Interactive comment on "Automated quantification of floating wood pieces in rivers from video monitoring: a new software tool and validation"

Anonymous Referee #2

Received and published: 3 February 2021

General comments:

In this work, the authors present a software that can automatically detect floating wood from video files of a river cross section. They introduce the processing steps of the software involving the formation of wood masks from training data followed by automatic detection and post-processing based on the recall rates of wood pieces. They discuss the factors that affect wood detection and present a quantitative analysis for automatically extracting wood from imagery. The automatic delineation of wood in imagery is not trivial and I believe this work would be of significant interest to readers of Earth Surface Dynamics and wood researchers. However, there are aspects of the manuscript that could be improved.

First, we want to thank referee #2 for his/her recommendations. Applying them can make this work smoother and more understandable. We corrected all comments in the manuscript. The changes are noted in blue in the revised version of the manuscript.

Overall, I believe the work would benefit from further clarification of the masking procedure and the post-processing steps. Given how important these steps are for accurate analysis, a more detailed description of these components would assist readers when using the software. For example, L151-153,283-285, how are images chosen for the annotation and how many are recommended?

We tried to clarify our explanations in different parts and precisely concerning the masking procedure. Therefore, the following correctios were applied to the text:

Lines 161-163: To set the algorithm parameters, pixelwise annotations of wood under all the observed lighting conditions were used to determine the mean (μ) and standard deviation (σ) of wood piece pixel intensity.

Lines 301-305: This annotation process is time-consuming, so a trade-off must be made regarding the purpose of the an-notated database and its required accuracy. Manual annotations are especially important when it is in-tended to be used within a training procedure, for which different lighting conditions, camera parameters, wood properties, and river hydraulics must be balanced. The rationale for manual annotations in the cur-rent study is presented in section 5.1.

Lines 563-571: The results from the current study were all taken from a single site in which a large database of manual annotations was available for developing the correction procedures. In future applications it is unlikely that such a large database would be available. In such cases it is recommended to first ensure that the images collected are of high quality by following the recommendations in (Ghaffarian et al., 2020; Zhang et al., 2021). As data are collected, the automatic algorithm can be run to identify periods of high wood flux. Manual review of other high-water periods is also recommended to assess whether lighting conditions were preventing the detection of wood. When suitable flood periods with floating wood are identified, manual annotations should be done to create the correction matrices.

However, it should be noted that in this study we only examine one field site and there was no way of knowing at this point what recommendations would be for another site.

A recommendation or assessment of the amount of training data would assist future users of the software. In the post-processing (L475) it was unclear to me how the precision matrices were applied to the day of interest. Are these matrices generated for each piece of wood detected?

To improve the precision, the eight parameters were calculated for each detected object, these values as four pairs were then used in this figure to calculate the precision for each pair. Finally, the total precision was the average of all four precisions in figure 14. Therefore, these parameters were calculated separately for each detected object. To clarify it in the text we add the following sentence to the text: "Having the value of the eight key parameters (four pairs of parameters in Fig 9) for each piece of wood, …"

What about the assessment of wood length between the manual annotations with the software?

This comment was an important part of the work to clarify how we estimate smaller pieces which software could not detect. Based on our observations the software works well for large pieces (O(10m), figure 14.b, c). Also, in some previous works (e.g., Ghaffarian et al. 2020) we observed that the length distribution is unique in a section of the river. Therefore, the length distribution of annotations was used as a benchmark to estimate smaller wood pieces. therefore, we add the following sentence to line [505]: "Based on this idea, the final step in the post processing is to estimate smaller wood pieces that were not detected by the software using the length distribution extracted by the annotations."

Also, at the times in the manuscript, it was unclear when you mention "it" or "object" what you were actually referring to- the software, a wood piece, the entire video set? I highlight some specific examples of this below.

Concerning the term "object" we used it as a detection which can be a piece of wood (TP) or a false detection (FP). However, it is important to clarify the reference of each term and this comment was taken into account more precisely in the revision of the text.

Specific comments:

I may have missed it, but a link to the software or the details of how to access the software would be useful.

The software will be freely available on GitHub soon and it will be indicated in the final version of the paper in section "Code/Data/Sample availability".

L164: I was confused on your use of "stable" here. Wouldn't "stable" wood detections be false positives as they would actually be vegetation or the bridge being detected?

The sentence should be corrected as follows:

Meanwhile, the intensity of pixels that keep on displaying wood tend to be rather stable

L150 and 168: You first define variable x as the pixel light intensity and then again as position in the image. Please revise.

Instead of "x", we use "i".

L429: How did you measure the correlation between these parameters?

To clarify the text the following sentence was added: "… by calculating each one of the eight parameters for all detections as one vector and then calculating the correlation between each pair of parameters."

L443: What is "it"?

"It" substituted by: "This spatial gradient in precision is likely …".

Technical comments:

L209: Replace "by" with "be"

It is corrected.

L222: should be "calculate"

It is corrected.

L241: Please revise sentence

It is revised.

L278: Should be "manual"

It is corrected.

L288-290: Some grammar issues, please revise.

It is corrected.

L452: Unclear what you mean by "many noises in a frame".

For clarification it is referred to **Error! Reference source not found.**.

Table 5: Is PR improvement also a percentage?

Yes, (%) is added.

L506: Replace "counting" with "counted"

It is corrected.

Fig 15: Horizontal axis label is cut off

It is corrected.