First of all, we appreciate the comments provided. Your expert opinion is helping this paper to improve its robustness and to reach more readers.

I will answer your comments in the following lines. These answers are in 'Italics' :

This paper talked about the upscaling of Drone image classification to open access satellite images, and applied to the coastal wetlands. I do think it is a good idea to overcome the disadvantages of low spatial resolution of open-access satellite images. It is a key step to achieve the continuous monitoring of coastal wetlands at a low cost.

However, I have a few concerns about this method. I suggest this manuscript needs to be major revised.

If I understand correctly, the author establishes an individual RF regression model for each community to each study site. My concern first concern is that have the authors united the plant fractional cover (PFC) of each landcover to make sure the sum value of the PFC of each community within each Sentinel pixel is equal to 1. If so, please highlight it or remind me where I can find it in the main text. If not, I suggest the authors to rescale the retrieved results and test the accuracy again. Some previous analyses have indicated that such rescaling can change the accuracy obviously (e.g., Immitzer et al., Remote Sensing of Environment, doi: 10.1016/j.rse.2017.09.031; Yang et al., Remote Sensing, doi: 10.3390/rs12193224).

We created a dataset for training – validation and testing based on PC values of each plant community that sum 1 (this is, 100% cover) within each space unit of MSI pixel. We will write a clearer sentence in the main text, section 2.5, as follows: "PFC was extracted for each plant community within each MSI pixel. PFC sums 100% for each plant community".

Here are some specific comments in the main text:

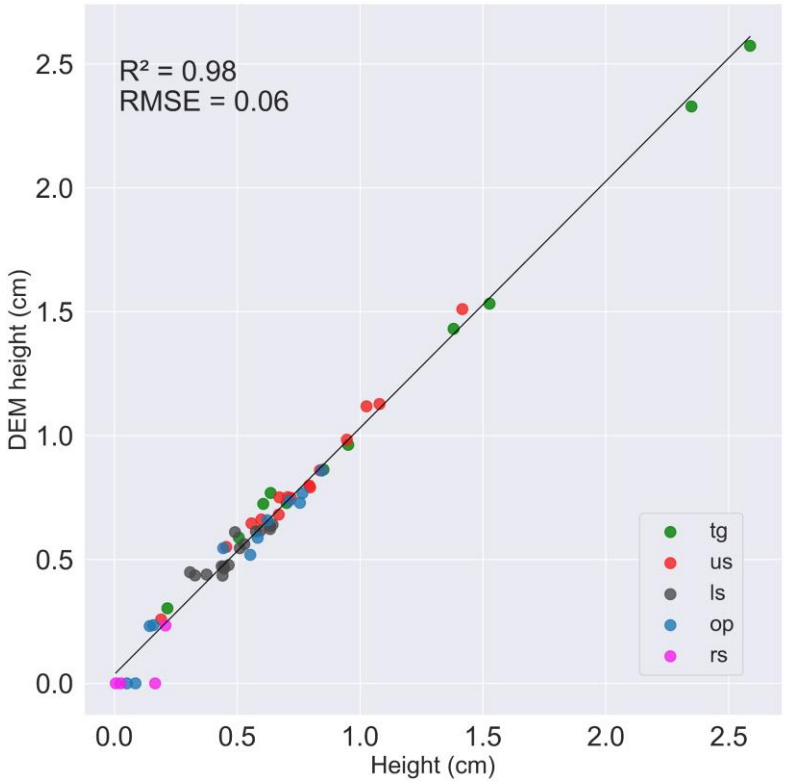
1. Do the authors think that tidal level would affect the accuracy of PFC estimation? Some previous works mentioned that tidal level can significantly affect the land cover classification (e.g., Kearney et al., 2009 Journal of Coastal Research, doi: 10.2112/08-1080.1). I suggest that the tidal level at the time of each drone image and Sentinal image acquisition should be reported

   I will add the following clarification: "In Estonian coastal wetlands, the tidal variation is negligible (0.02 m) and the range of plant communities is maintained through low-intensity grazing (Ward et al. 2016)"

2. The authors examine the accuracy of applications with and without elevation data provided by DEM**.** Can authors specify that did you use the same points to train and

test your two applications? I think it is important to show that adding DEM is useful. In addition, can the authors show the plots of DEM in the main text or supporting information?

Definitely, we used the same points for both applications (models). Thus, we will add the following sentence: "Both dataframes were trained with the same training samples". In addition, we will include a plot comparing the DEM values and ground truth measured by a differential GPS which is going to be described in the Methodology section. The figure is shown as follows:



Where "DEM height" shows the elevation values from the Digital Elevation Model and "Height" is the elevation measured with the differential GPS in the field. Both have units of centimeters (cm).

3. The introduction section needs to be reshaped to introduce the topic step by step. For example, you talked about remote sensing in line 20, and talk about wetlands, and then move to remote sensing again. Another example is that Lines 33- 40 are likely to appear in the study areas section.

Thank you, we will clarify this part of the introduction

4. Line 39. There are two references, i.e., Ward et al., 2016a and Ward et al., 2016. I did not find Ward et al., 2016a in your reference list. Another obvious error is Line 370. So I suggest the author please check this and also other references.

   This is reviewed and clarified.

5. Line 70. Please define VI, although you have defined it in the abstract. In addition, there are too many abbreviations, some of which were just used a few times, making the manuscript difficult to read and understand. So please remove unuseful abbreviations. And I also suggest the authors construct a table to explain each abbreviation.

   We will add the list of abbreviations to clarify them.

6. Line 90 - 95. Please specify the manufacture of your drone here. Please also specify the procedure of your radiometric correction, parameters and models used here. I think they are also useful to other researchers to do similar things.

   The manufacturer is senseFly and they do not provide the internal functioning of radiometric correction. We follow the best practices from senseFly to carry out the radriometric correction by using the Airnov panels.

7. Line 95 – 100. Can you also please show the confusion matrix for your classifications here? This would help to show that your selection of RF makes more sense.

   It is possible to find in the referenced article (Prentice et al. 2021) the metrics of accuracy, kappa and comparison metrics between the two classifiers that the authors used in that article. The present study does not aim to show all the details of the referenced one but we will show the exact numbers of metrics.

8. Line 123. Can you please explain more about the accuracy of the DEM used here. From my point of view, the accuracy of lidar-dem over wetlands is a bit low. So I think the specification is useful to show the robustness of your method.

   We will provide the RMSE values between ground-truth DEM values and DEM model used in this section with the figure mentioned in point number 2 of this document.

9. Table 2. I am not sure I lost something. But I do not know what @ means in this table. Please explain it in the table caption. And please show the unit (probably nm) of each band.

We will add the term "nm" for nanometers. The symbol @ means the bandwith and we will specify it too.

10. I think the first row of each table can be highlighted, making it easier to read and review.

    We will do this change.

11. I also suggest that Figs. A1 and A2 can appear in the main text to better display the accuracy.

    We considered including these figures in the main text before but as we have already eight figures and four tables, we keep them in the annexes.

12. In the discussion part, I suggest the authors explain the value of the proposed method in future research.

    This part will be clarified and extended in order to explain more about the impact of this study.

Again, thank you for the comments. I hope we have solve them. We will work to make these changes in the main manuscript.