Journal: Hydrology and Earth System Sciences
Title: 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
Article Iteration: revision 1

# 1 Overall assessment

The aim of the study essential is to quantify the impact of human activities (mostly in terms of water consumption) vs climatic changes on the Lake Urmia water balance. Even though 3 of the referees provided detailed reviews and pointed out to several shortcomings of the paper mostly on model setup and uncertainty, the authors' revision is minimal and in fact insufficient as many comments are effectively ignored.

All reviewers except Chaudhari took issue with the experiment design, particularly the model set up, input data time period, and lack of adequate model calibration and evaluation. Yet, authors have not changed the experiment design, and only added two performance metrics. No uncertainty or sensitivity analysis was conducted whatsoever, which is a common analysis required for any hydrological modeling study. This is even more serious as the manuscript is in fact inconsistent on the issue of uncertainty. While authors discussed the limitations of the work such as parameter uncertainty, no account of the hydrogeology of the lake, assuming constant bathymetry, among others; not only they have not accounted for these uncertainty sources by even a simple uncertainty/sensitivity analysis, they kept pushing that their study is "*a holistic and reliable modelling approach*", "*we are confident that human water use reduced lake inflow that would have occurred without human water use during 2003-2013 by about 41%*", and "*This study proved 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*", among other instances of false overpromises.

The manuscript is filled with redundant discussions (either well established in the literature or not relevant to the core research question of the paper), and is written in poor language with several typos. It is a waste of the editor's time and reviewers to resubmit a manuscript that has not been proof-read especially that most reviewers pointed this out. Further, some of the sources are either not peer-reviewed or in Persian. While local knowledge can be useful, the credibility of a non-peer reviewed source is always questionable. In such cases, authors should provide reason and demonstrate clearly instead of *just* referring to the source.

In my evaluation, the manuscript is **rejected**. If authors wish to resubmit the work, they must revise the modeling experiment to sufficiently address the comments by reviewers, and by addressing I specifically mean to change their modeling setup by providing a more robust calibration than a simple and

insufficient trial and error, transparently explaining the modeling setup to ensure (somewhat) the reproducibility of the study, and perform some sensitivity or uncertainty analysis. Also, remove all the redundant discussions from the manuscript and focus their discussion on the relevance of the results to the lake given the uncertainties. While most of the comments by reviewers still hold, here I highlights a few urgent ones.

### 1.1. Problem description

The problem description (i.e. drivers of the lake desiccation) particularly in the introduction as well as the later discussion of the results is problematic.

It is misleading and inaccurate to stack the climatic and human drivers together (e.g. page 29 lines 1-10). *Khazaei et al.* [2019] disentangled these two: they compared the influence of these two sets of drivers for the lake drying and demonstrated the regional human activities (including water management, but not limited to) are the primary drivers compared to climatic changes (including atmospheric droughts). *AghaKouchak et al.* [2015] also argued that droughts cannot be the primary driver. These two studies, among others, are based on directly analyzing the data themselves, without relying on a model of the system which in most cases are inadequate. While these studies have their own shortcomings, as any scientific study has, your modeling results are inadequate to challenge them. As pointed out by the reviewers your modeling setup has several issues. Inadequate models, regardless of the extent of their inputs and their results, are inadequate.

Authors said: "*This study proved 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*".

This conclusion is in direct contradiction with *AghaKouchak et al.* [2015] conclusion that "*a satellite-based gauge-adjusted climate record… of Lake Urmia basin's Standardized Precipitation Index… indicates no significant trend in droughts over the past three decades at the 0.05 significance (95% confidence) level... In fact, the region has experienced more severe drought events in the past (e.g., 1997–2002) that did not lead to a substantial change in the lake's surface area. Thus, we caution against overrating the role of droughts in the disruption of the lake's water balance to the extent that would cause such a massive shrinkage*". Given that *AghaKouchak et al.* [2015] directly analyzed the historic data of the lake without relying on any inadequate model of the lake system, it is reasonable to say this contradiction indicates the shortcoming and (un)reliability of the model set-up in this study. Notwithstanding the unscientific language of this sentence. Science is not in the business of proving anything. In science we demonstrate and approximate. This is more so the case when it comes to hydrological modeling with numerous types and sources of uncertainty including both model structure and data.

Also, what do you mean by “climate variations”? Are you referring to only *natural climate variability* or *climatic changes* which include both natural variability and human-induced changes?

### 1.2. Introduction

The introduction is very long, has redundant sections:

- Figure 3 has already been published and discussed by Aghakouchak and been repeated many times in the literature. (also pointed out by the referee 1)
- Citation to Zarghami (2011) on page 29 is irrelevant.
- Last 2 paragraphs are “method” material and not introduction

Another issue is that some of the sources are not peer-reviewed. Whether right or wrong, it is hard to rely on such sources. While local knowledge and literature may be a valuable source, it has to treated with caution, not to propagate any errors. So, I am hesitant to accept such discussions.

### 1.3. Modeling setup and results

**Time period:** 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.

**Model setup:** as pointed out by R1 comment 2 (and a few other comments), the model setup is not transparent and justifiable. Surely, documenting model setup (even as supplement) is more necessary than many redundant discussions. As a test, an adequate model setup should work against the updated observation (2013-present). Can the authors demonstrate this?

**Standard model:** inclusion and discussion of the standard model, as a reviewer pointed out, in unnecessary given that this model is not calibrated for this catchment.

**Model calibration and evaluation:** the model calibration is also questionable. It is only based on trial and error so the identified parameter sets are not reliable. Also, it is not clear how sensitive the model results are to these parameters. There is issue about over-parametrizing the model as in each variant new data is added (issue about correction factors pointed out by reviewers). A single model run for each variant is not sufficient, even if the calibration was done through an automatic parameter space search scheme. This is well-established in the literature and a model ensemble is required to account for the uncertainties, even tough partly. If the model is run on a daily basis, authors should be transparent and present the daily results too/

Monthly and annual performance metrics are usually high for most models. The devil is in the details though. On table 4, the flow is only calibrated on an annual scale and not evaluated at all. There is significant seasonality in this region. It is quite possible that seasonal errors are just canceling each other out and give seemingly good annual results. This can be seen on figure 8 where the model exhibits unrealistic seasonality which is not in the observed data, e.g. in Fig 8e, there is generally a negative bias in the first half of the RS_Q_GW_NA simulation, and positive bias in the second half (red line is first below the black line systematically, and then above it).

**Uncertainties:** as Referee 1 pointed out in their comment 6 and elsewhere, the uncertainties are playing a crucial role here. While authors added a section of uncertainty discussion, it is an ad-hoc discussion. While the discussion obviously undermines the experiment design and the results, they have not revised the experiment design to account for these uncertainties, and also they keep overpromising about the reliability of their results. Further, their discussion of uncertainty shows a lack of understanding about the area of model uncertainty. For instance, they said "*Model parameter uncertainty was reduced by the comprehensive multi-observation calibration*". Parameter uncertainty will not be reduced by *just* adding more data to the model; data uncertainty matters, "*garbage in, garbage out*" [*Kuczera et al.*, 2010]. The authors must demonstrate how adding input to the model reduced parameter uncertainty, while justifying the credibility of the data themselves. They have not done any uncertainty or sensitive analysis whatsoever.

# References


AghaKouchak, A., H. Norouzi, K. Madani, A. Mirchi, M. Azarderakhsh, A. Nazemi, N. Nasrollahi, A. Farahmand, A. Mehran, and E. Hasanzadeh (2015), Aral Sea syndrome desiccates Lake Urmia: Call for action, *Journal of Great Lakes Research*, *41*(1), 307-311.

Khazaei, B., S. Khatami, S. H. Alemohammad, L. Rashidi, C. Wu, K. Madani, Z. Kalantari, G. Destouni, and A. Aghakouchak (2019), 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.

Kuczera, G., B. Renard, M. Thyer, and D. Kavetski (2010), There are no hydrological monsters, just models and observations with large uncertainties!, *Hydrological Sciences Journal*, *55*(6), 980-991.