Please find our response in italics.

RESPONSE TO ANONYMOUS REFEREE #1

This paper presents a useful dataset of remotely sensed surface temperature (Landsat series) over French water bodies, including also some ancillary data. The dataset is described in detail, and one of the most remarkable things is the detailed discussion on the problems related to the direct comparison between temperatures retrieved from satellites (skin) and water temperatures measured by contact at different depths. The paper also shows a rigorous validation of the dataset. Therefore, this paper address the main topics of this journal, and my recommendation is to publish the paper after some minor comments:

GENERAL COMMENTS

As a general comment, I think the validation part could be improved (if not now, then in a future work). I would be nice to measure with a thermal radiometer over particular lakes and then compared contact temperatures with radiometric temperatures. Radiometric temperatures are also better indicator of the performance of the surface temperature retrieval from satellite data. Another option is to perform an intercomparison between standard and well-validated remote sensing SST/LST products. In this case most of the products available are at low spatial resolution (around 1-km), so this intercomparison may be restricted to the largest lakes.

We totally agree with reviewer #1 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.

It should be also justified why Landsat-8 is not used, since the straylight problem in the TIR bands was partially solved.

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.

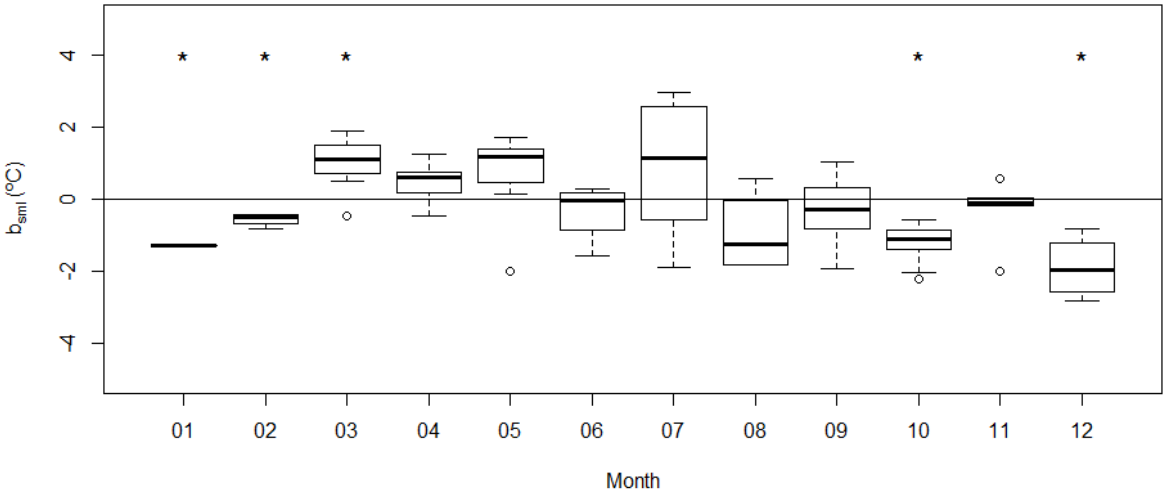
SPECIFIC COMMENTS:

-Section 3.2, page 6, line 11: If I am not wrong the algorithm used by Simon et al 2014 is the same than the algorithm presented by Jimenez-Munoz & Sobrino 2009, so it is not a new version.
That line would be more properly expressed as "The algorithm was implemented by Simon et al. (2014) and used for producing the LakeSST data set."

-Section 5.3, page 13: Please include some comments about the SCL-off problem in Landast-7, and how this problem is addressed in the presented dataset. If Landsat-8 is not used for any technical resason (e.g. straylight problem), then some sentence could be also added in this section.
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.

-Section 5.4, page 14, lines 31-32: "... we applied the algorithm also when clouds were partially present." This is a critical issue, because all the surface temperature retrieval algorithms working with Thermal-Infrared data are developed to be applied under clear sky conditions. This is well-known, so I think it has no sense to apply the algorithm in the presence of clouds. I would remove the data points contaminated by clouds and redo again the analysis to assess which are the main variables contributing to the seasonal bias. May be it is related to different atmospheric water vapor contents (?)

Still, the current analysis allows to determine the limit of cloudiness under which the algorithm can be applied (approx. clearness index k_t >0.6). When the data points contaminated by clouds are removed, a seasonal bias pattern is still present (see figure below). There is still a statistically significant relation between bias and solar radiation and air temperature. However, there is no statistically significant relation between atmospheric vapour content and bias.



-Table 1: The header of the table should be more informative.

We suggest replacing the previous heading by "Comparison of satellite-based temperature estimations made using the algorithm by Jiménez-Muñoz et al. (2009) ($T_{0.01\ m}$) to in situ measurements (temperature at 0.50 m, $T_{0.50\ m}$, and average temperature of the surface mixed layer, T_{sml})."

-References: in page 6 line 7-8 the authors refer to Sobrino 2004 with a strange symbol (#2738), but Sobrino 2004 is not included in the references list. Please correct.

Ok.