Referee 1

Overall, the manuscript entitled 'Interactions between biogeochemical and management factors explain soil organic carbon in Pyrenean grasslands' would have potential to be of great interest for the readers of Biogeosciences Journal. It provides interesting results on the effect of different drivers on soil carbon stocks in Pyrenean grasslands. However, I have noticed some important points that need to be addressed before this manuscript can be considered for publication.

We really appreciate your revision. Your comments definitively will improve our manuscript.

Concerning the abstract, I think that the scope and objectives of the study need to be better defined. After reading it, we do not have a clear idea of what factors have been tested. I have the same feeling after reading the introduction. Overall, we understand that there are many factors which can influence soil C stocks at different scales, but it is difficult to understand what are the real objectives of the study. Is the objective to determine which factors influence the most the soil C stocks, is this analysis done for different scales?

We have revised the abstract and the introduction sections, following your specific comments. Under our view, the scope and objectives are now more understandable. In a nutshell, the scope is to study the relative effects, including interaction effects, of geophysical and biochemical SOC drivers, and also to pinpoint how grazing management regulates the effects of other SOC controls.

In the material and methods section, the main issue that I noticed concerns the statistical approach. It is not clear for me why two separate approaches were done. It adds a certain complexity to the article and it needs to be better presented according to the objectives for each approach. Are both the approaches really relevant for the paper? The links between the objectives and the chosen modelling approach needs to be better defined. Also, concerning the calculation of soil C stocks, it would have been appropriate to correct soil C stocks according to the equivalent soil mass approach to account for possible differences in bulk density values (Ellert and Bettany, 1995; Ellert et al., 2008).

We explained in the specific comments why we think both statistical approaches are complementary and important. We revised the manuscript to emphasize and make clear this point. However, if our arguments neither convince nor the editors nor the referees, we are open to put the BTR model in supplementary material or even suppress it completely. Note that although it has been argued that the usefulness of using both approaches was not clear, referee 1 made several specific questions about the differences between their results, precisely about the points we consider interesting. We also commented the point about fixed mass approach for calculating SOC stocks in the corresponding specific comment.

Concerning results and discussion, even if the ideas are, overall, well supported by relevant references and the limits are underlined, I think that the organization will be improved after the clarification of the objectives

and the corresponding analyses. Also I noticed repetitions of results in the 'results' section and in the 'discussion' section so I would suggest to group all the results and discussion in one section if the journal guidelines allow it.

We think separate sections for results and discussion are important, since this is useful for separating the raw statistical results from results discussion and interpretations. We truly believe that the manuscript is going to be easier to read and understand if we maintain this structure. The statistical methods presented here could seem complex, and reading the results separately could help to their understanding since is the shortest and simplest section. Anyway, we followed your advice and we revised the manuscript to make it less repetitive. Your specific comments where greatly valuable for this task. The most important change is we suppressed the first paragraph in the discussion section, which was actually a summary of the result section. We also revised the paragraph about the modeling procedure, and we believe now is more clear.

We think the rest of the subsection titles in the discussion were useful to structure the text. Under our view, every sub-section was justified. However, we grouped the second and third subsections (Geophysical, biochemical and grazing management factors driving SOC stocks) as both referees asked us to reduce the number of sections. The idea is that first section gives an idea of the right way of interpreting the models. The second section answers the questions formulated at the end of the introduction; 1: "what are the relative and interaction effects of the geophysical and biochemical SOC controls?" and 2: "How grazing management regulate the effects of other SOC drivers? The fourth subsection discuss particular implications of our results in grazing management, a point which we consider noticeably important. Finally, we separated and revised the conclusion section following the indications of referee 2.

Of course, if after this revision, the referees and the editor consider that results and discussion section must be combined, we could do it without a problem.

Finally, it would be important that the manuscript be reviewed for the English. Some corrections might be necessary.

We did so.

In the next paragraph, I developed some detailed comments that will help the authors to improve the manuscript.

L 53-54 "at small spatial scales" instead of "at detailed spatial scales".

Change done. (L54)

L 56-57 I am not sure that it is a good reason to do a study... What is the objective of the study by

using this set of data?

To clarify this, we rewrote this sentence as follows:

“Taking advantage of the high variety of environmental heterogeneity in the Pyrenees, we built a dataset that comprise a wide range of environmental and management conditions.” (L56)

L 58 Do the authors have an explicative purpose or a predictive purpose? That is not clear for me, as they also use the ‘predictors’ term.

The study has an explicative purpose. We have changed “predictors” by “drivers” or “factors” in all the text to avoid misinterpretations.

L 59 This factor should be better defined.

We specified it in the following way:

“We found that temperature seasonality (difference between mean summer temperature and mean annual temperature; TSIS) was the most important geophysical driver of SOC in our study.” (L 60)

L 65 I think that the coma is not necessary.

The comma was removed. (L 67)

L. 95-96 I think that these variables should be better described. Also "be" should be removed.

These factors are not studied or they are not factors with a relevant impact in other studies?

This phrase was rewritten as follows, to clarify that these factors were not even considered in these previous studies and the meaning of climate seasonality:

“However, climate regular annual variations, represented by seasonality variables, are be commonly neglected when considering possible SOC drivers affecting SOC in broad-scale models, in spite of being some important factors for plant primary production or enzymatic activity of soil microorganisms.” (L 97 - 99)

L.112 Same question than earlier: are they omitted because they do not impact the SOC stocks?

The phrase was rewritten as follows:

However, these variables are commonly omitted as possible drivers of SOC in the broad-scale SOC studies, especially if those focusing on predictive models instead of explicative ones” (L 115 - 116)

L. 113 “focusing” instead of “focus”

Change done. (L 116)

L. 116 Overall, for the whole manuscript, the authors need to specify if it is SOC stock or concentration.

In is SOC stocks. That is now specified in multiple parts of the text.

L.127 What type of management do you consider?

In our case (natural grasslands), we consider livestock management. However, according to the cited article (Wiesmeier et al., 2019), management effects on grassland SOC in general are poorly understood. We rephrase the sentence:

"Apart from these factors, livestock management effects on grassland SOC is…" to clarify that in this paper we only refer to livestock management, which is from far the main management done in natural grasslands." (L 130).

L. 136 And what was their conclusion in regards of your objectives?

The conclusion in regards of our objectives was that grazing must be considered as a variable that can interact with many variables at multiple scales (as it is represented in Fig. 1). We reordered this paragraph and completed this particular sentence to clarify this point:

"It is known that herbivores can affect SOC through different paths, such as regulating the quantity and quality of organic matter returned to soil (Bardgett and Wardle, 2003), or affecting soil respiration and nutrients by animal trampling or soil microbiota alteration (Lu et al., 2017). Moreover, several studies describing interactions of grazing with other SOC drivers at diverse scales have been published (Abdalla et al., 2018; Eze et al., 2018; Lu et al., 2015, 2017; Zhou et al., 2017). Hence, grazing management on grasslands may be considered a unique SOC driver, because it has effects at multiple levels of the driver hierarchy (Fig. 1), both affecting other SOC drivers and interacting with them. However, most of the studies investigating grazing effects on SOC focus on grazing intensity, in spite of evidence pointing to a greater role of grazer species in determining vegetation and SOC (Chang et al., 2018; Sebastia et al., 2008)." (L 131-143)

L. 140 Among which drivers? There are many factors that can interact or be correlated together. We need to know which drivers will be tested. The authors should be clearer on the objectives of this study.

We modified this paragraph as follows:

"In this study, our goal was to identify the main drivers of SOC stocks and their interactions on Pyrenean mountain grasslands. For this purpose, we considered a wide set of regional, landscape, soil geophysical and biochemical and herbage quality factors, together with grazing management factors. Mountain grasslands comprise a wide range of all these conditions that make carbon stocks highly variable (Garcia-Pausas et al., 2007, 2017). For this reason data analyzed here comprise a wide range of environmental conditions, comparable to studies on SOC developed at continental or even worldwide scales (Table 1).

Additionally, we consider an exceptionally broad compilation of drivers (Table 1). To deal with correlations and interactions between SOC drivers, we developed an exhaustive modelling approach based on the controls over function hypothesis (de Vries et al., 2012)." (L 145 - 155)

L. 141 To asses

This sentence was changed and this word does not appear.

151-153 Do the authors want to study the effects of various factors, their links between them, the importance of the factors...?

We rewritten the questions as follows to put that point clear:

"1) what are the relative and interaction effects of the geophysical and biochemical SOC controls? 2) How grazing management regulate the effects of other SOC drivers?" (L 162)

L.175 grazer type instead of grazing management

Change done

L 189-190 Are the soil samples from the 4 quadrats composited to form one soil sample per depth for each grassland patch?

Yes, they are. Following the advice of referee 2, we made many clarifications about sampling design, including a new supplementary figure (Fig. S2).

L. 192-193 I think this paragraph should appear before...

We appreciate this comment, and we also recon that this paragraph could appear at the beginning of the methods section. However, we still find clearer to explain first how the sampling was performed and second how the samples were processed in order to get the environmental variables.

L. 194 There should be a coma between landscape and livestock

Change done (L 208)

L. 199 But you don't speak of mean summer temperature before...

We added MST where climate variables were introduced, in the second paragraph of the section 2.2: Regional and Landscape environmental drivers:

"Regional variables included climate variables and bedrock. Climate variables were determined from Worldclim 2.0 (Fick and Hijmans, 2017). We selected Mean Annual Temperature (MAT), Mean Summer Temperature (MST), Mean Annual Precipitation (MAP) and Mean Summer Precipitation (MSP)." (L 210 - 213)

L. 200-201 How did you appreciate that? We need to have more details on this factor.

During preliminary modelling exercises, two climatic variables appeared repeatedly in all tested models, always significantly contributing to the models with coefficients of similar magnitude but opposite signs: mean annual precipitation and mean summer temperature. In those initial models, even when those two variables were included in interaction with other drivers, this pattern was maintained; that is, the interactions were of opposite sign and the coefficients of the interactions were similar in magnitude. In this way, this is how the TSIS index initially emerged. This index has been found a significant driver of different variables of interest in the PASTUS database, including SOC and plat diversity (Rodríguez et al. 2018).

L. 218 For each patch considered?

Those grasslands are usually managed communally, and the livestock type and units are based on the number of animals, and type, sent to graze a given area during the grazing season. The unit area is usually related to the municipalities, although this situation might change a little depending on the mountain range. Grazing in the high-altitude grasslands in the Pyrenees is usually free-range.

L. 229 For determination of bulk density?

Yes. We modified the sentence as follows to clarify this point:

“Soil samples were air-dried and weighted to obtain its bulk density. “ (L 244)

L. 233-234 This sentence is not clear.

We rephrased this sentence to clarify it:

“We combined 0-10 and 10-20 cm values for obtaining the whole top 20 cm soil layer.” (L 249 - 250)

L. 243 It should have been important to correct soil C stocks according to the equivalent soil mass approach.

We decided to use a fixed depth approach for calculating SOC stocks due to the following reasons. First, the main advantage of fixed mass approaches is that they account and correct differences in bulk density due to temporal changes or when comparing different land uses (Haden et al. 2020). We do not consider variations in time, and neither have contrasting management regimes, as mentioned in the title of Ellert & Bettany´s paper (1995). We highlight that in our work samples came from natural mountain grasslands, where grazing intensity is always low to moderate, and moreover, herbivore presence is seasonal. Therefore, we do not expect important changes in bulk density due to land use. Second, we always used the same methods in our samplings (so we could not take advantage of fixed mass approaches for correcting biases due to different probe diameters, as suggested by Sharma et al. (2020). Finally, fixed mass approaches often have more technical difficulties than fixed depth measures even in the most modern procedures (Haden *et al.* 2020). On the other hand, Rovira et al. (2015) proposed a fixed mass approach

which, as expected, was found to deal properly with bulk density changes but not with stoniness differences. We did not find any other reference dealing with this point.

To clarify this point, we added the following lines to the text:

“Soil organic carbon (SOC) stocks in the upper 20 cm soil layer were then estimated taking into account the organic C concentration in the sample and its bulk density, and subtracting the coarse particle (> 2 mm) content, following García-Pausas et al. (2007). Despite recent studies suggested that fixed mass SOC stocks estimates are preferable to fixed depth methods because they would be more robust to temporal and land use changes in bulk density (Ellert & Bettany 1995), we chose a fixed depth method for measuring SOC stocks. This decision was based on the fact that our work samples came from natural mountain grasslands, where grazing intensity is always low to moderate, and moreover, herbivore presence is seasonal. Therefore, we do not expect important changes in bulk density due to land use. Additionally fixed mass approaches presented the disadvantages of implying more technical difficulties than fixed depth measures, even in the most modern procedures (Haden *et al.* 2020), and could not deal well with differences in stoniness.” (L 264 - 273)

L. 249 What was the vegetation: grassland species etc.

We added the following paragraph to provide that information:

“Almost all of the plant species in the grasslands from the PASTUS database are perennial (Sebastià, 2004), and plant biodiversity is highly heterogeneous as are the environmental conditions (Rodríguez et al., 2018).”

As this is not bromatological information we added this paragraph in lines 178 - 179 where describing the sample site conditions.

L. 267 The size of the police is not the same for all this paragraph. Does this paragraph of NIRS analysis refer to the analyses presented in the previous paragraph? It is not clear.

We changed some sentences in these paragraphs so now the relationship between these two methods is clear. Basically, bromatological analysis were done for training NIRS models and getting the remaining values using NIRS spectrum. (L 275)

L. 293 Among which variables?

Among SOC and all the considered drivers. To clarify, we modified the sentence as follows:

“Including all potential SOC drivers, we fitted a model with BRT to identify the most important ones affecting SOC.” (L 321)

L.301 “Firstly” instead of “First”

Change done. (L 330)

L. 306-307 What is this new set of variables?

This is standard procedure, according to (Elith *et al.* 2008). As it is explained, it refers to the variables that improve BTR model performance, the model set showed in Fig. 2.

L. 314-316 Why choosing these two models, on which hypothesis did you decide these two groups?

The geophysical variables are those commonly used in the literature, and are the first source of variation according to the hierarchy of controls over function hypothesis (Manning et al., 2015). Choosing these two models allows us to discuss the effects of geophysical variables on SOC without deleting some effects because of the inclusion of other variables (especially soil nutrients) whose effects may include those of geophysical variables, because geophysical variables could act trough other variables at smaller spatial scales (in this case, the biochemical variables). We consider Geophysical Model is interesting for discussion, since it allow comparisons with previous literature. Additionally, we reported which terms of the Geophysical Model were substituted by the biochemical variables, which suggests that those effects could affect SOC trough biochemical variables, while the other effects probably acted trough other mechanisms too. Finally, we believe that Geophysical Model has interest for future studies aiming to predict SOC in similar environmental conditions. As we mentioned before, these studies usually use what we call here geophysical variables, because they are easy and cheap to measure or obtain (Manning et al., 2015). We modified the referred sentence as follows, to emphasize some of these points:

“We built two models (Fig. S4), one model only based on geophysical drivers and grazing management (Geophysical Model), and another model by adding to the former the biochemical drivers: soil nutrients and herbage quality drivers (Combined Model). With this approach we aim to avoid ignoring significant effects of the geophysical variables, the original source of variation of SOC stocks according to the hierarchy of controls over function hypothesis (Manning et al., 2015), by masking them with the inclusion of biochemical drivers.” (L 343 - 349).

We also added the following modifications in the “Geophysical drivers driving SOC stocks” subsection of the discussion section to a better explanation of the usefulness of the Geophysical model:

“Considering the difficulties of modelling SOC in a widely heterogeneous mountain environment (Garcia-Pausas et al., 2017), the Geophysical Model provided important information about SOC drivers in the Pyrenees. This information could be useful not only for a better understanding of SOC patterns in mountain grasslands, but to future modelling studies aiming to predict SOC, since geophysical variables are easier and cheaper to obtain and measure compared do biochemical ones (Manning et al., 2015).” (L 470 – 476)

L. 316-320 Maybe it should be more appropriate in the introduction...

We think this is appropriate for the methods as it contributes to the understanding of the modeling procedure. However, we modified the last paragraph of the introduction, to specify these aspects too. In overall we believe that this important paragraph has been widely improved thanks to your comments and suggestions. This text is now as follows:

"In this study, our goal was to identify the main drivers of SOC stocks and their interactions on Pyrenean mountain grasslands. For this purpose, we considered a wide set of regional, landscape, soil geophysical and biochemical and herbage quality factors, together with grazing management factors. Mountain grasslands comprise a wide range of all these conditions that make carbon stocks highly variable (Garcia-Pausas et al., 2007, 2017). For this reason data analysed here comprise a wide range of environmental conditions, comparable to studies on SOC developed at continental or even worldwide scales (Table 1). Additionally, we consider an exceptionally broad compilation of drivers (Table 1). To deal with correlations and interactions between SOC drivers, we developed an exhaustive modelling approach based on the controls over function hypothesis (de Vries et al., 2012). To facilitate the formulation of our hypothesis, we classified SOC drivers into three main groups (Fig. 1): i) geophysical factors, which include regional and landscape factors and are supposed to be the first sources of variation, ii) biochemical factors, which include soil nutrients and herbage factors and could be conditioned by geophysical factors, and iii) grazing management factors, which could affect SOC through multiple interactions with the rest of the variables at multiple scales. In particular, the specific questions of this study are 1) what are the relative and interaction effects of the geophysical and biochemical SOC controls? 2) How grazing management regulate the effects of other SOC drivers?" (L144 - 164)

L.374 Why there are not all the predictors described in the introduction in this model? Grazing management for example?

Because they were discarded in the BTR modelling procedure. Note that three-based methods can have difficulties in modeling some functions (Elith *et al.* 2008). We answered about why using both methods three comments below.

L.381 Why these two variables are not selected in the model?

This point was discussed in the "Considerations about the modelling procedure" subsection inside the discussion section. "As a regression tree machine learning technique, the BTR model identified a set of SOC stocks predictor drivers (Fig. 2) avoiding some of the linear model disadvantages, like guarding against the elimination of good predictor drivers correlated to others or automatically modelling non-linear effects (Cutler et al., 2007; Elith et al., 2008).Thus, the BRT model included some SOC predictors, like a positive logarithmic

like effect of aboveground biomass or soil K on SOC (Fig. S7), which could be masked by the effects of other variables in our linear models (Yang et al., 2009)."

Basically, multiple predictor variables can not only be correlated but also have true cause-effect relationships between them (i.e. precipitation and aboveground biomass), what means that in a linear model, some drivers could be discarded not because they have no effects on the response variable, but because their effects were already included in other variable. In other words, some variables, like aboveground biomass, soil K or silt were not included in the linear models probably because they were correlated with other drivers which were included in the models. The advantage of including BTR analysis is that we could detect some of these variables.

L.411 Some repetition from the results section...

We deleted this paragraph as it is repetitive.

L.444-447 I wonder if the BRT model is really relevant for the manuscript. . . Also, Are you sure it is table S3???

It is table S5 (change done). To summarize, the BTR model is relevant insofar it provides information about the effects of the variables not included in the linear models due to correlation. Is this information relevant enough for the manuscript? We think it is. For instance, if BTR model were not included, one question one referee or regular reader would ask would be: "How you can explain that aboveground biomass was not included in your models?" "Does it means that aboveground biomass had no effects on SOC?" The answer is that aboveground biomass had effects on SOC, but in the GLMs these effects were masked by other variables which explain more variation than aboveground biomass, and probably affect SOC thought affecting aboveground biomass. Note that BRT model also is mentioned in the discussion of topography effects, as it provided information about potential paths through which topography would be exerting its effects on SOC. However, as both referees have the similar questions about the BTR model, we would be opened to move it to supplementary materials or even suppress it you find that our explanations and the information provided by the model are not relevant enough for this manuscript.

L. 487 SOC decrease with increase of slope

Change done. (L 511)

L.489 Not clear. . .

To clarify this sentence, we changed it as follows:

"In addition, high TSIS values compensated SOC stocks decrease with increase of slope, which could be due to reduced carbon inputs and increased carbon losses induced by steeper slopes" (L 511 - 513)

L.491 What I see is that SOC stocks are lower under low intensity of grazing for low values of TSIS. . .

"We changed the sentence as follows:

"Increases in grazing pressure elevated SOC stocks under low TSIS values (Fig 3D)." (L 514)

L.494-499 It is not really clear.

We changed some sentences in this paragraph to make it clear:

"Increases in grazing pressure elevated SOC stocks under low TSIS values (Fig 3D). This would be a surprising result according to recent meta-analyses, which conclude that grazing has a commonly decreasing effect on SOC (Abdalla et al., 2018; Eze et al., 2018; Mcsherry and Ritchie, 2013). However these effects were strongly context-specific, depending on other factors like climate and soil type vegetation (Abdalla et al., 2018; Eze et al., 2018; Mcsherry and Ritchie, 2013). Moreover, light and medium grazing intensities can increase SOC inputs by dung deposition and promoting aboveground and root biomass production (Franzluebbers et al., 2000; Zeng et al., 2015). Considering that in our natural grasslands all grazing intensities are relatively low (see methods), our medium and high stock rates may increase soil carbon inputs in low seasonality locations by enhancing aboveground and belowground productivity." (L 514 - 525)

L.507 high soil water contents?

To clarify, we changed the sentence as follows:

"In our study, high soil water contents caused by high MAP may inhibit decomposition if a shortage of oxygen supply occurs (Xu et al., 2016b)." (L 533 - 534)

L.525 "which might be explained by" instead of "which is an indicator"

Change done.