

Review of the paper: “DAHITI – An Innovative Approach for Estimating Water Level Time Series over Inland Water using Multi-Mission Satellite Altimetry“ by Schwatke et al.

The authors have improved their manuscript significantly and all issues of the major revision have been addressed. I still see some small issues with respect to investigating the improvements from individual parts of the processing chain. Additionally, the statement of providing better water heights (except from the fact of an improved temporal resolution) compared to other databases is not confirmed comprehensively (see below). Therefore, I recommend the paper for publication after considering the minor issues below.

### **General Comments**

The manuscript focus is much clearer now. The figures and tables are significantly improved.

The (minor) role of the Kalman filter for improving the water level heights is clearer now. Furthermore, the authors mention that the major improvement is due to the various outlier detection processes (see also specific comments).

I agree with the decision to remove the grid from the general computation process while still keeping it as an option for further investigations.

The analysis of the improvements due to individual aspects of the processing chain is more extensive than before. In my opinion there might still be room for further testing, e.g. the improvement due to the first outlier detection, retracking etc., but this would be a little bit out of the scope of this paper. The comparison with a median filter revealed that the major improvements is due to the outlier detection. However, it is still mentioned several times in the text that the Kalman filter is the “central part” of the computation of the DAHITI time series (e.g. Line 415 or 627). I would suggest to rephrase these parts since it gives the impression that the Kalman filter in its current state is mainly responsible for the improved quality of the water heights.

The differences to the other water level databases are explained better, e.g. that no outlier correction is performed. However, to confirm the superior quality (e.g. line 1100 “significant accuracy improvement ...”) of the DAHITI time series, a comparison between the same points (e.g. based on the same Jason data) from the external databases and DAHITI would have to be done.

### **Specific Comments**

---- Title ----

I think the new title better reflects the contents of the manuscript. However, the authors might consider a slight change from “Inland Water” to “Inland Waters” since the database provides water levels for numerous lakes and rivers.

---- Text ----

Line 63: “to be passed” → this should be rephrased to something like “to be crossed”

Lines 120 -121: change to “... have been developed that provide water level time series over inland

waters to the international community.”

Lines 340 – 343: The information about the special file format might be omitted.

Line 408: “black dotted line” → the line in the plot is cyan and dashed

Line 415: maybe rephrase “central part” as it gives the impression that the Kalman filter is the major reason for the improved water heights (see also general comments).

Lines 455 – 475: Since now usually all computations are referred to one “gridpoint” (so just one point), the title of the subsection should be rephrased.

Lines 485-487: This part first gives the reader the impression that there are vectors of  $n$  entries and then it is stated that  $n=1$ . What I mean is that the description tries to sell more than there is actually happening. Of course I understand that the vector notation has to be kept in case a grid is used, but I think that rephrasing of this part would make it more clear.

Lines 491-494: “assigned to them” → rephrase

Line 517: “5cm<sup>2</sup>” and also in line 568. The authors wrote in their response to the first comments that this value is chosen from experience. In that case they should also at least mention it in the text and explain how they arrived at that number.

Lines 579 – 581 and lines 602 – 604: The difference between these two confidence levels is unclear. I suppose one refers to the first outlier detection (before computing the mean height) and one to the outlier detection in the water height time series? This should be made more clear.

Line 692: Maybe rephrase “peak-to-peak” to “seasonal water ranges” which makes it more clear.

Line 763: I think the term “non-retracked” is not correct here unless the authors use the tracker range provided by the altimetry instrument. If standard ocean ranges are utilized, these are retracked with an ocean retracker.

Lines 773-778: The differences between the first outlier detection and the second one should be made more clear. The same applies to Section 4.3.2 and 4.3.3.

Lines 797 – 799: The comparison is not on an equal basis (see general comments). However, this is mentioned in the following.

Line 983: Rephrase “first” to “The first column shows ...”. This makes it more clear that the authors are talking about the table.

Line 1077: Is “der” part of the name of the institute? Otherwise it should be changed to “of the”

Table 3: I think the table could be improved by providing the actual error values used for the outlier criteria instead of the check marks.

Table 3: Some of the outlier percentages are quite high  $>30\%$ . I think this should be mentioned in the

text more clear, since rejecting a lot of measurements automatically improves the correlation and RMS compared to other databases.

## Technical Comments

I have probably not caught everything, so someone with better knowledge of the English language should check it, too.

Generally the authors always write “in situ” but I think “in-situ” is correct (eg. lines 615, 618, 648, ...). This should be changed.

Line 298: comma “... input data, whereas ...”

Line 546: comma “In parallel, the ...”

Line 559: “start” → “initial”

Line 587: comma “In a last step, an ...”

Line 597: remove “as”

Line 623: comma “... South America, we focus ...”

Line 649: remove “the”

Line 666: rephrase to “... extent by a factor of up to 20.”

Line 673: comma “... North America, two ...”

Line 677: “party” → “partly”

Line 682: comma “... mountains, the ...”

Line 683: “in winter” → “during winter”

Line 685: comma “... Lake Buenos Aires, in-situ ...”

Line 697: “and” → “to”

Line 701: remove “the”

Line 707: maybe rephrase “borne” to “kept”

Line 750: change to: “... 1992 to 2014. The ...”

Line 833: comma “... time series, the same ...”

Line 865: “with” → “the”

Line 925: “unequal” → “non-uniform”

Line 941: comma “... time series, as well as ...”

Line 1095: comma “Also, the ...”

Line 1115: “preprocessing” → “pre-processing”