysis. Such ~~comprehensive modelling was done by~~statistical analyses may be misleading about reasons for certain temporal changes. For example, when there is no trend in precipitation but a significant trend in streamflow, it may be concluded that human activities are the dominant case of streamflow reduction; most of the trend studies for Lake Urmia suffer from such a hasty conclusion. In hydrological modelling, more detailed information such as the depth of precipitation in each event, the interval between rainfall events (represented in soil moisture) and other involved elements to generate runoff are considered. All modelling studies (except Chaudhari et al. (2018) who used an uncalibrated model), i.e., Shadkam et al. (2016), Farokhnia et al. (2018) and our study, found that the impact of climatic variations could not be ignored over the basin, while, trend and correlation analysis studies such as Khazaei et al. (2019) and Ghale et al. (2018) stated the climate contribution is negligible compared to anthropogenic impacts. We suggest to do trend analysis of daily precipitation distinguishing different intensity classes (e.g. Bavil et al. 2018). ~~by 5-6 cells only was not able to simulate well the flows and storages in the basin. For example, simulated annual inflow into the lake was estimated to be 3,700·10^6 m^3 in 2003 (their Fig. 8),~~ , ~~while observed inflow was much higher, 5,835·10^6 m^3. In 2009, observed inflow, with 1,036·10^6 m^3, was only half of the simulated one. Therefore, the very high human the lake volume decline of 86% determined by Chaudhari et al. (2018) may arise from the poor performance of the uncalibrated model.~~

As a final word, the irrigated area used in this study obtained from the official report of ULRP (Kamali and Youneszadeh Jalili 2015). However, Chaudhari et al. (2018) estimated the irrigated area significantly less than the irrigated area used in the current study (Figure S3 compared to Figure 9 in their study). They used September for estimation of the irrigated area while the crops are completely matured in July and August in the basin. As a result, some crops are harvested in September. Therefore, it could be the main reason for such a significant underestimation of irrigated areas in the basin by Chaudhari et al. (2018). Also, Alizade Govarchin Ghale et al. (2019) estimated the irrigated area in the basin. Although their result is much closer to Kamali and Youneszadeh Jalili (2015) relative to Chaudhari et al. (2018), they used April and August to the estimated irrigated area, while Kamali and Youneszadeh Jalili (2015) used July and August that lead to some differences. Also, month April that was used by Alizade Govarchin Ghale et al. (2019) includes both irrigated and rainfed farms, the distinction between irrigated and rainfed cultivation may also make some differences. However, due to the fact, Kamali and Youneszadeh Jalili (2015)'s report was approved by the ULRP; we believe that the use of the official report from ULRP would be more reliable than other sources. However, the data reported by Kamali and Youneszadeh Jalili (2015) surly suffer some uncertainties that are inevitable.

4.3 **Limitations**

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). Given the low spatial model resolution (0.5°×0.5°), the model results are preferably aggregated to the basin as a whole as 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 the 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 adopted in this study by scaling it by basin-wide correction factors to better capture the temporal development of irrigation. Calibrated modelling results are also affected by the uncertainties of the observation data. GRACE TWSA data are more reliable for larger (100,000 km^2 according to Landerer and Swenson, 2012) areas than the basin area of 52,000 km^2. 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 and uncertain specific yields. The "observed" lake water volume decline likely underestimates the actual decline as a constant bathymetry was assumed when deriving lake water volume decline from remote sensing of lake water level elevation and lake water area (Tourian et al. 2015). However, there was an increase in the elevation of the lake bottom due to sedimentation and salt precipitation (Shadkam et al., 2016).; Sima and Tajrishy 2013; Karimi et al. 2016).

We determined that the results of the naturalized run with and without reservoirs for annual inflow into the lake differ by less than 2%, whereas Fazel et al. (2017) and Ghale et al. (2018) stated that dams have a significant impact on the lake shrinkage. However, Shadkam et al. (2016) showed the role of dams in the reduction of inflow into the lake did not exceed 5% due to evaporation from reservoirs. Moreover, in this study, the inflow into the lake was assessed on an annual scale, and there is no correlation between the dams' operation and annual inflow in the basin (Fathian et al. 2014). Therefore, the error from this source to our result should be negligible. Also, in this study, it is assumed that there is no significant direct relationship between the lake and groundwater. However, the hydrologic connectivity between the lake and groundwater remains an under-studied aspect of the lake dynamics (Danesh-Yazdi and Ataie-Ashtiani 2019). Finally, the study period 2003-2013 does not include some of the years with significant changes in the dynamics of the lake and the basin (i.e., years 2000 and 2001 that identified as the change point of the lake by Khazaei et al. (2019) and Fazel et al. (2017), respectively) due to data availability. Therefore, our results cannot be generalized to previous decades.

5 Conclusions

This study investigated the differential impact of human water use and climate variations on total, groundwater and lake water storage in the Lake Urmia basin as well as on inflow into the lake during 2003-2013. This was done by utilizing the information contained in multiple types of observation data to calibrate, specifically for the Lake Urmia basin, the global hydrological model WGHM, which takes into account the impact of human water use and man-made reservoirs on flows and storages. Observations include remote sensing data (for irrigated area, TWSA, and lake volume), in-situ streamflow observations (for of lake inflow), groundwater well data (for deriving groundwater storage anomalies) and statistical data on water use in the basin. A time series of observed lake volume was used for evaluation. Using the ensemble of best-performing models where all available observations were used for model calibration, the impact of human water use was determined by comparing the output of naturalized run, with human water use assumed to be zero, with the runs with the historical water use. To understand the value of different observational data types for calibration, WGHM was calibrated in six variants (two auto-calibrated and four manually calibrated) to different combinations of observational data types.

We found that the time series for water demand by irrigation, as assumed in the standard WGHM version, had to be adjusted using MODIS data such that the modification of seven model parameters could result in a good fit to observed GRACE TWSA. Consideration of these remote sensing data somewhat improved the dynamics of both inflow into Lake Urmia and lake water storage, but lake inflow was still overestimated by 66% and the seasonality of groundwater storage was strongly shifted. Additional calibration against observed inflow into the lake did not affect TWSA simulation and slightly improved the simulation of the lake water storage anomaly. Only by using monthly time series of mean groundwater level variations in the basins for calibration, we could adjust the fractions of human water use taken from groundwater and surface water such that seasonality of groundwater storage was simulated correctly. Only then it was possible to simulate the observed groundwater loss, and loss of lake volume was no longer overestimated. Statistical information on sectoral water withdrawals in the basin for one year as well as estimates for sectoral return flow fractions further improved the model, but only slightly. We recommend to include, in case of relevant trends in observations, the difference between observed and simulated trends as one of the calibration criteria, not only differences between time series of daily, monthly or annual values.

The calibration exercise showed that the calibration variant for which the highest number of observational data types were used, WGHM variant RS_Q_GW_NA, showed the best fit to all observations. Certainly, no general conclusions on the worth of specific observation data types for model calibration, including trade-offs among fit to multiple data types, can be derived from this study. Lake Urmia basin is particular with respect to 1) draining into a large end lake that dominates TWSA, 2) the strong impact of human water use and 3) the fact that the standard WGHM version estimates a net recharge to the groundwater due to surface water irrigation, which had to be corrected to a net abstraction. 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~~level 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, such data can be crucial for a correct understanding of the freshwater system.

Based on the good fit of WGHM variant RS_Q_GW_NA to four types of observational data, we found that human water use reduced lake inflow that would have occurred without human water use during 2003-2013 by about 39-45%. About 52-57% of the total water storage loss in Lake Urmia basin and only 39-43% of lake water loss during this time period was due to human water use, and the 43-48% and 57-61%, respectively, to climate variations. 87-90% 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.

GRACE TWSA data indicate an increasing trend in water storage in the basin during 2014-2017 due to both less water use due to water management (ULRP, 2015b) and the wet years 2015/2016. This trend is about half as strong as the decreasing trend during 2003-2013. Further strengthening of efforts for decreasing human water use in the basin should be undertaken, while at the same time, global-scale mitigation of climate change by reducing greenhouse gas emissions to prevent strong decreases of precipitation and runoff. Our study has shown that the 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.

Data availability. In-situ data from "Iran Water Resources Management Company" including groundwater levels, precipitation and temperature are available upon request from the corresponding author. All other data are available in supplementary. Also, 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 from the corresponding author.

Author contributions. S.M.H.M performed the modelling and writing. S. A., M. J. T., and K. E. contributed to editing the manuscript. P. D. contributed to analysis of the results, the discussion, and editing the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. The authors are grateful to Urmia Lake Restoration Program and Iran Water Resources Management Company for providing in-situ data. They also acknowledge the Center for Space Research (CSR) and Potsdam Institute for Climate Impact Research that provide GRACE data and global climate data, respectively. The authors graciously acknowledge Felix T. Portmann and Hannes Müller Schmied from Goethe University Frankfurt for their valuable contributions throughout the WaterGAP calibration, and thank the editor and reviewers for their comments and suggestions that helped to improve the paper.

References

Abbaspour, M., and Nazaridoust, A.: Determination of environmental water requirements of Lake Urmia, Iran: an ecological approach, International Journal of Environmental Studies, 64, 161-169, doi:10.1080/00207230701238416, 2007.

AghaKouchak, A., Norouzi, H., Madani, K., Mirchi, A., Azarderakhsh, M., Nazemi, A., Nasrollahi, N., Farahmand, A., Mehran, A., and Hasanzadeh, E.: Aral Sea syndrome desiccates Lake Urmia: Call for action, Journal of Great Lakes Research, 41, 307-311, doi:10.1016/j.jglr.2014.12.007, 2015.

Ahmadzadeh, H., Morid, S., Delavar, M., and Srinivasan, R.: Using the SWAT model to assess the impacts of changing irrigation from surface to pressurized systems on water productivity and water saving in the Zarrineh Rud catchment, Agricultural Water Management, 175, 15-28, doi:10.1016/j.agwat.2015.10.026, 2016.

Alborzi, A., Mirchi, A., Moftakhari, H., Mallakpour, I., Alian, S., Nazemi, A., Hassanzadeh, E., Mazdiyasni, O., Ashraf, S., Madani, K., Norouzi, H., Azarderakhsh, M., Mehran, A., Sadegh, M., Castelletti, A., and AghaKouchak, A.: 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.

Alizadeh-Choobari, O., Ahmadi-Givi, F., Mirzaei, N., and Owlad, E.: Climate change and anthropogenic impacts on the rapid shrinkage of Lake Urmia, International Journal of Climatology, 36, 4276-4286, doi:10.1002/joc.4630, 2016.

Alizade Govarchin Ghale, Y., Baykara, M., and Unal, A.: Investigating the interaction between agricultural lands and Urmia Lake ecosystem using remote sensing techniques and hydro-climatic data analysis, Agricultural Water Management, 221, 566-579, doi:1016/j.agwat.2019.05.028, 2019.

Amiri, V., Nakhaei, M., Lak, R., and Kholghi, M.: Geophysical, isotopic, and hydrogeochemical tools to identify potential impacts on coastal groundwater resources from Urmia hypersaline Lake, NW Iran. Environmental Science and Pollution Research, 23(16), 16738-16760, doi: 10.1007/s11356-016-6859-y, 2016.

Ashraf, B., AghaKouchak, A., Alizadeh, A., Baygi, M. M., Moftakhari, H. R., Mirchi, A., and Madani, K.: Quantifying anthropogenic stress on groundwater resources, Scientific reports, 7(1), 12910, doi: 10.1038/s41598-017-12877-4, 2017.

Azarnivand, A., and Banihabib, M. E.: A Multi-level Strategic Group Decision Making for Understanding and Analysis of Sustainable Watershed Planning in Response to Environmental Perplexities, Group Decision and Negotiation, 26, 629–648, doi: 10.1007/s10726-016-9484-8, 2017.

Azarnivand, A., Camporese, M., Alaghmand, S., and Daly, E.: Simulated response of an intermittent stream to rainfall frequency patterns. Hydrological Process, 34, 615-632, doi: 10.1002/hyp.13610, 2020.

Banihabib, M. E., Azarnivand, A., and Peralta, R. C.: A new framework for strategic planning to stabilize a shrinking lake, Lake and Reservoir Management, 31, 31-43, doi: 10.1080/10402381.2014.987409, 2015.

Bavil, S. S., Zeinalzadeh, K., and Hessari, B.: The changes in the frequency of daily precipitation in Urmia Lake basin, Iran, Theoretical and Applied Climatology, 133(1-2), 205-214, doi:10.1007/s00704-017-2177-7, 2018.

Beven, K., and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of hydrology, 249, 11-29, doi:10.1016/S0022-1694(01)00421-8, 2001.

Boroughani, M., Hashemi, H., Hosseini, S. H., Pourhashemi, S., and Berndtsson, R.: Desiccating Lake Urmia: A New Dust Source of Regional Importance, IEEE Geoscience and Remote Sensing Letters, 1-5, doi: 10.1109/LGRS.2019.2949132, 2019.

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.

Danesh-Yazdi, M., and Ataie-Ashtiani, B.: Lake Urmia Crisis and Restoration Plan: Planning without Appropriate Data and Model Is Gambling. Journal of Hydrology, 576, 639-651, doi: 10.1016/j.jhydrol.2019.06.068, 2019.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A. M. T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE transactions on evolutionary computation, 6(2), 182-197, doi:1089-778X(02)04101-2 , 2002.

Delju, A. H., Ceylan, A., Piguet, E., and Rebetez, M.: Observed climate variability and change in Urmia Lake Basin, Iran, Theoretical and Applied Climatology, 111, 285-296, doi:10.1007/s00704-012-0651-9, 2012.

Döll, P., Douville, H., Güntner, A., Müller Schmied, H., and Wada, Y.: Modelling freshwater resources at the global scale: Challenges and prospects. Surveys in Geophysics, 37, 195-221, doi:10.1007/s10712-015-9343-1, 2016.

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, Journal of Geodynamics, 59, 143-156, doi:10.1016/j.jog.2011.05.001, 2012.

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

Eimanifar, A., and Mohebbi, F.: Urmia Lake (northwest Iran): a brief review, Saline Systems, 3, 5, doi:10.1186/1746-1448-3-5, 2007.

Farajzadeh, J., Fakheri Fard, A., and Lotfi, S.: Modeling of monthly rainfall and runoff of Urmia lake basin using “feed-forward neural network” and “time series analysis” model, Water Resources and Industry, 7-8, 38-48, doi:10.1016/j.wri.2014.10.003, 2014.

Farokhnia, A., Morid, S., Abbaspour, K., and Delavar, M.: Development of SWAT-LU model for simulation of Urmia lake water level decrease and assessment of the proposed actions for its restoration; (role of anthropogenic and climatic factors on hydrological change of the basin and lake), Iranian Journal of Irrigation and Drainage, 12, 1041–1058 (in Persian with English abstract), 2018.

Fathian, F., Morid, S., and Kahya, E.: Identification of trends in hydrological and climatic variables in Urmia Lake basin, Iran, Theoretical and Applied Climatology, 119, 443-464, doi:10.1007/s00704-014-1120-4, 2014.

Forootan, E., Rietbroek, R., Kusche, J., Sharifi, M. A., Awange, J. L., Schmidt, M., Omondi, P., and Famiglietti, J.: Separation of large scale water storage patterns over Iran using GRACE, altimetry and hydrological data, Remote Sensing of Environment, 140, 580-595, doi:10.1016/j.rse.2013.09.025, 2014.

Ghale, Y. A. G., Altunkaynak, A., and Unal, A.: Investigation Anthropogenic Impacts and Climate Factors on Drying up of Urmia Lake using Water Budget and Drought Analysis, Water Resources Management, 32, 325-337, doi:10.1007/s11269-017-1812-5, 2018.

Hasemi, M.: A socio-technical assessment framework for integrated water resources management (IWRM) in Lake Urmia Basin, Iran, 2011.

Hassanzadeh, E., Zarghami, M., and Hassanzadeh, Y.: Determining the main factors in declining the urmia lake level by using system dynamics modeling, Water Resources Management, 26, 129-145, doi:10.1007/s11269-011-9909-8, 2012.

Hesami, A., and Amini, A.: Changes in irrigated land and agricultural water use in the Lake Urmia basin, Lake and reservoir management, 32, 288-296, doi:10.1080/10402381.2016.1211202, 2016.

Hosseini-Moghari, S.M., Araghinejad, S. and Ebrahimi, K.: Monthly Precipitation Assessment: a misleading tool for understanding the effects of climate change, 8th Global FRIEND-Water Conference, Beijing, China, 6-9 November 2018.

Jalili, S., Kirchner, I., Livingstone, D. M., and Morid, S.: The influence of large-scale atmospheric circulation weather types on variations in the water level of Lake Urmia, Iran, International Journal of Climatology, 32, 1990-1996, doi:10.1002/joc.2422, 2012.

Karimi, N., Bagheri, M. H., Hooshyaripor, F., Farokhnia, A., and Sheshangosht, S.: Deriving and evaluating bathymetry maps and stage curves for shallow lakes using remote sensing data. Water Resources Management, 30(14), 5003-5020, doi:10.1007/s11269-016-1465-9, 2016.

Kelleher, C., McGlynn, B., and Wagener, T.: Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding, Hydrology and Earth System Sciences, 21, 3325–3352, doi:10.5194/hess-21-3325-2017, 2017.

Khatami, S., Peel, M. C., Peterson, T. J., and Western, A. W.: Equifinality and Flux Mapping: A new approach to model evaluation and process representation under uncertainty, Water Resources Research, 55(11), 8922-8941, doi:10.1029/2018WR023750, 2019.

Khazaei, B., Khatami, S., Alemohammad, S. H., Rashidi, L., Wu, C., Madani, K., and Aghakouchak, A.: Climatic or regionally induced by humans? Tracing hydro-climatic and land-use changes to better understand the Lake Urmia tragedy, Journal of Hydrology, 569, 203-217, doi:10.1016/j.jhydrol.2018.12.004, 2019.

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

Madani, K., AghaKouchak, A., and Mirchi, A.: Iran's socio-economic drought: challenges of a water-bankrupt nation, Iranian Studies, 49, 997-1016, doi:10.1080/00210862.2016.1259286, 2016.

Mirjalili, S.: Genetic algorithm. Evolutionary Algorithms and Neural Networks, Springer, 43-55, 2019.

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.

Pengra, B.: The drying of Iran's Lake Urmia and its environmental consequences, UNEP-GRID, Sioux Falls, UNEP Global Environmental Alert Service (GEAS), 2012.

Saemian, P., Faridzad, M., Youneszadeh, S., Emadzadeh, M., Jalilvand, E., Farahmand, H., and Taheri, M.: Estimation of Evapotranspiration, water saving and agricultural water consumption at Urmia Lake Basin, 2010, Report of Urmia Lake Restoration Program (ULRP)(In Persian), 2015.

Scrucca, L.: GA: a package for genetic algorithms in R, Journal of Statistical Software, 53(4), 1-37, doi:10.18637/jss.v053.i04, 2013.

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.

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.

Sima, S., and Tajrishy, M.: Using satellite data to extract volume–area–elevation relationships for Urmia Lake, Iran, Journal of Great Lakes Research, 39, 90-99, doi:10.1016/j.jglr.2012.12.013, 2013.

Tabari, H., and Talaee, P. H.: Analysis of trends in temperature data in arid and semi-arid regions of Iran, Global and Planetary Change, 79, doi:1-10, 10.1016/j.gloplacha.2011.07.008, 2011a.

Tabari, H., and Talaee, P. H.: Temporal variability of precipitation over Iran: 1966–2005, Journal of Hydrology, 396, 313-320, doi:10.1016/j.jhydrol.2010.11.034, 2011b.

Tourian, M. J., Elmi, O., Chen, Q., Devaraju, B., Roohi, S., and Sneeuw, N.: A spaceborne 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.

Tsou, C. S.: nsga2R: Elitist Non-dominated Sorting Genetic Algorithm based on R, R package version 1.0, Available at: https://CRAN.R-project.org/package=nsga2R, 2013.

ULRP: Evaluation and deduction of Urmia Lake Restoration Program need at least 3 years, Available at: http://ulrp.sharif.ir/en/news/news-1435-evaluation-and-deduction-urmia-lake-restoration-program-need-least-3-years, 2015b.

ULRP: The trend of Lake Urmia changes, Available at: http://ulrp.sharif.ir/fa (In Persian), 2015a.

ULRP: Urmia Lake - Causes of shrinkage and potential threats. Available at: http://ulrp.sharif.ir/fa (In Persian), 2015c.

Vaheddoost, B., and Aksoy, H.: Interaction of groundwater with Lake Urmia in Iran. Hydrological processes, 32(21), 3283-3295, doi:10.1002/hyp.13263, 2018.

Werth, S., and Güntner, A.: Calibration analysis for water storage variability of the global hydrological model WGHM, Hydrology and Earth System Sciences, 14, 59, doi:10.5194/hess-14-59-2010, 2010.

Zeinoddini, M., Tofighi, M. A., and Vafaee, F.: Evaluation of dike-type causeway impacts on the flow and salinity regimes in Urmia Lake, Iran, Journal of Great Lakes Research, 35, 13-22, doi:10.1016/j.jglr.2008.08.001, 2009.