Answers to reviewer

Thank you for your concise reviews and valuable suggestions. We have carefully considered each of them while revisiting the manuscript. Our responses to the comments and suggestions are provided in the following lines.
Italic font type: Author's replies.

The authors test the possibility to estimate the relative cover of different coastal communities by coupling UAV image and public-available but coarse image. It is interesting for me and of high importance for future coastal monitoring. And I appreciate the authors try to solve my concerns. I think it is good to product if the following concerns solved.

My main concern is, for me, Fig. 8 and Fig. 10 looks very similar. I suppose the authors may use them mistakenly. So I suggest the authors please double check and correct them if it is needed. After that, I suggest that the Fig. 10 can go to supporting information.

Thank you very much for this comment, we have fixed this issue. Figure 10 shows the lower performance of RF regressions trained with DF2 after rescaling the predicted PFC. Figure 8 shows the prediction errors of models trained with DF1.

There are some minor issues from my point of view.

1) Line 7: I suppose Vegetation Indices (VIs) makes more sense than vegetation indices (VI) here. And all 'VI' over the main test may be plural.

Indeed, it is easier for the reader to use the abbreviation VI for Vegetation Index, and then, use VIs for the plural form. Accordingly, we have modified this throughout the text and in the Table 1.

2) Line 227: which models were not used map? I did not understand. Can you please make it clearer?

The models that were not used to map the PFC over the study areas were those trained with DF1 and rescaled values from DF2. Thus, only models with relative values of PFC from DF2 were used to map the PFC over the study areas. We have clarified this point and simplified the lines 224 to 231.

3) Lines 254-257: I agree with the fact that the mixture is one of the reasons for uncertainties of RF. But do you think it also can be attributed to the nature of the RF. The RF regression makes prediction by using the average value of each tree inside, thus avoiding the appearance of extremely low or high values. Do you think this can be solved by including dominant and mixed samples? It is just my speculation. What do you think about that?

In Random Forest modeling, dealing with uncertainty and ensuring the model is interpretable is a challenging task. Indeed, the averaging of individual tree predictions during training influences Random Forest regression predictions. We overcome the uncertainties with 1) sampling a balanced dataset that represents the entire spectrum of PFC values and 2) an approach to hyperparameter tuning with the Grid Search Cross Validation.

Before the under-sampling to balance the dataset, we studied the distribution of all the PFC values in the initial dataset (DI), having an unbalanced distribution towards the extreme values of PFC due to a

completely or partially absence of plant communities in a grid (S2 pixel) or high cover. We performed various Random Forest regressions keeping these values in order to test the stability of the model under different training scenarios. Models exhibited good scores during training but suboptimal test scores, suggesting the possibility of overfitting. Thus, we considered that keeping the unbalanced extreme values was not appropriate to train the models. In order to keep samples that are more mixed, we also tried reducing the interval of the bins and this led to worse performance of the models. The best approach was used in the present study.

We agree that RF may also incorporate a certain degree of uncertainty in its architecture. In our opinion, not only reducing the uncertainties may also depend on complexity and dataset characteristics but also the model complexity.

We have addressed this discussion in lines 250 -259.

4) Line 173-174: how do you carry out the under-sampling? Just randomly select some? Can you please specify here? One question from curiosity, do you think the under-sampling emphasize the above-mentioned over/under-estimation problem?

The under-sampling was done by selecting the S2-pixels randomly with PFC values contained in a bin. The number of S2-pixels per bin equals the number of values in the minority bin.

The under-sampling strategy was done to balance the initial dataset (DI) to avoid the model to learn from one bin more than another, avoiding potential overfitting. However, we appreciate the question from the reviewer and we think that over or under-estimation could be improved by constructing an exhaustive training dataset from field survey plots equal to the area of S2 pixels. However, this is high time consuming due to logistic issues.

This clarification is added in lines 169 to 175 and discussed in lines 255-259, because they align with the previous comment.

5) Line269: As for the importance analyses. I am curious about that VIs represent spectral information, or different fusions of spectral bands. So different VIs, although calculated via different equations, contain somehow similar information. So I think it is a bit arbitrary to say DEM is the most important variable for some cases, although importance value from RF is very high. However, if we consider all VIs are spectral information, your analyses show that spectral information is more important than elevation information in all cases. How about your thoughts about that?

First of all, the Digital Elevation Model, although its coarse spatial resolution in this study, still represents the microtopography in these type of wetlands as seen in the results. Plant communities are strongly dependent on this variation.

In relation to the spectral information from Vegetation indices, the Random Forest algorithm evaluates the individual contribution of each Vegetation Index and Digital Elevation Model to a single tree's total impurity reduction, meaning that it calculates the importance independently. This is a relative measure of the model performance. The lack of additivity is due to the complex interactions and non-linear relationships considered by the Random Forest algorithm, which can lead to the different Vegetation Indices being more or less important independently of their calculation. In this way, the Digital Elevation

Model is relatively more important for the model to estimate the PFC. In other words, it makes more accurate predictions within the ensemble of decision trees that make up the Random Forest.

Since variable importance in RF is not an additive variable, we cannot not really claim that spectral information is more important than the DEM. In a similar manner, VIs are not additive representations of the electromagnetic spectrum. In that regard, we should not add them together and take it as "spectral information". We hope this answer is satisfactory and clarifies the idea of VIS and variable importance in RF.

In order to summarize all these points, we have added a clarification in the lines in Methodology section 189 – 191, and in the Discussion section in lines 280-284.