

Review for Yang et al. 2023's paper

‘Similar importance of inter-tree and intra-tree variations in wood density observations in Central Europe’

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.

Detailed comments:

Throughout the paper, the figures should appear after being called.

Introduction:

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

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.

Line 36: has the potential to have a positive impact: consider : may impact positively the ...

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

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.

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...

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 55: I would rather say : ‘fast growth due to low competition for light and space’ . Gyr (and not Gyr) 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.

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

Line 60: ‘a novel dataset of wood density measurements’

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

Line 68: ‘Compare the extent of inter-tree and intra-tree variation in wood density for the tree species or forest plots analysed’.

Line 71: What do you mean by ‘wood density formulas’?

‘only the older trees’

Line 72: ‘In many previous works’: references are needed.

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’?

Line 74: ‘collected in Poland’

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

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

Line 87: ‘...add more complexity’: I would not say it is more complex. These are just additional criteria worth exploring.

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 classes’

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

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’

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).

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.

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?

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 procedure 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.

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.

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.

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

Results

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

Line 145: ‘... is lower ~~compared to~~ than the density of...’

Line 148: ‘...slightly lower’ : Is it statistically significant?

...’significantly lower’: You should systematically test the significance of the differences between means.

Figure 2: ‘Eight tree species... “unnecessary as it is stated in the text.

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?

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?

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.

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?

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.

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

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

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

Line 221: The multiple comparison test should be presented in 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.

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’.

Line 261: variability rather than heterogeneity?

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.’

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

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

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

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.

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).

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.~~’