Interactive comment on "A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry" *by* Tim Busker et al.

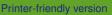## Anonymous Referee #1

Received and published: 22 March 2018

General comments

This paper explores the use of the JRC global surface water dataset, and the DAHITI satellite altimetry database to estimate hypsometry relationships for a reasonable number of lakes across the globe. The paper should be of interest to a people working in water resources, and potentially is a publishable paper. At the moment however, the paper is a fairly simple data analysis with insufficient statistics (i.e. uncertainties) warrant publications as it is. Might improve the paper if the authors consider what they learnt from analysing the dataset, and what limits would they place on the size of the dam might suit using this approach, rather than given vague qualitative statements.

While the paper appears to be overall well written, there are some issues with the

material presented (see comments below), and it would be good to have estimates of the uncertainty in the regressed coefficients (maybe indicated through confidence bounds on the fitted functions shown in the plots would be best?). I think the paper needs some revision before being ready for publication.

Specific comments

1. Page 1, lines 19-25: the average r is given across 18 lakes. Would be good to know what the standard deviation is also as this would at least give the reader some idea of the scatter.

2. Page 4, lines 25-28: The definition of a large lake (ocean-like conditions) is a little vague. Might be useful to have a quantitative definition of what a large lake and a small lake are? Maybe something related to the minimum width of the widest part of the lake?

3. Page 10, Figure 4: there are large departures in the plot for Lake Nasser – what could cause these? How significant are they?

4. Page 11, lines 14: It would be good to give some information on how the uncertainty was obtained, and how the no data pixels were treated in estimating the points shown in Figure 6. Are the red points likely to be lower bounds on the lake volume? From the Figure, these seems to be the case. If they are lower bounds, then the red shaded area seems to span between this low bound and an estimated upper bound. How are the individual pixels within the MWE converted to an area? Is this simply adding up the number of pixels with a detection of a water surface? Appears to be so based on what I can see, in which case this is a very simplified approach, and better estimates of the upper and lower bounds could be made by considering the part of wet pixels at other times (I note that

some discussion on this appears in page 18, reinforcing my interpretation that the simplified approach has been used).

5. Page 14, lines 8-13: Information on data sources seems to be incomplete. Sources for US, Spain and Sudan are given, what about the source for the 2 lakes in Australia?

6. Page 14, lines 13-14: Seems strange to make the statement that the NRMSE is relatively low taking into account all sources of uncertainty, but there is no discussion about what the sources of uncertainty are that have been considered, or the magnitude of the overall uncertainty. Does this statement mean that the NRMSE is a lot smaller than would be expected given the estimated uncertainty? If so, it suggests a possible error in the uncertainty quantification (e.g. ignoring the impact of serial correlation between the different component uncertainties).

7. Page 15, line 1: Yes, extrapolating beyond the limits of the data will result in higher errors. This is why the uncertainty in the regressed coefficients should be reported. Even then, the uncertainty estimated from the regressed quantities will be a lower bound on the uncertainty in the extrapolation as the estimate is based on the assumption that the fitted function still holds. Possible explanation for the over-estimation of the extrapolated storage for Lake Mead shown in Figure 8 (regressed coefficients are time dependent, or relationship is not as linear as was originally thought), or is the red line shown there within the uncertainty bounds for the original regression?

8. Page 19, lines 2-5: A non-linear hypsometry relationship shouldn't mean the lake volumes are unreliable. Just that more care and some more maths is needed to derive the volumes. The main issue would be the choice of fitted function, and how this behaves under extrapolation. Given the result shown in Figure 12, a hyperbolic function that becomes roughly constant as area decreases, and linear

Interactive comment

as area increases would likely be a much better function to fit than a quadratic or a cubic.

Minor comments

1. Page 5, line 26 (and elsewhere): might be better to have all acronyms in capital letters.

2. Page 9, line 11: "Lakes Powell, Kariba, Mead and Nasser"

3. Page 12, line 6 (and elsewhere): km^3 is not a standard SI unit. The equivalent SI unit would be TL (teralitres). Is km^3 acceptable?

4. Page 12, line 23: "during which time it lost approximately 30 km^3"?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-21, 2018.