To Reviewer #1:

Thank you for your time and effort to review our manuscript. Each request or comment is repeated below in black, and our responses are in blue. Additional or altered text that has appeared in the manuscript is marked in red.

[General Comment] Using a feature selection approach and random forest model, the study conducted by Yang et al. identify key predictors of wood density, including satellite-based vegetation indexes, topographic variables, and soil sand content. The database used is both important and impressive. It offers a robust statistical perspective on wood density. However, the paper has some weaknesses that require partial rewriting. The methodology is insufficiently presented, with some factors (e.g., NDVI, topographic metrics) appearing in the results section without proper introduction or details about their origin and precision. Some elaborate statistics are not and should be introduced in the methods section. In the results section, the statistical significance of the results should be systematically presented. The discussion section is short and miss some important aspects, like the relevance and bias of the sampling which are barely addressed. The authors do not place their findings in a broader perspective and, in particular, do not clearly highlight what aligns with what is already known versus what is new. They may discuss the implications of their results on the understanding and modeling of the carbon cycle. Some parts of the paper (introduction and discussion, in particular) are not very well written. A careful reading and corrections are necessary. Some references should be added. Some are cited in the text but missing in the references list. In conclusion, the paper deserves to be published, but only after addressing these issues and being partly rewritten.

[Response] Thank you for your constructive and positive comments. We appreciate your efforts to improve our manuscript. We provided answers and clarified all the points that you raised, and revised our manuscript accordingly. Detailed responses can be found below.

[Detailed comments] Throughout the paper, the figures should appear after being called

[Response] Corrected.

[Introduction]:

Line 30: ...is a fundamental trait which describes the carbon...

Line 38: ...Kraft et ak. 2010). For example, Chao...

Line 52: ...and lower elevation (Sungpalee et al. 2009) and the soil water availability, which has a positive impact on wood density in dry biomes and a negative impact in wet biomes (Rocha et al. 2020).
Line 60: ‘a novel dataset of wood density measurements’
Line 68: ‘Compare the extent of inter-tree and intra-tree variation in wood density for the tree species or forest plots analysed’.
Line 74: ‘collected in Poland’
[Response] Corrected.

Line 31: a key indicator for various ecological and physiological processes, such as: you cannot say that hydraulic properties is an example of an ecological or physiological process.
[Response] The sentence is now written as: “It serves as a key indicator for various ecological and physiological processes, such as tree growth, mortality rates, and the vulnerability to hydraulic failure.” (Page: 2; Lines: 40-41)

Line 36: has the potential to have a positive impact: consider : may impact positively the ...
[Response] Thanks. The sentence has been revised as you suggested.

Line 39: Chao et al. ‘s paper shows that tree mortality rates are related to wood density but, to my understanding, not due to wood density.
[Response] Thank you for pointing it out. The sentence has been revised as follow: “Chao et al. (2008) found contrasting tree mortality rates in eastern and western Amazon forests, which were related to variations in wood density, with lower density associated with higher mortality rates.” (Page: 2; Lines: 47-49)

Line 50: ‘It is important to note that these influencing factors can vary from one region to another’. It is so obvious that I do not think it is worth mentioning...
[Response] The sentence has been deleted.

Line 55: I would rather say : ‘fast growth due to low competition for light and space’ . Gryc (and not Gyrc) et al 2011 is missing from the list of references. I guess it is : DOI: 10.5194/egusphere-2023-2691 You should make clear that this conclusion was reached for coniferous trees.

[Response] Thank you for your comment and suggestion. The sentence has been revised to: "For example, Gryc et al. (2011) found that in coniferous trees, thin cell walls resulting from fast growth due to lower competition for light and space are typically associated with low wood density. Conversely, thick cell walls, which result from slower growth, are related to high wood density." (Page: 2; Lines: 62-64). Additionally, the reference of Gryc et al. (2011) (Thanks for your correction) has been added to the reference list.

Line 57: How does the elements above suggest anything about tree development stages?

[Response] We have re-written this part as follows: "Additionally, the growth rate of individual trees can vary over their lifespan, leading to variations in wood density. Generally, young trees grow quickly while mature trees grow steadily (Bowman et al., 2012). Thus, the growth strategy between trees and the development stage of an individual tree's lifespan can play a role in shaping wood density gradients." (Page: 2; Lines: 65-68)

Line 66: northern or southern discs? How can a disk have an orientation?

[Response] We apologize for confusion caused by our previous wording. Each disc was divided into two parts, i.e., the northern and southern part. We have clarified this in the revised manuscript.

Line 71: What do you mean by 'wood density formulas'? 'only the older trees'

Line 72: 'In many previous works': references are needed.

[Response] Sorry for our overly ambitious sentences. We have rewritten this part as follows: "Many previous studies have assessed the relationship between environmental conditions and wood density using data from limited the mature forest plots (Baker et al 2004; Dias et al 2018; Phillips et al 2019)." (Page: 3; Lines: 126-127)

Line 73: What do you mean: 'The relationships between 'the changes in environmental conditions and variations in wood density' or 'the relationship between environmental conditions and wood density'?

[Response] We have revised the sentence as "the relationship between environmental conditions and wood density." (Page: 3; Line: 126)

[Methods]

Line 79: ‘wood density samples...’: I suggest: ‘the density measurements of more than 48000 samples, from 2920 trees from 391 forest plots in Poland, carried out....’

Line 90: ‘analysis belong to three Plant functional types (PFTs) categories..’

Line 94: ‘Regarding the age of the trees, all trees can be classified into nine age classes: I suggest: ‘The tree population were divided arbitrarily into nine age clases’

Line 96 and fwd: please replace ‘can be classified’ by ‘were classified’.

Line 97 and 98 : The height and DBH of trees are not ‘defined as ’. you may consider ‘divided into the following categories’

[Response] Thanks. These sentences have been revised as you suggested.

Line 85: fertile/ low fertile soil: On what criteria is this separation based?

Line 87: What is the typical elevation of low and high plots?

[Response] Thank you for your comment. We have clarified this in the revised manuscript: “Low plots are typically located in lowlands, ranging from 0 to 300 m asl. In this study, field plots were selected from elevations not exceeding 100 m asl. In contrast, high plots begin at 300 m asl, encompassing both uplands and mountainous areas. In Polish conditions, these high plots extend up to 1600 m asl.” (Page: 3; Lines: 142-145)

Line87:’...add more complexity’: I would not say it is more complex. These are just additional criteria worth exploring.

[Response] The sentence has been revised as follows: “These distinctions within the species-specific plots allow for a more detailed exploration of the impacts of environmental conditions on wood density.” (Page: 4; Lines: 193-194)

Line 95: ‘bottom of a trunk’: What height?

[Response] Thank you for your question. Unfortunately, we cannot provide the exact height from the base of the trunk where the rings were counted to determine the tree’s age. According to the project protocol, each tree was divided into three equal parts. Discs were cut from the middle section of these three logs for wood density measurements. This approach ensures consistency in the biological and physico-mechanical properties of wood samples, regardless of the tree's height or age. The final disc for counting annual rings was cut from the bottom section of the lower log. The height at which this disc was taken depended on the tree's diameter and safety considerations during felling with a chainsaw.

Figure 1: What about: ‘a) Distribution of samples (solid colour bars), trees (light colour bars), and forest plots (transparent bars) utilized for density measurements across the eight species; b) location of the 391 forest plots in Poland (symbols as in panel a).

[Response] We has revised the caption of Figure 1 as you suggested.

Line 103: The sampling procedure is unclear. You mention taking samples with an increment borer and then refer to discs. Do you mean that you took cores from the discs? If so, please make it clear.

[Response] Yes, thanks for pointing it out. We now make it clear in the revised manuscript.

Line 106: To my understanding: ‘bottom”, ‘middle’ and “top” can correspond to very different height depending on the total height of the tree. Therefore are these categories meaningful?

[Response] The reviewer is correct that the ‘bottom’, ‘middle’ and ‘top’ categories correspond to different heights for different trees, making direct comparison across trees less meaningful. However, our approach was to compare the wood density between ‘bottom’, ‘middle’ and ‘top’ categories for each individual tree (as shown in Figure 5). We believe these intra-tree comparisons are valid, and using the terms ‘bottom’, ‘middle’ and ‘top’ provides readers with an intuitive understanding of the vertical profile of wood density within each tree.

Line 107: What do you mean here by sampling (the sampling was conducted from the inner ...’. Do you mean the density measurements? How was the wood density determined? The proocedure used should be described in the method section. The data should be made available in a repository, along with supplementary materials that include relevant statistics (e.g., mean diameter, height, elevation for each species) and detailed information about vegetation indexes, water content, etc.

[Response] Thank you for your comments.

Firstly, we have added the detail of procedures for collecting wood density measurements in the revised manuscript, as follows: “Each disc was cut from north to south to obtain a strip of wood. The samples were divided and numbered into two rays: north and south, starting from the core to the peripheral part. This method allowed for the estimation of variation in the radial density of wood. The number of samples obtained for each disc varied depending on the width of the disc, but each disc typically yielded more than 10 samples along these radial directions. Standardized wood density samples, measuring 2×2×3 cm, were cut from the strips, which were dried in a dryer at temperature of 103 ± 2 ºC to an absolutely dry state. After the samples

cooled down in the desiccator, the linear dimensions of the samples were measured using an electronic caliper, and their weight was measured on a laboratory scale. The stereometric density was then calculated from the classical mass/volume formula." (Page: 5; Lines: 232-241)

Secondly, regarding making the data publicly available, we respectfully request a delay in its release. Our colleagues in Poland have invested significant effort in building this complex database and would like to retain priority for its use. There are ongoing research projects based on this database, and once those are published, the data will be made available in a public repository.

Thirdly, while we cannot make the full wood density database public, we have added a new table in the supplementary materials that includes the mean, minimum, maximum, 25th, and 75th percentiles of wood density for each species.

Table S1. The mean, minimum, maximum, and the 25th and 75th percentiles of wood density for eight tree species.

WD (g cm^{-3})	Species							
	Pinus sylvestris	*Picea abies*	*Abies alba*	*Larix decidua*	*Quercus robur*	*Fagus sylvatica*	*Betula pendula*	*Alnus glutinosa*
Min	0.31	0.29	0.30	0.32	0.45	0.57	0.45	0.38
Q_{25}	0.41	0.38	0.38	0.46	0.64	0.66	0.55	0.47
Mean	0.44	0.41	0.40	0.50	0.67	0.68	0.57	0.49
Q_{75}	0.47	0.44	0.43	0.54	0.70	0.70	0.60	0.52
Max	0.58	0.56	0.63	0.62	0.80	0.78	0.68	0.62

Line 115-fwd: The trees have varying diameters, and some species are more represented than others (e.g., Pinus sylvestris versus Alnus glutinosa). Therefore, is the average wood density truly representative of the population? If not, does it matter? Could this be corrected by using a weighted mean? These points should be presented and discussed.

[Response] Good point. Thank you for this very constructive suggestion.

As the reviewer mentioned, we calculated the average of sample-level wood density to represent the tree-level wood density. Firstly, to determine if the average value is a good representation of population, we compared the mean and median values within each individual trees. As shown in Figure R2a, the difference between mean and median values are minor (slope: 1.02, R^2 = 0.99). This suggests that the wood density samples within individual tree tend to follow a normal distribution, and the tree-level wood density can be effectively represented by the average of the sample-level values. Secondly, we examined whether the intra-tree variability of wood density differs among tree species. Figure R2b shows that, despite varying tree diameters, the magnitude of intra-tree variability is similar across tree species, with no significant differences observed.

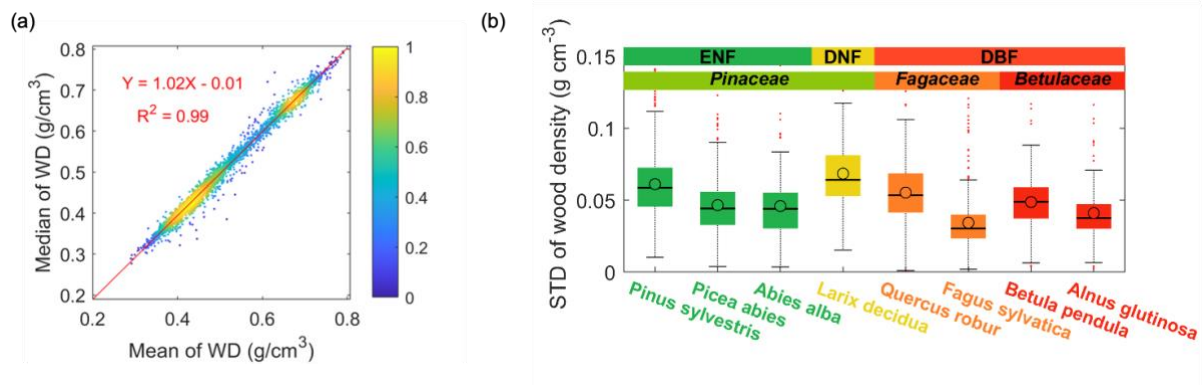

Figure R1. (a) Comparison of the mean and median values of wood density within the individual trees. (b) The standard deviation (STD) of the sample-level wood density within the individual trees. Boxplots indicate the median, mean, minimum, maximum, and the 25th and 75th percentiles of wood density for eight tree species.

We also added the following text: "Tree-level wood density was calculated by the average of all samples within each individual trees. This method is used because there was no significant difference between the mean and median values of the samples (Figure S1a), indicating that wood density within an individual tree typically follows a normal distribution. Furthermore, the magnitude of intra-tree variability is consistent across eight tree species (Figure S1b)." (Page: 6; Lines: 274-277)

Line 123: 'Less than 100 or 500 m, grid sizes of 0.05 or 0.1°'. The results indicate no consistent differences when comparing distances or grid sizes. Therefore, you may explain that both

criteria (100 and 500; 0.05 and 0.1) were tested and yielded similar results. For simplicity, you may present the results for only one distance and one grid size.

[Response] We has revised the related texts in the Method and Results as you suggested.

Line 133: Where is Table 1?? Where do the covariates come from? How reliable are they?

[Response] We apologize for the oversight. Table 1 has been added, and the sources of covariates are now listed. We have verified that none of these data have reported data quality issues in Central Europe.

Table 1. The predictor covariates used in the random forest model for inter-tree variations in wood density. The original 8-daily values of NDVI and NDWI were aggregated into a median (P50) and a standard deviation (STD) for the entire period.

Variables	Description	Unit	Original resolution	Source
SNDPPT	Weight percentage of the sand particles (0.05–2 mm)	%	250 m	SoilGrids database
NDWI	8-daily Enhanced Vegetation Index (EVI) generated using the gridded daily surface reflectance product.	1	0.083º	MOD13A2
NDVI	8-daily Normalized Difference Vegetation Index (NDVI) generated using the gridded daily surface reflectance product.	1		
Geomorphons	a pattern recognition approach to classification and mapping of landforms from digital elevation models (DEMs)	-	30m	Jasiewicz & Stepinski (2013)

[Results]

Line 141-142: 'For the analysis... 2920 trees": Already said in the method section.

[Response] The sentence has been deleted.

Line 145: '... is lower compared to than the density of...'

Figure 2: 'Eight tree species... "unnecessary as it is stated in the text.

[Response] Corrected.

Line 148: '...slightly lower' : Is it statistically significant? ...'significantly lower': You should systematically test the significance of the differences between means.

[Response] We have added the significance of differences in between tree families.

Line 163: 'the large difference in wood density observed among species tend to diminish when considering the averages within geographical locations'. Or does it show that location is just not a discriminating factor?

Line 165: 'attributed to the relatively even distribution of the eight primary species': I do not interpret it this way, as the species are evenly distributed. If the spatial distribution were uneven, there could be a potential confusion between species and location, but that is not the case here. There is likely no spatial bias, and the small percentage of the variance explained indicates that location is not a predictor for wood density. Aren't the fertility and elevation criteria more or less included into the location criteria (grid cell)? What is the scale for the difference ? What does the black line represent?

[Response] Firstly, we agree with the reviewer that location is not a discriminating factor, and we have clarified this in the revised manuscript. "The reason is that the location does not differentiate the tree species in Poland; as different tree species with varying wood density are distributed across similar locations in Poland, as shown in Figure 1b." (Page: 8; Lines: 332-333)

Secondly, we agree with the reviewer's insightful point. It is indeed correct that soil, elevation, and vegetation properties can account for the variation not explained by location. This is why we found that the random forest model, even without species information, still achieved a high R^2 value (0.91) when using only these vegetation, soil, and topography-related covariates. We have added the sentence in the revised manuscript: "This variance in wood density, unexplained by location, could be related to local environmental conditions such as vegetation characteristics, soil properties and topography." (Page: 8; Lines: 333-335)

Additionally, the black line is used to bracket the labels related to location.

Line 174-fwd: How are height, DBH and age related to one another? They are probably not independent. How do you separate the effect on wood density of height or DBH from age?

[Response] We agree with the reviewer that height, DBH and age are inter-related and not independent variables. Generally, height and DBH increase with age, especially in the younger trees, while the growing rates of mature trees are relatively low. Since the combined contribution of tree age, height, and DBH to wood density variance are very low (less than 6%), we don't believe it is necessary to separate their individual contribution. Nevertheless, in the revised manuscript, we clarify that the collinearity among height, DBH and age may result in an overestimation of the contribution of each individual factor, as follow: "Note that height, DBH and age are interrelated and not independent variables, and their collinearity may result in an overestimation of the contribution of each individual factor." (Page: 9; Lines: 360-3631)

Line 177: Taking into account the uncertainties, can you say that there is a difference of behaviour between broadleaf and needleleaf base on the age. In addition, the lack of VIII and IX class for needleleaf trees, makes the statement irrelevant.

[Response] Thanks for pointing it out. We have rewritten the related texts as follows: "regarding the impacts of tree ages (Figure 3a), for broadleaf trees, wood density tends to increase with tree age up to approximately 140 years (class VII), after which it stabilizes. For needleleaf trees, wood density also exhibits an increase in wood density with age up to 140 years (class VII). Since there are no observation of wood density from older needleleaf trees, it remains unclear whether wood density would continue increase or stabilized beyond 140 years." (Page: 9; Lines: 362-366)

Line 179: 'for broadleaf the impact of height is more pronounced for taller trees': what does it mean? The graph indicates that taller trees tend to have a higher average wood density, but the standard deviation is substantial. The significance of the difference between the means, such as between class III and VII, should be evaluated. What are the relevant statistics?

[Response] Our previous statement was mainly based on the mean values for each classes, and we understand the reviewer concerns on the no difference in standard deviation. We have rewritten the related texts as follows: "both needleleaf and broadleaf trees show an increase in mean wood density with height and DBH classes, especially for tall broadleaf trees (height class $\geq$ 4, i.e., tree height $\geq$ 20m). However, the variance in wood density within these height and DBH classes are large, resulting in no statistically significant difference in wood density distributions among classes (Figure 3b-c)." (Page: 9; Lines: 366-370)

Line 190's: The description of the random forest approach, of the factors tested (including where they come from, and their accuracy) should be provided in the method section.

[Response] These points related to the random forest model have been added into the method section.

Line 200: link between NDVI and NDWI values and carbon and water contents respectively: references are needed.

[Response] The references have been added.

Figure 4: What are the y-axis in b. Is the shap values for density in %?

[Response] The y-axis in Figure 4b represents the absolute values of wood density, measured in g cm^{-3}. This has been clarified in the caption of Figure 4.

Line 210: With a $p<0.1$ the evidence is weak.

[Response] We re-do the analysis using a threshold of $p < 0.05$ for significance, and the updated results are generally consistent with the previous findings. We have also revised the corresponding text accordingly.

The corresponding texts have been revised as follow: "To examine the differences in wood density among the bottom, middle and top parts of trees, we conducted an ANOVA analysis. The results indicate that about 35% of all the trees exhibit significant vertical variations (p-value < 0.05) in wood density (Figure 5a). The percentage of trees with significant vertical variations (p-value < 0.05) varied among species, with Alnus glutinosa and Pinus sylvestris species having the highest percentage (around 50%), while Betula pendula species had the lowest percentage of 8%." (Page: 11; Lines: 408-412)

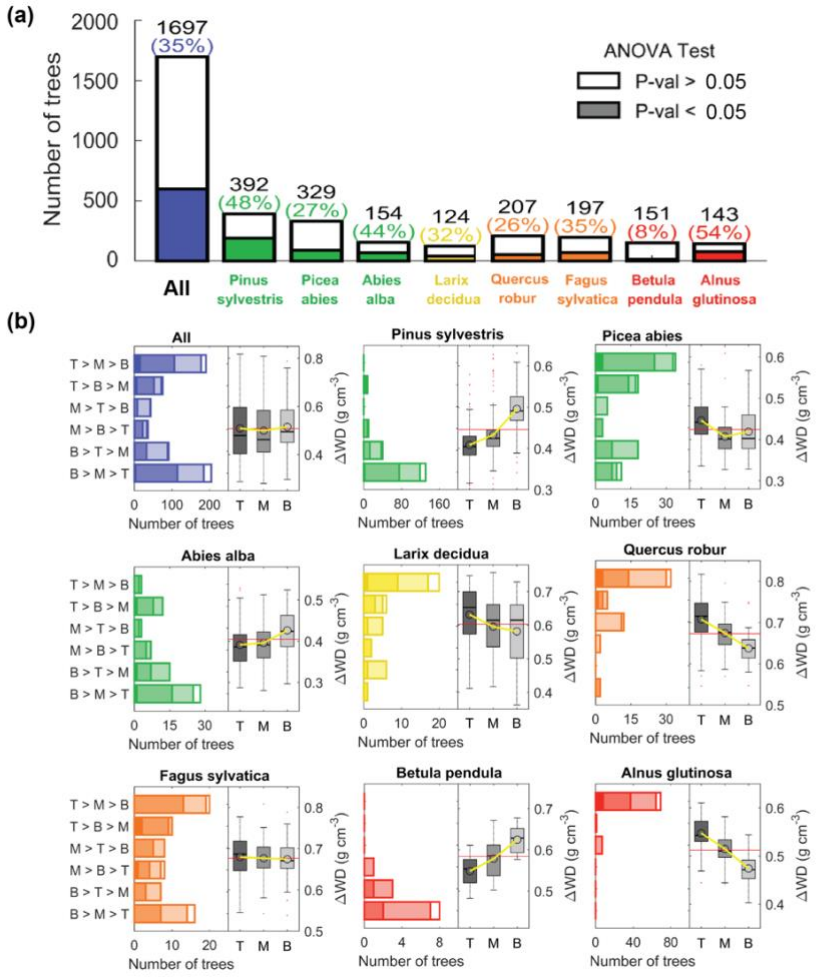

Figure 5. Comparison of wood density of the discs from the top (T), middle (M) and bottom (B) part of tree. (a) Fractions (The number) of trees with significant differences (P-val < 0.05) in wood density among three levels (top, middle, and bottom) discs via ANOVA analysis. (b) The relation among the top, middle and bottom discs, in comparison to the mean values of wood density, via the multiple comparison test. The color depth in the left-hand panels presents the pairs of two discs with significant difference. The darkest colors indicate that any two of three discs have significant difference (P-val < 0.05), while the lightest colors indicate that none of two discs have significant difference. The right-hand panels show the boxplots of wood density of top, middle, bottom discs of trees, respectively. The yellow curves indicate the changes in the averaged wood density among top, middle and bottom discs. And the red lines present the averages of all wood density records.

Line 221: The multiple comparison test should be presented in the method section

[Response] Thanks for pointing it out. We have moved the text related the multiple comparison test into the method section.

Line 230 ‘The magnitude of radial variations in wood density is typically larger at the bottom disc compared to the middle and top discs ‘: The older part of the tree is located in the bottom disk. How might this affect the results?

‘However, ….’ Should come before this sentence, as it refers to the outer-inner gradients in the various species.

[Response] Thank you so much for your comments. We think this pointing raised by the reviewer is a possible explanation for the large magnitude of radial variations at the bottom discs. We have added the sentence: “This greater variation at the bottom discs could be associated with the presence of older and earliest growing parts, which are only located in the bottom sections of the trees.” (Page: 11; Lines: 429-430). Additionally, the word of ‘However’ has been deleted.

[Discussion]

Line 259: I propose: ‘Large-scale variations in wood density have been reported to correlate with climatic variability. However, in this study, tree-level variations in wood density are more strongly linked to vegetation indices than to climatic variables’.

[Response] Thanks. These sentences have been revised as you suggested.

Line 261: variability rather than heterogeneity?

Line 268: ‘exhibit greater complexity’: Than what?

[Response] Corrected.

Line 265: I propose: ‘Therefore, the utilization of satellite-based NDVI and NDWI can effectively predict a substantial portion of the variations in wood density.’

[Response] Thanks. These sentences have been revised as you suggested.

Line 269-fwd: Your point is not very clear. What implicit relationship do you draw between NDVI and leaf size?

[Response] We have rewritten this part as follows: “However, our findings reveal an opposite relationship between satellite-based vegetation index (NDVI, which usually indicate canopy greenness and cover) and wood density when controlling for factors such as vegetation water content, landform types, and soil texture. Specifically, trees with high NDVI (indicating large canopy coverage) exhibit high wood density.” (Page: 14; Lines: 488-490)

Line 289: ‘Specifically, valleys tend to have less fertile soils. In addition, trees on ridges may receive more sunlight compared to those in valleys.’

[Response] Corrected.

Line 295: For instance, pine trees and fir trees (Alnus glutinosa and Pinus sylvestris species), Please give latin name consistent with vernacular ones!

Line 297: ‘while alder, spruce, oak, and larch trees (Picea abies, Larix decidua, Quercus robur, and Alnus glutinosa species) ‘: same order for vernacular and latin name, please.

[Response] Thanks for pointing it out. We have corrected all of vernacular names.

Line 304: ‘The species is characterized by a significant presence of large vessels in the outer zones, closer to the bark, leading to a lower wood density’. Where is this from? Not in Woodcock and Shier, 2002.’These authors rather link radial increase to early successional status (and change of carbon allocation when trees reach the canopy and are subjected to more wind).

[Response] We have added the reference.

Line 311: Please correct: ‘Across all eight species analyzed, the results consistently demonstrate that the variations within individual trees exhibit larger magnitudes compared to the inter-tree variations ~~between trees~~.’

[Response] Corrected.