We would like to thank the reviewer for their quite positive and helpful feedback regarding this paper. We give our responses to each reviewer query in italic typeface below each comment (reviewer comments in regular typeface).

<u>Reviewer comments:</u>

Two things come to mind regarding the method. In many cases this type of river has substantial range in water turbidity that may correlate generally with discharge but also due to other conditions. It would be helpful to know whether the 'light water' /'dark water' categories account for this or if a wider range of RGB values might be expected related to this characteristic of the water.

The reviewer is correct to note that proglacial braided rivers exhibit large ranges of turbidity, both within the same river over time and across similar rivers. In this case, the monitoring section is very close to the glacial terminus (~3.1km), and as such the sediment load is fairly constant, the river well mixed, and sediment relatively unsorted. Given these conditions, the light/dark classes do cover nearly all the turbidity values observed in the Isortoq River after image similarity filtering. In rivers with more variable turbidity, more water/non-water classes might be needed to adequately cover the range of observed sediment loads. As requested, we will add this description of this issue in our revisions.

Also at low flow, in fairly clear water, shallow flows may look different because of the bed being visible under the water which may cause different RGB histograms – to what extent is there a need for a range of filters for these conditions?

Again, this is a good point raised by the reviewer about conditions that occur in braided rivers, but do not occur in the Isortoq where the sediment load is too high for there to be any clear water. In such clear water situations, there are two ways to accurately classify water. The first depends on the similarity filtering, and if none of these clear water images pass this filter, then none will be used to assess W_e. The second way to handle this situation would be to include an additional water category (classification class) and gather training data for that class in addition to the other two water classes. We will also address this comment along with discussion of the turbidity point above, and feel this helps broaden the applicability of the technique.

To what extent does the filtering/selection of images reduce the total number of usable images and the sample size over the range of discharges?

The filters substantially reduce the number of images available for W_e extraction from image collection to classification. For instance, the automated environmental filtering removed 9,487 images with sun glint, shadowing, or winter conditions, leaving 840 images for further operations. The similarity filtering further reduced the image pool to 168 images that were ultimately passed to classification and W_e extraction. This is a high percentage of images removed, but this stringent filtering leaves only very high quality images that are easily classified using the semi-supervised approach. This culling still leaves images with daily (or better) temporal resolution available for parameter extraction. If hourly or better resolution images are needed, then the similarity filtering would need to be performed on iterative batches of images- removing sets of images with different characteristics and creating different training data and classification sets for each group. As requested, we will clarify our writing to underscore this point: while all of this information was already included in the paper, we agree that adding a direct discussion (as in this paragraph) will be beneficial to readers.