Author´s response

Dear editor:

In the following lines you can find the answers to your comments. As we did a new deep revision of the text, we also have added an updated version of the answers to the reviewers.

Regards,

The Authors

Editor comments:

Comments to the Author:

Dear authors,

Thanks a lot for responding to the reviewer comments. Both reviewer found that this manuscript presents an interesting data set and makes a significant contribution in exploring factors of SOC storage in grasslands. However, both reviewer noted that there are several formal aspects, which can be clarified as outlined in your responses. However, the overall writing has to be improved considerably.

We really appreciate the comments and suggestions the reviewers and you have done. The manuscript has definitively improved a lot. We have done an especial effort in improving the overall writtling.

In addition to the reviewer, I have the following comments:

1. In the Abstract make clear on how many sites your study is based on and to which depth samples have been taken.

Changes done. (L 87)

2. Take care that all abbreviations are explained properly. E.G. BRT might be known by statisticians but not by all soil scientists.

We checked the remaining abbreviations in the text, and under our view they are well explained. In the case of BRT models, we added the following information.

"We applied two different modelling procedures: Boosted Regression Trees (BRT) and General Linear Models (GLM). BRT is an automatic technique that combines insights from traditional statistical modelling and machine learning traditions (Elith et al., 2008). GLM allowed us to design a hypothesis-based modelling procedure, ensuring that only effects with biological meaning where included; whereas BRT provided information about the data that could be neglected, if only a GLM approach was followed." (L331-336)

3. Tables. Explain all abbreviations (especially the not well known ones, e.g. TSIS). Tables and Figures should be self-explanatory

We added abbreviation explaining in all table and figure captions, including supplementary material.

4. Final model. Is it meaningful to include soil N? More than 90% of N are bound to soil organic matter and consequently N must correlate with C, at least their concentrations. Due to this inherent relationship, I would argue that N is a controlling factor. In the model you have used to explain stocks (whose estimate includes bulk density), but you have to provide the relation between C and N concentrations and then the reader may judge how meaningful soil N is.

We agree with this comment at 90%, and we find it useful and interesting. We would say that although it is true that SOC and soil N must be correlated as they constitute soil organic matter, due to the wide range of conditions and the randomized sampling design of the PASTUS database, raw correlation between soil N and SOC was something discrete (r = 0.88; p-value = 0.001; R^2 = 0.09) when comparing with other studies (i.e. Yan *et al.* 2020). What our models proposes is that soil N modulates SOC response to certain drivers (grazing management and NDF). In other words, grazing management and NDF effects on SOC differ depending on soil N conditions. So soil N would not be completely a controlling factor, although is not as the rest of SOC predictors because of the reasons you exposed.

To explain this, we modified the corresponding paragraph in the discussion section as follows:

"A positive relationship between SOC and soil N was also expected, since most of the soil N is in combined form with organic matter (Cambardella and Elliott, 1994). However, in this case, due to the wide range of conditions and the randomized sampling design of the PASTUS database, the raw correlation between soil N and SOC was somehow discrete (r = 0.297; p-value = 0.001; R2 = 0.088), in comparison to other studies (i.e. Yan et al. 2020). However, the novelty revealed by our model is that soil N could modulate the effects of certain SOC drivers, including livestock type and herbage NDF." (L 629-635)

Could you so provide support that the final model indeed allows that many factors (e.g. by showing AIC)?

We consider that the number of samples (128) is high enough for allowing the number of factors in the Combined Model. Additionally, as we explained in the manuscript, we did not use AIC in the modelling procedure, but the corrected version AICc, which penalizes much more the inclusion of additional terms, hence ensuring that new included factors contribute with really significant information (Burnham & Anderson 2002; Burnham *et al.* 2011). As far as we know, AIC and AICc of the whole model do not give information about if the number of factors is appropriate, but about the information explained by the model in relation to other models. Anyway, AICc of the Combined model was -14.16466. Additionally, below this lines you can find a table showing AICc changes (ΔAICc) in the Combined model when removing each main variable or interaction. Results for main variable effects show changes when removing the main effect and all the interactions using that variable. F test and p-value columns show the anova test between the Combined model and the model resulting of removing that variable or interaction. In this table you can appreciate how excluding any term implies an important information loss (ΔAICc > 2; (Burnham & Anderson 2002; Burnham *et al.* 2011)); and that those models are significantly different according to anova test between models. We decided not to include this table because we find it redundant with Table 3, which provides significance test for each model term and also the coefficients of the model. We could include them as supplementary material or even in the main document if you consider it necessary.

Variable or interaction term	ΔAICc	F test	p-value	
TSIS	26.58	9.26	0.000	***
Slope	4.89	4.65	0.012	*
MAP	3.97	5.93	0.017	*
Soil N	62.15	20.87	0.000	***
Log(C/N)	66.05	77.03	0.000	***
Grazing species	10.86	4.28	0.001	***
Grazing intensity	10.52	4.98	0.001	**
NDF	12.30	8.14	0.001	***
ADL/NH	10.65	5.79	0.001	**
Soil N x Grazing management	15.95	9.94	0.000	***
TSIS x Slope	6.31	8.03	0.006	**
Soil N x NDF	10.92	12.28	0.001	***
TSIS x Grazing species	5.29	4.83	0.010	**
Grazing species x ADL/NH	3.78	4.14	0.019	*

5. How did you consider that factors are autocorrelated e.g. pH and Mg?

We are not sure of understanding if this question refers to spatial autocorrelation or to correlations between variables. Anyway, as our response variable was not spatially autocorrelated according to Morans´s I test (z-score = -1; p-value = 0.3), we do not consider that the fact that explanatory variables were autocorrelated or not could have any noticeable impact for modelling procedure. In other words, spatial autocorrelation could be a problem if it were presented by the response variable (soil organic carbon), not the explicative variables. Additionally, the randomnized sampling design of the PASTUS database probably reduced the possibility of having spatially autocorrelated variables.

Concerning the correlations between explanatory variables, the randomized sampling design of the PASTUS database and the GLM modelling procedure were designed to reduce and control that issue. However, GLM procedure did not selected Mg or pH not because they were highly correlated with other variables, but because the effect they could have on soil organic carbon were surely included in other variables with more explanatory power. As BRT is a modelling procedure that includes explanatory variables regardless of the information provided by other variables, we can detect that Mg and pH had some effects on soil organic carbon. To identify which concrete variable(s) produce pH and Mg displacement in the GLM procedure would imply an exhaustive exercise of forward-backward modelling.

In the text, this is explained as follows:

“Furthermore, BRT model provided some valuable information, identifying some relevant SOC drivers which were discarded during the GML modelling, like aboveground biomass, or soil silt and K (Fig. 2 and S8). The effects of those drivers were probably masked by the effects of other variables in our linear models (Yang et al., 2009), indicating that these factors were presumably pathways through which other variables drove SOC (de Vries et al., 2012). These variables, identified by BRT and discarded by GLM, should be considered as potential SOC drivers in further studies, particularly when more detailed and difficult to obtain biochemical variables, present in our database, are not available.” (L 473-476)

We made several changes in the explanation of this point in the methods and discussion, and we truly believe now is clearer and more understandable than before.

6. Table 1. Is it really meaningful to show what has been measured in which study but not showing the actual outcome of these studies? For instance SOC stocks, main controlling factors, etc.

The aim of including that table was to highlight the wide variety of climate conditions that the PASTUS database contains and the different soil organic carbon drivers that includes. In a previous study about species richness done with dese database (Rodríguez *et al.* 2018), we get some criticism from certain journals because the argued that the range of study was “too regional”. This time, we really appreciate the recognition that both referees had done to the effort of compiling this amount of data.

We added some of the information to the table. In the caption, we specified which studies considered other response variables instead of soil organic carbon (total carbon stocks, soil organic carbon concentration etc.). However, adding more information is not an easy task. Soil organic carbon ranges are difficult to summarize because each study considers different soil depths and there would be necessary an intensive work to get comparable values. We included the types of explanatory variables included in each study, but specific variables and effects and so heterogeneous that are difficult to include in a table, and could lead to the reader to misleading conclusions about these papers.

7. Figure 2: Clarify that the relative importance is related to explained variance.

The way that relative importance of explanatory variables is not exactly related with the concept of explained variance. According to Elith et al. (Elith *et al.* 2008):

“The measures are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman & Meulman 2003). The relative influence (or contribution) of each variable is scaled so that the sum adds to 100, with higher numbers indicating stronger influence on the response.”

If you consider it appropriate, we can include this information on the methods, although we thing it could be a little bit excessive. In the caption of Fig. 2 we added: “Higher numbers indicate stronger influence on SOC stocks (Elith et al., 2008).”

8. The explanation of interaction seems very speculative.

We revised once again both the manuscript and the literature for making our discussion less speculative. Finding experimental studies to support our results is a difficult task because experimental studies assessing interactions between SOC drivers are not as frequent as one could expect. This is a problem that affects not only to SOC, but to soil properties in overall. For instance Rillig *et al.* (2019) revised 1228 experimental studies about global change drivers of diverse soil properties, published between the years 1957 and 2017, and they found that 80% looked just at the effects of one single factor and 19% at the interaction of two factors. This pattern did not changed over the time, so this shortage of experimental studies about interactions is also manifested in current publications. To make things less favourable, in the mentioned revision fertilization was one of the tested factors in more than half of the experiments assessing interaction effects, which suggest that natural or at least extensively managed grassland conditions are poorly represented in experimental studies. Being conscious of these limitations, we are convinced that of work will provide highly valuable information. Interaction experiments are expensive, since they require high sample sizes. Patterns observed in studies like ours can be useful to propose hypothesis to test instead of raising hypothesis blindly.

We added some of this information in the discussion section:

“Those results must be interpreted cautiously, because they are based on observational data, but can contribute to generate testable hypotheses for later studies about some complex and untested relationships between SOC and its drivers. Interaction experiments concerning soil properties are expensive and rare in the literature (Rillig et al., 2019).” (L 621-625)

However, if after reading our revised manuscript you find that some parts of the discussion still need an improvement, we would be pleased of listening your suggestions. We find that our manuscript has improved a lot with this revision.

9. Language should be carefully checked. It can be smoothed at many places.

We have intensively checked language. However no one of the authors is an English native speaker. Our institution has a language revision service, but tis August was closed because of COVID. If you still find that language has to be improved after this revision, we would be able to send it to that service as of September.

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.

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 subsections 2-4 (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? 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 revised the English.

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. (L 59)

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 (n = 128) that comprises a wide range of environmental and management conditions. This was used to understand the relationship between SOC stocks and their drivers considering multiple environments." (L 61)

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 <u>(difference between mean summer temperature and mean annual temperature</u>; TSIS) was the most important geophysical driver of SOC in our study." (L 65-66)

L 65 I think that the coma is not necessary.

The comma was removed. (L 71)

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 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 104 - 108)

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 studies, especially in those studies focusing on predictive rather than explicative models" (L 122 - 125)

L. 147 "focusing" instead of "focus"

Change done. (L 123)

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

It 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 have rephrased the sentence:

"In addition to 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 138).

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). Several studies confirmed the interaction between grazing and other SOC drivers at diverse scales (Abdalla et al., 2018; Eze et al., 2018; Lu et al., 2015, 2017; Zhou et al., 2017). Hence, grazing management may be considered a SOC driver with 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 139-151)

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 which 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 152 - 161)

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 170)

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 216)

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 218 - 221)

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).

We added the following explanation:

"The difference between mean annual and mean summer temperature emerged as a relevant explanatory factor of soil organic carbon stocks during previous modelling efforts by one of the co-authors (M-T. Sebastià). Later attempts to improve models by substituting this variable with other temperature indices from climatic databases (Fick and Hijmans, 2017) showed that, for the PASTUS database, this variable provided higher explanatory power than other temperature seasonality indices. Thus, we decided to keep it and here we name it Temperature Seasonality Index of Sebastià (TSIS from now on)" (L 2221-228).

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:

"To obtain bulk density, we air-dried and weighed the soil samples: we then sieved each sample to 2 mm to separate stones and gravels from the fine earth fraction. " (L 242)

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 247 - 248)

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 suggesting 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 259 - 271)

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)." (L 185-187)

As this is not bromatological information we added this paragraph in lines 209 - 211 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 272)

L. 293 Among which variables?

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

"Including all SOC potential drivers, we fitted a model with BRT to identify the most important ones affecting SOC." (L 340)

L.301 "Firstly" instead of "First"

Change done. (L 348)

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. S5), one model based only on geophysical drivers and grazing management (Geophysical Model), and another model including, in addition to the former drivers, the biochemical drivers: soil nutrients and herbage quality (Combined Model). With this approach we aimed 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 360 - 373).

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 broad-scale and topographic SOC drivers in the Pyrenees. This information could be useful not only for a better understanding of SOC patterns in mountain grasslands, but also for future modelling studies aiming to predict SOC, since geophysical variables are easier and less expensive to acquire and measure compared to biochemical variables (Manning et al., 2015)." (L 484 – 490)

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 in 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, which make carbon stocks highly variable (Garcia-Pausas et al., 2007, 2017). For this reason, data analysed here include a wide range of environmental conditions, comparable to studies on SOC developed at continental or even worldwide scales (Table 1). Additionally, we considered 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 specific questions to answer in this study, 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 does grazing management regulate the effects of other SOC drivers?" (L152 - 172)

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. "Furthermore, BRT model provided some valuable information, identifying some relevant SOC drivers which were discarded during the GML modelling, like aboveground biomass, or soil silt and K (Fig. 2 and S8). The effects of those drivers were probably masked by the effects of other variables in our linear models (Yang et al., 2009), indicating that these factors were presumably pathways through which other variables drove SOC (de Vries et al., 2012)." (L 473-478).

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 a referee or regular reader would ask would be: "How can you 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 522)

L.489 Not clear. . .

To clarify this sentence, we changed it as follows:

"In addition, SOC stocks decreased with increase of slope, which may be attributed to reduced carbon inputs and increased carbon losses induced by steeper slopes" (L 522 - 523)

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:

"At low TSIS values, SOC stocks increased under moderate to high grazing pressure; this effect disappeared as TSIS values increased (Fig. 3D)" (L 527-528)

L.494-499 It is not really clear.

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

"At low TSIS values, SOC stocks increased under moderate to high grazing pressure; this effect disappeared as TSIS values increased (Fig. 3D). Recent meta-analyses concluded that intensive grazing commonly has decreasing effects 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 including climate and soil type vegetation (Abdalla et al., 2018; Eze et al., 2018; Mcsherry and Ritchie, 2013). Moreover, moderate grazing intensities can increase SOC inputs by dung deposition, and aboveground and root biomass production (Franzluebbers et al., 2000; Zeng et al., 2015). In our study, grazing intensity was relatively moderate (see methods), therefore in our study increasing stocking rates may increase soil carbon inputs in moderate seasonality locations by enhancing aboveground and belowground productivity." (L 527 - 537)

L.507 high soil water contents?

To clarify, we changed the sentence as follows:

"high MAP may inhibit decomposition if a shortage of oxygen supply occurs (Xu et al., 2016b)." (L 541)

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

Change done.

Referee 2

General comments

The manuscript aims to understand how environmental and management factors affect SOC in mountain grasslands. And fitted a set of models with explicative purposes using data that comprise a wide range of environmental and management conditions to find the most important driver of grassland SOC. The authors are to be commended on the framing of an interesting study, the collection of a reasonable set of ancillary environment and management data and soil data in what appears to be good quality piece of research. The workload of this article is very huge.

However, too many sections and repetitive statements in this article. Be better structured and more concise to attract readers.

Please, see our answers to referee 1 about the modifications in the text structure.

Deep discussion and comparison of your work is needed in an international context. In discussion section, some discussion on the mechanism of environmental and management factors should be added.

We would really appreciate it if you could specify more about which mechanisms need more discussion. Referee 1 found that discussion section was “overall, well supported by relevant references and the limits are underlined”. We recon you have a point concerning biochemical or management species effects on SOC: the mechanisms are not widely explained but, as we explain in the text, that is a difficult task since there are few publications addressing these issues. We revised the published works from this manuscript was sent to Biogeosciences until now and, under our view, no remarkable novelties have appeared in these topics. However, we found a bibliometric study which concluded that interaction experiments concerning soil properties are expensive and rare in the literature (Rillig et al., 2019). We would appreciate it if suggestions about ideas or publications we could omit were made in order to improve the manuscript.

I suggest you add a conclusion section, a concise and clear conclusion will make your article more eye-catching and let readers understand the conclusion of this article more quickly and easily.

We separated the conclusions from the discussion section, and we changed that paragraph to make it as much clear and concise as possible, focusing on the main contributions of our manuscript to scientific knowledge.

As the manuscript contains some uncertainties in description of the methods, results, and English writing, I suggest a moderate revision necessary before it can be acceptable for publication in this journal.

We corrected the uncertainties in the text. The specific comments of both reviewers were really helpful and we really appreciate them.

Specific comments

Line 75 “Soil organic carbon plays key roles in the terrestrial ecosystems.” It sounds strange.

We rephrased this sentence as follows:

“Soil organic carbon (SOC) is crucial for the functioning of terrestrial ecosystems.” (L 84)

Line 179 At least one to two replicates of each patch type were sampled. What are the types of the patch?

To clarify this point, we rephrase this sentence as follows:

“Grassland patches were then listed by type and arranged within each list randomly to determine sampling priority. At least one to two replicates of each patch type defined by the stratification variables were sampled.” (L 199 - 201)

Line 155 Not clear sampling design description. Showing a figure with sampling design would be helpful. Add a schematic of experimental design to make it clearer.

We added figure S2, which illustrates sampling design.

Line 192 The abbreviation for soil organic carbon had appeared in line 75, here only need to write SOC.

Change done. (L 214)

Line 193 There are 30 variables written in table S1, but here you have written 29 independent environmental variables. Are the two management variables belong to environmental variables? Please check these numbers.

Change done. (L 215)

Line 194 These variables were grouped into Regional, landscape, livestock management,

soil nutrient stocks, and herbage variables? If so, replace “:” with “,”.

Change done. (L 216)

Line 201 MTS?

M-T. Sebastià. We changed this to make it clear. (L 223)

Line 220 Here used livestock stocking rates which measured as livestock units ha-1 to determine grazing intensity. But the feed intake of different types of livestock is different. For example, the intake of cattle is higher than that of sheep. So, can't simply use the livestock units/ha-1 as livestock stocking rates, you need use standard livestock unit.

We used a standard transformation index where 8 small ruminants correspond to 1 big ruminant. This is standard and provided by the Catalan Government for the region.

Line 314 Geophysical model based on geophysical predictors and grazing management? There haven't grazing management in Figure S4.

Now Figure S4 has grazing management.

Line 371 Authors need to better describe statistics of SOC.

We added some information about the statistics of SOC. However, we do not know what more to add apart of basic descriptive statistics we already show. We will really appreciate it if you could specify what statistics you miss in this part of the text. (L 420 - 423)

Line 375 Generally, a part of the sample is used for modeling, and the other part is used for validation. Please describe clearly in here and in Line 279.

Concerning the line 375 (now 425) (BRT model) we did not validate the whole model with a fraction of the dataset, because our BTR model was fitted by cross validation (CV; it is used to select the number of trees with the best performance). Note that according to Elith et al. (2008), results of this cross validation procedure are often very similar to those obtained with independent datasets. Additionally, note that each tree was actually fitted with 66% of the data (out of the bag fraction parameter), so our procedure properly dealt with stochasticity too. All these are standard methods explained by Elith et al. (2008), so we prefer just to refer to this publication instead of extending our methods section, and to focusing on other parts of the statistical procedure that need to be clearer as possible. However, if you think that some of the standard aspects of the BRT procedure deserves to be explained in our manuscript, we will follow your advice.

We detailed the herbage-bromatological analysis (L 279 (Now 272) and so on). 130 samples where used for the validation of NIRS equations.

Line 379 Silt in here, loam in fig.2. Use consistent terminology of silt, loam, etc? Use one, Please!

Change done. Silt is now the only name used.

Line 382 Why TSIS was the most relevant selected climate predictor? In figure 6s, Soil C/N has a higher relative importance.

TSIS was the most relevant of the climate predictors (without considering other variable types). To clarify this point, we rephrased this sentence as follows:

“TSIS was the most relevant among the climate drivers considered.” (L 432)

Line 383 Please confirm this sentence and the quoted figure. I didn’t find TSIS in figure S5 and S6. In table s1, TSIS described as MST-MAT. In figure s8, MMT also described as MST-MAT Use consistent terminology of MMT, TSIS, etc? Use one, Please!

Change done. MMT is a previous nomenclature. TSIS is the proper one.

Line 381 Aboveground biomass and silt had a high relative contribution in the final BRT model obtained, why not selected them in the linear models?

This was also true for soil K and silt. This point was discussed in the “Considerations about the modelling procedure” subsection inside the discussion section. “Furthermore, BRT model provided some valuable information, identifying some relevant SOC drivers which were discarded during the GML modelling, like aboveground biomass, or soil silt and K (Fig. 2 and S8). The effects of those drivers were probably masked by the effects of other variables in our linear models (Yang et al., 2009), indicating that these factors were presumably pathways through which other variables drove SOC (de Vries et al., 2012). These variables, identified by BRT and discarded by GLM, should be considered as potential SOC drivers in further studies, particularly when more detailed and difficult to obtain biochemical variables, present in our database, are not available.” (L 473 - 481)

Basically, as 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. There is more about BTR models in some answers to referee 1.

Line 1121 Please add the fitting equation in figure 3 and 4. It is hard to distinguish which trend line belongs to which grazing species or grazing intensity. You can distinguish by color, or add the legend.

We changed all the plots to the main document to color plots, so lines and dots are more distinguishable than before. We also added the sentence “The estimates on Table 2-3 were those used to elaborate these plots.” so the equation values can be easily found.

Line 25 in SUPPLEMENT Figure S1: points indicate sampling location, sampling location means the sample patches? Please add the legend of the points in this figure.

As we explained in the methods section, each sampling patch contains a sampling location, located in the middle of the grassland patch. Sampling location were added in the legend of this figure. As you suggested in some lines above, we added the figure S2 to clarify the sampling design, and the legend of the points in Fig. S1.

Line 39 in SUPPLEMENT There is no reference of Figure S3 in the text.

We added the reference in the "general linear models" subsection, in the material and methods section:

"We designed and executed a modelling procedure based on general linear models (Legendre and Legendre, 1998) and a hierarchy of controls over function (Diaz et al., 2007; de Vries et al., 2012). We log-transformed SOC using natural logarithm to prevent a breach of the normality assumption by the residuals of the models (Fig. S4)." (L 357; Fig. S3 is now S4)

Interactions between biogeochemical and management factors explain soil organic carbon in Pyrenean grasslands

Authors

Antonio Rodríguez[1,2], Rosa Maria Canals[3], Josefina Plaixats[4], Elena Albanell[4], Haifa Debouk[1,2], Jordi Garcia-Pausas[5], Leticia San Emeterio[6,7], Àngela Ribas[8,9], Juan José Jimenez[10], M.-Teresa Sebastià[1,2]

Afiliations

1: Group GAMES, Department of Horticulture, Botany and Landscaping, School of Agrifood and Forestry Science and Engineering, University of Lleida, Lleida, Spain

2: Laboratory of Functional Ecology and Global Change (ECOFUN), Forest Sciences Centre of Catalonia (CTFC), Solsona, Spain

3: Grupo Ecología y Medio Ambiente and ISFood Institute, Universidad Pública de Navarra, Campus Arrosadia, Pamplona, Spain.

4: Group of Ruminant Research (G2R), Department of Animal and Food Sciences, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain

5: Forest Sciences Centre of Catalonia (CTFC), Solsona, Spain

6: Research Institute on Innovation & Sustainable Development in Food Chain (ISFOOD), Universidad Pública de Navarra,31006 Pamplona, Spain.

7: Departamento de Agronomía, Biotecnología y Alimentación, Universidad Pública de Navarra, 31006 Pamplona, Spain.

8: Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain.

9: Centre for Ecological Research and Forestry Applications (CREAF), 08193, Bellaterra, Spain.

10: ARAID/IPE-CSIC, 22700, Huesca, Spain

ORCID

Antonio Rodríguez http://orcid.org/0000-0002-0536-9902

Elena Albanell https://orcid.org/0000-0002-6158-7736

Jordi Garcia-Pausas https://orcid.org/0000-0003-2727-3167

Leticia San Emeterio http://orcid.org/ 0000-0002-8063-0402

Àngela Ribas https://orcid.org/0000-0002-5938-2408

M.-Teresa Sebastià http://orcid.org/0000-0002-9017-3575

Author contributions

Antonio Rodríguez designed the statistical procedure, carried out the statistical analyses and wrote the original draft.

Rosa M Canals was responsible for field monitoring, lab analyses and acquisition of information for the database implementation in the Western Pyrenees (Navarra). She also reviewed the draft.

Elena Albanell designed the NIRS study and reviewed the draft.

Haifa Debouk sampled and processed some of the data in the PASTUS Database and reviewed the draft.

Jordi García-Pausas processed some of the data in the PASTUS Database and reviewed the draft.

Josefina Plaixats carried out the chemical analyses of herbage samples for NIR calibration and validation equations and reviewed the draft.

Leticia San Emeterio designed methodology and data collection, performed soil and vegetation sampling. She also reviewed the draft.

Àngela Ribas sampled and processed some of the data in the PASTUS Database and reviewed the draft.

Juan José Jiménez collaborated in the fieldwork and reviewed the draft.

M.-Teresa Sebastià contributed to the conception, design and development of the PASTUS database. In addition, she ensured funding and coordinated the projects whose data are included in PASTUS. Finally, she contributed to initial modelling, supervised the development of the paper, read and reviewed the drafts.

Abstract

Grasslands are one of the major sinks of terrestrial soil organic carbon (SOC). Understanding how environmental and management factors drive SOC is challenging because they are scale-dependent, with large-scale drivers affecting SOC both directly and through drivers working at small scales. Here we addressed how regional, landscape and grazing management, soil properties and nutrients, and herbage quality factors affect 20 cm depth SOC stocks in mountain grasslands in the Pyrenees. Taking advantage of the high variety of environmental heterogeneity in the Pyrenees, we built a dataset (n = 128) that comprises a wide range of environmental and management conditions. This was used to understand the relationship between SOC stocks and their drivers considering multiple environments. 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, depending on topography and management. TSIS effects on SOC increased in exposed hillsides, slopy areas, and relatively intensively grazed grasslands. Increased TSIS probably favours plant biomass production, particularly at high altitudes, but landscape and grazing management factors regulate the accumulation of this biomass into SOC. Concerning biochemical SOC drivers, we found unexpected interactive effects between grazer type, soil nutrients and herbage quality. Soil N was a crucial SOC driver as expected, but modulated by livestock species and neutral detergent fibre contentin plant biomass; herbage recalcitrance effects varied depending on grazer species. These results highlight the gaps in the knowledge about SOC drivers in grasslands under different environmental and management conditions. They may also serve to generate testable hypotheses in later/future studies directed to climate change mitigation policies.

Keywords

SOC, natural grasslands, grazer type; grazing management, herbage quality; climate change, soil nutrients; topography; temperature seasonality; TSIS

Introduction

Soil organic carbon (SOC) is crucial for the functioning of terrestrial ecosystems (Lal, 2004a). SOC enhances soil and water quality and biomass productivity, and has an important role in relation to climate change (Lal, 2004b). Although grasslands have small aboveground biomass compared to other ecosystems, their SOC stocks can be comparable to those in forest ecosystems (Berninger et al., 2015). This is due to their high root biomass and residues, which are a substantial carbon source and can contribute to water retention in soil. This creates favourable conditions for the accumulation of organic matter (Von Haden and Dornbush, 2014; Yang et al., 2018). These attributes, together with the high extent of grassland global cover, make grasslands store around 34% of the terrestrial carbon, mostly in their soils (White et al., 2000).

SOC accumulation results from a complex equilibrium between primary production and organic matter decomposition which depends on multiple environmental factors such as climate, soil texture and nutrients, or land management (Jenny, 1941; Schlesinger, 1977). Understanding how these scale-dependent environmental factors drive SOC is challenging because large scale drivers affect also those working at fine spatial scales. This has been described as a hierarchy of controls over SOC (Fig. 1; Manning et al., 2015).

Climate is known to be the main SOC driver at broad (global and regional) scales; mean annual precipitation (MAP) and mean temperature (MAT) being the most frequent climate indicators (Wiesmeier et al., 2019). However, climate annual variations represented by seasonality variables are commonly neglected when considering possible SOC drivers in broad-scale models, in spite of being important drivers of plant primary production and enzymatic activity of soil microorganisms (Fernández-Alonso et al., 2018; Garcia-Pausas et al., 2007; Puissant et al., 2018). Other regional and landscape factors like bedrock or topography are also considered to be at the top of the hierarchy because they influence multiple geophysical and biochemical factors affecting SOC, including soil texture and water flow paths (Gray et al., 2015; Hobley et al., 2015). Next in the hierarchy after regional and landscape factors, are several soil geophysical properties, like pH and texture, which are controlled by climate, bedrock, and which affect SOC through both plant primary production and microbial activity and the capacity to stabilise the SOC (Deng et al., 2016; Xu et al., 2016a).

Soil macro and micronutrients are in the next level of the hierarchy, as their abundance is determined by multiple factors, including climate, soil pH, water content or clay content (Hook and Burke, 2000; de Vries et al., 2012). They play an essential role in primary production and herbage quality, and act as resources for microbes to mineralise SOC (Aerts and Chapin, 1999; Vitousek and Howarth, 1991). However, these variables are commonly omitted as possible drivers of SOC in the broad-scale studies, especially in those studies focusing on predictive rather than explicative models (Gray et al., 2015; Manning et al., 2015; Zhang et al., 2018). This kind of variables is less frequently available and more difficult to measure than the other indicators used for SOC modelling (Manning et al., 2015). Furthermore, the use of soil nutrients as SOC drivers in linear models can be challenging, as they are often strongly linked to SOC dynamics. This may mask the effect of other drivers acting at larger spatial scales (Bing et al., 2016; Cleveland and Liptzin, 2007; Tipping et al., 2016).

Vegetation represents another group of SOC drivers, affected by climate, topography and soil properties and nutrients (Fernández-Martínez et al., 2014; de Vries et al., 2012; Zhu et al., 2019). Plant biomass is the main input of organic carbon into the soil (Shipley and Parent, 1991). However, a not so frequently considered factor is plant litter quality, which can determine decomposition rates and patterns, and hence soil carbon sequestration (Ottoy et al., 2017; Yan et al., 2018, 2019).

In addition to these factors, livestock management effects on grassland SOC is a noteworthy issue since they are poorly understood (Wiesmeier et al., 2019). 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). Several studies confirmed the interaction between grazing and other SOC drivers at diverse scales (Abdalla et al., 2018; Eze et al., 2018; Lu et al., 2015, 2017; Zhou et al., 2017). Hence, grazing management may be considered a SOC driver with 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).

In this study, our goal was to identify the main drivers of SOC stocks and their interactions in 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, which make carbon stocks highly variable (Garcia-Pausas et al., 2007, 2017). For this reason, data analysed here include a wide range of environmental conditions, comparable to studies on SOC developed at continental or even worldwide scales (Table 1). Additionally, we considered 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 specific questions to answer in this study, 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 does grazing management regulate the effects of other SOC drivers?

Material & methods

2.1 Location and sampling design

The set of data used in this study has been extracted from the PASTUS Database (http://ecofun.ctfc.cat/?p=3538), which was compiled by the Laboratory of Functional Ecology and Global Change (ECOFUN) of the Forest Sciences Centre of Catalonia (CTFC) and the University of Lleida (UdL). We sourced a wealth of data of 128 grassland patches distributed across the Pyrenees (Fig. S1), and including topographical, climate, soil, herbage and management variables. The elaboration of the PASTUS Database concerning this study is summarised in Fig. S2). The sampled area encompasses a wide variety of temperate and cold-temperate climates, with different precipitation conditions, depending on altitude and geographical location from Mediterranean to Continental and Boreo-Alpine environments (de Lamo & Sebastià, 2006; Rodríguez et al., 2018; Table 1). Almost all of the plant species in the grasslands from the PASTUS database are perennial (Sebastià, 2004), and plant diversity is highly heterogeneous as are the environmental conditions (Rodríguez et al., 2018).

Sampling in the PASTUS database was designed according to a stratified random scheme, where samples were selected at random within strata. This process was done using the software ArcMap 10 (ESRI, Redlands, CA, USA). The basis for randomization was the map of habitats of Catalonia 1:50000 (Carreras and Diego, 2006) for the Eastern and Central sectors of the Pyrenees, the map of habitats of Madres-Coronat 1:10000 (Penin, 1997) for the North-Eastern sector and the land use map of Navarra 1:25000 (Gobierno de Navarra, 2003) for the Western sectors. Four variables were initially considered for sampling stratification within each sector: altitude (< 1800 m; 1800-2300 m; > 2300 m), slope (0-20º; 20-30º; > 30º), macrotopography (mountain top/southern-facing slope; valley bottom/northern-facing slope) and grazer type (sheep; cattle; mixed). Accordingly, we determined a set of homogeneous grassland patches by crossing the stratification variable layers. Grassland patches were then listed by type and arranged within each list randomly to determine sampling priority. At least one to two replicates of each patch type defined by the stratification variables were sampled.

In each sampled grassland patch, a 10 x 10 m^2 plot was systematically placed in the middle of each homogeneous grassland patch, including a particular plant community. We collected soil and vegetation samples, and assessed environmental variables within each 100 m^2 plot (see Rodríguez et al., (2018) for additional details about sampling design). Local variables were assessed inside the 100 m^2 plots. Aboveground biomass was estimated from herbage cut at ground level in four 50 x 50 cm^2 quadrats placed in a

2 x 2 m^2 subplot inside the 100 m^2 plot. Herbage from two of the four quadrats were dried and sent to the laboratory for duplicated chemico-bromatological analysis. In addition, in each quadrat, a 20-cm depth soil core was extracted with a 5 x 5 cm probe after herbage was removed. The soil sample in the probe was separated into two soil layers: 0-10 and 10-20 cm.

2.2 Regional and landscape environmental drivers

In order to investigate the relationship between SOC and potential environmental drivers, 30 independent environmental variables were initially considered (Table S1). These variables were grouped into five sets: regional, landscape, livestock management, soil nutrient stocks, and herbage variables.

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). The difference between mean annual and mean summer temperature emerged as a relevant explanatory factor of soil organic carbon stocks during previous modelling efforts by one of the co-authors (M-T. Sebastià). Later attempts to improve models by substituting this variable with other temperature indices from climatic databases (Fick and Hijmans, 2017) showed that, for the PASTUS database, this variable provided higher explanatory power than other temperature seasonality indices. Thus, we decided to keep it and here we name it Temperature Seasonality Index of Sebastià (TSIS from now on).

Bedrock type was determined in the field and confirmed by the geographical maps mentioned above. Bedrock was categorized into three categories: basic (marls and calcareous rocks), acidic (mostly sandstones and slates) and heterogeneous.

Landscape variables included topography and soil type variables. Topography variables included Slope, Aspect, Macrotopography and Microtopography. Slope and Aspect were determined in the field by clinometer and compass respectively. Macrotopography and microtopography were determined visually in the field. Preliminary modelling efforts suggested the reduction of the four macrotopographical positions initially identified in the field into two: Mountain top and south-facing slopes were classified as exposed positions and valley bottoms and north-facing slopes as protected macrotopographical positions. Microtopography included three positions: convexities, concavities and smooth areas. Soil type variables are described in the following.

2.3 Soil physicochemical analysis

To obtain bulk density, we air-dried and weighed the soil samples: we then sieved each sample to 2 mm to separate stones and gravels from the fine earth fraction. The fine fraction was sent to the laboratory for further physicochemical analysis. Standard physicochemical soil analyses were performed in the upper 0-10 cm soil layer of all grasslands. Some analyses were also performed on samples from the 10-20 cm soil layer, including soil organic carbon and total nitrogen. For those variables, we combined 0-10 and 10-20 cm values to obtain the whole top 20 cm soil layer.

All soil physicochemical analyses were performed on the fine earth, according to standard soil analysis methods. Textural classes were determined by the Bouyoucos method (Bouyoucos, 1936). Soil pH (measured in water), total organic carbon (TOC) total nitrogen (TN), Calcium content (Ca), Extractable phosphorus (P), magnesium (Mg) and potassium (K) were measured on air-dried samples (Schöning et al., 2013; Solly et al., 2014). Soil carbonates were determined using the Bernard calcimeter. Total carbon and nitrogen (N) contents of the fine earth were determined by elemental auto-analyser. The organic C fraction was determined by subtracting inorganic C in the carbonates from the total C. Available phosphorus (P) was extracted by the Olsen method (Olsen, 1954) Ca, Mg and K were extracted by ammonium acetate (Simard, 1993) and measured by flame Atomic Absorption Spectroscopy (AAS) (David, 1960)). 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 suggesting 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.

2.4 Herbage chemical and bromatological analysis, and NIRS analysis

All four herbage samples per plot were oven-dried at 60ºC to constant weight to determine aboveground biomass and converted into g m^{-2}. Two out of the four samples were sent to the laboratory for herbage quality analysis. Dried samples were ground to pass a 1 mm stainless steel screen (Cyclotec 1093 Sample mill, Tecator, Sweden) and stored at 4ºC until it was needed for use.

To develop NIRS prediction models, a random subset of 130 samples was used and analysed in duplicate according to the reference methods mentioned further. Procedures described by AOAC were used to determine dry matter (DM) and ash content or mineral matter (MM). Crude protein (CP) was determined by the Kjeldhal procedure (N x 6.25) using a Kjeltec Auto 1030 Analyser (Tecator, Sweden). Samples were analysed sequentially for neutral detergent fibre (NDF), acid detergent fibre (ADF) and acid detergent lignin (ADL) in accordance with the method described Van Soest et al. (1991) using the Ankom 200 Fibre Analyser incubator (Ankom, USA). The fibre analysis were determined on an ash-free basis and without alpha amylase. We calculated two additional herbage quality indexes often used in the bibliography: NDF/CP and ADL/NH

(Stockmann et al., 2013). For each subsample the C and N content (CH and NH)were determined by the Dumas dry combustion method, using an Elemental Analyzer EA1108 (Carlo Erba, Milan, Italy).

Afterwards, a total of two hundred herbage samples were scanned as described below to collect their NIRS spectra. We estimated chemical and bromatological variables according to the equations derived from the previous calibrations on the initial 130 random samples.

NIRS data were recorded from 1,100 to 2,500 nm using a FOSS NIRSystem 5000 scanning monochromator (Hillerød, Denmark). Separate calibration equations were generated for grassland samples. Reflectance (R) data were collected in duplicate every 2 nm. A WinISI III (v. 1.6) software (FOSS, Denmark) was employed for spectra data analysis and development of chemometric models. Prior to calibration, log 1/R spectra were corrected for the effects of scatter using the standard normal variate (SNV), detrend (DT) and multiple scatter correction (MSC) and transformed into first or second derivative using different gap size (nm) and smoothing interval. For each sample, the mean of the spectra from the two lectures were used. Modified partial least square (MPLS) was the regression method used for calibration development and cross validation was undertaken using the standard methodology in the NIRS software program. The performance of the model was determined by the following statistical tools: standard error of calibration (SEC), standard error of cross validation (SECV); coefficient of determination for calibration (R^2) and cross validation (r_{cv}^2); the ratio of performance to deviation (RPD) described as the ratio of standard deviation for the validation samples to the standard error of cross validation (RPD=SD/SECV) should ideally be at least three; and the range error ratio (RER=Range/SECV) described as the ratio of the range in the reference data to the SECV should be at least 10 (Williams and Sobering, 1996; Williams et al., 2014).

2.5 Livestock management variables

The management variables (grazer type) initially used for sampling stratification were determined from records available in the municipalities of the study area. Once the specific grassland patches to be sampled were determined, we carried out a detailed analysis of the management where the patches were located. To this effect, we carried out detailed surveys among farmers, shepherds and land managers. Sometimes the information collected was modified according to visual records in the field (e.g., cattle and/or cattle dung found in supposedly ungrazed areas). Information from municipalities was usually the most imprecise.

We considered two management variables: Grazing intensity and Grazer type. Grazing intensity was determined estimating livestock stocking rates measured as livestock units ha^{-1} (LU ha^{-1}), and treated as a semi-quantitative variable with three categories (): low (1; lower than 0.2 LU ha^{-1}), medium (2; between 0.2-0.4 LU ha^{-1}) and high (3; above 0.4 LU ha^{-1}). Grazer type was categorised into three main types: sheep, cattle and mixed. Mixed grazing included associations comprising small and big livestock species, mainly sheep and cattle, and more rarely horses. Sheep flocks always included some goats.

2.7 Statistical analyses

We applied two different modelling procedures: Boosted Regression Trees (BRT) and General Linear Models (GLM). BRT is an automatic technique that combines insights from traditional statistical modelling and machine learning traditions (Elith et al., 2008). GLM allowed us to design a hypothesis-based modelling procedure, ensuring that only effects with biological meaning where included; whereas BRT provided information about the data that could be neglected, if only a GLM approach was followed.

All the statistical analyses were performed with the software R ver. 3.4.3 (R Core Team, 2017), at 95% significance level when appropriate.

Boosted regression trees global model

Including all SOC potential drivers, we fitted a model with BRT to identify the most important variables affecting SOC. BRT uses two algorithms: regression trees and boosting. Regression trees are from the decision tree group of models, and boosting builds and combines a collection of models (Elith et al., 2008). We chose this method because BRT can handle multiple variables better than other techniques as GLM, and can detect automatically curvilinear relationships and interactions, ignoring non-informative ones. We used the gbm and dismo packages (Greenwell et al., 2019; Hijmans et al., 2017), which provide several functions to fit these models.

Firstly, fitted a model with all the drivers (Table S1), configured with 15 folds, a Gaussian distribution of the error, a tree complexity of 5, a learning.rate of 0.005, a bag.fraction of 0.666, and 5 minimum observations by node. Secondly, we reduced the number of drivers by the method described in Elith et al., (2008). We estimated the change in the model´s predictive deviance dropping one by one each driver, and re-fitted the model with the set of variables which actually improved model performance (Fig. S3). We checked the relative importance of the drivers and the shape and size of the effects by partial effect plots.

General linear models

We designed and executed a modelling procedure based on general linear models (Legendre and Legendre, 1998) and a hierarchy of controls over function (Díaz et al., 2007; de Vries et al., 2012). We log-transformed SOC using natural logarithm to prevent a breach of the normality assumption by the residuals of the models (Fig. S4). We built two models (Fig. S5), one model based only on geophysical drivers and grazing management (Geophysical Model), and another model including, in addition to the former drivers, the biochemical drivers: soil nutrients and herbage quality (Combined Model). With this approach we aimed 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. We considered that the geophysical factors that potentially affect SOC were regional and landscape (topography and soil type drivers), as they have been widely used in previous studies to model and predict SOC from landscape to continental scales (Manning et al., 2015; Wiesmeier et al., 2019). In addition to soil nutrients and herbage variables, we included again the livestock management variables in the Combined Model and looked for interactions involving these variables and biochemical drivers of SOC.

For model building (Fig. S5A), we added driver groups following a sequential order. For fitting the Geophysical Model, we started adding regional, landscape and grazing management drivers, and subsequently included soil properties. Afterwards, we sequentially included soil nutrients and herbage drivers to obtain the Combined Model. We added Management variables from the beginning of the modelling process and re-included the discarded ones in each step to guarantee the detection of interactions between Management variables and the rest of the drivers. Each time we added a set of drivers, we first considered their main effects and some quadratic terms which were found by preliminary analyses with the scatterplot.matrix function in the R package car (Fox et al., 2018); afterwards we included possible level 2 interactions between all the selected drivers.

At every step we selected several candidate terms by a semi-automatic procedure (Fig. S5C) using a genetic algorithm included in the R package glmulti (Calcagno, 2015). We used SOC as response variable at the first step, and the residuals of the previous model in the remaining steps (Fig. S5B). This semi-automatic process began by obtaining a best subset of models using the corrected Akaike information criterion (AICc), appropriate when n/k is less than 40, n being the sample size and k the number of parameters in the most complex model (Symonds and Moussalli, 2011). We selected the best model and its equivalents obtained by the genetic algorithm, which were those within 2 Akaike information criterion-corrected (ΔAICc) values of the best model, as a ΔAICc < 2 indicates that the candidate model is almost as good as the best model (Burnham and Anderson, 2002).

For this set of models, we built averaged models using the MUMIn package (Barton, 2015). We calculated partial standardized coefficients, obtained by multiplying the unstandardized coefficient in the model by the partial standard deviation of the variable, which is a function of the standard deviation of the variable, the sample size, the number of variables in the model and the variance inflation factor of the variable (Barton, 2015). We selected all the variables with significant effects (alone or in interaction with each other) in the conditional average model, which was preferred over the full average model because we wanted to avoid excessive shrinkage effects at this moment of the modelling procedure (Grueber et al., 2011).

Then, we added these terms to the consolidated model, and made a selection through a backward forward procedure. We used several methods to compare and determine the final model, including the AICc, the adjusted determination coefficient R^2 (R_{adj}^2) and model comparison techniques with the “anova()” function in R, using Chi-square tests to test whether the reduction in the residual sum of squares was statistically significant. Once we had the final model we assessed the significance of each term by removing it and performing an F test. For estimating the significance of the main effects we also removed the interaction terms in which they were involved, to avoid transferring the effects of the main terms to the interaction terms (de Vries et al., 2012). We estimated the variance explained by the models through the adjusted determination coefficient R^2 (R_{adj}^2).

Finally, we estimated the importance of the terms of each model by the lmg method in the relaimpo package (Grömping, 2006), and drew partial effect plots making predictions with the R package emmeans (Lenth et al., 2019).

Results

SOC stocks of the upper 20 cm layer ranged between 2.6 and 23 kg m^{-2}, with a median and a mean value of 9.1 and 9.6 kg m^{-2} respectively. Standard deviation of the mean was 3.15 (n= 125). Minimum, maximum, median and mean values of the continuous drivers are shown in Table S2.

3.1 Relative importance of SOC stocks drivers

The final BRT global model achieved a good goodness of fit, with a cross-validated correlation value of 48% and an explained deviance of 88.31%. The most important variables explaining SOC stocks (Fig. 2) were soil N (18.3 %), soil C/N (14.4%) and Clay (13%) although other variables such as aboveground biomass (7.3%), ADL (6.4%) or silt (6.1%) were also relevant for explaining SOC storage. Three important variables in the BRT model, aboveground biomass, silt and soil K, were not selected in the linear models (Tables 2 & 3). Although accounting for a lower importance value than the previous variables (5%), TSIS was the most relevant among the climate drivers considered. TSIS was also noticeably important in both linear models (Fig. S6), especially in the Geophysical Model, not only as main effect, but in interaction with other variables (Img: 4-10%). According to the Combined linear model, soil nutrient and herbage variables were other important SOC stocks drivers(Fig. S7), but many of these effects occurred in interaction with grazer type.

3.2 Geophysical, biochemical and grazing management effects on SOC stocks

The Geophysical Model (Table 2) explained 34% of the total variance (R^2_{Adj}). Overall, SOC stocks increased with TSIS under certain conditions: exposed hillsides, high slopes and low stocking rates (Fig. 3A, 3B & 3D). On the other hand, Clay had a positive relationship with SOC under low MAP values (Fig. 3C), which turned into negative at high MAP values.

Adding nutrient and herbage variables to the previous Geophysical Model to build the Combined model (Table 3) increased the total variance (R^2_{Adj}) up to 61%. Macrotopography, and Clay effects described by the Geophysical model were removed by the new model terms included. SOC increased with C/N (Fig 4A). Soil nitrogen modulated the effects of livestock type and NDF on SOC. Cattle-grazed grasslands stored more SOC than mixed and sheep grazed grasslands under low soil N conditions, whereas the reverse occurred at high soil N levels (Fig. 3B). NDF had negative effects on SOC stocks at high soil N values but had no effect under low soil N levels Fig. 4C). Finally, herbage ADL/NH had positive effects on SOC under mixed and sheep grazing regimes, but there was no effect under cattle management (Fig. 4D).

Discussion

3.1 Considerations about the modelling procedure

Unsurprisingly, the SOC drivers selected and their main effects in both of the modelling approaches (BRT and GLMs) were highly congruent (Figs. 2-4; S8). Consequently, we preferred to focus on the results from the linear models because this approximation allowed us to build models under a hierarchy of controls over function hypothesis (Manning et al., 2015). Hence, although it is not possible to unequivocally establish the causal links between SOC drivers (Grace, 2006; Grace and Bollen, 2005), with our GLMs procedure we guarantee that the effects of the biochemical variables added in the Complete Model on SOC stocks have not been exclusively induced by geophysical drivers (de Vries et al., 2012). If this was the case, soil nutrient and herbage quality drivers wouldn’t have entered the Complete Model as significant terms. This happened with aboveground biomass, which is assumed to be a very important SOC driver, and indeed aboveground biomass was relevant in the BRT model, but in the GLM was substituted by other, more meaningful, variables. In addition, our GLM modelling approach enabled us the selection of biologically meaningful interactions (Manning et al., 2015; de Vries et al., 2012), which cannot be done with a fully automatic approach like BRT. This GLM sequenced modelling procedure, looking for the primary sources of variation, together with the stratified sampling design, is useful as it led us to select a set of lowly correlated drivers for our linear models (Table S5). Furthermore, BRT model provided some valuable information, identifying some relevant SOC drivers which were discarded during the GML modelling, like aboveground biomass, or soil silt and K (Fig. 2 and S8). The effects of those drivers were probably masked by the effects of other variables in our linear models (Yang et al., 2009), indicating that these factors were presumably pathways through which other variables drove SOC (de Vries et al., 2012). These variables, identified by BRT and discarded by GLM, should be considered as potential SOC drivers in further studies, particularly when more detailed and difficult to obtain biochemical variables, present in our database, are not available.

3.2 Geophysical, biochemical and grazing management factors driving SOC stocks

Considering the difficulties of modelling SOC in a widely heterogeneous mountain environment (Garcia-Pausas et al., 2017), the Geophysical Model provided important information about broad-scale and topographic SOC drivers in the Pyrenees. This information could be useful not only for a better understanding of SOC patterns in mountain grasslands, but also for future modelling studies aiming to predict SOC, since geophysical variables are easier and less expensive to acquire and measure compared to biochemical variables (Manning et al., 2015).

Most studies on soil carbon usually pinpoint mean temperature and precipitation as the most important climate drivers of SOC (Hobley et al., 2015; Manning et al., 2015; Wiesmeier et al., 2019). Climate regulates large-scale patterns of aboveground net primary production (Chapin et al., 1987). In our study, temperature seasonality (TSIS) was a key driver of SOC, modulated by macrotopography, slope and grazing intensity (Table 2; Fig. 3). The highest variation of TSIS in our database, that is, the broadest temperature seasonality, occurred in cold environments, as compared to mild climates (Fig. S9). In mountain grasslands, cold climates imply a short phenological period of development for plants (Gómez, 2008). Hence, the positive effect of TSIS on SOC could be associated with a higher biomass accumulation in cold locations with more favourable temperatures during summer, this fact reducing geophysical stress for plants and broadening their growth period (Garcia-Pausas et al., 2007; Kikvidze et al., 2005). This increase in soil organic matter inputs during summer would overcome an eventual increase of soil organic matter decomposition rates related to high temperatures (Sanderman et al., 2003) which in those cold environments with contrasted temperature seasonality would not occur.

The interactive effects of TSIS on SOC stocks with macrotopography and slope illustrate the capacity of landscape factors to modulate macroclimate effects on soil (Hook and Burke, 2000). Induced microclimate changes are often the explanation for the effects of topography in SOC (Lozano-García et al., 2016). In our case, SOC stocks increased with temperature seasonality, particularly in exposed locations, including south-facing hillsides and hillside tops (Fig. 3A; Table 2). In protected locations, including shady hillsides and valley bottoms, the hypothesized positive effect of increased TSIS values on plant productivity could be mitigated due to reduced solar radiation, long snow-covered periods and freezing episodes (Garcia-Pausas et al., 2007; López-Moreno et al., 2013). Additionally, differences in SOC between exposed and protected sites may also occur through other mechanisms, for instance the alteration of soil physico-chemical properties (Zhang et al., 2018), or differences in vegetation (Sebastià, 2004). Since we used a hierarchy of controls approach (Manning et al., 2015), these indirect topographical effects on SOC stocks could be behind the exclusion in the linear models of some drivers selected in the BRT model, like silt or pH (Figs. 2 & 3). In addition, SOC stocks decreased with increase of slope, which may be attributed to reduced carbon inputs and increased carbon losses induced by steeper slopes (Yuan et al., 2019 and refferences therein). However, we found that increased temperature seasonality (TSIS) values partly compensated negative slope effects on SOC.

The effect of temperature seasonality on SOC stocks was also modified by grazing management. At low TSIS values, SOC stocks increased under moderate to high grazing pressure; this effect disappeared as TSIS values increased (Fig. 3D). Recent meta-analyses concluded that intensive grazing commonly has decreasing effects 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 including climate and soil type vegetation (Abdalla et al., 2018; Eze et al., 2018; Mcsherry and Ritchie, 2013). Moreover, moderate grazing intensities can increase SOC inputs by dung deposition, and aboveground and root biomass production (Franzluebbers et al., 2000; Zeng et al., 2015). In our study, grazing intensity was relatively moderate (see methods), therefore in our study increasing stocking rates may increase soil carbon inputs in moderate seasonality locations by enhancing aboveground and belowground productivity.

Soil texture also showed interactive effects on SOC stocks with climatic variables. In particular, clay effects on SOC stocks became negative as MAP values increased, (Fig. 3C; Table 2). Both MAP and clay content are widely assumed to be positively correlated to SOC (Wiesmeier et al., 2019) but high soil water content caused by high MAP may inhibit decomposition if a shortage of oxygen supply occurs (Xu et al., 2016b). Furthermore, fine texture soils could be waterlogged frequently, leading to inhibition of root growth and soil C allocation belowground (Mcsherry and Ritchie, 2013).

3.2 Geophysical, biochemical and grazing management factors driving SOC stocks

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 also for future modelling studies aiming to predict SOC, since geophysical variables are easier and less expensive to acquire and measure compared to biochemical ones (Manning et al., 2015).

TSIS was a key driver of SOC with a varying effect depending on macrotopography, slope and grazing intensity (Table 2; Fig. 3). This result contrasts with most of the previous studies addressing soil carbon, which usually pinpoint mean temperature and precipitation as the most important climate drivers of SOC (Hobley et al., 2015; Manning et al., 2015; Wiesmeier et al., 2019). Climate regulates large-scale patterns of aboveground net primary production (Chapin et al., 1987). In the case of mountain grasslands, cold climates imply a short phenological period of development for plants (Gómez, 2008). Cold Sites characterised by low mean temperatures presented a wider spectrum of TSIS values than warm sites, presenting both the lowest and the highest TSIS values (Fig. S9). Hence, the positive effect of TSIS on SOC could be associated with a higher biomass accumulation in cold locations with more favourable temperatures during summer, this fact reducing geophysical stress for plants and broadening their growth period (Garcia-Pausas et al., 2007; Kikvidze et al., 2005). This rise in soil organic matter inputs during summer would overcome an eventual increase of soil organic matter decomposition rates due to high temperatures (Sanderman et al., 2003), which could even be diminished if microbial biomass decreases as a result of soil moisture reduction (Puissant et al., 2018).

The interactions of TSIS with macrotopography and slope illustrate the capacity of landscape factors to modulate macroclimate effects on soil (Hook and Burke, 2000). Induced microclimate changes are often the explanation for the effects of topography in SOC (Lozano-García et al., 2016). In our case, SOC stocks increased with temperature seasonality, particularly at mountain-exposed areas (Fig. 3A; Table 2). In protected sites, located in shady slopes and valley bottoms, the hypothesized positive effect of high TSIS values on plant productivity could be mitigated due to lower solar radiation, longer snow-covered periods and freezing episodes (Garcia-Pausas et al., 2007; López-Moreno et al., 2013). Conversely, negative effects of low TSIS values on plant productivity could be compensated thanks to the more humid conditions in protected sites compared to the exposed sites (Garcia-Pausas et al., 2007). Additionally, it is important to take into account that differences in SOC between exposed and protected sites may also occur through other mechanisms, for instance the alteration of soil physico-chemical properties like pH, soil texture or stoniness (Zhang et al., 2018), or differences in vegetation (Sebastià, 2004). Since we used a hierarchy of controls approach (Manning et al., 2015), these topography indirect effects on SOC stocks could be behind the exclusion in the linear models of some drivers selected in the BRT model, like silt or pH (Figs. 2 & 3).

In addition, high TSIS values compensated SOC stocks decrease with a greater slope, which may be attributed to reduced carbon inputs and increased carbon losses induced by steeper slopes (Yuan et al., 2019 and refferences therein). Increases in grazing pressure elevated SOC stocks under low TSIS values (Fig. 3D). This was a surprising result according to recent meta-analyses, which concluded that grazing has commonly decreasing effects 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.

Interestingly, clay content and precipitation presented interacting effects on SOC (Fig. 3C; Table 2). Both MAP and clay content are widely assumed to be positively correlated to SOC (Wiesmeier et al., 2019). High MAP would increase SOC inputs by promoting plant productivity (Author et al., 2000; Hobley et al., 2015). Clay positive effects are often attributed to a larger contact surface of soil particles (Kennedy et al., 2002), the absorption of negatively charged organic matter, high soil water retention and the exclusion of decomposer organisms due to their low pore size (Krull et al., 2001). 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). However, as MAP values increased, clay effect on SOC became negative. To explain low SOC values at high MAP and high clay content, McSherry and Rithchie (2013) hypothesized that finer texture soils could be waterlogged more frequently, leading to inhibition of root growth and soil C allocation belowground.

The addition of soil nutrient and herbage variables to our Geophysical Model implied substitution of terms, including clay content and macrotopography, by newly added variables (Tables 2 & 3). This highlights the importance of indirect effects of these variables on SOC through other small scale drivers (Leifeld et al., 2015; Xu et al., 2016b; Zhu et al., 2019). The Combined Model was complex and included unfrequently tested effects involving interactions between grazer type, soil nutrients and herbage quality variables (Table 3, Fig 4). Those results must be interpreted cautiously, because they are based on observational data, but can contribute to generate testable hypotheses for later studies about some complex and untested relationships between SOC and its drivers. Interaction experiments concerning soil properties are expensive and rare in the literature (Rillig et al., 2019).

For this reason, SOC increased with the C/N ratio (Fig 4A), which may be explained by the difficulty of soil organic matter decomposition by soil microbes, decreasing decomposition rates of SOC with increasing soil C/N (Wanyama et al., 2019; Xu et al., 2016b). A positive relationship between SOC and soil N was also expected, since most of the soil N is in combined form with organic matter (Cambardella and Elliott, 1994

However, in this case, due to the wide range of conditions and the randomized sampling design of the PASTUS database, the raw correlation between soil N and SOC was somehow discrete ($r = 0.297$; p-value = 0.001; $R^2 = 0.088$), in comparison to other studies (i.e. Yan et al. 2020). However, the novelty revealed by our model is that soil N could modulate the effects of certain SOC drivers, including livestock type and herbage NDF.

Cattle-grazed grasslands stored more SOC than mixed- and sheep-grazed grasslands, but only under low soil N conditions (Fig. 4B). Chang et al. (2018) found that in a N poor semi-arid grassland, sheep decreased SOC content in comparison to cattle due to vegetation changes caused by their feeding preference for highly palatable forbs (Sebastia et al., 2008), thus promoting less palatable grasses which supported less root biomass. In overall, under low soil N conditions, palatable plants are expected to contribute to SOC inputs through the stimulation of C allocation in forb roots (Ågren and Franklin, 2003; Warembourg et al., 2003) and the increase in the overall plant productivity due to legume atmospheric N fixation (Van Der Heijden et al., 2008).

However, these processes could decline under high soil N contents. For instance, legume atmospheric N fixation could be reduced since it requires additional energy in comparison to nitrogen acquisition from the soil (Ibañez et al., 2020; Minchin and Witty, 2005). Additionally, sheep selective feeding habits could shift plant leaf traits in the community towards nutrient-conservative leaf traits, which commonly induce fungal-based soil food webs, with slow nutrient–cycling and high SOC storage due to low decomposition rates (Orwin et al., 2010).

Additionally, grasslands with mixed grazed regimes stored even more SOC than sheep-grazed grasslands under high soil N conditions (Fig. 4B, Table 3). This result emphasises that mixed livestock assemblages deserve particular attention, because mixed grazing can affect plant composition distinctly from single grazing species regimes, and alter traveling and trampling behaviour of grazing animals (Aldezabal et al., 2019; Chang et al., 2018; Liu et al., 2015).

NDF was negatively related to SOC at high soil N values (Fig 4C). NDF proportion represents the amount of structural compounds on litter, and hence is inversely related to non-structural compounds content (Goering and Van Soest, 1970). The latter are the main source of organic matter formation at the early stages of decomposition, and they are incorporated into microbial biomass in a highly efficient way (Cotrufo et al., 2013). However, if microbial necromass was recycled by microbes before its incorporation to mineral-associated organic matter (Córdova et al., 2018), it could be respired and mineralized instead of stored. Thus, our model suggests that incorporation of labile and fast metabolized non-organic compounds to soil organic matter could be a pathway of SOC allocation at high soil N in Pyrenean grasslands.

On the other hand, the ADL/NH ratio was positively related to SOC in sheep and mixed grazed grasslands (Fig. 4D). The ADL/NH ratio is a commonly used indicator for the resistance of litter to degradation, particularly at later stages of decomposition (Taylor et al., 1989). Hence, the increase of SOC stocks with ADL/NH should be related to the physical pathway of soil organic matter incorporation, forming coarse particulate organic matter (Cotrufo et al., 2015). Moreover, our model suggests that this pathway would be inhibited under cattle grazing, presumably because of their higher digestive efficiency, and thus less recalcitrant faeces (Wang et al., 2018); and their less selective diet compared to sheep, as the latter would avoid plants with high lignin content, promoting recalcitrant litter (Rosenthal et al., 2012; Sebastià et al., 2008).

Our results concerning interactions between grazer type and herbage quality provide some evidence of grazing effects not only through alterations of plant communities that were reported by previous studies in the region (Canals and Sebastià, 2000; Sebastià et al., 2008), but also through interactions with them. Although grazing effects were not the most important factors affecting SOC stocks, this is by far the easiest component to manipulate in order to increase or maintain SOC in soils and face climate change (Komac et al., 2014). Considering our results, we suggest conducting more experiments to investigate grazer type effects on SOC under different soil nutrient conditions, and within plant communities with contrasting herbage quality parameters. Grazing management also has other advantages such as preventing the accumulation of aboveground C, and reducing the risk of forest fires (Nunes and Lourenço, 2017).

One key point of our results is that reinforce the idea that grazer type might be at least as important as grazing intensity in regulating grassland ecosystem dynamics (Tóth et al., 2018), and highlight the need for a more thorough research effort in disentangling not only grazing intensity but also grazer type effects on grassland soil organic carbon and nutrient cycling, under different environmental circumstances. Complete Model provided some evidence supporting that grazing may affect SOC not only through alterations of plant communities (Canals and Sebastià, 2000; Sebastià et al., 2008), but also through interactions with them. Although grazing effects were not the most important factors affecting SOC stocks, this is by far the easiest component to manipulate in order to increase or maintain SOC in soils and face climate change (Komac et al., 2014). Despite the need of a precise knowledge on the effects of different land uses on ecosystems for climate change mitigation (Lo et al., 2015), studies addressing grazer type effects on SOC are scarce (i.e. Zhou et al., 2017; Chang et al., 2018). Considering our results, we suggest conducting more experiments which investigate grazer type effects on SOC under different soil nutrient conditions, and within plant communities with contrasting herbage quality parameters.

Conclusion

The models presented here show a series of novel broad-scale and local patterns concerning SOC stocks and their geophysical, biochemical and grazing management drivers. Factors driving SOC stocks often interacted in complex ways, within and between spatio-temporal scales. Temperature seasonality (TSIS) was the most critical geophysical factor, affecting SOC through interactions with topographical drivers and grazing intensity. This illustrates how not only climate mean annual conditions should be considered when modelling SOC drivers, but also seasonal patterns. Concerning biochemical factors, we found that the expected positive effect of soil N was modulated by livestock species and herbage NDF; and herbage recalcitrance effects on SOC varied depending on grazer type. Overall, we found a number of interactions highlighting the need to expand knowledge on grassland SOC drivers under different conditions, specially grazing. The latter is the most easily tractable factor affecting SOC. In conclusion, we provided valuable information for further studies dealing with SOC predictions at broad several scales, and laid out the basis to generate new testable hypotheses for future studies, which may be useful for designing improved policies to palliate climate change.

DATA ACCESSIBILITY

Data are not public as the PASTUS database is currently being used for other research projects. Please contact one of us by e-mail for queries concerning the data used in this study.

Acknowledgements

We would like to express our thanks to the many people who collaborated in fieldwork, sample processing and data analysis over time. Research in this paper is based on the PASTUS database, which was compiled from different funding sources over time, the most relevant being: the EU Interreg III- A Programme (I3A- 4- 147- E) and the POCTEFA Programme/Interreg IV- A (FLUXPYR, EFA 34/08); the Spanish Science Foundation FECYT- MICINN (CARBOPAS: REN2002- 04300- C02- 01; CARBOAGROPAS: CGL2006- 13555- C03- 03 and CAPAS: CGL2010- 22378- C03- 01); the Foundation Catalunya- La Pedrera and the Spanish Institute of Agronomical Research INIA (CARBOCLUS: SUM2006- 00029- C02- 0). L. San Emeterio was funded through a Talent Recruitment grant from Obra Social La Caixa - Fundación CAN. ARAID foundation is acknowledged for support to J.J. Jiménez. This work was funded by the

Spanish Science Foundation FECYT- MINECO (projects BIOGEI: GL2013- 49142- C2-1- R; and IMAGINE: CGL2017-85490-R) and the University of Lleida (PhD Fellowship to AR).

References

Abdalla, M., Hastings, A., Chadwick, D. R., Jones, D. L., Evans, C. D., Jones, M. B., Rees, R. M. and Smith, P.: Critical review of the impacts of grazing intensity on soil organic carbon storage and other soil quality indicators in extensively managed grasslands, Agric. Ecosyst. Environ., 253(November 2017), 62–81, doi:10.1016/j.agee.2017.10.023, 2018.

Aerts, R. and Chapin, F. S.: The Mineral Nutrition of Wild Plants Revisited: A Re-evaluation of Processes and Patterns, Adv. Ecol. Res., 30(C), 1–67, doi:10.1016/S0065-2504(08)60016-1, 1999.

Ågren, G. I. and Franklin, O.: Root:shoot ratios, optimization and nitrogen productivity, Ann. Bot., 92(6), 795–800, doi:10.1093/aob/mcg203, 2003.

Aldezabal, A., Pérez-López, U., Laskurain, N. A. and Odriozola, I.: El abandono del pastoreo afecta negativamente a la calidad del pasto en pastizales atlánticos ibéricos ; Grazing abandonment negatively affects forage quality in iberian atlantic grasslands, Rev. Ecol. Montaña, 174, 42, doi:10.3989/pirineos.2019.174002, 2019.

Author, V., Jobbagy, E. G., Jackson, R. B., Jobbagyi, E. G. and Jackson} ', R. B.: The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate and Vegetation, Source Ecol. Appl. Ecol. Appl., 10(102), 423–436, doi:10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2, 2000.

Bardgett, R. D. and Wardle, D. A.: Herbibore-mediated Linkages Between

Abobeground and Belowground Communities, Ecology, 84(9), 2258–2268, doi:10.1890/02-0274, 2003.

Barton, K.: MuMIn: Multi-model inference. R package version 1.9.13, Version, 1, 18, doi:citeulike:11961261, 2015.

Berninger, F., Susiluoto, S., Gianelle, D., Bahn, M., Wohlfahrt, G., Sutton, M., Garcia-Pausas, J., Gimeno, C., Sanz, M. J., Dore, S., Rogiers, N., Furger, M., Eugster, W., Balzarolo, M., Sebastià, M. T., Tenhunen, J., Staszewski, T. and Cernusca, A.: Management and site effects on carbon balances of european mountain meadows and rangelands, Boreal Environ. Res., 20(6), 748–760, 2015.

Bing, H., Wu, Y., Zhou, J., Sun, H., Luo, J., Wang, J. and Yu, D.: Stoichiometric variation of carbon, nitrogen, and phosphorus in soils and its implication for nutrient limitation in alpine ecosystem of Eastern Tibetan Plateau, J. Soils Sediments, 16(2), 405–416, doi:10.1007/s11368-015-1200-9, 2016.

Bouyoucos, G. J.: Directions for making mechanical analysis of soils by the hydrometer method, Soil Sci., 42(3), 225–2230, 1936.

Burnham, K. P. and Anderson, D. R.: Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed)., 2002.

Calcagno, V.: Package glmulti 1.0.7, Title Model selection and multimodel inference made easy, Community Ecol. Packag., 1–20, 2015.

Cambardella, C. A. and Elliott, E. T.: Carbon and Nitrogen Dynamics of Soil Organic Matter Fractions from Cultivated Grassland Soils, Soil Sci. Soc. Am. J., 58(1), 123–130, doi:10.2136/sssaj1994.03615995005800010017x, 1994.

Canals, R. M. and Sebastià, M. T.: Analyzing mechanisms regulating diversity in rangelands through comparative studies: A case in the southwestern Pyrennees, Biodivers. Conserv., 9(7), 965–984, doi:10.1023/A:1008967903169, 2000.

Carreras, J. and Diego, F. C.: Cartografia dels hàbitats de Catalunya 1:50000, edited by Generalitat de Catalunya, Barcelona., 2006.

Chang, Q., Wang, L., Ding, S., Xu, T., Li, Z., Song, X., Zhao, X., Wang, D. and Pan, D.: Grazer effects on soil carbon storage vary by herbivore assemblage in a semi-arid grassland, edited by S. Mukul, J. Appl. Ecol., 55(5), 2517–2526, doi:10.1111/1365-2664.13166, 2018.

Chang, R., Wang, G., Fei, R., Yang, Y., Luo, J. and Fan, J.: Altitudinal Change in Distribution of Soil Carbon and Nitrogen in Tibetan Montane Forests, Soil Sci. Soc. Am. J., 79(June), 1455–1469, doi:10.2136/sssaj2015.02.0055, 2015.

Chapin, F. S., Bloom, A. J., Field, C. B. and Waring, R. H.: Plant Responses to Multiple Environmental Factors, Bioscience, 37(1), 49–57, doi:10.2307/1310177, 1987.

Cleveland, C. C. and Liptzin, D.: C:N:P stoichiometry in soil: is there a “Redfield ratio” for the microbial biomass?, Biogeochemistry, 85(3), 235–252, doi:10.1007/s10533-007-9132-0, 2007.

Córdova, S. C., Olk, D. C., Dietzel, R. N., Mueller, K. E., Archontouilis, S. V. and Castellano, M. J.: Plant litter quality affects the accumulation rate, composition, and stability of mineral-associated soil organic matter, Soil Biol. Biochem., 125, 115–124, doi:10.1016/J.SOILBIO.2018.07.010, 2018.

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K. and Paul, E.: The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter?, Glob. Chang. Biol., 19(4), 988–995, doi:10.1111/gcb.12113, 2013.

Cotrufo, M. F. F., Soong, J. L. J. L., Horton, A. J. A. J., Campbell, E. E., Haddix, M. L. M. L. L., Wall, D. H. D. H. and Parton, W. J. W. J.: Formation of soil organic matter via biochemical and physical pathways of litter mass loss, Nat. Geosci., 8(10), 776–779, doi:10.1038/ngeo2520, 2015.

David, D. J.: The determination of exchangeable sodium, potassium, calcium and magnesium in soils by atomic-absorption spectrophotometry, Analyst, 85(1012), 495–503, doi:10.1039/AN9608500495, 1960.

Deng, X., Zhan, Y., Wang, F., Ma, W., Ren, Z., Chen, X., Qin, F., Long, W., Zhu, Z. and Lv, X.: Soil organic carbon of an intensively reclaimed region in China: Current status and carbon sequestration potential, Sci. Total Environ., 565, 539–546, doi:10.1016/j.scitotenv.2016.05.042, 2016.

Díaz, S., Lavorel, S., De Bello, F., Quétier, F., Grigulis, K. and Robson, T. M.: Incorporating plant functional diversity effects in ecosystem service assessments, Proc. Natl. Acad. Sci. U. S. A., 104(52), 20684–20689, doi:10.1073/pnas.0704716104, 2007.

Duarte-guardia, S., Peri, P. L., Amelung, W., Sheil, D., Laffan, S. W., Borchard, N., Bird, M. I. and Peri, P. L.: Better estimates of soil carbon from geographical data : a revised global approach, , 355–372, 2019.

Elith, J., Leathwick, J. R. and Hastie, T.: A working guide to boosted regression trees, J. Anim. Ecol., 77(4), 802–813, doi:10.1111/j.1365-2656.2008.01390.x, 2008.

Eze, S., Palmer, S. M. and Chapman, P. J.: Soil organic carbon stock in grasslands: Effects of inorganic fertilizers, liming and grazing in different climate settings, J.

Environ. Manage., 223(June), 74–84, doi:10.1016/j.jenvman.2018.06.013, 2018.

Fernández-Alonso, M. J., Díaz-Pinés, E., Ortiz, C. and Rubio, A.: Disentangling the effects of tree species and microclimate on heterotrophic and autotrophic soil respiration in a Mediterranean ecotone forest, For. Ecol. Manage., 430(August), 533–544, doi:10.1016/j.foreco.2018.08.046, 2018.

Fernández-Martínez, M., Vicca, S., Janssens, I. A., Sardans, J., Luyssaert, S., Campioli, M., Chapin, F. S., Ciais, P., Malhi, Y., Obersteiner, M., Papale, D., Piao, S. L., Reichstein, M., Rodà, F. and Peñuelas, J.: Nutrient availability as the key regulator of global forest carbon balance, Nat. Clim. Chang., 4(6), 471–476, doi:10.1038/nclimate2177, 2014.

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, Int. J. Climatol., 37(12), 4302–4315, doi:10.1002/joc.5086, 2017.

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D. and Baud-Bovy, G.: Package “car,” R Doc., 146, doi:10.1093/nar/gkq363, 2018.

Franzluebbers, A. J., Stuedemann, J. A., Schomberg, H. H. and Wilkinson, S. R.: Soil organic C and N pools under long-term pasture management in the Southern Piedmont USA, Soil Biol. Biochem., 32(4), 469–478, doi:10.1016/S0038-0717(99)00176-5, 2000.

Garcia-Pausas, J., Casals, P., Camarero, L., Huguet, C., Sebastià, M. T., Thompson, R. and Romanyà, J.: Soil organic carbon storage in mountain grasslands of the Pyrenees: Effects of climate and topography, Biogeochemistry, 82(3), 279–289, doi:10.1007/s10533-007-9071-9, 2007.

Garcia-Pausas, J., Romanyà, J., Montané, F., Rios, A. I., Taull, M., Rovira, P. and

Casals, P.: Are Soil Carbon Stocks in Mountain Grasslands Compromised by Land-Use Changes?, in High Mountain Conservation in a Changing World, edited by J. Catalan, J. M. Ninot, and M. M. Aniz, pp. 207–230, Springer International Publishing, Cham., 2017.

Gobierno de Navarra: Mapa de usos del suelo de Navarra 1:25000., 2003.

Goering, H. K. and Van Soest, P. J.: Forage fiber analyses (apparatus, reagents, prcedures, and some applications), USDA Agr Handb [online] Available from: http://agris.fao.org/agris-search/search.do?recordID=US201301156229 (Accessed 18 April 2019), 1970.

Gómez, D.: Aspectos ecológicos de los pastos, in Pastos del Pirineo, edited by F. Fillat, R. García-González, D. Gómez, and R. Reiné, pp. 61–73, Consejo Superior de Investigaciones Científicas, CSIC, Madrid, Spain., 2008.

Gray, J. M., Bishop, T. F. A. and Wilson, B. R.: Factors Controlling Soil Organic Carbon Stocks with Depth in Eastern Australia, Soil Sci. Soc. Am. J., 79(6), 1741, doi:10.2136/sssaj2015.06.0224, 2015.

Greenwell, B., Boehmke, B. and Cunningham, J.: Package “gbm,” [online] Available from: https://orcid.org/0000-0002-3611-8516 (Accessed 15 February 2019), 2019.

Grömping, U.: R package relaimpo: relative importance for linear regression, J. Stat. Softw., 17(1), 139–147, doi:10.1016/j.foreco.2006.08.245, 2006.

Grueber, C. E., Nakagawa, S., Laws, R. J. and Jamieson, I. G.: Multimodel inference in ecology and evolution: Challenges and solutions, J. Evol. Biol., 24(4), 699–711, doi:10.1111/j.1420-9101.2010.02210.x, 2011.

Von Haden, A. C. and Dornbush, M. E.: Patterns of root decomposition in response to soil moisture best explain high soil organic carbon heterogeneity within a mesic, restored prairie, Agric. Ecosyst. Environ., 185, 188–196, doi:10.1016/j.agee.2013.12.027, 2014.

Van Der Heijden, M. G. A., Bardgett, R. D. and Van Straalen, N. M.: The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems, Ecol. Lett., 11(3), 296–310, doi:10.1111/j.1461-0248.2007.01139.x, 2008.

Hijmans, R. J., Phillips, S., Leathwick, J. and Maintainer, J. E.: Package “dismo” Type Package Title Species Distribution Modeling. [online] Available from: https://cran.r-project.org/web/packages/dismo/dismo.pdf (Accessed 15 February 2019), 2017.

Hobley, E., Wilson, B., Wilkie, A., Gray, J. and Koen, T.: Drivers of soil organic carbon storage and vertical distribution in Eastern Australia, Plant Soil, 390(1–2), 111–127, doi:10.1007/s11104-015-2380-1, 2015.

Hook, P. B. and Burke, I. C.: BIOGEOCHEMISTRY IN A SHORTGRASS LANDSCAPE: CONTROL BY TOPOGRAPHY, SOIL TEXTURE, AND MICROCLIMATE, Ecology, 81(10), 2686–2703, doi:10.1890/0012-9658(2000)081[2686:BIASLC]2.0.CO;2, 2000.

Ibañez, M., Altimir, N., Ribas, A., Eugster, W. and Sebastià, M. T.: Phenology and plant functional type dominance drive CO2 exchange in seminatural grasslands in the Pyrenees, J. Agric. Sci., 1–12, doi:10.1017/S0021859620000179, 2020.

Jenny, H.: Factors of soil formation: A System of Quantitative Pedology., Dover Publications., 1941.

Kennedy, M. J., Pevear, D. R. and Hill, R. J.: Mineral surface control of organic carbon in black shale., Science, 295(5555), 657–60, doi:10.1126/science.1066611, 2002.

Kikvidze, Z., Pugnaire, F. I., Brooker, R. W., Choler, P., Lortie, C. J., Michalet, R. and Callaway, R. M.: Linking patterns and processes in alpine plant communities: A global study, Ecology, 86(6), 1395–1400, doi:10.1890/04-1926, 2005.

Komac, B., Domènech, M. and Fanlo, R.: Effects of grazing on plant species diversity and pasture quality in subalpine grasslands in the eastern Pyrenees (Andorra): Implications for conservation, J. Nat. Conserv., 22(3), 247–255, doi:10.1016/J.JNC.2014.01.005, 2014.

Krull, E., Baldock, J. and Skjemstad, J.: Soil Texture Effects on Decomposition and Soil Carbon Storage, in Net Ecosystem Exchange Workshop, pp. 103–110. [online] Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.121.3042 (Accessed 13 May 2019), 2001.

Lal, R.: Soil Carbon Sequestration Impacts on Global Climate Change and Food Secruity, Science (80-.)., 304(56777), 1623–1627, 2004a.

Lal, R.: Soil carbon sequestration to mitigate climate change, Geoderma, 123(1–2), 1–22, doi:10.1016/J.GEODERMA.2004.01.032, 2004b.

de Lamo, X. and Sebastià, M. T.: Role of biogeographical , ecological and management factors on plant diversity in grasslands, in Sociedad Española para el Estudio de los Pastos (SEEP), vol. 11, edited by J. Lloveras, A. González-Rodríguez, O. Vázquez-Yañez, J. Piñeiro, O. Santamaría, L. Olea, and M. J. Poblaciones, pp. 832–834, Sociedad Española para el Estudio de los Pastos (SEEP), Madrid, Spain., 2006.

Legendre, P. and Legendre, L.: Numerical Ecology, 2nd ed., edited by Elsevier, Amsterdam., 1998.

Leifeld, J., Meyer, S., Budge, K., Sebastia, M. T., Zimmermann, M. and Fuhrer, J.: Turnover of grassland roots in mountain ecosystems revealed by their radiocarbon signature: Role of temperature and management, PLoS One, 10(3), 1–13, doi:10.1371/journal.pone.0119184, 2015.

Lenth, R., Singmann, H., Love, J., Buerkner, P. and Herve, M.: Package ‘emmeans,’ , doi:10.1080/00031305.1980.10483031, 2019.

Liu, C., Wang, L., Song, X., Chang, Q., Frank, D. A., Wang, D., Li, J., Lin, H. and Feiyue Du, |: Towards a mechanistic understanding of the effect that different species of large grazers have on grassland soil N availability, J. Ecol., 106, 357–366, doi:10.1111/1365-2745.12809, 2018.

Liu, J., Feng, C., Wang, D., Wang, L., Wilsey, B. J. and Zhong, Z.: Impacts of grazing by different large herbivores in grassland depend on plant species diversity, , doi:10.1111/1365-2664.12456, 2015.

Lo, Y. H., Blanco, J. A., Canals, R. M., González de Andrés, E., San Emeterio, L., Imbert, J. B. and Castillo, F. J.: Land use change effects on carbon and nitrogen stocks in the Pyrenees during the last 150 years: A modeling approach, Ecol. Modell., 312, 322–334, doi:10.1016/j.ecolmodel.2015.06.005, 2015.

López-Moreno, J. I., Revuelto, J., Gilaberte, M., Morán-Tejeda, E., Pons, M., Jover, E., Esteban, P., García, C. and Pomeroy, J. W.: The effect of slope aspect on the response of snowpack to climate warming in the Pyrenees, Theor. Appl. Climatol., 117(1), 207–219, doi:10.1007/s00704-013-0991-0, 2013.

Lozano-García, B., Parras-Alcántara, L. and Brevik, E. C.: Impact of topographic aspect and vegetation (native and reforested areas) on soil organic carbon and nitrogen budgets in Mediterranean natural areas, Sci. Total Environ., 544, 963–970, doi:10.1016/J.SCITOTENV.2015.12.022, 2016.

Lu, X., Yan, Y., Sun, J., Zhang, X., Chen, Y., Wang, X. and Cheng, G.: Carbon, nitrogen, and phosphorus storage in alpine grassland ecosystems of Tibet: effects of grazing exclusion., Ecol. Evol., 5(19), 4492–504, doi:10.1002/ece3.1732, 2015.

Lu, X., Kelsey, K. C., Yan, Y., Sun, J., Wang, X., Cheng, G. and Neff, J. C.: Effects of grazing on ecosystem structure and function of alpine grasslands in Qinghai-Tibetan Plateau: a synthesis, Ecosphere, 8(1), e01656, doi:10.1002/ecs2.1656, 2017.

Manning, P., de Vries, F. T., Tallowin, J. R. B., Smith, R., Mortimer, S. R., Pilgrim, E. S., Harrison, K. A., Wright, D. G., Quirk, H., Benson, J., Shipley, B., Cornelissen, J. H. C., Kattge, J., Bönisch, G., Wirth, C. and Bardgett, R. D.: Simple measures of climate, soil properties and plant traits predict national-scale grassland soil carbon stocks, J. Appl. Ecol., 52(5), 1188–1196, doi:10.1111/1365-2664.12478, 2015.

Mcsherry, M. E. and Ritchie, M. E.: Effects of grazing on grassland soil carbon: A global review, Glob. Chang. Biol., 19(5), 1347–1357, doi:10.1111/gcb.12144, 2013.

Minchin, F. R. and Witty, J. F.: Respiratory/carbon costs of symbiotic nitrogen fixation in legumes, in Plant respiration, pp. 195–205, Springer., 2005.

Nunes, A. N. and Lourenço, L.: Increased vulnerability to wildfires and post fire hydro-geomorphic processes in Portuguese mountain regions: What has changed?, Open Agric., 2(1), 70–82, doi:10.1515/opag-2017-0008, 2017.

Olsen, S. R.: Estimation of available phosphorus in soils by extraction with sodium bicarbonate, US Dept. of Agriculture., 1954.

Ottoy, S., Van Meerbeek, K., Sindayihebura, A., Hermy, M. and Van Orshoven, J.: Assessing top- and subsoil organic carbon stocks of Low-Input High-Diversity systems using soil and vegetation characteristics, Sci. Total Environ., 589, 153–164, doi:10.1016/J.SCITOTENV.2017.02.116, 2017.

Penin, D.: Cartographie des habitats naturels (typologie Corine Biotope) du massif Madres-Coronat., 1st ed., AGRNN/CPIE du Conflent., 1997.

Peri, P. L., Rosas, Y. M., Ladd, B., Toledo, S., Lasagno, R. G. and Pastur, G. M.: Modelling soil carbon content in South Patagonia and evaluating changes according to climate, vegetation, desertification and grazing, Sustain., 10(2), doi:10.3390/su10020438, 2018.

Puissant, J., Jassey, V. E. J., Mills, R. T. E., Robroek, B. J. M., Gavazov, K., De Danieli, S., Spiegelberger, T., Griffiths, R., Buttler, A., Brun, J.-J. J. and Cécillon, L.: Seasonality alters drivers of soil enzyme activity in subalpine grassland soil undergoing climate change, Soil Biol. Biochem., 124(June), 266–274, doi:10.1016/j.soilbio.2018.06.023, 2018.

R Core Team: R: A Language and Environment for Statistical Computing, 2017.

Rillig, M. C., Ryo, M., Lehmann, A., Aguilar-Trigueros, C. A., Buchert, S., Wulf, A., Iwasaki, A., Roy, J. and Yang, G.: The role of multiple global change factors in driving soil functions and microbial biodiversity, Science (80-.)., 366(6467), 886–890, doi:10.1126/science.aay2832, 2019.

Rodríguez, A., de Lamo, X., Sebastiá, M.-T. and Sebastià, M.-T.: Interactions between global change components drive plant species richness patterns within communities in mountain grasslands independetly of topography, edited by B. Collins, J. Veg. Sci., 29(August), 1029–1039, doi:10.1111/jvs.12683, 2018.

Rosenthal, G., Schrautzer, J. and Eichberg, C.: Low-intensity grazing with domestic herbivores: A tool for maintaining and restoring plant diversity in temperate Europe, Tuexenia, 32(1), 167–205, 2012.

Sanderman, J., Amundson, R. G. and Baldocchi, D. D.: Application of eddy covariance measurements to the temperature dependence of soil organic matter mean residence time, Global Biogeochem. Cycles, 17(2), n/a-n/a, doi:10.1029/2001GB001833, 2003.

Schlesinger, W. H.: Carbon Balance in Terrestrial Detritus, Annu. Rev. Ecol. Syst., 8(1), 51–81, doi:10.1146/annurev.es.08.110177.000411, 1977.

Schöning, I., Grüneberg, E., Sierra, C. A., Hessenmöller, D., Schrumpf, M., Weisser, W. W. and Schulze, E. D.: Causes of variation in mineral soil C content and turnover in differently managed beech dominated forests, Plant Soil, 370(1–2), 625–639, doi:10.1007/s11104-013-1654-8, 2013.

Sebastia, M.-T., Kirwan, L. and Connolly, J.: Strong shifts in plant diversity and vegetation composition in grassland shortly after climatic change, J. Veg. Sci., 19(3), 299–306, doi:10.3170/2008-8-18356, 2008.

Sebastià, M.-T.: Role of topography and soils in grassland structuring at the landscape and community scales, Basic Appl. Ecol., 5(4), 331–346, doi:10.1016/j.baae.2003.10.001, 2004.

Sebastià, M.-T., de Bello, F., Puig, L. and Taull, M.: Grazing as a factor structuring grasslands in the Pyrenees, Appl. Veg. Sci., 11(2), 215–222, doi:10.3170/2008-7-18358, 2008.

Shipley, B. and Parent, M.: Germination Responses of 64 Wetland Species in Relation to Seed Size, Minimum Time to Reproduction and Seedling Relative Growth Rate, Funct. Ecol., 5(1), 111, doi:10.2307/2389561, 1991.

Simard, R. R.: Ammonium acetate-extractable elements, in Soil sampling and methods of analysis, vol. 1, pp. 39–42, Lewis Publisher FL, USA., 1993.

Solly, E. F., Schöning, I., Boch, S., Kandeler, E., Marhan, S., Michalzik, B., Müller, J., Zscheischler, J., Trumbore, S. E. and Schrumpf, M.: Factors controlling decomposition rates of fine root litter in temperate forests and grasslands, Plant Soil, 382(1–2), 203–218, doi:10.1007/s11104-014-2151-4, 2014.

Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., Courcelles, V. de R. de, Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., Lehmann, J., O'Donnell, A. G., Parton, W. J., Whitehead, D. and Zimmermann, M.: The knowns, known unknowns and unknowns of sequestration of soil organic carbon, Agric. Ecosyst. Environ., 164(2013), 80–99, doi:10.1016/j.agee.2012.10.001, 2013.

Symonds, M. R. E. and Moussalli, A.: A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion, Behav. Ecol. Sociobiol., 65(1), 13–21, doi:10.1007/s00265-010-1037-6, 2011.

Taylor, B. R., Parkinson, D. and Parsons, W. F. J.: Nitrogen and Lignin Content as Predictors of Litter Decay Rates: A Microcosm Test, Ecology, 70(1), 97–104, doi:10.2307/1938416, 1989.

Tipping, E., Somerville, C. J. and Luster, J.: The C:N:P:S stoichiometry of soil organic matter, Biogeochemistry, 130(1–2), 117–131, doi:10.1007/s10533-016-0247-z, 2016.

Tóth, E., Deák, B., Valkó, O., Kelemen, A., Miglécz, T., Tóthmérész, B. and Török, P.: Livestock Type is More Crucial Than Grazing Intensity: Traditional Cattle and Sheep Grazing in Short-Grass Steppes, L. Degrad. Dev., 29(2), 231–239, doi:10.1002/ldr.2514, 2018.

Vitousek, P. M. and Howarth, R. W.: Nitrogen Limitation on Land and in the Sea: How Can It Occur?, 1991.

de Vries, F. T., Manning, P., Tallowin, J. R. B., Mortimer, S. R., Pilgrim, E. S., Harrison, K. A., Hobbs, P. J., Quirk, H., Shipley, B., Cornelissen, J. H. C., Kattge, J. and Bardgett, R. D.: Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities, Ecol. Lett., 15(11), 1230–1239, doi:10.1111/j.1461-0248.2012.01844.x, 2012.

Wanyama, I., Pelster, D. E., Butterbach-Bahl, K., Verchot, L. V, Martius, C. and Rufino, M. C.: Soil carbon dioxide and methane fluxes from forests and other land use types in an African tropical montane region, Biogeochemistry, 143(2), 171–190, doi:10.1007/s10533-019-00555-8, 2019.

Warembourg, F. R., Roumet, C. and Lafont, F.: Differences in rhizosphere carbon-partitioning among plant species of different families, Plant Soil, 256(2), 347–357, doi:10.1023/A:1026147622800, 2003.

White, R., Murray, S. and Rohweder, M.: Pilot Analysis of Global Ecosystems: Grassland Ecosystems., 2000.

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H.

J. and Kögel-Knabner, I.: Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales, Geoderma, 333(July 2018), 149–162, doi:10.1016/j.geoderma.2018.07.026, 2019.

Williams, P. C. and Sobering, D. C.: How do we do it: a brief summary of the methods we use in developing near infrared calibrations, in Near infrared spectroscopy: The future waves, pp. 185–188., 1996.

Williams, T. E., Chang, C. M., Rosen, E. L., Garcia, G., Runnerstrom, E. L., Williams, B. L., Koo, B., Buonsanti, R., Milliron, D. J. and Helms, B. A.: NIR-Selective electrochromic heteromaterial frameworks: A platform to understand mesoscale transport phenomena in solid-state electrochemical devices, J. Mater. Chem. C, 2(17), 3328–3335, doi:10.1039/c3tc32247e, 2014.

Xu, X., Shi, Z., Li, D., Rey, A., Ruan, H., Craine, J. M., Liang, J., Zhou, J. and Luo, Y.: Soil properties control decomposition of soil organic carbon: Results from data-assimilation analysis, Geoderma, 262, 235–242, doi:10.1016/j.geoderma.2015.08.038, 2016a.

Xu, X., Shi, Z., Li, D., Rey, A., Ruan, H., Craine, J. M., Liang, J., Zhou, J. and Luo, Y.: Soil properties control decomposition of soil organic carbon: Results from data-assimilation analysis, Geoderma, 262, 235–242, doi:10.1016/J.GEODERMA.2015.08.038, 2016b.

Yan, J., Wang, L., Hu, Y., Tsang, Y. F., Zhang, Y., Wu, J., Fu, X. and Sun, Y.: Plant litter composition selects different soil microbial structures and in turn drives different litter decomposition pattern and soil carbon sequestration capability, Geoderma, 319, 194–203, doi:10.1016/J.GEODERMA.2018.01.009, 2018.

Yan, Y., Tian, L., Du, Z., Chang, S. X. and Cai, Y.: Carbon, nitrogen and phosphorus stocks differ among vegetation patch types in a degraded alpine steppe, J. Soils Sediments, 19(4), 1809–1819, doi:10.1007/s11368-018-2191-0, 2019.

Yang, Y., Chen, Y., Li, Z. and Chen, Y.: Land-use / cover conversion affects soil organic-carbon stocks : A case study along the main channel of the Tarim River , China, , 1–14, 2018.

Yuan, Z.-Q., Fang, C., Zhang, R., Li, F.-M., Javaid, M. M. and Janssens, I. A.: Topographic influences on soil properties and aboveground biomass in lucerne-rich vegetation in a semi-arid environment, Geoderma, 344, 137–143, doi:10.1016/J.GEODERMA.2019.03.003, 2019.

Zeng, C., Wu, J. and Zhang, X.: Effects of grazing on above-vs. below-ground biomass allocation of alpine grasslands on the northern tibetan plateau, PLoS One, 10(8), 1–15, doi:10.1371/journal.pone.0135173, 2015.

Zhang, X., Liu, M., Zhao, X., Li, Y., Zhao, W., Li, A., Chen, S., Chen, S., Han, X. and Huang, J.: Topography and grazing effects on storage of soil organic carbon and nitrogen in the northern China grasslands, Ecol. Indic., 93, 45–53, doi:10.1016/J.ECOLIND.2018.04.068, 2018.

Zhao, Y., Ding, Y., Hou, X., Li, F. Y., Han, W. and Yun, X.: Effects of temperature and grazing on soil organic carbon storage in grasslands along the Eurasian steppe eastern transect, PLoS One, 12(10), 1–16, doi:10.1371/journal.pone.0186980, 2017.

Zhou, G., Zhou, X., He, Y., Shao, J., Hu, Z., Liu, R., Zhou, H. and Hosseinibai, S.: Grazing intensity significantly affects belowground carbon and nitrogen cycling in grassland ecosystems: a meta-analysis, Glob. Chang. Biol., 23(3), 1167–1179, doi:10.1111/gcb.13431, 2017.

Zhu, M., Feng, Q., Qin, Y., Cao, J., Zhang, M., Liu, W., Deo, R. C., Zhang, C., Li, R. and Li, B.: The role of topography in shaping the spatial patterns of soil organic carbon, CATENA, 176, 296–305, doi:10.1016/J.CATENA.2019.01.029, 2019.

Table captions

Table 1: Considered factors affecting SOC stocks in some recent studies. V: the study considers this variable type; -: the study does not consider this variable type.

Table 2: Results of the Geophysical Model for soil organic carbon (R^2_{Adj} = 0.34).

Table 3: Results of the Combined model for soil organic carbon (R^2_{Adj} = 0.61).

Figure captions

Figure 1: Conceptual scheme used in this work to investigate potential environmental drivers with SOC. We assume that drivers may affect soil organic carbon (SOC) both directly or hierarchically through another driver. Interactions between factors acting at different scales and belonging to different categories could also drive SOC. Grazing management has a special status because it may be acting at different scales, landscape and local.

Figure 2: Relative contributions (%) of driver variables in the final BRT model obtained. Soil N: soil nitrogen; Soil C/N: soil carbon to nitrogen ratio, Clay: clay content; Abiom: aboveground biomass; ADL: acid-detergent lignin; Loam: loam content; K: soil potassium; TSIS: temperature seasonality; NDF: neutro detergent fibre; pH: soil pH; CH: carbon in the herbage; Mg: soil magnesium; Slope: terrain slope; MAP: mean annual precipitation; ADF: acid detergent fibre. See Table S1 for more details about variables.

Figure 3: Relationship between SOC, and regional and landscape scale factors in the Geophysical Model. In A) solid lines and circles represent exposed hillsides, and dotted lines and triangles indicate protected hillsides. In D) solid lines and circles indicate low grazing intensity, dotted lines and triangles indicate medium grazing management intensity and dashed lines and squares indicate high grazing management intensity. In A-D line and plane values are predictions of the model across the corresponding

predictors´ range according to estimated marginal means. Grey areas around regression lines indicate standard errors. In A) and D) points indicate actual values.

Figure 4. The relationship between SOC, and biochemical and herbage factors in the Combined model. In B) and D) solid lines and circle points represent cattle-grazing, dashed lines and square points indicate sheep-grazing, and dotted lines and triangle points indicate mixed-grazing. In A-D line and plane values are predictions of the model across the corresponding predictors´ range according to estimated marginal means. In A-D line and plane values are predictions of the model across the corresponding predictors´ range according to estimate marginal means. Grey spectrum indicate 95% confidence intervals. In A) and D) points indicate actual values.

Tables

Table 1: Considered factors affecting SOC stocks in some recent studies. V: the study considers this variable type; -: the study does not consider this variable type.

1: It considers SOC concentrations
2: It considers total carbon stocks
3: It considers total carbon stocks and its fractions.

Article	Location	LAT (º)	LONG (º)	MAP (mm)	MAT (°C)	Topography and bedrock	Grazing Management	Soil properties	Soil nutrients	Herbage
Present study	***Pyrenees***	***42.14 – 43.3***	***-1.22 – 2.26***	***964 – 1586***	***1.1 – 9.9***	✓	✓	✓	✓	✓
Duarte-guardia et al., 2019	Worldwide	-51.72 – 80.23	-163.95 – 158.25	65 – 5115	-21.2 – 30	✓	-	✓	-	✓**
Abdalla et al., 2018	Worldwide	-45.85 – 51	-114 – 120.7	150 – 1650	0 – 21	-	✓	✓	-	✓
Eze et al., 2018	Worldwide	-44 – 65	-121 – 175	120 – 2000	-4.8 – 26.8	-	✓	✓	✓*	✓**
Peri et al., 2018[1]	South Patagonia	- 52 – -45	-73.5 – 65.5	139 – 865	4.2 – 11	✓	✓	-	-	✓
Zhang et al., 2018	Northern China	103.5 – 124.16	32.5 – 42.5	500 – 1000	8 – 14	✓	✓	✓	-	-
Zhao et al., 2017	Mongolia	41.95 – 53.93	108.28- 116.2	150 – 400	-1.3 – 2.1	-	✓	✓	-	✓
Zhou et al., 2017[2]	Worldwide	-42.1 – 52.3	-121 – 175	200 – 600	0 – 10	-	✓	-	-	✗
Deng et al., 2016	Eastern China	28.71 – 30.45	120.87 – 122.43	940 – 1720	16.86 – 18.57	✓	-	✓	-	✗
Gray et al., 2015	Eastern Australia	-16.7 – -43.5	-31.8 – -28.7	500 – 2000	10 – 30	✓	✗	✗	-	✓
Lu et al., 2017	Qinghai-Tibetan Plateau	27 – 32	83 – 108	37 – 718	-4.04 – 6.3	-	✓	✗	-	-
Chang et al., 2015[1]	Tibet	Not Reported	Not Reported	397 – 1910	1.7 – 15.5	✓	-	-	-	✓
Manning et al. 2015[3]	England	50.77– 54.58	-4.43 – 0.87	596 – 3191	6.5 – 10.9	-	✓	✓	-	✓
McSherry & Ritzie 2013	Worldwide	Not reported	Not reported	180 – 950	Not reported	-	✓	✓	-	✓
Garcia-Pausas et al. 2007	Pyrenees	-7 – 2.2	42.5 – 42.9	1416 – 1904	-0.7 – 5	✓	-	✓	-	-

*Fertilizer effects.
** Only aboveground and/or belowground biomass index.

Table 2: Results of the Geophysical Model for soil organic carbon (R^2_{Adj} = 0.34). MAP: mean annual precipitation; TSIS: temperature seasonality; Slope: terrain slope; Exposed: Exposed position according to Macrotopography; Clay: clay content; Low and medium intensity: Low and medium Grazing intensity.

Model term	Estimate	SE	t-value	P-value	
Intercept	-0.525	1.802	-0.291	0.771	
Climate variables					
MAP	0.003	0.001	4.560	<0.001	***
TSIS	-0.098	0.228	-0.429	0.669	
Topography variables					
Slope	-0.339	0.095	-3.569	0.001	***
Exposed	-3.130	0.936	-3.344	0.001	**
Soil type variables					
Clay	0.121	0.027	4.500	<0.001	***
Management variables					
Low intensity	-5.013	1.196	-4.192	<0.001	***
Medium intensity	2.012	1.168	1.722	0.088	
Interactions					
TSIS x Exposed	0.417	0.124	3.358	0.001	**
TSIS x Slope	0.044	0.013	3.452	0.001	***
MAP x Clay	0.000	0.000	-4.637	<0.001	***
TSIS x Low intensity	0.655	0.159	4.110	<0.001	***
TSIS x Medium intensity	-0.262	0.156	-1.684	0.095	

Table 3: Results of the Combined model for soil organic carbon (R^2_{Adj} = 0.61). MAP: mean annual precipitation; TSIS: mean summer temperature minus mean annual temperature; Slope: terrain slope; Cattle and Mixed: Cattle and mixed management according to grazing species; Low and medium intensity: Low and medium intensity according to Grazing intensity; Soil C/N: soil carbon to nitrogen ratio; soil N: soil nitrogen; NDF: neutro-detergent fibre; ADL/NH: acid-detergent lignin to nitrogen in the herbage ratio.

Model term	Estimate	SE	t-value	P-value	
Intercept	-0.290	1.458	-0.199	0.843	
Climate variables					
MAP	-0.001	0.000	-2.434	0.017	*
TSIS	-0.004	0.181	-0.022	0.982	
Topography variables					
Slope	-0.225	0.078	-2.868	0.005	**
Management variables					
Cattle	0.487	0.101	4.834	<0.001	***
Mixed	-0.289	0.093	-3.106	0.002	**
Low intensity	-3.249	1.014	-3.204	0.002	**
Medium intensity	1.666	1.073	1.553	0.123	
Soil nutrient variables					
Log(Soil C/N)	0.665	0.076	8.777	<0.001	***
Soil N	3.302	0.617	5.349	<0.001	***
Herbage variables					
NDF	0.014	0.006	2.525	0.013	*
Herbage ADL/NH	0.026	0.009	2.987	0.003	**
Interactions between variable types					
TSIS x Slope	0.030	0.010	2.833	0.006	**
TSIS x Low intensity	0.423	0.136	3.104	0.002	**
TSIS x Medium intensity	-0.214	0.143	-1.495	0.138	
Soil N x Cattle grazing	-0.736	0.168	-4.380	<0.001	***
Soil N x Mixed grazing	0.493	0.175	2.813	0.006	**
Soil N x NDF	-0.039	0.011	-3.505	0.001	***
Cattle x Herbage ADL/NH	-0.030	0.010	-2.872	0.005	**
Mixed x Herbage ADL/NH	0.014	0.011	1.252	0.213	

Figures

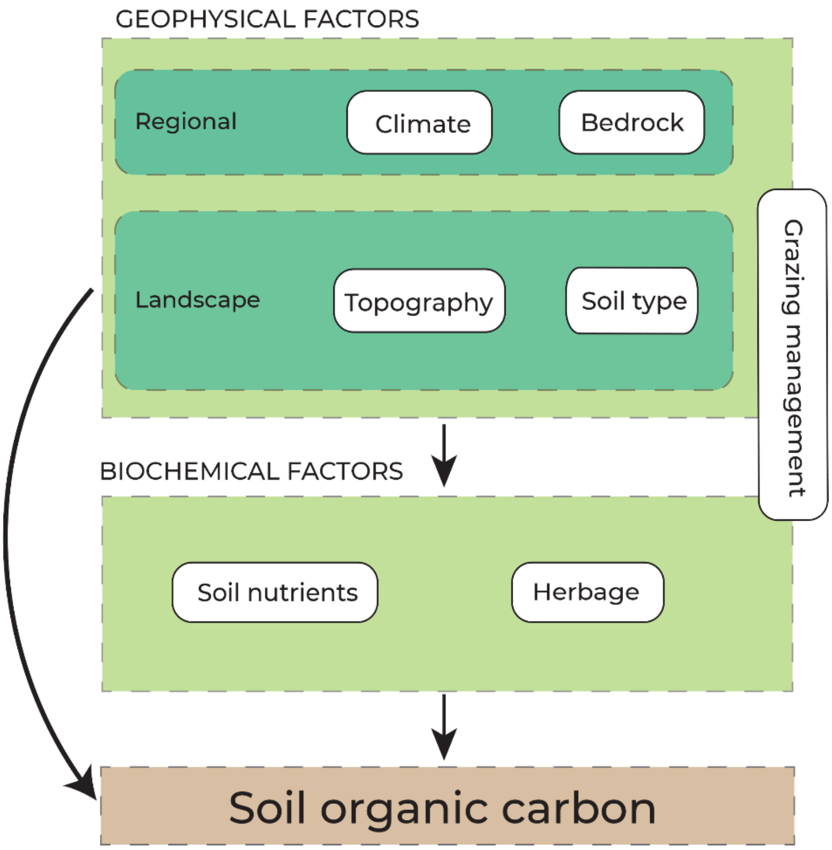


Figure 1: Conceptual scheme used in this work to investigate potential environmental drivers with SOC. We assume that drivers may affect soil organic carbon (SOC) both directly or hierarchically through another driver. Interactions between factors acting at different scales and belonging to different categories could also drive SOC. Grazing management has a special status because it may be acting at different scales, landscape and local.

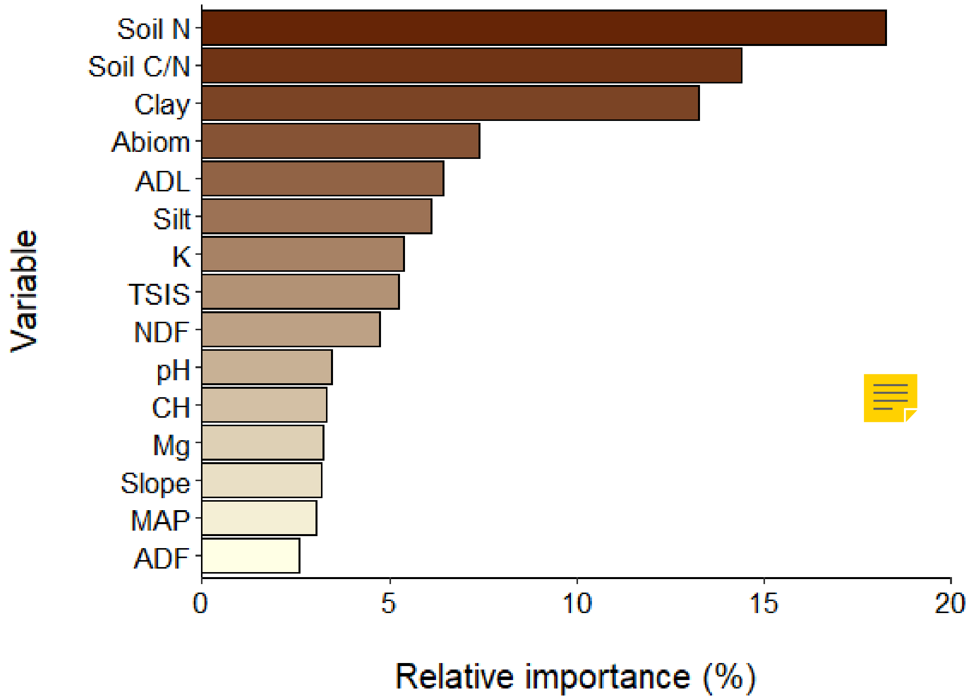


Figure 2: Relative contributions (%) of driver variables in the final BRT model obtained. Soil N: soil nitrogen; Soil C/N: soil carbon to nitrogen ratio, Clay: clay content; Abiom: aboveground biomass; ADL: acid-detergent lignin; ~~Loam: loam~~ content; K: soil potassium; TSIS: temperature seasonality; NDF: neutro detergent fibre; pH: soil pH; CH: carbon in the herbage; Mg: soil magnesium; Slope: terrain slope; MAP: mean annual precipitation; ADF: acid detergent fibre. See Table S1 for more details about variables.

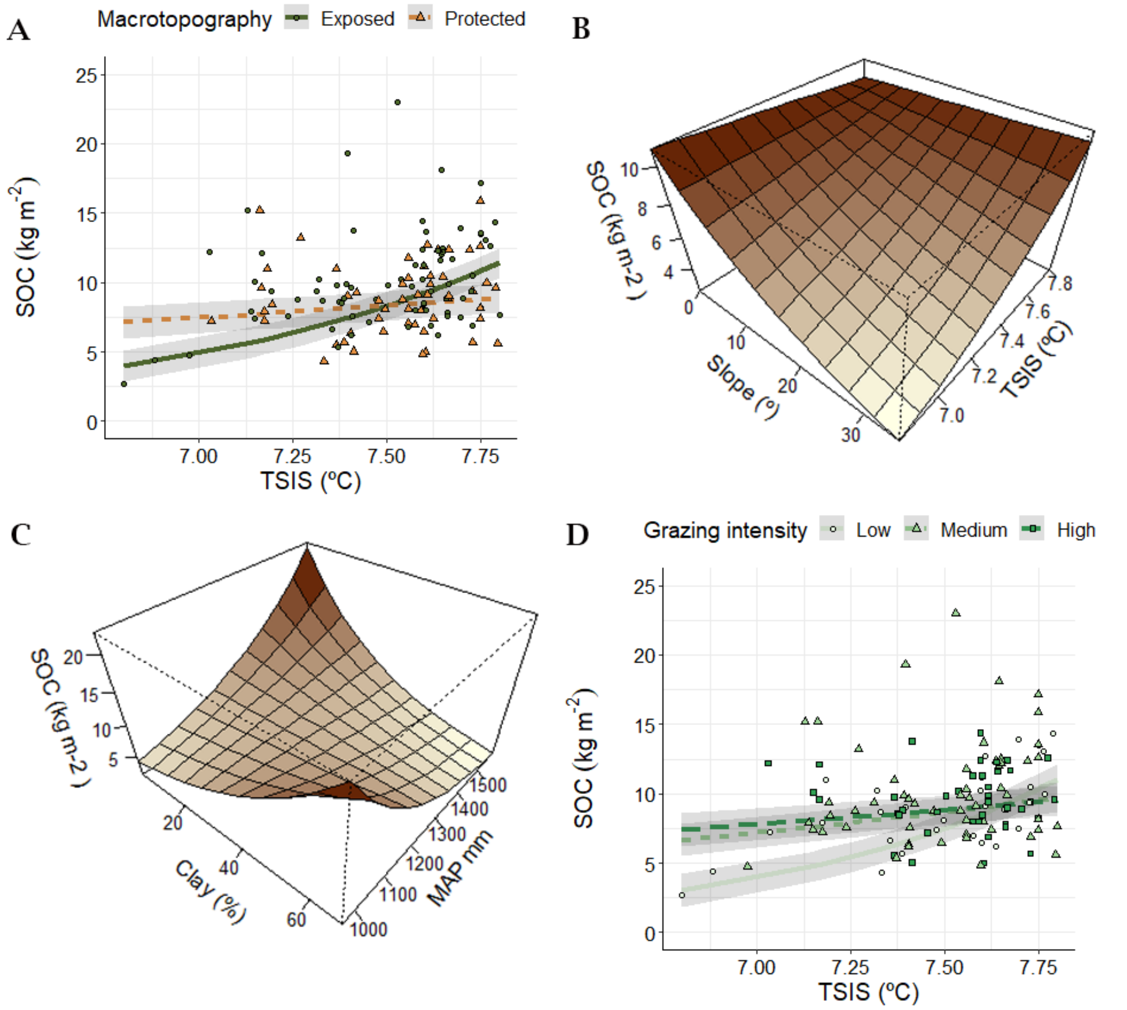


Figure 3: Relationship between SOC, and regional and landscape scale factors in the Geophysical Model. In A) solid lines and circles represent exposed hillsides, and dotted lines and triangles indicate protected hillsides. In D) solid lines and circles indicate low grazing intensity, dotted lines and triangles indicate medium grazing management intensity and dashed lines and squares indicate high grazing management intensity. In A-D line and plane values are predictions of the model across the corresponding predictors´ range according to estimated marginal means. Grey areas around regression lines indicate standard errors. In A) and D) points indicate actual values.

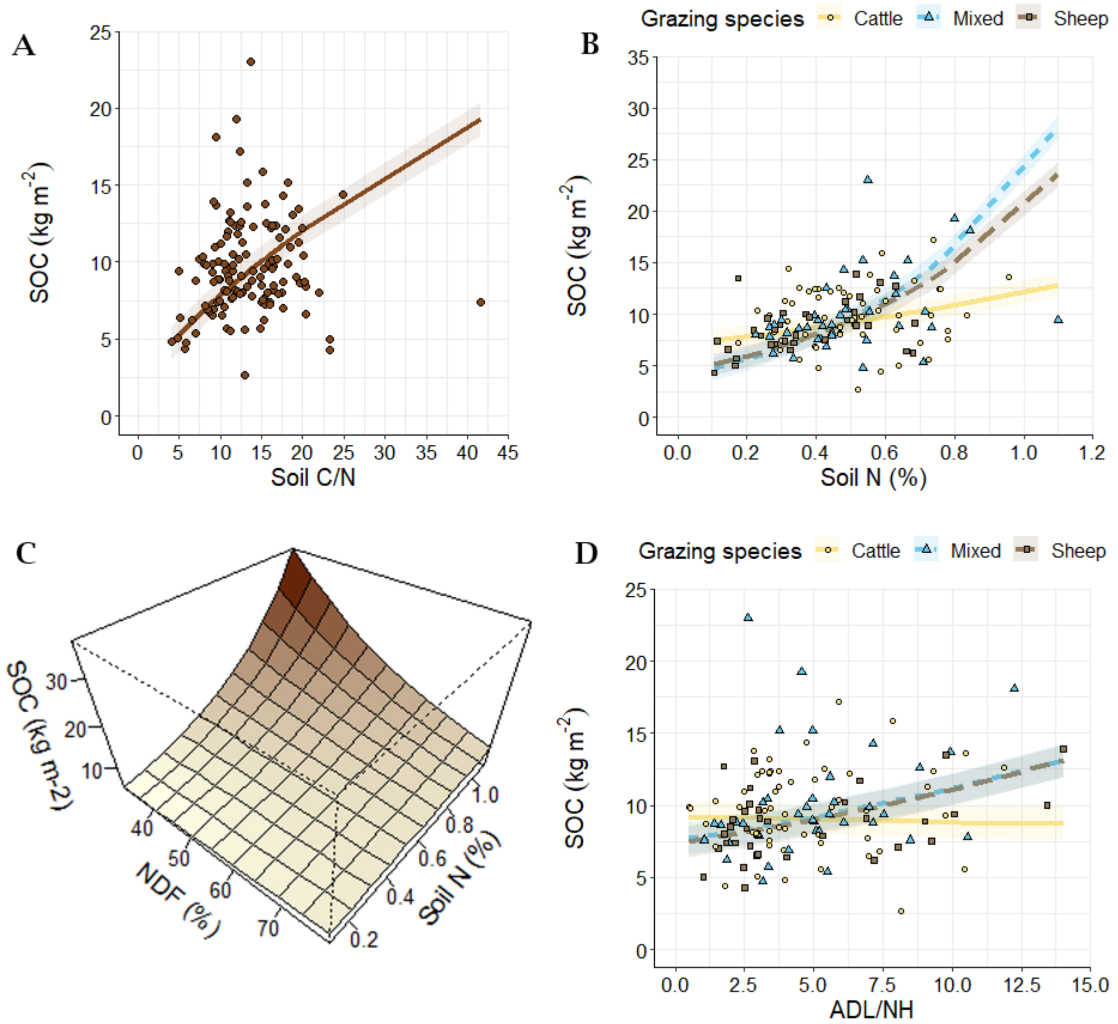


Figure 4. The relationship between SOC, and biochemical and herbage factors in the Combined model. In B) and D) solid lines and circle points represent cattle-grazing, dashed lines and square points indicate sheep-grazing, and dotted lines and triangle points indicate mixed-grazing. In A-D line and plane values are predictions of the model across the corresponding predictors´ range according to estimated marginal means. In A-D line and plane values are predictions of the model across the corresponding predictors´ range according to estimate marginal means. Grey spectrum indicate 95% confidence intervals. In A) and D) points indicate actual values.

Supplementary material

Table S1: Variables considered in this study.

Variable	Description
Regional variables	
Climate variables	
MAP	Mean Annual Precipitation, mm.
MSP	Mean Summer Precipitation, mm.
MAT	Mean Annual Temperature, ºC.
MST	Mean Summer Temperature, ºC.
TSIS	MST-MAT.
Bedrock	3 cathegories : Basic (marls and calcareous rocks), Acidic (mostly sandstones and slates) or Mixed.
Landscape variables	
Topographical variables	
Slope	Pendent, º.
Aspect	Cos(º)
Macrotopography	Protected; north-facing slopes; Exposed, south-facing slopes.
Microtopography	Flat areas, convexities or mounds, and concavities, convexities or smooth areas.
Soil type variables	
Sand10	Percentage of sands in the 10 cm upper layer (%).
Clay	Percentage of clays in the 10 cm upper layer (%).
Loam	Percentage of loams in the 10 cm upper layer (%).
pH	pH value in soil 10 cm upper layer.
Management variables	
Management	Grazer type : Cattle, Sheep, Mixed
Grazing	Grazing intensity, (units of big grazer (UBG ha-1) low (1; lower than 0.2 UBG ha-1), medium (2; between 0.2-0.4 UBG ha-1) and high (3; up to 0.4 UBG ha-1).

Soil nutrient variables	
Soil N	N in soil 20 cm upper layer. (%).
C/N	Soil C/N ratio
P10	Cations of P10 in soil 10 cm upper layer. (ppm).
K10	Cations of K10 in soil 10 cm upper layer. (ppm).
Herbage	
Abiom	Avoveground biomass in g/m^2
ADL	Lignin concentration by the acid detergent lingin method (%/DM).
ADF	Fiber concentration by the acid detergent fiber method (%/DM).
NDF	Fiber concentration by the neutro detergent fiber method (%/DM).
NH	Nitrogen in the herbage (%/DM).
CH	Carbon in the herbage (%/DM)
CH/NH	CH/NH
ADL/NH	ADL/NH
NDF/CP	NDF/CP (CP: crude protein)
SOC20	Soil Organic Carbon stocks in the 20 cm upper layer (kg m^{-2}).

Table S2: Minimum, maximum, median and mean values of the continuous predictors of this study. Units are shown in Table S1. MAT: mean annual temperature; MST: mean summer temperature; TSIS: mean summer temperature minus mean annual temperature; MAP: mean annual precipitation; MSP: mean summer precipitation; Slope: terrain slope; Aspect:; Sand: sand content; Loam: loam content; Clay: clay content; pH: soil pH; Soil N: soil nitrogen; Soil P: soil phosphorus; Soil C/N: soil carbon to nitrogen ratio; Soil Mg: soil magnesium; Soil K: soil potassium; NDF: neutro-detergent fibre; ADF: acid-detergent fibre; ADL: acid-detergent lignin; NH: nitrogen in the herbage; CH: carbon in the herbage; CH/NH: carbon to nitrogen ratio in the herbage; Abiom: aboveground biomass; NDF/CP: neutro-detergent fibre to crude protein ratio; ADL/NH: acid-detergent lignin to nitrogen in the herbage ratio.

	Minimum	Maximum	Median	Mean
MAT	1.08	9.90	4.72	4.96
MST	7.88	16.93	12.23	12.47
TSIS	6.80	7.80	7.58	7.51
MAP	964	1586	1252	1242.91
MSP	169.00	258.00	235.00	228.90
Slope	0.00	35.00	16.50	16.88
Aspect	1.00	3.00	1.84	2.05
Sand	3.10	72.20	32.80	32.67
Loam	13.60	73.50	38.60	39.80
Clay	2.90	68.60	27.25	27.53
pH	3.90	7.80	5.47	5.74
Soil N	0.11	1.10	0.46	0.47
Soil P	4.00	54.00	11.00	12.98
Soil C/N	4.13	41.60	12.47	13.39
Soil Mg	2.89	5.99	4.99	4.92
Soil K	3.40	6.84	4.99	5.03
NDF	31.20	78.90	52.45	52.08
ADF	17.70	46.60	29.55	30.07
ADL	1.16	12.72	6.32	6.63
NH	0.48	3.03	1.66	1.63
CH	22.60	49.10	45.15	44.53
CH/NH	13.90	97.20	26.60	31.14

Abiom	64.52	1224	308.32	341.91
NDF/CP	2.15	17.20	4.77	5.71
ADL/NH	0.50	14.02	3.92	4.78

Table S3: Chemical composition of herbage samples used for NIRS calibration. DM: dry matter; MM: mineral matter or ash content; CP: crude protein; NDF: neutro-detergent fibre; ADF: acid-detergent fibre; ADL: acid-detergent lignin; NH: nitrogen in the herbage; CH: carbon in the herbage.

Parameter, %	N	Min.	Max.	Mean	SD
DM	67	91.60	96.73	93.48	1.39
MM (Ash)	67	3.58	19.73	10.10	3.98
CP	67	5.50	14.67	9.29	1.90
NDF	67	36.82	73.11	55.42	9.27
ADF	67	21.95	41.97	30.00	4.70
ADL	67	3.35	12.52	6.18	2.08
NH	55	0.75	2.10	1.44	0.31
CH	55	36.83	51.13	45.10	2.99

Table S4: Calibration and cross validation statistics for predicting the chemical composition parameters in herbage samples by NIRS analysis. DM: dry matter; MM: mineral matter or ash content; CP: crude protein; NDF: neutro-detergent fibre; ADF: acid-detergent fibre; ADL: acid-detergent lignin; NH: nitrogen in the herbage; CH: carbon in the herbage.

Parameter	Math[a] treatment	Scatter[b] correction	R^2	r^2	SEC	SECV	RPD
DM	2,4,4,1	DT	0.92	0.85	0.392	0.539	2.58
Ash	2,4,4,1	MSC	0.83	0.70	1.583	0.830	4.80
CP	2,4,4,1	SNV	0.97	0.94	0.331	0.451	4.21
NDF	2,4,4,1	DT	0.83	0.72	3.756	4.728	1.96

ADF	2,4,4,1	DT	0.81	0.70	2.031	2.548	1.84
ADL	2,4,4,1	MSC	0.80	0.66	0.900	1.178	1.77
N	2,4,4,1	MSC	0.97	0.95	0.055	0.068	4.56
C	2,4,4,1	MSC	0.97	0.95	0.422	0.581	5.15

[a]Designations: derivate order, gap, first smoothing, and second smoothing; [b]Standard Normal Variance (SNV), Detrend (DT) and Multiplicative Scattering Correction (MSC) transformations.

R^2 = coefficient of determination for calibration. r^2 = coefficient of determination for cross validation. SEC = standard error of calibration. SECV = standard error of cross validation. RPD = ratio of performance to deviation (RPD=SD/SECV).

Table S5: Variance inflation values for the continuous predictors included in the GLMs. Values under 5 are considered non-problematic (Heiberger, 2017). MAP: mean annual precipitation; TSIS: mean summer temperature minus mean annual temperature; Slope: terrain slope; Clay: clay content; Soil C/N: soil carbon to nitrogen ratio; soil N: soil nitrogen; NDF: neutro-detergent fibre; ADL/NH: acid-detergent lignin to nitrogen in the herbage ratio.

Predictor	MAP	MMT	Slope	Clay	Log(soil C/N)	Soil N	NDF	ADL/NH
Geophysical model	1.26	1.16	1.27	1.22	-	-	-	-
Complete model	-	1.26	1.32	-	1.58	1.82	1.32	1.67

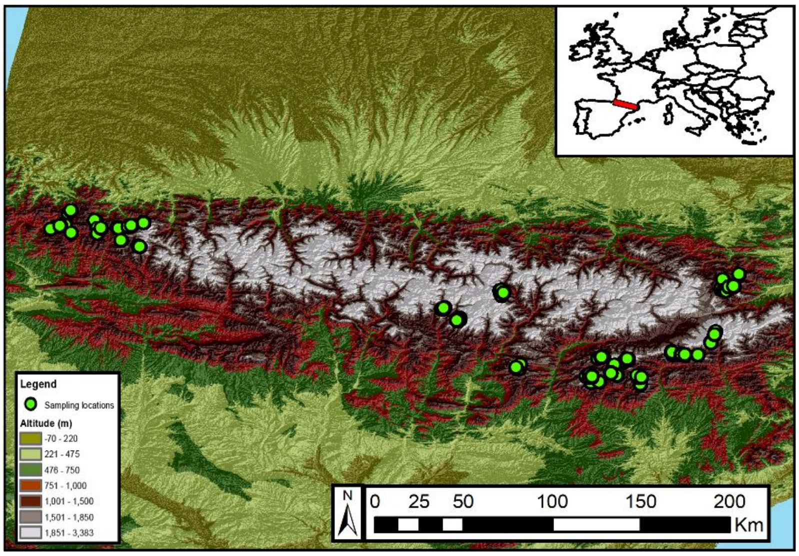
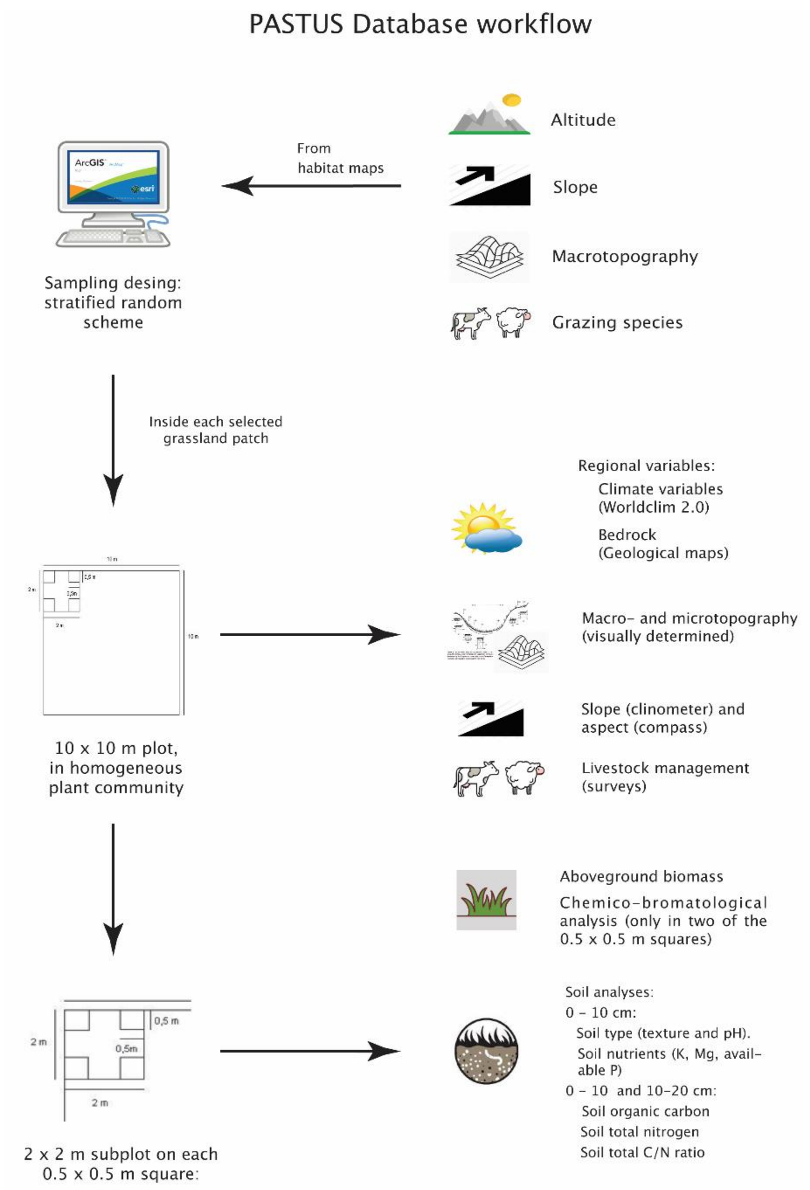


Figure S1: Map of the study area. Points indicate sampling locations.

43 Figure S2: Scheme of the sampling procedure for building the PASTUS database



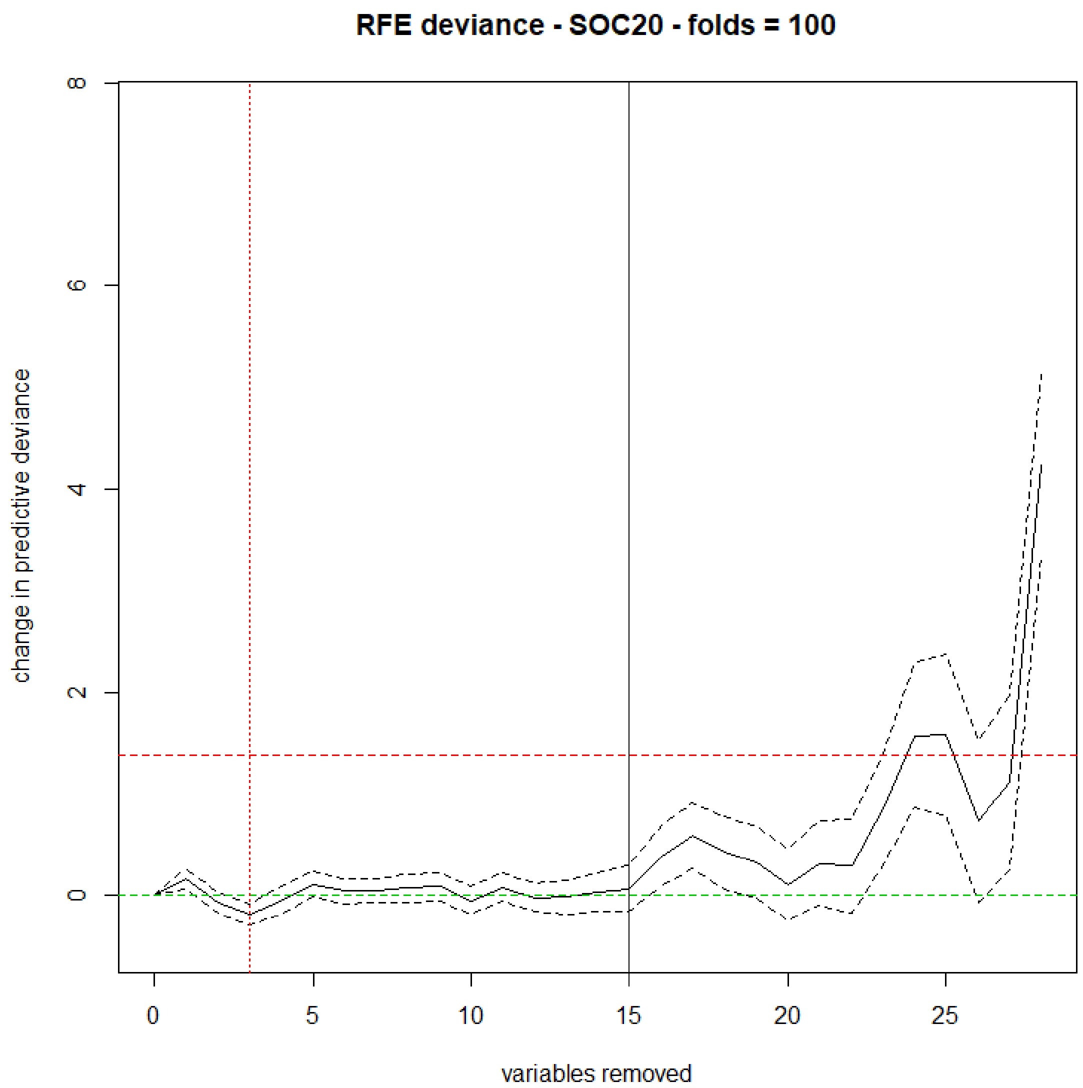


Figure S3: Changes in the predictive deviance of BRT models by backward removal of its predictors. The solid line indicates the mean change in predictive deviance, and the dotted line the standard error, calculated over the 10 folds of the cross-validation. Solid vertical line indicates the variables removed for the second fit. Dotted vertical line indicates minimum change in predictive deviance. Dotted horizontal line indicates mean change in predictive deviance.

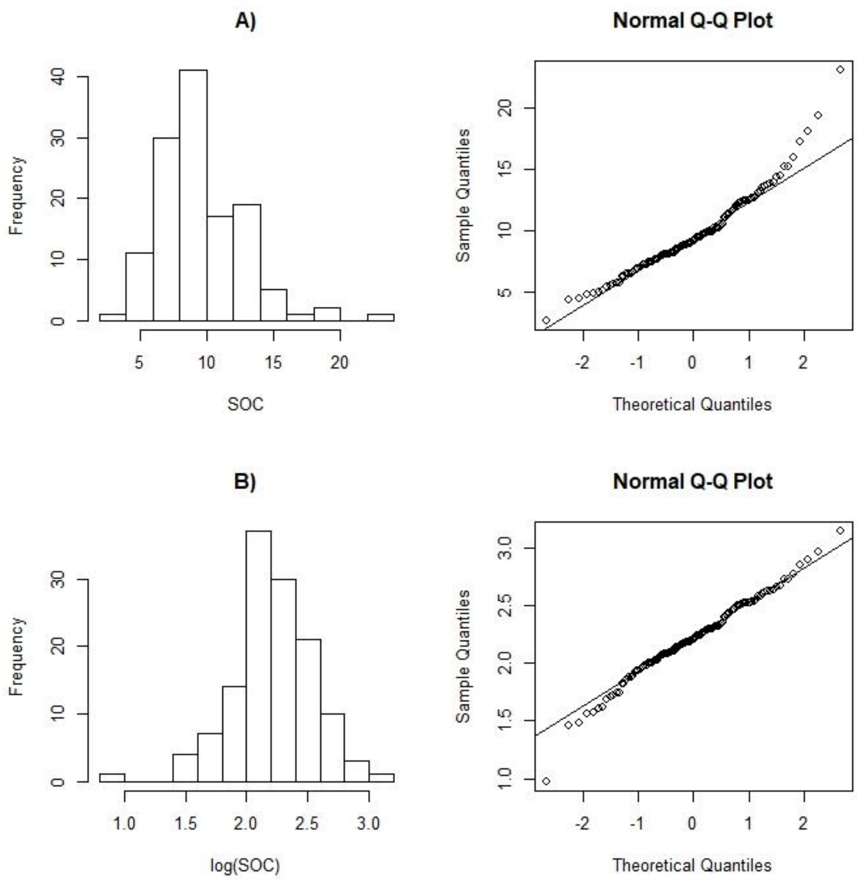


Figure S4: Histogram and normal Q-Q plot of A) SOC and B) log(SOC). Result of Shapiro Wilk W test result were W = 0.948; p-value < 0.001 and W = 0.99; p-value = 0.18 respectively. SOC: soil organic carbon.

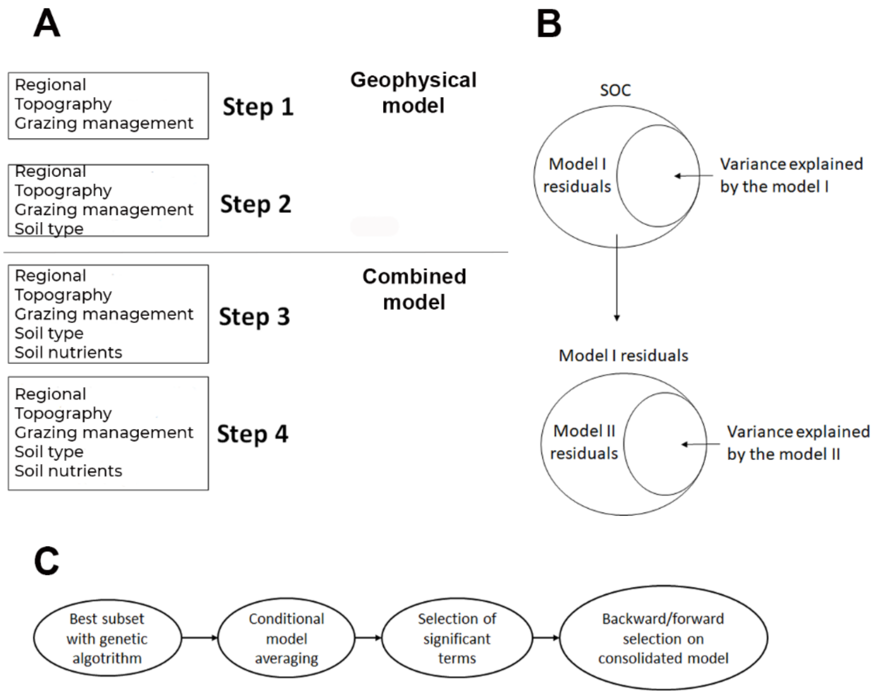


Figure S5: Linear modelling procedure. A) Variables introduced in each step. The first linear model (Geophysical model) is fitted until Step 2 and the second linear model (Complete Model) is fitted until Step 4. B) For selecting the candidate predictor terms on each step, residuals of the model obtained in the previous step are used as response variables in C. C) Procedure to select candidate terms on each step. First, genetic algorithm was used to obtain a set of best models. Second, these models were averaged and the significant terms were selected as candidates for backward forward selection in the main/consolidated model.

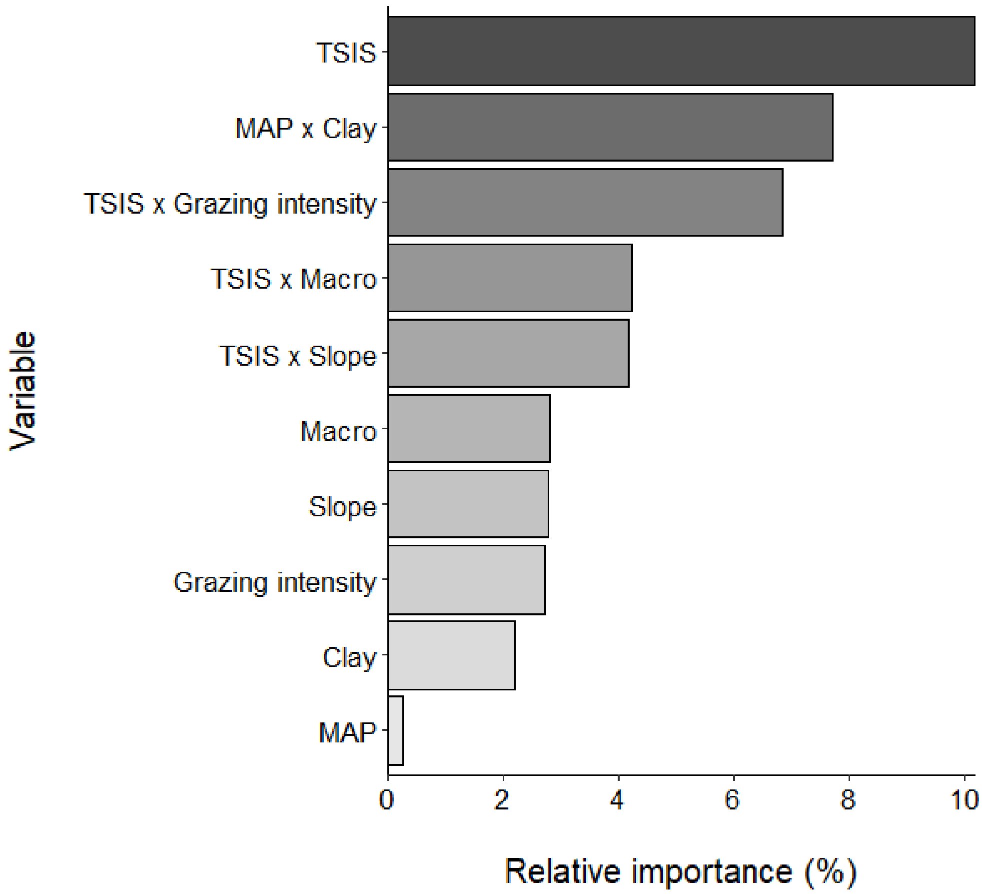


Figure S6: Relative contributions of variable groups in the linear model explaining Soil Organic Carbon, using regional, landscape and management predictors. MAP: mean annual precipitation; TSIS: mean summer temperature minus mean annual temperature; Slope: terrain slope; Exposed: Exposed position according to Macrotopography; Clay: clay content; Low and medium intensity: Low and medium intensity according to Grazing intensity.

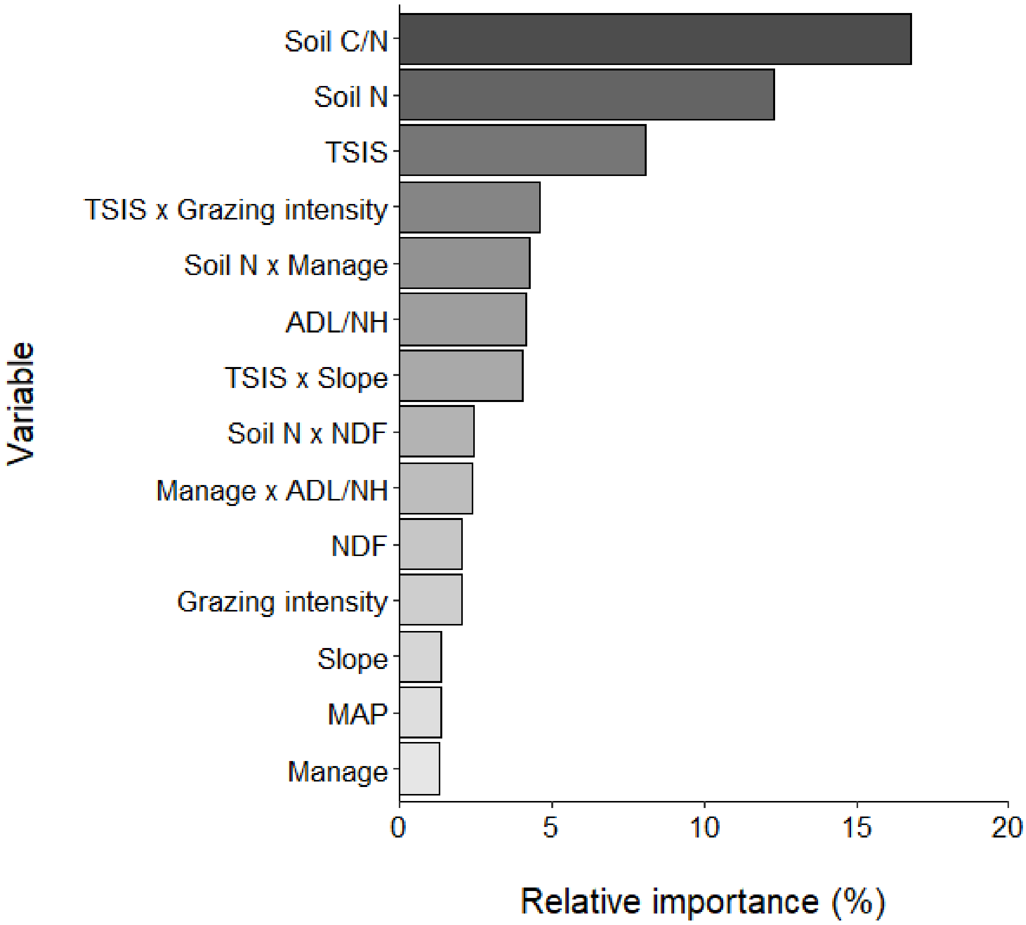


Figure S7: Relative contributions of variable groups in the linear model explaining Soil Organic Carbon using regional, landscape, management and biochemical predictors. MAP: mean annual precipitation; TSIS: mean summer temperature minus mean annual temperature; Slope: terrain slope; Cattle and Mixed: Cattle and mixed management according to grazing species; Low and medium intensity: Low and medium intensity according to Grazing intensity; Soil C/N: soil carbon to nitrogen ratio; soil N: soil nitrogen; NDF: neutro-detergent fibre; ADL/NH: acid-detergent lignin to nitrogen in the herbage ratio.

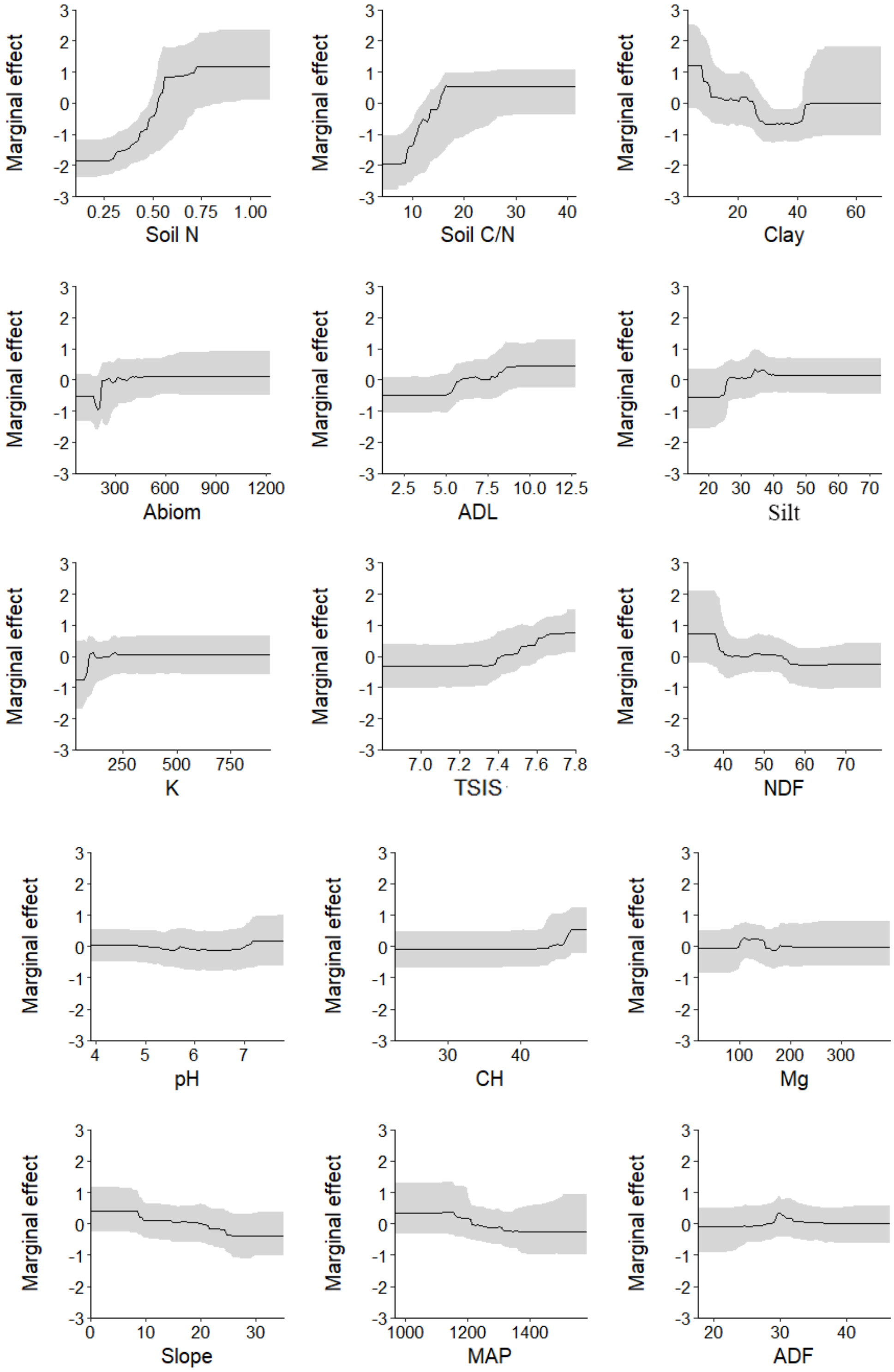


Figure S8: Partial dependence plots for the 15 selected predictors in the BRT model. Y axes are centred to have zero mean over data distribution. Values (solid lines) are predictions of the model across the predictor´s range maintaining the rest of the predictors at their average values. Grey areas around prediction lines indicate 95% bootstrap confidence intervals. Soil N: soil nitrogen; Soil C/N: soil carbon to nitrogen ratio, Clay: clay content; Abiom: aboveground biomass; ADL: acid-detergent lignin; Silt: silt content; K: soil potassium; TSIS: mean summer temperature minus mean annual temperature; NDF: neutro-detergent fibre; pH: soil pH; CH: carbon in the herbage; Mg: soil magnesium; Slope: terrain slope; MAP: mean annual precipitation; ADF: acid detergent fibre. See Table S1 for more details about variables.

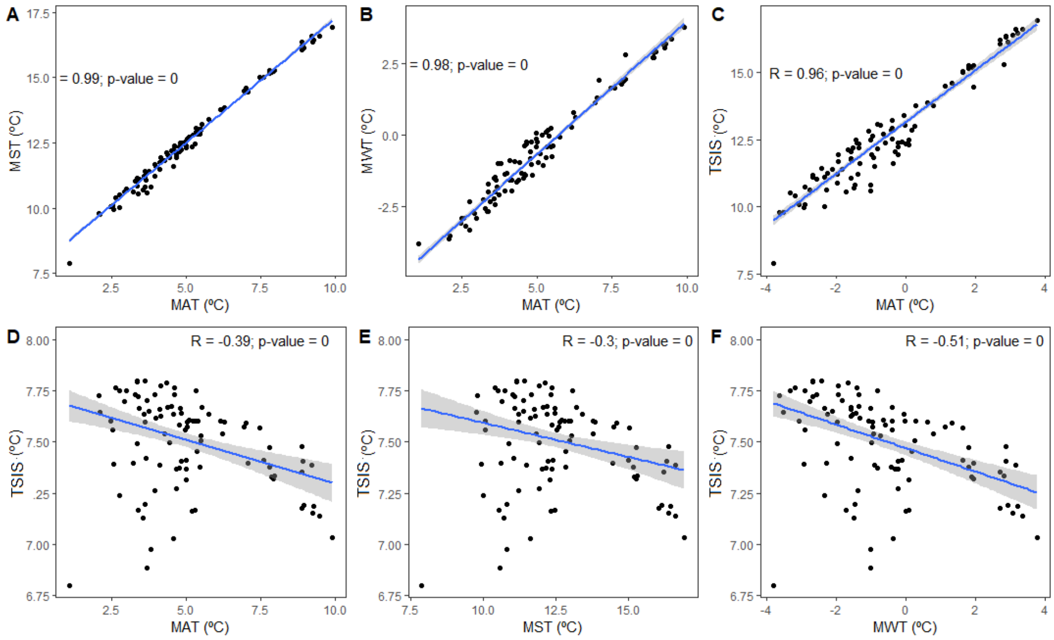


Figure S9: Pairwise Pearson´s correlations between climate variables. MST: mean summer temperature; MWT: mean winter temperature; MAT: mean annual temperature; TSIS: inter-annual seasonality measured as MST-MAT.