Response to reviewer 1 :

"Quantification of the Beauce's Groundwater Contribution to the Loire River Discharge Using Satellite Infrared Imagery" uses Landsat TIR images to determine groundwater contributions to the Loire River using a simple energy budget approach and compares this to a groundwater budget approach. A method for determining groundwater contributions to rivers over space and time is presented, however there were many different assumptions and acknowledged errors in data utilized, calculations completed, or comparisons made that undermine the potential impact of the study. "

Despite the uncertainties, this study shows that extracted temperature profiles nevertheless remain in agreement with known areas of groundwater discharge along the Loire River.

A quantification of the uncertainty associated to the heat budget method has been added to the revised version of the manuscript (part 3.5 – revised manuscript). We show that uncertainties are not very likely to undermine the major findings of this study.

We also choose to present in the revised version of the manuscript a deterministic process-based hydrogeological model of the Loire River basin (part 3.4 – revised manuscript). This model allows the quantification of the daily groundwater discharge along the Loire River. It is therefore better suited to the comparison with the heat budget than the groundwater budget previously used. We find that both methods (groundwater budget and hydrogeological model) give similar results and that they are in agreement with the heat budget.

Number of pixels spanning the channel that were included within the analysis. In general, there were 3 pixels spanning the channel, but at times these were mixed pixels (water and land). The mixed pixels were still included within the analysis (2053 line 3-7).

We do not use mixed pixels in the study (i.e. composed of land and water). All the pixels used are water pixels only. However, the number of water pixels across the stream is variable and at times lower than 3. That means, we do use pixels that are not pure (i.e adjacent to mixed pixels but still composed of water only). In our terminology, pure water pixels stands for water pixels that are situated more than a pixel away from any mixed pixels. The manuscript has been clarified in this regard (line 167 – revised manuscript).

No atmospheric corrections of the satellite TIR (2053 line 11) and shade influences from clouds were removed (2053 line 1-2). There was no explanation of how cloud influences were removed.

No atmospheric corrections are done in this study. However, this is mostly an issue when considering temperature variations over distances covering several satellite images (Handcock et al., 2012). In the current study, the studied river length is only 135 km and included in a single image. The river flows over a flat landscape. On the days when the images are taken, the sky was clear over the whole area and atmospheric conditions were therefore expected to be homogeneous. Furthermore, the Loire River is discretized in sections that do not exceed 30 km in length. It is therefore expected that atmospheric influences over the infrared radiations emitted from the water do not play a significant role in explaining the temperature variations observed along each river section. A comment was added in the manuscript in this regard (line 176-178 – revised manuscript).

It is nevertheless true that a global shift of each Loire temperature longitudinal profile by a constant value is to be expected after taking into account atmospheric corrections (Handcock et al., 2012). However, this shift is likely to be small (<1°C), since the average difference between temperature measured in-situ and temperature estimated from the non-atmospherically corrected TIR images does not exceed 1°C. Overall, the error made while not taking into account atmospheric correction is therefore of the same order of magnitude as the error made while not taking into account groundwater temperature variability, i.e. 10 to 30% (see next comments on the sensitivity analysis).

Clouds and their shades on the ground surface are detected visually using the TM8 band and the corresponding pixels from the TM6 band are removed manually from the analysis. Overall, clouds are few as only images with under 10% of cloud cover are selected.

Tributaries and power plant influences were considered negligible even though their influence was difficult to separate (2051 line 24-25) and can be close to 1_C in the winter (2051 line 10-16).

No warming of the Loire River temperature was observed downstream of Dampierre and Saint-Laurent des Eaux, based on the TIR images (line 572-576 – revised manuscript). We do not possess in-situ measurements of the water warming in the vicinity of the power stations. Reports from EDF show that, at Dampierre, in July 2010, the mean temperature increase is 0.1°C, while the maximum temperature increase is 0.18°C. Such a low temperature increase can not necessarily be identified with the satellite TIR images. EDF uses cooling towers to reduce the temperature of the water that is released into the Loire River. A 1°C maximum temperature increase was reported in winter, but only at low flow (i.e. below 100 m^3/s). Such flows were not observed during the acquisition period of the TIR images, in winter. The choice was therefore made not to take into account the influence from the power plants, as the induced water temperature changes are small.

Influence from the tributaries was not considered in this study (line 123-132 – revised manuscript). In the case of the main tributary, the Loiret River, its influence is not separated from that of the groundwater because it is very short in length (less than 10 km) and its water is mainly of groundwater origin. Thus, we consider the Loiret discharge as groundwater discharge. Temperature variations along the Loire River, which can be attributed to the main groundwater discharge area (close to La Chapelle Saint-Mesmin), start upstream of the confluence with the Loiret River (see Figure 7). This shows that the Loiret River is not the only reason behind the temperature variations observed around river kilometer 635. All the other tributaries have flows under 1 m^3/s and temperatures close to the Loire River temperature. Their influences on the Loire River temperature profile is therefore expected to be small and were not observed on the TIR images.

Weir influences along the river (2050 line 25-27) were not accounted for.

Weirs influences were discussed briefly in the discussion part of the manuscript (lines 577-582 – revised manuscript). Temperature differences between the 1 km upstream reach and the 1 km downstream reach of the main weirs remain small (less than 0.1°C). It is therefore concluded that no significant temperature change along the water course could be related to a weir, based on the TIR images.

Surface area estimates within the heat budget calculations were based on the pixels selected for the analysis. These did not cover the entire channel surface area (2054 line 18-22). The potential 20% error in surface area translates into increased error in heat budget calculations because this value scales all surface flux estimates (S in eqn. 3).

The choice was made to consider the water pixels from the TM61 band of the LANDSAT images to estimate the Loire River surface area, since we do not possess aerial images of finer spatial resolution at the date of the satellite images. This technique allows taking into account variations in the extent of the Loire River with time. The error in the surface estimate we discussed about is estimated by comparing, over each Loire River section, the area calculated using the water pixels from the TM 61 band (30 m) and the area calculated using the TM 8 band with a better spatial resolution (15 m). A description of this comparison has been added in the manuscript (lines 334-338 – revised manuscript).

On average, considering all dates and all river sections, the surface is underestimated by 33%. We found that the resulting average bias in our estimation of the groundwater discharge is of about 25%, with a maximum bias of 50% (see next comments on the sensitivity analysis).

Groundwater temperatures were assigned for summer and winter based on a data base (2054 line 16). No information was provided regarding the data or variability in these values.

ADES is a French database on groundwater data. It notably gathers most of the groundwater temperature measurements carried out by the different surveying agencies and water companies. The temperatures are measured irregularly over time. The precision of the temperature measurements is +-0.1°C. Data from the piezometers situated close to the Loire River is gathered for the period 1991-2011 (292 measurements). Looking at the measured temperatures, it appears that 80% of the temperatures are comprised between 11.5°C and 14°C in summer and between 11°C and 13.5°C in winter. These details have been added in the manuscript (lines 212-215 – revised manuscript).

The influence on the computed groundwater discharge of such a variability in the groundwater temperature can be assessed, considering that surface water temperatures varies between 4.5°C and 6°C in winter and between 20°C and 26°C in summer (see next comments on the sensitivity analysis). Taking into account these temperature variations, we found that the groundwater discharge can fluctuate between 90% and 130% of the previously computed groundwater flow, based on mean groundwater temperatures. The highest errors in the calculation of the groundwater discharge are likely to occur in winter, when the river temperature is high and when the difference between surface water temperature and groundwater temperature is therefore low.

Inaccurate estimates of river temperature from TIR when compared to river temperatures. At times differences were > 3 _C different (Figure 2) and on average they were +0.3 _C in winter and -1 _C in summer (2056 line 5). Some of the "sharp" changes in temperature used to estimate groundwater influences were 0.5 _C (2057 line 19), which is a small or possibly insignificant change relative to the errors observed.

Longitudinal temperature profiles varied less than 2 _C when the variability was at its highest (Figure 3).

Temperature accuracy (bias) should be differentiated from temperature uncertainty (Handcock et al., 2012). This has been clarified in the manuscript (part 4.1 – revised manuscript).

Temperature accuracy is the average difference between the temperature estimated from the TIR images and the temperature measured in-situ. Temperature accuracy from the TIR images is 1°C on average in summer and 0.3°C on average in winter.

Temperature uncertainty is the temperature variability observed in an area that should have a homogeneous temperature (i.e. repeatability of measurement). Temperature uncertainty is therefore reduced, by averaging temperature over 200 m long sections and by using a moving average to smooth the temperature profile. The study of the longitudinal evolution of the difference between TIR images based temperature and in-situ measurements may give some ideas about the uncertainty (see next comments on the sensitivity analysis and figure 2). On average, the temperature difference variation remains below 0.8°C over the 100 km reach Dampierre – Saint-Laurent-des-Eaux (mean variation of the temperature difference of 0.0072°C/km). Sharp temperature changes need to be compared with the uncertainty and not with the accuracy. The sharpest temperature changes observed on the longitudinal profiles are comprised between 0.04°C/km and 0.1°C/km (mean of 0.074°C/km). The sharpest temperature changes are therefore at least one order of magnitude higher than the changes that are to be expected from the uncertainty. They are therefore likely to be meaningful in terms of physical processes.

The overarching concern with these combined assumptions and errors are the influences on the findings within the paper. It is unclear if there is enough variability in the longitudinal temperatures to confidently back out groundwater influences and needs to be further investigated. There are many questions and concerns regarding the influence of the assumptions or treatment of data. What are the errors in the satellite based TIR data and what is the influence of not correcting for atmospheric conditions that will vary throughout the study reach and over different times of year? Torgersen et al. 2001 states that 10 pixels are required to avoid the influences of banks emission and to get accurate river temperatures. It does not seem that 3 pixels are adequate, particularly when they are mixed pixels. Given these issues and additional uncertainty in other foundational data used in the heat balance approach (e.g., assumed groundwater temperature and incorrect surface area estimates), the confidence in groundwater estimates are likely low.

We previously discussed the influence atmospheric corrections would have on our study. It would have an influence on the temperature accuracy but not on the temperature uncertainty.

Torgersen et al. (2001) chose arbitrarily 10 pixels in each thermal image and took the median temperature value. Temperature longitudinal profiles were then drawn using these median values. This method can only be employed when using multiple images (mostly for airborne campaign). However, our method is similar in that we average river temperatures by sections of 200 m to draw the longitudinal profiles. This is a spatial extent of the same order of magnitude as the usual ground coverage of a TIR image taken from an airborne campaign. The advantage of our method is that we consider all the water pixels from the water course. There could therefore be more than 10 pixels in the 200 m sections. Then, uncertainty is further reduced through a moving average smoothing of the data over +-2 km.

We carried out sensitivity tests to estimate the overall uncertainty in our groundwater discharge estimation using the heat budget. Details about these tests have been added in the new manuscript (lines 413-419 and lines 484-489 – revised manuscript). One figure is added in the manuscript to show the confidence interval of the groundwater discharge estimation at two dates, one in summer and one in winter (Figure 6).

The current comparison with the groundwater budget that has long averaging times, similar uncertainties, and is vaguely described does not provide the type of validation needed to illustrate the potential of this approach. In order for this paper to have an impact within the remote sensing and groundwater communities, more information regarding a quantitative understanding of the accuracy of the proposed methodologies is necessary. Some additional information that validate the findings is also needed.

To validate further the findings, we replace the groundwater discharge calculated using the groundwater budget by the groundwater discharge calculated using a deterministic process based groundwater model over the entire Loire River basin. Using this model, the groundwater discharge to the Loire River can be calculated at each date and at every river kilometers. Uncertainty in the model prediction of the Loire River flow is known and low (Nash criteria of 0.98). Details about the uncertainty in the groundwater discharge estimated through modeling have been added in the manuscript (lines 499-501 – revised manuscript). The groundwater model was developed by Fulvia Baratelli and Nicola Flipo. They are included in the new authors list of the manuscript.

We found that the newly calculated groundwater discharge remains in agreement with the groundwater discharge previously calculated with the groundwater budget (see Figure A below). The highest groundwater discharges calculated by both methods are situated between river kilometers 620 and 660. However, on average, groundwater discharge rates calculated using groundwater modeling are higher than the groundwater discharge rates estimated with the groundwater budget. Higher groundwater discharge rates are also estimated in winter than in summer, which is in agreement with what was found using the heat budget. This remark has been added to the manuscript (lines 543-544 – revised manuscript).

Two figures are added in the manuscript to show the groundwater discharge calculated by the groundwater model (Figure 5; Figure 6).

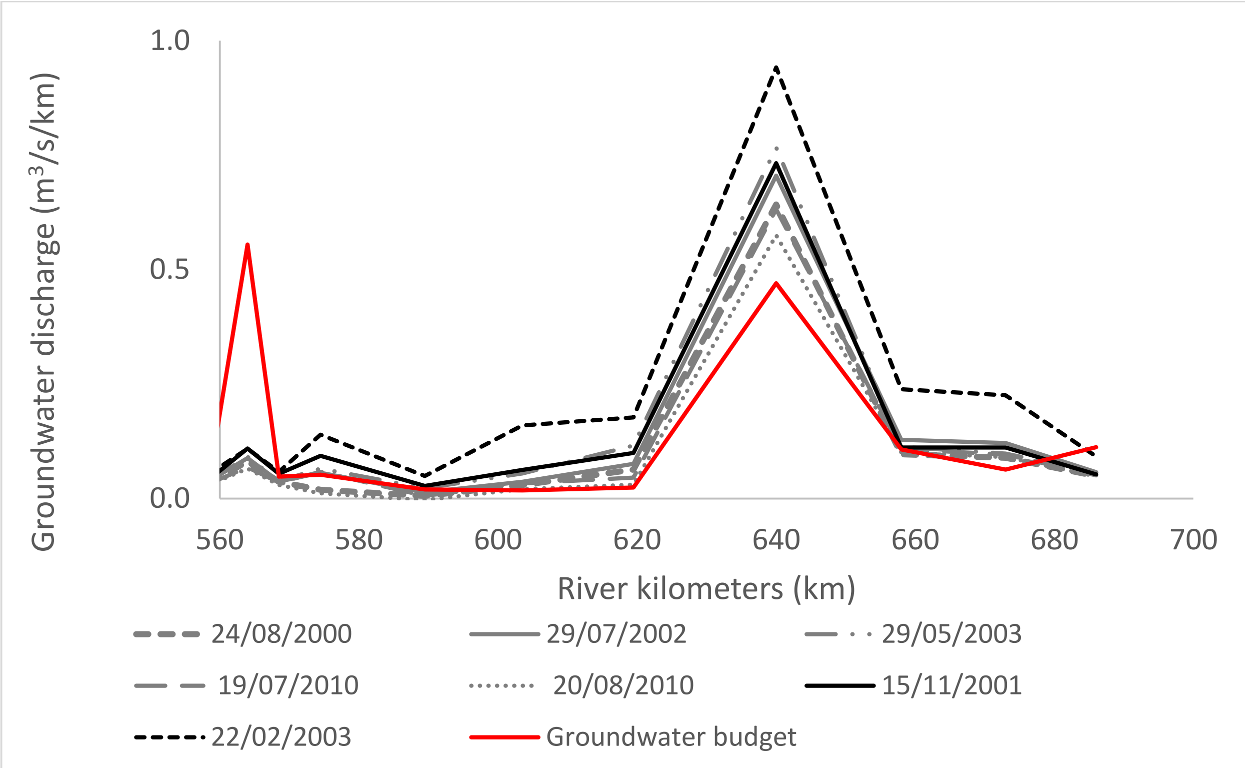

Figure A: Groundwater discharge estimated using a groundwater budget over successive Loire River groundwater catchment areas and using groundwater modeling over the entire Loire River basin.

Response to reviewer 2 :

This manuscript presents interesting results on how Landsat imagery in the TIR band can be used to map water temperature in a large river synoptically over hundreds of kilometers. This approach has been used in other large rivers, but the Loire River is particularly interesting because it is influenced by relatively high-volume groundwater inputs and is quite narrow (in places) for using satellite TIR imagery. Furthermore, the seasonal differences in river temperature provide an important perspective on thermal heterogeneity experienced by riverine biota. The paper could significantly improve our understanding of riverine thermal regimes and spatial patterns at broad scales, and it could be a useful contribution to the literature on thermal remote sensing of rivers, but unfortunately its presentation is quite poor. It is confusingly written from the standpoint of scientific English, and its organization requires significant revision to highlight the strengths and weaknesses of the study. For example, the data on the accuracy assessment need to be presented in more detail. The only data presented on the accuracy of the method are in Figure 2, which only presents means, which are not very useful. The authors need to present box and whisker plots perhaps to show the reader how variable the differences were. Furthermore, the authors mention that linear regression was used to evaluate kinetic and radiant temperatures, but these linear regressions and their statistics are not shown or reported. It would seem that the remote sensing part of this study would alone be a nice contribution but would require more more detail for the reader to truly evaluate the data. I am not qualified to evaluate the methods for estimating groundwater discharge, but it appears that this part of the manuscript is poorly

developed. The main objectives of the paper pertain to the TIR data and how they can be used to locate thermal anomalies associated with groundwater at different times of the year. The authors may wish to reconsider how important the actual calculations of discharge are for this paper.
Many small modifications have been made to improve the readability of the manuscript. They are visible in the marked-up manuscript.

Comments on the accuracy and uncertainty have been added to the manuscript (see previous comments).

Linear regression was not used to correct radiant temperature from in-situ measurements of kinetic temperature. Linear regression does not work well, although radiant temperature tends to overestimate kinetic temperature in winter and to underestimate it in summer (see Figure 2).

Linear regression was used to compare, when this was possible, temperatures extracted from the pure water pixels and temperatures extracted from the non-pure water pixels, in order to assess the robustness of the method (see next comments).

We found that the calculation of the groundwater discharge is an important part of this work. One of the findings of this study is that, despite all the uncertainties associated to the use of satellite TIR images, the main groundwater discharge area in the Loire River can still be identified. Moreover, the calculated groundwater flow remains credible in regard to what was found in previous studies and to what we find using a groundwater flow budget over the successive catchment areas and groundwater modeling. Quantification of groundwater discharge using TIR images has already been conducted in the past (Loheide and Gorelick, 2006) but it has been used on a much smaller river. It is therefore interesting to see if Landsat images could also be employed.

Title: Specify "thermal IR" not just IR. Also, write out Beauce Aquifer because most readers won't know what the "Beauce" is.
The corrections have been made.

Page 2048, Line 20: Throughout the manuscript, the authors write "Thermal InfraRed". Just write "thermal infrared (TIR)" and use standard terminology as in the papers that are cited in the references.
The corrections have been made.

Page 2049: Check spelling of "Burckholder". I think it doesn't have a "k". Also, the word "evolution" doesn't make sense as it is used throughout this manuscript.
In fact, it should have been written “Burkholder”. It has a “k” but no “c”. The corrections have been made. The word “evolution” was replaced by “variations”.

Page 2050: The authors need to say something about the presence of large wood, boulders, and gravel bars because they can also be a cause for mixed pixels, not just the banks.
In the Loire River, there are no boulders, as the sediments are mostly composed of sand and gravel. The gravel and sand bars are detected using the TM 8 band from the Landsat images. They are considered in

the same way as the river banks and pixels from TIR images are therefore discarded when overlapping sand bars. Trees in the water, as well as very small sand bars, are not likely to be detected due to the resolution of the TM 8 band pixels (15*15 m^2 pixels). But, it is therefore assumed that these obstacles do not cover an important area within the 60*60 m^2 water pixels from the TM 6 band.

Page 2052, Line 18: This is confusing because the authors refer to the near IR data before they even describe the TIR data from the satellite. In fact, the authors don't identify the spatial resolution of the IR and TIR bands in the methods. Please check your methods. They are not presented in a logical order and they need to provide more detail.

Comments have been added in the manuscript for better clarity (lines 159-172 – revised manuscript). Resolutions of the IR and TIR bands are described.

Page 2053: The fact that the authors use data where there are only three pixels across the width of the stream is quite surprising, given what papers have described. It is really important for these data to be fully reported. After reading this paper, I am somewhat convinced that < 3 pixel may work in certain instances, but I need more data to be convinced.

The choice to use all the water pixels was made since we could otherwise not have covered the full length of the selected river reach. However, we made sure that the resulting bias was not too important (see next comments).

Page 2056: Where are the results and plots for the regression analysis?

We add one figure in the manuscript showing the comparison between temperatures extracted from non-pure water pixels and temperatures extracted from pure water pixels, over all the 200 m sections of the Loire River where pure water pixels could be found (Figure 3). We found that there is no significant shift between temperatures extracted from pure water pixels and temperatures extracted from non-pure water pixels. The non-pure water pixels do not particularly overestimate the water temperature in summer (Figure 3a in the manuscript), as it was expected from the high river banks temperatures. The slope of the regression line is 0.99 and the coefficient of determination is 0.98. In winter, a slight underestimation of water temperature within the non-pure water pixels could be seen (Figure 3b in the manuscript), with a slope of the regression line of 0.72. However, the coefficient of determination is quite low ($R^2 = 0.69$) and we lack data to conclude (the range of variation of water temperature is much smaller in winter than in summer). These results are added in the manuscript (lines 334-338 – revised manuscript). Considering both summer and winter data, the slope of the regression line is 1 with a regression coefficient of 1 (see Figure B below).

The difference between temperatures extracted from pure water pixels and from non-pure water pixels usually remains in the +-0.5°C interval (for over 98% of the 200 m sections). This 0.5°C gap corresponds to the approximate sensor resolution of the satellite camera (see Figure C below).

As we consider in our analysis both pure and non-pure water pixels, and since we use a moving average over +-2 km to smooth the temperature profile, we expect the bias resulting from the use of non-pure water pixel to remain relatively low.

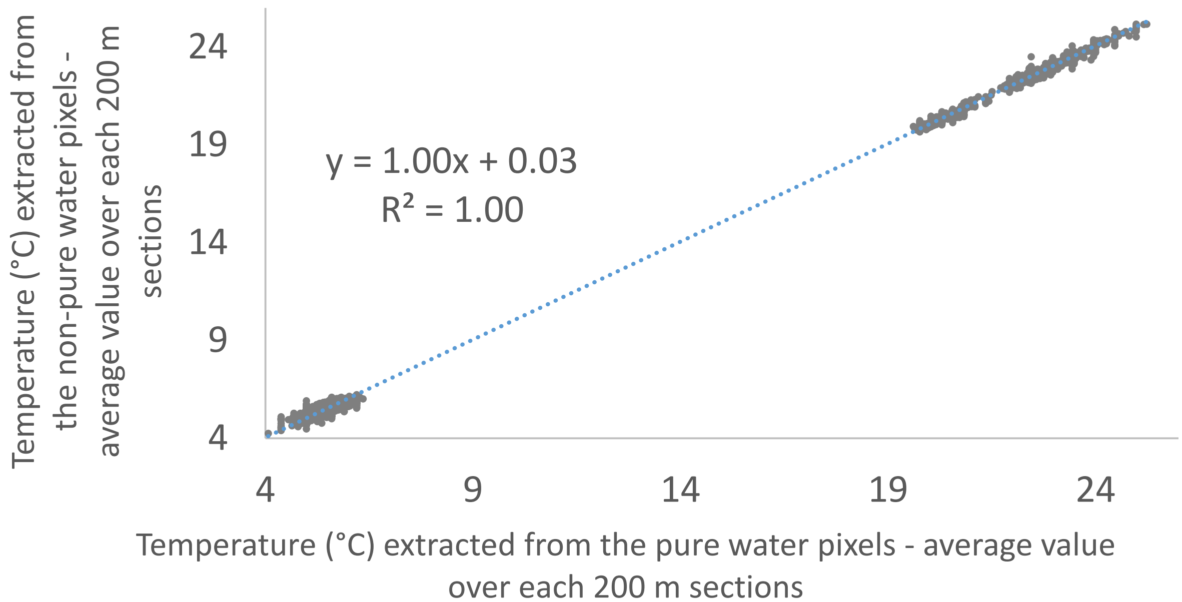

Figure B: Relation between the temperature extracted from the non-pure water pixels and the temperature extracted from the pure water pixels. Temperature values of both pixels types are averaged over the successive 200 m sections where pure water pixels exist. Both winter and summer temperature values are represented.

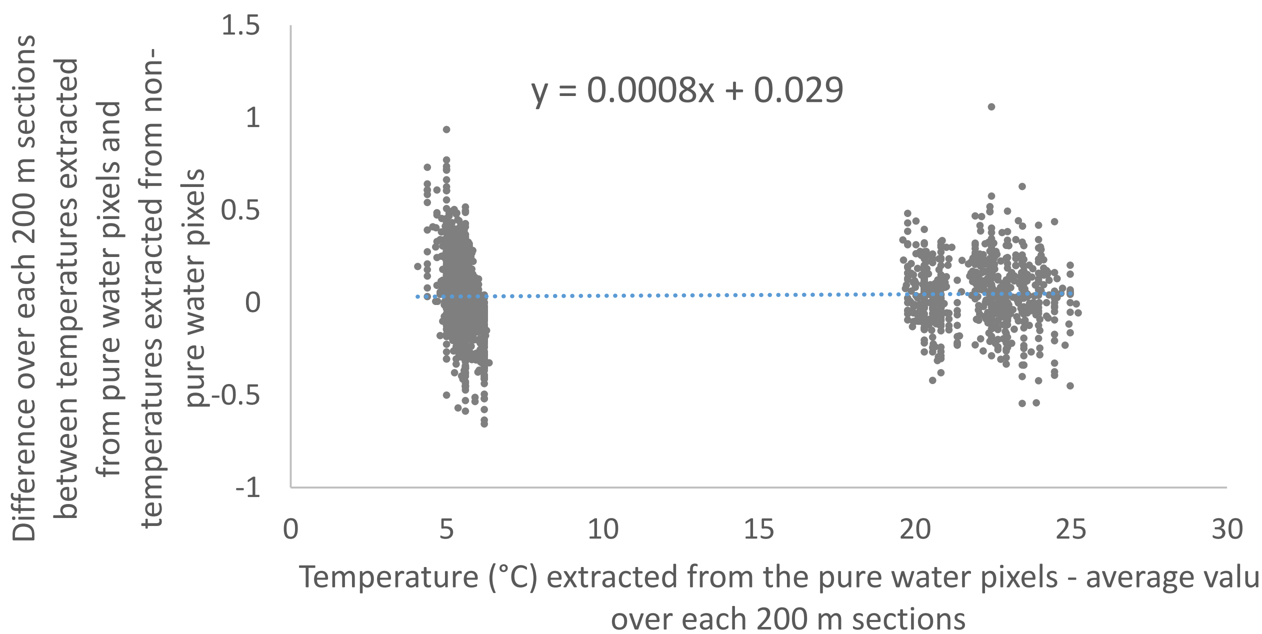

Figure C: Difference between the temperature extracted from the pure water pixels and the temperature extracted from the non-pure water pixels. Temperature values of both pixels types are averaged over the successive 200 m sections where pure water pixels exist.

Table 1: What time were these temperature data collected? I think it says this in the methods, but you should probably have it in the table as well. Standardize the significant digits in these numbers.
The temperatures were collected at 11:30 LT in winter and 12:30 LT in summer (lines 138-139 – revised manuscript). The significant digits have been standardized.

Table 3: Which sections? All sections? How many sections?

All the 200 m sections of the Loire River are included in this analysis. The legend has been modified accordingly.

Figure 1: The symbols on this map are difficult to see. The triangles and the crosses are too faint. Also, the river km numbers need to be moved slightly so they are not on top of other symbols. Note that the town of Saint Laurent has a symbol that gets in the way of other symbols, and it is hard to read the text of the name. The font size is generally too small throughout this figure. Need to show groundtruth locations if possible. What is the light grey area? This needs to be stated in the caption.
The map has been modified.

Figure 2: The y-axis label is too long. Shorten and provided clarification in the text. Don't use "ones" in the label; this is not good scientific writing. Are these mean differences? I think it would be better to have box and whisker plots of these so you can see variation.

The term "ones" has been removed from the figure.

This figure shows difference between in-situ measurements of water temperatures and temperatures estimated from the longitudinal temperature profiles obtained from the TIR images. They are therefore a kind of mean differences. The figure caption has been modified for better clarity.

We present here four graphs showing box and whiskers plots of the temperatures extracted from all the water pixels from the TIR images, in the vicinity of Dampierre and Saint-Laurent des Eaux (see Figures below D, E, F, G). In the case of Dampierre, all the water pixels situated between river kilometers 570 and 576 were considered. In the case of Saint-Laurent des Eaux, all the water pixels situated between river kilometers 668 and 672 were considered. These graphs show that, in most cases, the temperature measured in-situ is comprised within the range of the temperatures observed at the neighboring water pixels from the TIR images. In these later cases, temperature discrepancies between the 2 methods could easily be explained by local temperature heterogeneities in the water course or satellite sensor's resolution. However, there are 3 cases where these phenomenon may not offer an adequate explanation. It occurs on the 29/05/2003 at Dampierre and on the 29/07/2002 at Dampierre and Saint-Laurent des Eaux. Taking into account the box and whiskers plots, water temperature is underestimated by the TIR images respectively by at least 1.5°C, 2.8°C and 1°C. Satellite resolution alone can not explain these temperature differences. This analysis is consistent with what was deduced from the previous comparison between the longitudinal temperature profiles and the in-situ temperature measurements. Figure 2 was therefore kept as it was in the revised manuscript. The discrepancies between temperatures measured in-situ and TIR images derived temperatures are taken into account in the uncertainty analysis.

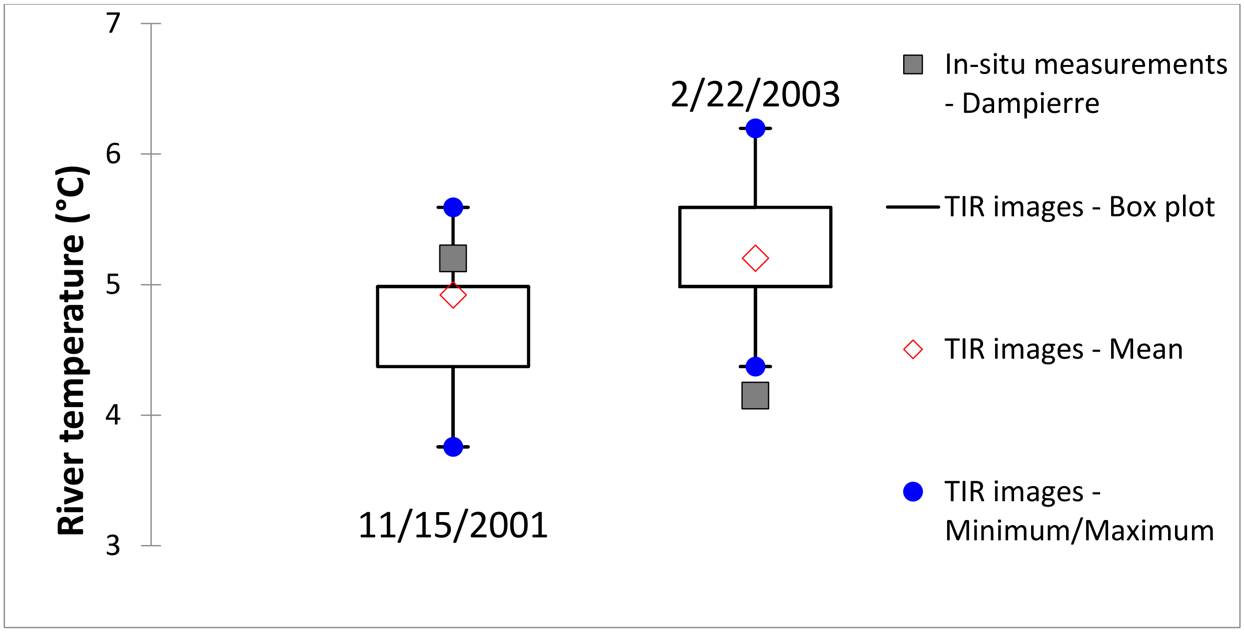

Figure D: River temperatures measured in-situ at Dampierre and estimated from the TIR images in a 6 km reach close to Dampierre. The temperature from the water pixels is reported in the box and whiskers plots. Winter temperatures are presented.

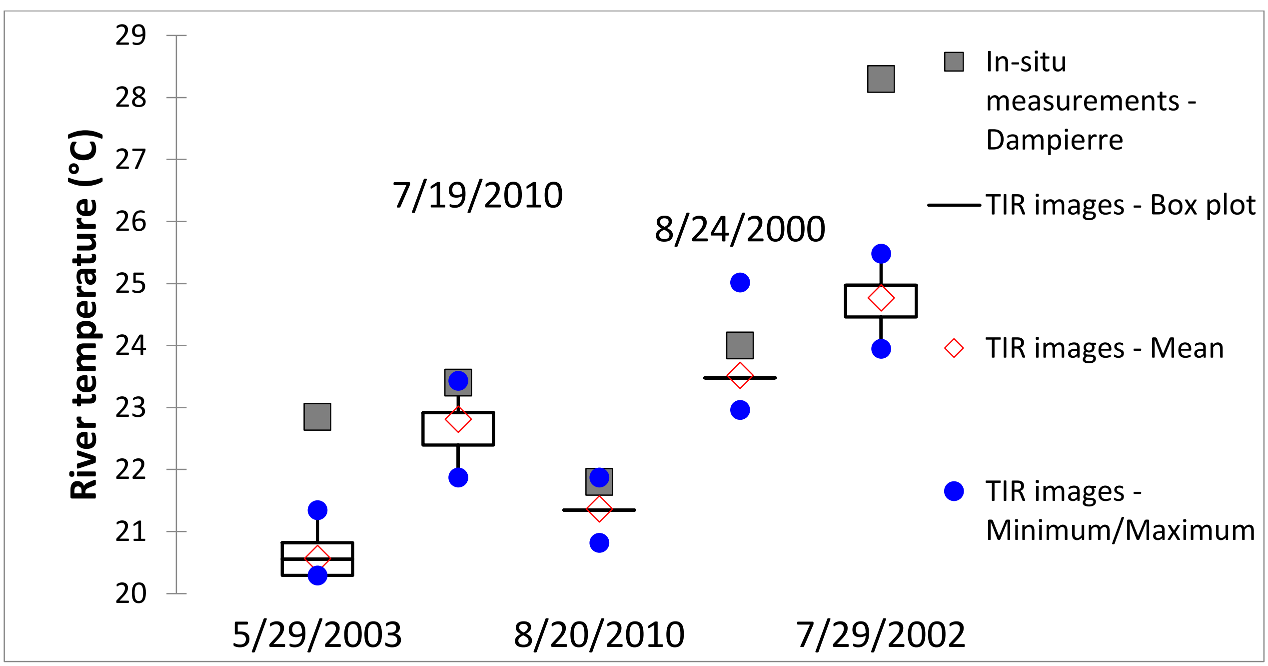

Figure E: River temperatures measured in-situ at Dampierre and estimated from the TIR images in a 6 km reach close to Dampierre. The temperature from the water pixels is reported in the box and whiskers plots. Summer temperatures are presented.

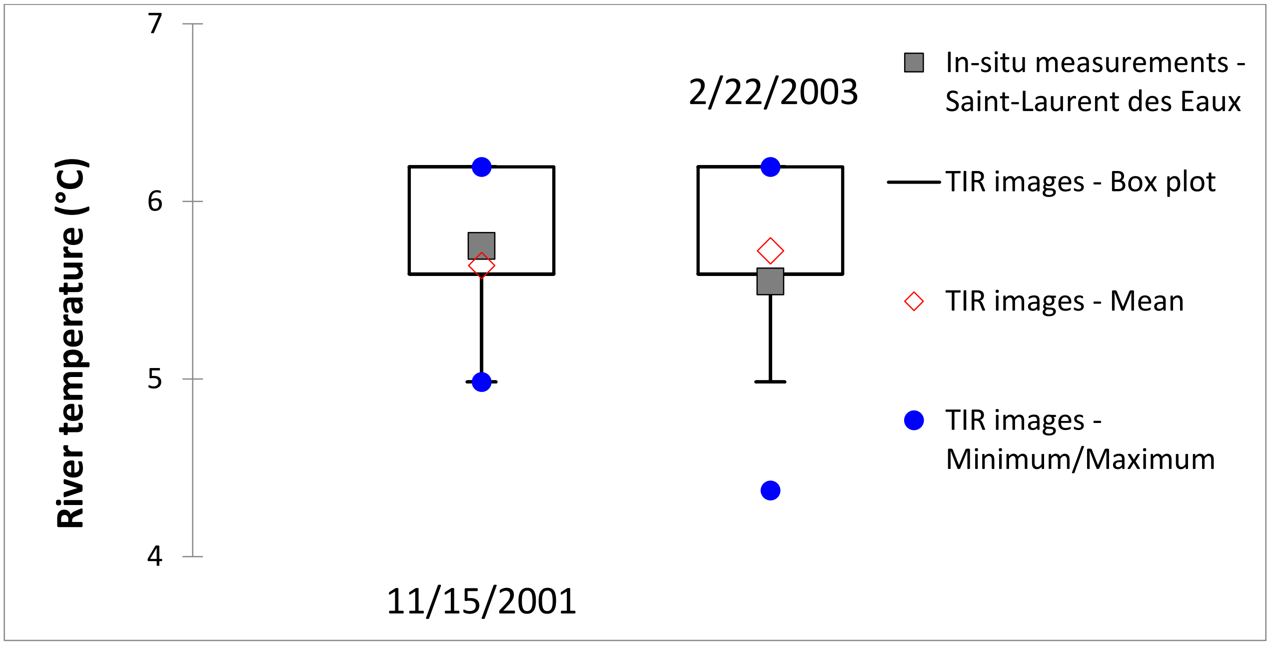

Figure F: River temperatures measured in-situ at Saint-Laurent des Eaux and estimated from the TIR images in a 4 km reach close to Saint-Laurent des Eaux. The temperature from the water pixels is reported in the box and whiskers plots. Winter temperatures are presented.

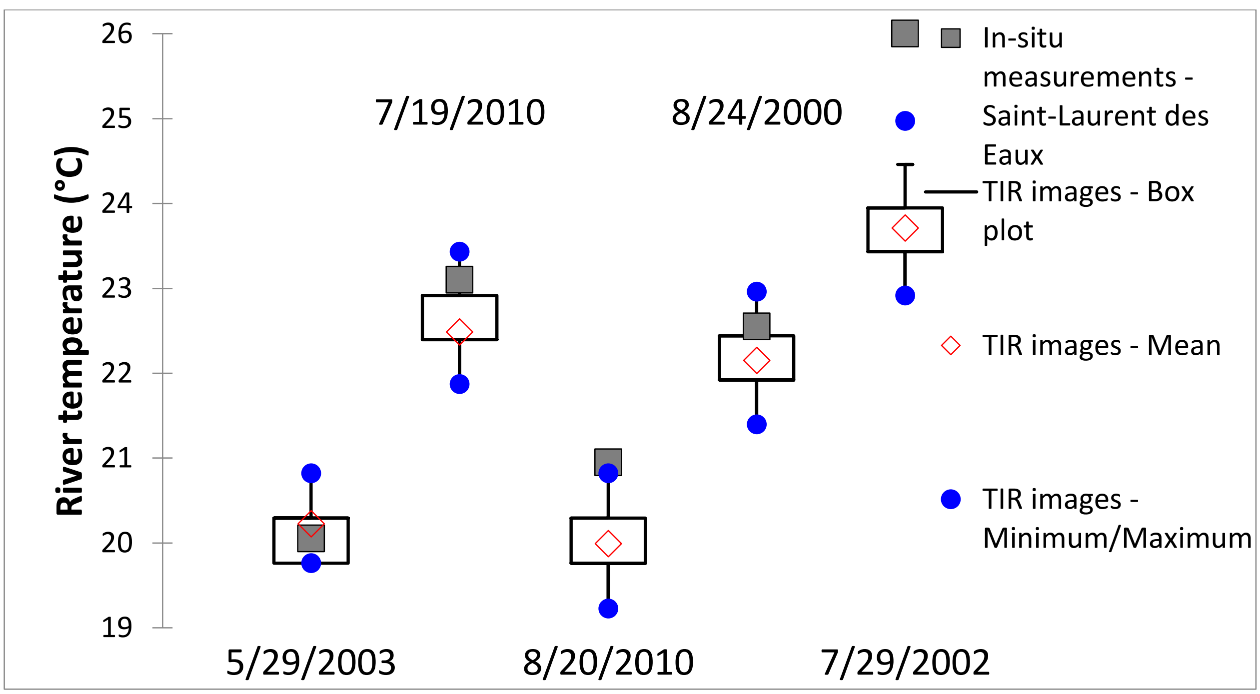

Figure G: River temperatures measured in-situ at Saint-Laurent des Eaux and estimated from the TIR images in a 4 km reach close to Saint-Laurent des Eaux. The temperature from the water pixels is reported in the box and whiskers plots. Summer temperatures are presented.

Figure 3: State that these are derived from satellite imagery. What does "removed" mean in the y-axis label? Move the x-axis at the bottom of the figure.

The corrections have been made.

Figure 4: I think it would be really helpful to have Figure 3 and Figure 4 be panels in the same figure.

It has been done.