

Please find our response in italics.

RESPONSE TO ANONYMOUS REFEREE #2

My recommendation is to publish the paper after some minor comments have been addressed:

Accuracy of 1.5 degrees is problematic. This is a very large discrepancy. Can you justify why this is sufficient for climate change studies?

An RMSE for $T_{0.50m}$ of 1.51 °C was calculated without any correction. If we remove the points affected by clouds as suggested by Reviewer #1 the RMSE decreases to 1.24 °C. If we further use an average correction of $\Delta T = -0.41$ °C ($\Delta T_c = -0.46$ °C, according to COARE results, and $\Delta T_w = 0.05$ °C, according to measurements at Bimont), the RMSE decreases to 1.16 °C. This is still a rather high uncertainty, but after the correction the satellite-based estimations of $T_{0.50m}$ are in average unbiased (mean absolute bias of 0.02 °C).

A high uncertainty means longer time series may be necessary to determine trends (especially the feeblest ones). After deseasonalizing the 1999-2016 time series for the five validation sites, we calculated the linear trends in the table below (we indicate in bold the trends with $p\text{-value} < 0.10$).

Although the time series is relatively short (18 years), temperature trends may already be detected in two of the water bodies. If climate change is having an effect on the surface temperature of the other water bodies, longer data series may be necessary to detect this effect. The usefulness of the LakeSST for climate change studies will increase in the future as more data becomes available.

<i>Lake code</i>	<i>Trend (°C/yr)</i>	<i>p-value</i>	<i>n</i>
<i>AUL13</i>	<i>-0.038</i>	<i>0.335</i>	<i>109</i>
<i>BIM13</i>	<i>-0.010</i>	<i>0.637</i>	<i>141</i>
<i>LPC38</i>	<i>-0.045</i>	<i>0.406</i>	<i>58</i>
<i>PAV63</i>	<i>0.049</i>	<i>0.030</i>	<i>171</i>
<i>SCR04</i>	<i>0.022</i>	<i>0.081</i>	<i>454</i>

P2. L6: “4” should be “four”. Any number less than 10 should be spelled out.

Ok.

Measurements with a thermal radiometer over the water bodies would be ideal. Then a more robust comparison of temperatures with radiometric temperatures is possible. Radiometric temperatures are a much better indicator of the performance of the surface temperature retrieval from satellite data.

We totally agree with reviewer #2 but we do not have surface temperature data measured by radiometry yet. We are reflecting on how to obtain such data either by acquiring radiometric measurements through collaborations or by obtaining funding for the necessary instrumentation and field work.

Include some comments about the SCL-off problem in Landsat-7, and how this problem is addressed in the presented dataset. Discussion about why was Landsat 8 not included is necessary.

On May 31, 2003 the Scan Line Corrector of the Landsat-7 satellite failed. As a result, the measurement scans cannot be corrected for the forward motion of the satellite and about 22% of an image data are lost. The gaps are less important in the centre of the image and increase towards the edges. Since this problem does not affect the radiometric and geometric corrections, SLC-off data could still be used for the creation of the dataset. No interpolation was applied to fill the data gaps. We recommend using the median SST provided in the dataset as an approximation of average lake skin surface temperature, since the median is not very sensitive to missing data and outliers.

When this study was started, the issues with Landsat 8 TIR bands made us discard Landsat 8 data. Besides, the algorithm correcting stray light effects was only implemented into the Landsat processing system on February 2017. At the moment, some of the authors, in collaboration with other researchers, are testing different algorithms for Landsat 8 TIR bands. Surface skin temperature derived from Landsat 8 thermal images will be included in future versions of the database.

I would highly recommend consultation with a native English speaker to clarify the language used. For example: P2. L13-15: this statement makes no sense.

Thanks for the suggestion. We can replace the statement in P2, L13-15 by the following text:

“Research projects may also provide useful data. But in spite of the current efforts to make research data more accessible, it is still difficult to locate and access these data. In addition, the use of different instrumentation, methodologies and formats may require the homogenisation and formatting of the data sets prior to their use, requiring long treatment times.”

Much effort is given to the evaluation of various algorithms but more effort is needed to justify the use of daytime satellite retrievals and the impacts of the skin effect.

Several practical reasons justify the use of Landsat thermal infrared images to study the surface temperature of inland water bodies:

- The use of satellite images allow the study of water bodies in a large spatial scale, even where field measurements are not possible or too costly.*
- The high resolution of Landsat data allows the study of small and or narrow water bodies that could not be studied using other satellite data.*
- The long period of availability of Landsat data allows the study of long term trends.*
- Nighttime satellite TIR products do not offer the high resolution and long data availability that Landsat offers.*

In addition, as mentioned above, we have decreased the RMSE of the surface bulk temperature estimations to 1.2 °C after removing the data affected by clouds and correcting the cool skin effect.

Further improvements could be made by correcting the warm layer effect. Since the performance of the COARE algorithm was not very good (coefficient of correlation of 0.45) we prefer not to use it to estimate the warm layer effect. Several one-dimensional hydrodynamic models are used to simulate water temperatures in the limnological community. However, their RMSE is typically of the same order of magnitude as the RMSE of the satellite-based estimations. In fact, because of the microclimatic conditions at inland water bodies, the use of an algorithm or model to correct the warm layer effect would require using in-situ (above the water) meteorological data that are rarely available.

At the reservoir of Bimont (BIM13), the average warm layer effect at 10 a.m. UTC (approximate time of satellite measurements) was 0.05 ± 0.21 °C. This accounts for a small part of the total error. Of course, these are results for a single study site and might not be applicable to other sites. However, we do not have the data to make a similar analysis for other sites. We could include a comment in the paper explaining this.

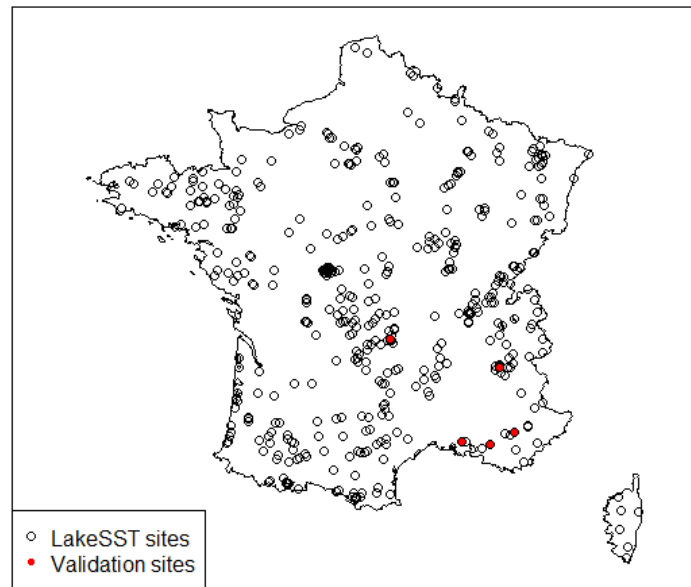
At present, limnologists often work with temperature uncertainties of the same order of magnitude as those observed in satellite-derived data. One-dimensional hydrodynamic lake models often have uncertainties of 1-2 °C. Temperature profiles are often measured at a single point in the water body, not taking into account the spatial variability of surface temperature. In fact, the average standard deviation of surface temperature in a given image was 0.3-0.5 °C in the five validation sites (see table below).

<i>Lake code</i>	<i>Average standard deviation (°C)</i>
<i>AUL13</i>	<i>0.37</i>
<i>BIM13</i>	<i>0.33</i>
<i>LPC38</i>	<i>0.33</i>
<i>PAV63</i>	<i>0.34</i>
<i>SCR04</i>	<i>0.50</i>

For these reasons, we think the data set may be used to study the thermal behaviour of inland water bodies. What is important in that case is to take into account the uncertainty of the measurements, to which end the quality analysis included in the paper will be useful. Still, we would be happy to make other efforts to justify the use of the LakeSST if it is possible with the available data. Is it there any analysis you would recommend?

Why include the numerous outlines of various french regions in Figure 1? These are not defined in the legend and only distract from the overall presentation of the figure.

We have deleted the outlines of French regions.



Fonts on the y-axis of Figure 4 are quite small and difficult to read.

We have improved the figure.

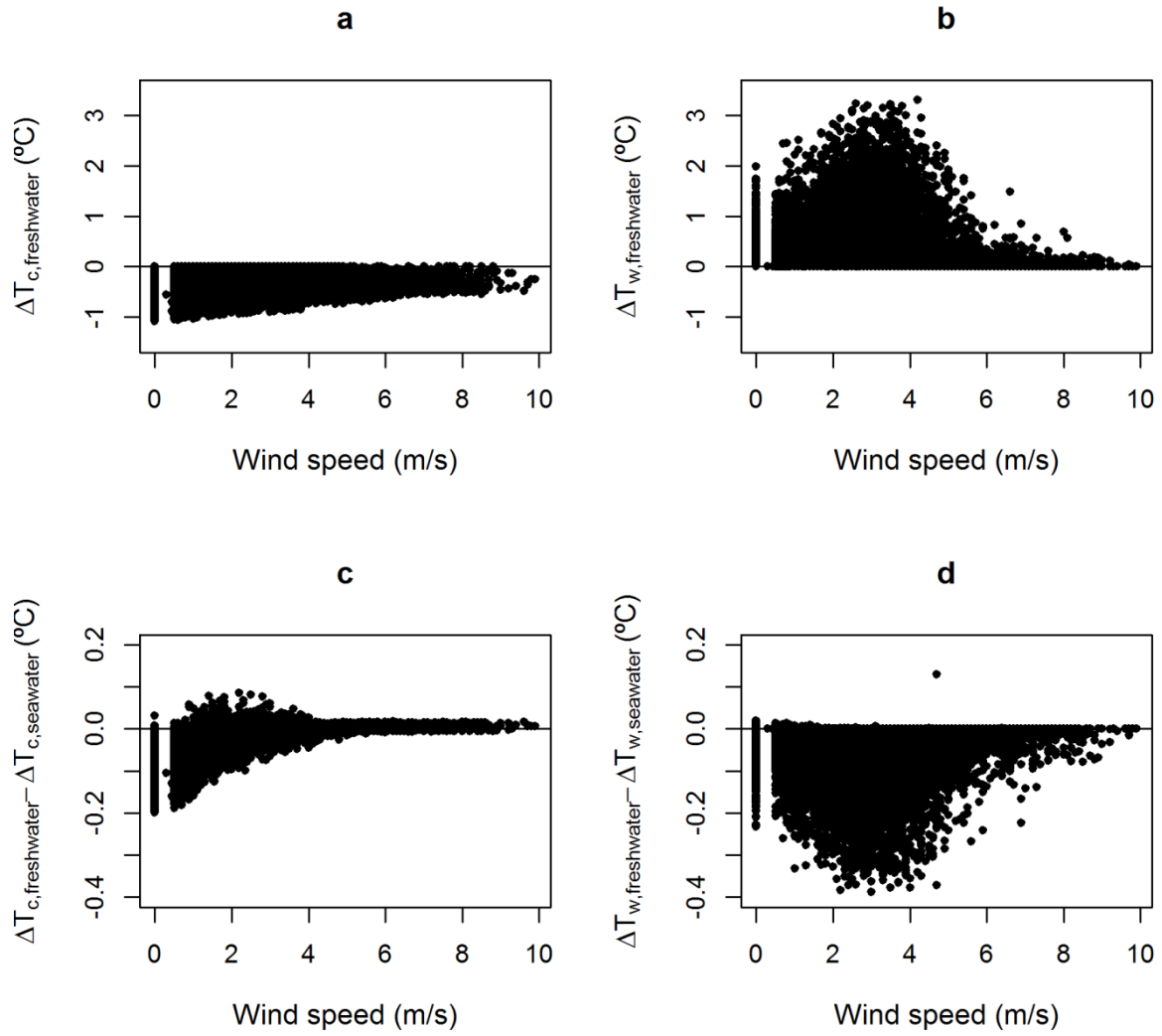


Figure 11 should include R2 and p values

We made no regression on the Figure 11. The diagonal line represents the 1:1 line (information that will be included in the figure legend in the next version of the manuscript). So, we do not include p-values, but we do include the RMSE and R^2 .

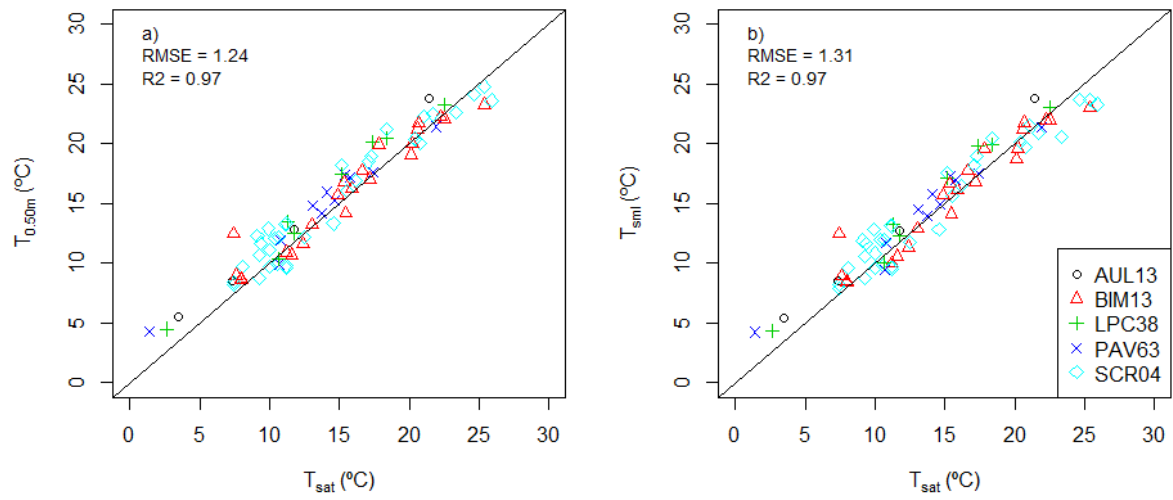


Figure 11: Comparison of satellite-based temperature estimations and *in situ* temperatures: a) temperature at 0.5 m, b) temperature of the surface mixed layer. The diagonal line indicates the 1:1 relation.