Anonymous Referee #2

Received and published: 1 June 2015

The manuscript by Schwatke et al. is of major importance since it presents new results for the monitoring of lakes and lakes that look to be a dramatic improvement (much lower rms in comparison with in-situ series) with respect to existing databases. Consequently, the paper deserve publication but it is not acceptable in the present state. Major points have to be improved.

The presentation of the results is not clear in the sense that their methodology merges three aspects:
1 - the use of a "homemade" retracking of the radar echoes
2 - replacement of the true data value by one predicted through a Kalman filtering.
3 - fine rejection of the outliers

And It is absolutely necessary that the part played by each of these three points in the final improvement is better stated since these 3 points more or less make the difference between the existing website cited in the paper (some use GDR retracking when other use their own retracking, some have a refined detection of the outliers in the raw data when others are more loose on this question, some publish all the values obtained after the post-processing step when other reject dubious values, ...). In other word, Authors must show how much improvement is due in the present study to the retracking of the radar waveform, how much is due to the Kalman filter ?, and how important is the choice of the "valid" points ?

We agree that analyzing the importance of the different aspects of our approach will improve the paper. Thus, we will show the impact of the different processing steps in more detail. We will separately analyze the impact of the pre-processing (i.e. retracking and outlier detection) and the kalman filter approach. We will not handle the retracking separately since we did no innovative work here: we simply apply the existing “Improved Threshold Retracker” to achieve reliable heights for missions such as Jason-1 and smaller water bodies. Furthermore, it is shown in Table 3 that the usage of the ocean product (e.g. Lake Argentino, Lake Buenos Aires, etc) will also improve the resulting RMS and R^2 compared to the other databases. This shows that the retracking has only a small impact on the improvements of the time series.

Planned updates in revised version:

- Height estimation by using a median approach instead of the and Kalman Filter will be performed for the three selected study cases to demonstrate the impact of the Kalman Filter on the resulting time series
- Additional explanation of the impact of different applied outlier criteria on the resulting time series will be added.
- Additional information which outlier criteria was applied will be added to Table 3 and 4.

According to the Title, the Kalman filter seems to be the major source of improvement. If it happens that the retracking algo and/or outlier detection play an important part in the improvement, maybe the title should be changed to take it into account.

After some investigation we can say that the current implementation of the Kalman Filter without any dynamic model is not the major source for the improvements but the applied outlier detection. Therefore, we decided to change the title of the paper.

Planned updates in revised version:

- The title will be probably updated to "*DAHITI – an innovative approach for estimating water level time series over inland water using multi-mission satellite altimetry*"

The question of the biases is not clear. Authors should -as least briefly- state how they evaluated the biases for the tracker that they use, for each mission (were they estimated globally prior the computation of the series?, are they evaluated separately for each series merging several missions?), And -in the case of global values- publish the values.

We apply range biases which are derived from a global multi-mission crossover analysis described in Bosch et al. (2014). For the multi-mission crossover analysis, the ocean products from the different altimeter missions were used. For our application, the radial errors were interpolated (due to missing crossover points) over land to provide an individual range bias correction for each altimeter measurement. The global mean range biases are already published in our manuscript (in Table 1).
This approach works quite well (not only over ocean but also over land) as long as the ocean product is used for the computation of water levels. However as soon as retracking is involved retracker offsets will occur. In order to minimize the relative offsets between different missions and altimeter tracks we use the same retracker (with identical parameters) for all measurements over one target. That minimizes the inter-mission biases. Remaining differences may still be present in the time series (e.g. due to mission-dependent or location-dependent effects such as waves). These inconsistencies will map into the quality of the time series and will show up by increased RMSE when comparing the altimetry product with in-situ data. This is a point which needs additional investigation in the future.

Planned updates in revised version:

- The explanation on handling range biases will be extended
- The problem of remaining inter-mission offset will be discussed additionally in the outlook

The authors mention in the Introduction the key point of the slant measurements (off-nadir measurements). But they do not explain how they deal with it. Do they ignore it ? If not, where is it corrected for? In the pre-processing step ? How is it modelled (best fitting parabola, parabola constrained by geometrical considerations?, other analytical expression?, etc ..)

In the current processing strategy, the hooking effect is not corrected but the effected measurements are excluded from further processing. The restriction of measurement latitude within the preprocessing step limits the usage of off-nadir observations. Furthermore, the additional computed input uncertainty for each altimeter measurement will increase for off-nadir observations. Thus, only few off-nadir measurements will be used within the approach and these will have larger uncertainties.
In addition, we are about to publish an advanced approach for correcting the hooking effect. We plan to integrate this approach in the DAHITI software in the future.

Planned updates in revised version:

- An explanation on the handling of the hooking effect will be added

Legends in the Tables are not complete. For example, in Table 4, what does the N stand for (number of cycles ?) ? Besides, it would be useful to indicate the mission (Jason-2, Envisat, ...). Also, the river widths do not correspond to actual widths of the reaches. Is it the length of the track segment ? Figures showing comparison of time series are not easy to read. For exp, in Figure 9, it is not possible to see if points are missing or hidden by others. Authors should seek for better way to show all the series in a single view (points in the background larger than the points in the foreground, or use different symbols ....)

Planned updates in revised version:

- The legends will be extended
- The used altimeter mission will be added to Table 3 and 4.
- A definition of the river width will be added (it is the length of the track segment based on Google Maps)
- The Figures 7, 8 and 9 will be updated to make them easier to read and more understandable

I don’t comment the English which is better than mine .