

Earth Syst. Sci. Data Discuss., referee comment RC2
<https://doi.org/10.5194/essd-2021-151-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on **essd-2021-151**

Anonymous Referee #2

Referee comment on "An integrated dataset of daily lake surface water temperature over the Tibetan Plateau" by Linan Guo et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-151-RC2>, 2021

The authors present a dataset with lake temperature data for the Tibetan Plateau (TP). They extrapolated lake surface water temperatures (LSWT) as measured by satellite sensors back in time to obtain a complete LSWT dataset also for the time prior to 2000 when such data became available from the MODIS sensor. They used a simple physical model for this purpose which was trained with MODIS LSWT and forced by air temperature measurements from the extremely sparse station network on the TP.

Both the data itself and the data article would benefit greatly from a more detailed analysis and description of uncertainties, model limitations, resulting bias/patterns etc. to prove that this data is reliable and may be used by others for the varied purposes mentioned in the paper introduction.

There are three major aspects I strongly recommend the authors to address:

1) Data presentation: This paper is about time series of temperature data, but it does not provide any time series plots, only maps where each time series is reduced to a single point. The maps are useful to get a spatial overview and reveal spatial patterns, but they are not at all suitable to assess/present any temporal or seasonal aspects and also simply don't show how the data looks like. This makes it impossible for the reader to judge whether this data would suit their purpose without downloading and analysing the data themselves - which ought to be the main purpose of a data article. I suggest additional data quality checks/analyses/comparisons over time, for different seasons and temperature ranges. It would be especially useful to compare the modelled data with independent measurements/data for the time span 1978-2000.

The data itself: A brief analysis of the T time series in the TPLxxx.csv datasets revealed the following points, all of which should be addressed in the article in order to properly

characterise the data:

- for most lakes, the model data does not differ much between different years, and for many lakes it even seems to consist of a series of nearly exact clones of annual T curves. These often have a characteristic kink in spring which deserves to be addressed to exclude bias/inappropriate model behaviour (these kinks are at ca. -2, so likely related to thawing processes, but then again this feature doesn't show up on all lakes and seemingly also not in the MODIS LSWT data).
- the modelled data are way more smooth in time than the MODIS LSWT data and don't reproduce any extreme values or variability.
- for the period before 2000, there seems to be a fixed annual maximum temperature for each lake that doesn't change.
- there are gaps in the MODIS LSWT data, but it is nowhere described where these come from.
- No data/temperature uncertainties are provided. An uncertainty range/measure should be included at least for the model data.

2) LSWT: To give the readers an idea on the nature and reliability of this data, the introduction should contain an explanation of LSWT - what it is, how it is measured, including a description of the different available satellite sensors/datasets (with special emphasis on the data and validation products used in this study), and what the limitations of such measurements are.

Also, you mention several other LSWT datasets (Wan et al., 2017, Zhang et al., 2014) and that yours is different - but different in what regard, and how does it compare to these datasets? I am missing a clear introduction on the differences and strengths/weaknesses/advantages of the existing datasets to justify this new dataset, and most of all to help users decide whether this is the right data for their purpose.

3) Model: More background on the air2water model would help to prove that this model is fit for the purpose and can produce reliable extrapolation results. This also includes the forcing data: are the sparse station measurements really fit for the purpose? From the information given right now, it cannot be judged whether this model was used as intended by the original authors - or taken out of context/pushed beyond its limitations.

Detailed comments (P=page, L=line)

Data:

- lake_info.csv: Please add the encoding of the .csv file (to the readme, for example) - common international encodings corrupt the chinese signs which may cause trouble for users

Abstract

- data for 160 lakes from RS and modelling: what RS data, what (kind of) model? Please state here.

- validated against in-situ observations for 2000-2017, extended back to 1978
- comparable to LSWT from AVHRR and ARC-Lake - has there been done a comparison, or you just claim that? If yes: give some numbers, if not: please do a comparison.
- "unique for its time span and resolution" - what is the resolution? Please state here.
- trend analysis: warming rate increases from 0.01 to 0.4 per decade - for which periods? Please clarify here.
- mostly in winter - what does "more evident" mean? Please clarify here.
- "integrated data and methods" - what methods? Please clarify already in the abstract.

Introduction

- What is lake surface water temperature, in this context? Please explain what the satellites measure, and what in-situ sensors measure (since your dataset is based on both). This is important background to understand the data. When comparing/validating with in-situ measurements, as in this paper, you need to consider that in-situ- and satellite-based sensors usually measure something different, and address this.
- p2, L38: but some TP lakes have in-situ data? Please add some more information on that, as this would provide a very important validation data source for this paper. Have you consulted this data? If not, why not?
- P2, L41: a reference from 2010 looks a bit outdated for a claim on "increasingly higher temporal and spatial resolutions" in 2021 as a lot has happened since that paper was written. Reconsider its relevance / correct placing in the text.
- P2, L141ff: What I am missing here is an overview over sensors that measure LSWT. Both the sensors/data sources in themselves (i.e. individual data/scenes), and for the datasets you mention here.
- P2, L45f: How can these datasets be global and nevertheless spatially incomplete on the TP? Please clarify.
- P2, L47: What sensors?
- P2, L61ff: OK, you go for the simplified model because you have barely any in-situ observations and this model still works with limited in-situ data - but what input is your model based on, then? This doesn't sound so convincing. Please rewrite with examples or some more information to clarify (for what area/setup has this model been tested/developed, for the TP or somewhere else?).
- P2, L66: LSWT from which sensor? and does the air2water model work with LSWT data, if it was developed for / is aiming at in-situ measurements?

Study area

- P3, L75: why "most" but not all, how did you select?

Methods

- P4, L90ff: What is the accuracy/uncertainty of this data? (How much can one trust these temperatures? Any danger for a consistent bias?)
- P4, L93: what do the times in brackets mean? Night equator overpasses? Please clarify.
- P4, L95f: what is the difference between MOD11A2 and -A1? Only A1 is explained.
- P5, L100f: different in what regard? Just the method, or also the temperatures you get - how do they compare, and why is your method better? Please clarify.
- P5, L106: Are there data gaps, and/or was there any filtering applied to the data, both in space and time?

- P5, L110: How are these surface temperatures measured? How accurate/reliable/representative are the measurements? This plays a huge role for your model setup and thus dataset, as very few stations are forcing the model for the majority of the lakes. And how long are the records?
- P5, L125: As far as I am aware, many of the TP lakes are very shallow. Is a deep water temperature of 4 degrees realistic at all times of the year?
- P5f, L126ff: please explain briefly in words what this model does, and what parameters play the strongest role here, for those users who are not that familiar to read/think in equations.
- P6, L138f: there are very few stations on the TP, and none whatsoever in 1/4 of the area you look at, and where the majority of the addressed lakes lie. At the same time, the TP has a very varied topography and greatly varying influence of different weather systems (I/E Monsoon, Westerlies). Interpolation between "nearby" stations doesn't seem valid at all to me, questioning the entire model setup and thus the entire dataset. Have you considered uncertainty in these data? How much do the modelled lake T change if you alter the interpolated air T input (sensitivity analysis)?
Reanalysis data are maybe not perfect either in this area (due to too little measurement input), but they consider way more physics in deriving the temperature spatial pattern than a simple spatial/elevational interpolation. Have you tried to downscale air T for the lake locations from reanalysis data? How does that change the modelled lake T?
- P6, L145f: Please introduce that dataset (and AVHRR) properly so that the readers can judge your comparison/validation.

Results and discussion

- P6, L151: The ARC dataset also needs to be introduced properly in order for the readers to be able to judge your comparison/validation. What about in-situ lake T measurements (which you mention in the introduction)? These would provide a really valuable validation dataset.
- P7, L156f: An R2 of 0.6 isn't great, and "highly comparable" doesn't sound convincing when the temperatures should be identical if both datasets were correct - and both datasets are from remote sensing (not completely different ways of measuring), so one wouldn't expect such big differences.
- P7, L157: How did you compute the bias, and which of the two datasets is warmer? Consistent biases of -3 to +5.6 degrees seem very much.
- P7, Fig. 3a: Some of the poorest agreement between your data and TPlake_Temp data are clustered on the NW TP. Why, have you looked into that?
- P7, Fig. 3b: all except one station have a positive bias, for many lakes, the bias is <2.5 degrees. This seems a lot. Why could that be?
- P7, L166ff: not sure whether I understand correctly: these 11 plots include all lakes within the ARC dataset within your study area? Which lakes are these? Adding a map as in Fig. 3 for the comparison with the ARC dataset would help greatly for the readers to compare the data more easily.
- P8, Fig. 4: The datasets agree well in a linear way up to a point at ca. 10-15 deg C ARC-Lake T, where the relationship suddenly changes. Why is that?
- P8, Fig. 4: How did you compute the bias here? As the fitting line is not 1:1, the bias increases/decreases with measured T, which should be addressed.

- P8, L176f: please state these numbers for the validation period, excluding the training period.

- P8, L177ff: again, what do you mean by bias? And how do the results compare for different times of the year, or different water temperatures? Or distance to the weather stations (forcing data)? These aspects reflect model parameters better than latitude/elevation etc.
- P9: what is the uncertainty of your temperature data?
- P9: it is hard to get a feeling of the data from these maps only. I am missing time series from a few example lakes with different location/elevation/temperature/size/performance.

- P9, L197: label these lakes in the map/figure (the same accounts for other lakes mentioned in the text)
- P10, L198ff: also here, some time series plots would be useful to show these trends. Are they visible in the entire data range, or just part of the data (for example the period where you have measurements)? How do they compare to the forcing air T trends, are the trends shown here maybe simply reproducing air T trends? This aspect deserves discussion to undermine the validity and reliability of your trend analysis.
- P10, L202ff: you argue that cold glacier water could have led to decreasing summer water T but state no source for this - and as a matter of fact, the glaciers in this part of the TP have been growing in recent decades. I thus suggest you remove speculations on the reason of the pattern (alternatively, you need to considerably extend this discussion and undermine it with literature).

Data availability

- P11, L216: Nice that you are sharing your code! A quick recap on what this contains (data preprocessing only, none of the modelling?) could be useful here. I quickly checked the code and see there's a variable error on line 14, you may want to fix that.

Conclusions

- P11, L226: "successfully": You never show a validation for the period 1978-2000, thus I disagree with that claim.

Figures

- All figures need to clearly state what data period they represent (not currently the case). Also, in some cases the used colour maps are rather little intuitive (usually, green = good and red = bad). Consider a single colour / two-colour diverging scale, varying symbols or symbol sizes to improve readability.