

Dear Editor Martin De Kauwe and referees,

We have thoroughly addressed all points raised by the 2 referees and prepared a new manuscript version with increased readability and a clearer storyline.

Please find below our new specific responses to the comments of referee 1 & 2. Original referee comments are marked as black text. Our answers to each comment are marked as blue text.

Please note in addition to the requested changes we also propose a new slightly different manuscript title which fits better to the storyline.

With best regards and on behalf of all co-authors,

Boris Sakschewski

## Response to the comments of referee 1 on Sakschewski et al. 2020: Variable tree rooting strategies improve tropical productivity and evapotranspiration in a dynamic global vegetation model

In this manuscript, the authors extended the LPJmL4.0 dynamic vegetation model to simulate variable rooting depth. Comparisons between different model versions for tropical South America showed that the novel approach improves most of the bench-marks used in the study. Overall, the study is well-conducted and solid. I agree with the authors that including variable rooting depth is important in DGVMs and the results and the presented model approach are therefore relevant for the modeling community. Previous models typically assumed PFT specific fixed rooting depth and/or a fixed soil depth globally. My main point is, however, that the manuscript reads like a collection of many interesting results while I was missing more specific aims or questions regarding for example implications of root diversity for the ecology or biogeography of the simulated vegetation types. Currently, the aims of the study are to 'describe an approach' and to 'evaluate its effect on' various model variables.

Dear referee,

Thank you for your positive evaluation and many constructive comments on our manuscript. We agree that the paper benefits from focusing on a specific research question which creates a stronger storyline. As we will now highlighted in the introduction in lines 101-103 we now focus on "the hypothesis that varying tree rooting depth is key to explain major patterns of evapotranspiration, productivity and the geographical distribution of tropical evergreen forests in South America" and structure the manuscript accordingly. In connection to comments of referee 2 we will shorten the manuscript and move parts of the results to the supplement. This also helps to create a stronger and clearer storyline.

L 25: PFT-specific instead of biome-scale?

Thank you for this comment. In the abstract we wanted to avoid specific terms like "PFT" to keep it easy to understand for readers which are not familiar with those terms, but still convey the main message. In the new version of the abstract there is no introduction of PFTs anymore.

L 29: realistically simulated

Thank you for this suggestion. We changed it accordingly, now in line 27.

L 41: delete quotation marks for evergreen and deciduous?

Thank you for this suggestion. We changed it accordingly, now in line 40.

L 47-48: Isn't it the other way round? Traits are aggregated to define PFTs and the fractional cover of different PFTs in simulations then defines the biome type?

Thank you for raising this point. We apologize for the sloppy usage of "biome" and "PFT". To clarify the text we now state "In general these models condense the diversity of such functional tree traits into so called plant functional types (PFTs), which represent average tree growing strategies on scales as large as biomes." in line 45-47.

L 54: reword 'different attempts were carried out', maybe 'different approaches were presented'?

Thank you for this suggestion. We changed it accordingly, now in line 52.

L 54: Schymanski et al. also developed root models (e.g. [www.hydrol-earth-syst-sci.net/12/913/2008/](http://www.hydrol-earth-syst-sci.net/12/913/2008/)). Although these models were not developed explicitly in the DGVM context, they might be relevant for the introduction or the discussion.

Thank you for this suggestion. We now refer to Schymanski et al. (2008) in the introduction in line 84 to acknowledge that root optimization models do exist, but the general knowledge base on phenotypic plasticity of plant traits remains too little to be applied in a DGVM. E.g. it remains unclear how fast plasticity can occur and what the true drivers of plasticity are.

L 55-56: 'study ... searched for...' reword, I would think that a study can's search.

Thank you for this correction. We changed the sentence to "In a pioneering study more than 20 years ago, Kleidon and Heimann (1998) systematically searched for rooting strategies which yield highest net primary productivity ..." in line 52.

L 74: 'allocation strategy' instead of 'direction'?

Thank you for this suggestion. We changed it accordingly, now in line 73.

L 97: in competition with other rooting strategies/plants/PFTs?

Thank you for this question. We changed the respective sentence to "all tree rooting strategies of this spectrum grow in competition..." in line 96-97 as it is meant to contrast this study with some of the former approaches. A more detailed sentence would lead to a full description of the competition scheme which can be found in the methods in Sect. 2.2.6.

L 100: this suggests to me that rooting depth is related to or a function of tree height.

Thank you for this comment. Yes as we clarified in the methods we used this approach in Sect. 2.2.5.

L 105-108: I think that such a general overview is not necessary and could be removed (also elsewhere in the methods section). Further, it only refers to some selected section and it is incomplete.

We agree with the referee and deleted this part.

L 115: simulates instead of employs?

Thank you for this remark. We changed it accordingly, now in line 114.

L 116: Should it be 'bioenergy functional type'?

Thank you for pointing out this mistake. We inserted "functional" now in line 114.

L 122: 'how different tree rooting strategies (implemented in the new scheme) compete'?

Thank you for this suggestion. We implemented it, now in line 124-125.

L 131: Does that mean that soil texture is similar for all layers in the soil column?

Thank you for this question. Yes. In LPJmL4.0 each grid cell has one soil texture information for its 3 m soil column only. We followed this approach for our larger soil columns as well. In fact a high resolution soil texture information in 3 dimensions for the study region is so far not available, or only partly available. We wanted to keep things simple and comprehensible also with regard to comparing results of LPJmL4.0-VR to LPJmL4.0. We now better clarify the approach in line 134-135 stating "Equal to LPJmL4.0 (Schaphoff et al., 2018), we use a grid cell specific soil texture information which is applied to the whole soil column".

L 142: PFT scale instead of biome scale?

Thank you for this suggestion. We changed it as suggested, now line 161.

L 144 and elsewhere: is it really the depth which is reached by 95% of the roots or does it rather mean that 95% of the fine root biomass is between the surface and D95\_max? Wouldn't 'reached by 95%' mean that 95% are deeper than D95\_max?

Thank you very much for pointing out that out. We changed the sentence to "...we here calculate the depth at which the cumulated fine root biomass from the soil surface downwards makes up 95% of all fine root biomass", now in line 162-163.

L 146: that's a question related to LPJmL4.0 and not to the VR version: why were different values for evergreen and deciduous selected if the resulting root profiles are essentially identical. Would model results differ for similar values?

Thank you for this question. The parameter values are average values for those 2 PFTs found in Jackson et al. (1996). At model development of LPJmL those values were the best available. We now specifically mention where those parameter values came from in line 165-166.

For now it is clear that increasing the difference between the beta values of the 2 PFTs would most probably enhance model performance. Even with only 2 PFTs present. Here the best way to go would probably be to increase the beta value of the evergreen PFT to reach a D95\_max of maybe 2-4 m, which would buffer against some dry season signal. Regarding the second question: If regarding beta values only, the answer is no, there would be no difference between the evergreen and deciduous PFT. Fortunately, they also differ in other functional traits such as specific leaf area (SLA) or leaf longevity (LL), determining their phenology strategy and therefore their performance under different climate regimes.

L 149: why 18m and not 20m or 95% of 20m?

Thank you very much for this question. As we have chosen a maximum soil depth of 20m we wanted to avoid a significant accumulation of fine roots in the last soil layers. A D95\_max of 20m in a soil column of 20 m would mean that the additional 5% of fine roots are also distributed between 19-20m. In principle it was a choice between a round number of soil depth or a round number at the largest value of D95\_max. In the current model version of LPJmL4.0-VR a D95\_max of 20m is not missing in the study area as distributions of D95\_max always flatten towards the 18m bin in grid cells where deep roots dominate. Nevertheless, it is possible to substantially increase soil depth and create more sub-PFTs with larger D95\_max values. This step might even be necessary in future versions of the model or other study areas. We now explain why we chose the largest D95\_max at 18m in line 169-170 stating "We chose 18m as the largest D95\_max value in order to avoid that roots of the respective sub-PFT significantly exceed the maximum soil depth of 20m (see also 2.2.4 and Fig. 2 right panel)."

L 154: a new carbon allocation scheme?

Thank you for this suggestion. We changed it accordingly, now in line 175.

L 163: 'sapwood ... proportional' which constant was used to describe this relation? And shouldn't root and stem sapwood be identical to be able to transport the same amount of water?

Thank you for pointing this out. Our explanation was misleading. We do not use a constant here. The term proportional is falsely used. It is as your second question suggests and as it is written in the description of Fig. 2. We now clarify the approach in the methods section in line 184-187 with "Root sapwood cross-sectional area in the first soil layer is equal to stem sapwood cross-sectional area, as all water must be transported through the root sapwood within this soil layer. In the following soil layers downwards, root cross-sectional area decreases by the relative amount of fine roots in all soil layers above (Fig. 2)."

L 180: 'to derive a functional relation between tree height and rooting depth'?

Thank you for this suggestion. We changed it accordingly, now in line 200-203.

L 203: check brackets in Huang et al

Thank you for pointing this out. We changed the brackets, now in line 226.

L 233: I suggest to delete 'The new features... direction.' and say 'In LPJmL4.0-VR, PFTs can...'

Thank you for this suggestion. We changed it accordingly, now in line 258-259.

L 234: 'formally', I assume this should be 'formerly' or 'in previous model versions'.

Thank you very much. We changed it accordingly, now in line 258.

L 233-240: An increase of mortality rates from 3% to 7% is quite substantial and more than twice as high. In addition, it is stated that observed rates do not exceed 6% which means that 7% is not in the real world boundaries as stated in L 237 but it overestimates this rate by >16%.

Thank you very much for raising this point. We agree that claiming 7% is in the boundary of 6% is overstated. The point we wanted to make was mixed up and lost by citing the study mentioning a observed maximum of 6% background mortality rate in 167 Amazon forest plots (Johnson et al., 2016) and comparing it to the maximum (not actual background) mortality of LPJmL4.0-VR. In fact observed background mortality rate and conceivable maximum mortality rate are not the same thing. The growth efficiency related mortality which is used in many DGVMs needs a maximum mortality rate, because otherwise the simulated mortality rate would rise to 100% when total respiration exceeds productivity. In the real world though plants would optimize and reallocate carbon pools under those circumstances as far as they can. Nevertheless, a mortality rate of 100% in the real world is conceivable as well. The maximum mortality rate of a forest under a hypothetical super extreme scenario (no rain for 12 month) could of course be 100%.

Accordingly, in a hypothetical model which truly captures all important mechanisms of tree survival and mortality, the maximum mortality rate could be set close to 100% as we tried to explain formerly in line 239-240. In other words this constant could eventually be obsolete. The road to such a model will be long and might never be ending, but until then, a maximum mortality rate at any value will always be necessary to achieve biomass values in acceptable ranges. We now deleted the reference of observed mortality rates of 6% in the manuscript as this number is not comparable to the maximum mortality rate set in the model. Moreover, we will describe the true purpose of a maximum mortality rate and why we changed it (as described in this answer) now in Sect. 2.2.7.

L 239: 'We regard ...' I don't understand this sentence. And what does 'right direction' refer to? I suggest to reword.

We deleted this sentence in accordance with our response above regarding "L 233-240".

L 243: Why were four different climate inputs used? Most simulation results are shown only for CRU anyway. Which climate variables were taken from the different datasets?

Thank you very much for those questions. We wanted to show that our results remain robust using various climate inputs and decided (also with regard to the amount of figures and respective manuscript text) to show this with regional rooting depth (Fig. 6) and regional evapotranspiration (corridors in Fig. 11). In connection to referee 2 requesting a substantial reduction of the manuscript size and amount of figures, it seems impracticable to produce all figures for all climate inputs. We also argue that, Fig. 6 & 11 show some proof that results for different climate inputs are very similar. There are local differences, but the big picture remains the same. As we do not want to go into detail regarding local scale differences we hope the current amount of results is enough.

We added 2 sentences to clarify our point in the discussion in line 495-498.

We now better clarify what climate variables were taken from the different data sets in Sect. 2.3.1 in line 265-266.

L 267: Reword to 'caused by the the presence or absence of variable rooting strategies'?

Thank you very much for this suggestion. We changed it accordingly, now in line 289-291.

L 269: Given that model spin-up was conducted for different time periods for different climate input datasets and at 278ppm. Is there a jump in CO<sub>2</sub> in the transition between spin-up and transient phase?

Thank you very much for this question. No there is a smooth transition between spin-ups and transient simulations regarding atmospheric CO<sub>2</sub> content. Before the year 1840 a constant value of 278ppm is used, while after this year values are rising according to a LPJmL protocol first introduced in Sitch et al. (2003) and eventually following the Mona Loa record of Tans and Keeling (2015). We now clarify this approach in the methods section in line 292-302.

L 280: I understood that replicate simulations were not conducted. I was wondering how robust or deterministic the selection of rooting strategies is? Would you expect substantial differences in the results when conducting replicate model runs?

Thank you very much for this question. Given the vast amount of different model test runs during model development, we can assure that results are very robust. We admit we do not fully proof this in this study. However, the 4 very similar rooting depths maps (formerly Fig. 6 now Fig. 5) provide some proof.

Since there are no real stochastic processes in LPJmL which could lead to a path dependency of vegetation dynamics, this model behavior was expected. Therefore, it is also a standard procedure to not conduct any replicate simulations when using LPJmL. We now clarify this point in the discussion in line 495-498.

L 285: which method was used for re-gridding?

Thank you for pointing out this lack of information. We now name the methods, software packages and underlying studies. Re-gridding of Avitabile et al. (2016) was conducted in R (R Core Team, 2019) using the "aggregate" function of the R-package "raster" (Hijmans, 2019), which aggregates to lower resolutions by taking the arithmetic mean excluding NAs. Re-gridded data of Saatchi et al. (2011) was taken from Carvalhais et al. (2014). This information is now provided in line 313-316.

L 289: check brackets in Brien

Thank you for pointing that out. In order to shorten the manuscript according to referee 2 we decided to remove the biomass comparisons based on inventory biomass. Therefore, the citation is now deleted.

L 287: I found it difficult to understand the description of the Rammig et al method and I had to go back to the original paper. I suggest to check the paragraph again for clarity.

Thank you for raising this point. In order to shorten the manuscript according to referee 2 we decided to remove the biomass comparisons based on inventory biomass from the manuscript.

L 304: check bar in 'average  $\bar{x}$ '

Thank you for pointing out this mistake.

In order to shorten the manuscript according to referee 2 we decided to remove the biomass comparisons based on inventory biomass from the manuscript.

L 311: I am skeptical when using gridded climate products to simulate local scale EC fluxes, because these products might not capture some local rainfall events (for example) that have strong impacts on the fluxes. Hence, models will fail to simulate the fluxes. I assume that there are there local scale meteorological data available for the flux sites that could be used for running the model or at least for comparing agreement between gridded data products and observation at EC sites.

We fully agree that using gridded climate products is not optimal to reproduce locally measured ET fluxes, because they can lack information of local weather events. We solely use gridded data in this study due to several reasons: 1) We want to stay consistent with the regional results of this study. As we e.g. show a simulated regional rooting depth maps (formerly Fig. 6 now Fig. 5) and plots of underlying local tree rooting strategies (formerly Fig. 5 now Fig. 4) as well as regional ET (formerly Fig. 11 now Fig. 9) which are all based on gridded climate data, changing the climate input for simulations at local scale seemed inconsistent. 2) Even though we apply statistical metrics to compare model vs. flux agreement, our focus was not on the effects of local weather events on ET, rather on the general climate signal, most importantly the presence or absence of a dry season and the effects on simulated rooting depth and ET and the differences between the different model versions. We fully agree that forcing the model with local climate data could in principle enhance model performance, but we never aimed for a perfect match of ET and NEE at all sites. 3) There is meteorological data available for different flux sites, but these data sets are often cluttered with gaps and are only available for a few up to 10 years only. Moreover, each site has its own limitations when it comes to model implementation. Taken together, this creates quite some problems for DGVM simulations. The LPJmL model needs continuous climate data and long time series. Jumping from a spin-up simulation into repeating 5-10 very similar years can cause artifacts which should be avoided. We would be happy simulate rainforest sites with real site meteorological data, but for this we would need longer time spans than currently available. There are approaches trying to solve those problems, but they are currently beyond the scope of this study. 4) Given that we compare monthly means of simulated and measured ET at the local scale and that simulation results appear to capture seasonal signals (formerly Fig. 9-10 now Fig. 7-8), we are convinced that our approach is sufficient to deliver the message of this manuscript.

We will insert a clarification of why we used gridded climate input data only in the discussion in line 584-588.

L 313: I suggest to state why NEE was only simulated for 3 sites, this information is currently hidden in the figure caption.

Thank you very much for this remark. Unfortunately, in the data sets accessible to us, continuous NEE data covering at least 2 years was only available for 3 sites. We now clarify why we compare NEE only at 3 sites in the methods sections in line 320-321 and in the caption of Fig. 8.

L 336: replace 'called' by ':'

Thank you very much. We changed it accordingly, now line 345.

L 376: replace 'over' by 'instead of'?

Thank you for pointing that out. We changed it accordingly, now line 382.

Generally the results section contains some statements or explanations that do not only describe the results but already go beyond and might be more more appropriate for the discussion.

We will check the results section and will transfer all potential interpretations into the discussion. We have moved and rewritten large parts of the results and discussion section in order to avoid interpretations in the results.

L 387, Fig 5: when looking at this figure, I was wondering if simulated distributions are always unimodal or if there the model can also simulate bi-modal or multi-modal distributions indicating that very distinct rooting strategies can coexist? I also suggest to add to the figure caption which site is wetter and which site is drier.

Thank you for this comment. Indeed these distributions can be bimodal, indicating that very distinct rooting strategies can co-exist. We have not systematically checked all grid cells, but a clear tri-modal distribution was not observed so far. This might also need different ways of detecting them as multiple modes might be hidden in a continuous distribution. So far we observed 2 cases where distributions can clearly be bi-modal. 1) In areas with dominant evergreen tree cover and a “medium” dry season, where shallow and deep rooting evergreen sub-PFTs can co-exist. 2) Areas with a substantial dry season where the evergreen and deciduous PFT co-exist. Here the deciduous PFT shows shallower roots and the evergreen PFT deeper roots. Multi-modal distributions are highly connected to the topic of niche segregation and have many ecological implications, which we want to avoid in this study. With regard to the comment of referee 2 to significantly shorten the manuscript, we regard this topic as beyond the scope of this study. It will definitely be in the focus of future studies.

We indicate the drier and wetter site directly in the respective figure (formerly Fig. 5 now Fig. 4) as suggested.

L 406: Fig 11 (not 9f?)

Thank you for pointing out this mistake. By coincidence the former Fig. 11 is now Fig. 9. The reference is now in line 414.

L 434: why 4m? Is there a reference for this value?

This short paragraph is a rough qualitative description of our simulation results of mean rooting depth in relation to climate variables. It is not a description of results from other studies. We now clarify this by inserting the word “simulation” in line 423.

L 445: ‘behavior: Whereas..’ Full stop or small w in whereas.

Thank you very much for this suggestion. The sentence is deleted in the new version of the manuscript.

L 453: what exactly does ‘reversely’ mean?

Thank you very much for pointing out this unclear formulation. We reformulated this paragraph into “At STM K77 (Fig. 7f) local circumstances show the influence of variable rooting strategies on ET in a different way. This former rainforest site was converted to pasture before Eddy covariance measurements began. This local land-use at STM K77 is not representative for the respective 0.5° grid cell, and thus all 3 LPJmL model versions simulate mainly natural vegetation instead of pasture.” now in line 442-445.

L 377: The text in this paragraph and the figures are mainly about PFTs, not biomes.

Thank you for pointing that out. We changed the heading of this section into “Distribution of plant functional types” now in line 463.

L 519: ‘uncertainty...is’ or ‘uncertainties...are’

Thank you for pointing that out. We decided to remove the whole paragraph from the manuscript in order to shorten it.



L 541: I agree that it's important to look at below ground biomass but comparing Fig 15 and Fig 13 suggest that the ratio between aboveground and below ground biomass is extreme in some areas with high aboveground biomass but low below ground biomass(300-400t/ha aboveground vs ca 20t/ha belowground). Are such ratios realistic in these regions and how can this be explained?

Thank you for this question. In a recent review Fearnside (2018) found that information on belowground biomass of trees in the Amazon region is still very sparse. Available empirical data for 3 sites showed a range of about 15.2 – 33.4 % and a mean of 23.7 % of tree biomass below-ground. No data seems to be available for the western Amazon. Especially in this area we simulate high AGB and low BGB. Here a ratio of 20/(350+20) yields 5.4% of total biomass below-ground. While we are not in the position to validate these values we agree that they might be too low. According to our approach, more root biomass might not be necessary for water uptake and conduction in those regions, but more root biomass might very well be necessary for the statics of trees. Structures ensuring a tree's stability are neglected by the LPJmL model. The implementation of tree statics would most likely increase the belowground biomass in regions which currently show a very low percentage of total biomass allocated to roots. Nevertheless those structures would be necessary for all sub-PFTs and the overall results presented in our manuscript would most likely not change significantly. We now critically discuss our findings of belowground biomass in this context in line 567-571.

L 557: according to figure caption in Fig 12, these are PFTs not biomes.

Thank you for pointing that out. We deleted this paragraph in order to shorten the manuscript.

L 565: 'where' instead of 'were'

Thank you for pointing out this mistake. We changed the word accordingly, now in line 504.

I was surprised that grasses and fire were only shortly mentioned in the discussion, given that the study region also includes seasonal areas with Cerrados and not only evergreen forests. How are these systems represented? Only by deciduous forest or does the model also simulate a grassy component and fire? Fire also has some impacts on biomass in these regions and it has been argued that lateritic layers constrain rooting depth and might thereby influence grass-tree coexistence in these regions.

Thank you for this comment. We agree that fire is an important driver forming the vegetation distribution especially outside tropical evergreen forests. For the current version of LPJmL4.0-VR we used the most simplistic fire module available in LPJmL (GlobFire, Thonicke et al., 2001), which calculates a fire return interval and burned area based on litter moisture only. A PFT dependent parameter for the fraction of killed individuals then determines the burned biomass. Fire-vegetation-feedbacks are therefore existent, but very simplistic. Future studies will incorporate LPJmL's recently updated and much more complex SPITFIRE module (Drüke et al., 2019) and enable to investigate those fire-vegetation feedbacks in a more comprehensive way.

Even though C3 and C4 grasses are explicitly simulated as PFTs in LPJmL4.0, LPJmL4.0 and therefore also LPJmL4.0-VR currently underrepresents the occurrence of grass. This is mainly due to the fact that grass PFTs compete with tree PFTs for area. In that way grass abundance is often highly underestimated when tree PFTs are present. This in turn has the effect that grass-fire feedbacks which naturally stabilize grasslands by reducing tree cover, are not simulated as desired even with a

better fire module. Current ongoing developments of LPJmL aim at allowing grass PFTs to grow under any tree PFT canopy. Here, grass PFTs would mainly be affected by the light reduction of trees. In that way grass fire feedbacks could transform areas currently dominated by the deciduous PFT to become more dominated by the grass PFT, i.e. comparable to savanna-like vegetation types. We now show grass PFT FPC in a new appendix figure (Fig. A6).

With regard to 1) the simplistic fire module used, 2) the underrepresentation of the grass PFTs, 3) our new research question, 4) the comment of referee 2 to shorten the manuscript, and 5) that we do not assess the mechanisms that influence the distribution or stability of potential natural vegetation (i.e. without land-use), we regard the topic of grass-tree coexistence as beyond the scope of this study. Nevertheless, we will discuss the aforementioned shortcomings of LPJmL4.0 and their implications in the discussion now in line 545-551.

L 606: the extent of evergreen forest has not been presented, but rather the extent of the evergreen PFT and the deciduous PFT.

Thank you for pointing that out. We now changed the wording “biome” into “PFT” throughout the text.

Further, Fig 12 shows that the extent of the evergreen PFT is very similar in the original and the VR version (although the FPC is much lower in the original version).

Thank you for this remark. We may have missed explaining this result in the manuscript. The former Fig. 12g-h which is now Fig. 10g-h show that the standard LPJmL4.0 model simulates a rather similar dominance of the evergreen and the deciduous tree PFT in the Amazon region (an almost 50/50 dominance of both PFTs in this region). This model behavior can be explained by the fact that LPJmL4.0 is not capable of simulating a true competitive exclusion over time. PFT establishment rates are not coupled to PFT performance and are in fact equal for all PFTs for every time step (even though the overall establishment rate can vary, but for all tree PFTs in the same way). In the former Fig. 12e-f now Fig. 10e-f we see clear dominance patterns of the evergreen and deciduous tree PFT even though they vastly deviate from the evaluation data (formerly Fig12a-b now Fig. 10a-b) as well. The underlying model LPJmL4.0-VR-base (just as LPJmL4.0-VR) does simulate a performance dependent PFT establishment as described in the methods in Sect. 2.2.6 (formerly 2.2.5). Therefore, the dominance pattern of evergreen and deciduous tree PFTs in Fig. 10e-f can be explained by competitive exclusion. A similar pattern would also be expected in standard LPJmL4.0 when a performance dependent PFT establishment would be implemented. The reason why standard LPJmL4.0's evergreen PFT (Fig. 10g) shows a similar geographical distribution towards the Southern and Eastern border of the Amazon region compared to the extent of the evergreen PFT in LPJmL4.0-VR (formerly Fig. 12c now Fig.10c) is simply because of the human land-use of 2001-2010 applied to all model versions used in our study. The similar extent marks the border of the arc of deforestation in this region. We now clarify this coherence in line 518-528.

I suggest to clarify or to classify vegetation into biomes based on the FPC of different PFTs. This would allow comparisons of biome cover in different scenarios.

Thank you very much for this suggestion. While we regard some of the PFTs used in LPJmL as representatives of biomes, (e.g. the “tropical broadleaved evergreen tree”, as the representative of the biome “tropical rainforest”) we agree not to use the word “biome” in the manuscript. When it

comes to accounting for what factors actually determine a biome the scientific community follows different definitions. Therefore, we also want to avoid classifying simulated vegetation in yet a new way as suggested. Moreover any classification would lead to a loss of information regarding the results of simulated geographical PFT distribution, especially for those familiar with PFTs and DGVMs. We rather follow the other suggestion and now better clarify what the tropical PFTs stand for in line 116-119 and avoid the word “biome” throughout the text.

L 627: ‘Expansion’ instead of ‘Extent’?

Thank you very much. The sentence containing this mistake is now deleted.

L 638: PFT instead of biome types? In the discussion I was missing some discussion of the results in the context of previous modeling studies, such as the studies cited in the introduction. As the study region also includes the Cerrados, the rooting niche separation ideas that explain grass-tree coexistence in savannas might be relevant for the discussion, e.g. Van Langevelde et al 2013 Ecology.

Thank you very much for pointing out that missing connection. We will discuss our results and studies mentioned in the introduction in the context of our new research question. As mentioned above in our answer to the comment regarding L565, grass-tree coexistence is hardly possible to realistically simulate with the current status of the LPJmL model. As stated above we now point out the model shortcomings in line 545-551. In this study we want to avoid the analyses on evergreen to deciduous PFT transition as this would require a whole new set of experiments, figures and statistics which is currently beyond the scope of the manuscript (also with respect to referee 2 requesting a significant shortening of the manuscript) and beyond the capabilities of the model settings. We here want to focus on the role of root adaptations for the broad scale distribution of tropical evergreen forest, ET and productivity in South America. We discuss the potential causes for PFT transitions in line 545-557, but don’t go into a in depth analyses. In fact we argue that analyses of e.g. natural forest-savannah transition and bi-stability should rather be based on simulations excluding human land use and human fire ignitions (which make up the majority of ignitions in South America), i.e. such analyses should be based on simulations of potential natural vegetation, which would bias simulated regional ET rates and their comparison to current regional ET rates based on remotely sensed data sets.

L 1083: Figure S3 not provided.

We are very sorry for this mistake. We now provide the figure as Fig. A5.

Just out of interest, can the model easily be adapted to global scale, and will these model developments be included in the global ‘default’ version of LPJmL4.0? Or would this lead to computational constraints?

Thank you for your interest. First tests of global scale simulations look very promising, but of course have to be evaluated in detail. The principles found in this study seem to apply to other regions as well. The model runs stable on a global scale and currently needs about 2-4 times longer when including 10 sub-PFTs for all 8 natural PFTs in LPJmL. It is therefore conceivable to make this model development (or a version with a somewhat reduced amount of sub-PFTs per PFT) a constant model feature or at least an optional feature in the future. This still has to be decided by the LPJmL model development community.

## References:

- Avitabile, V., Herold, M., Heuvelink, G. B. M., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., Bayol, N., Berry, N. J., Boeckx, P., de Jong, B. H. J., Devries, B., Girardin, C. A. J., Kearsley, E., Lindsell, J. A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E. T. A., Nagy, L., Qie, L., Quinones, M. J., Ryan, C. M., Ferry, S. J. W., Sunderland, T., Laurin, G. V., Gatti, R. C., Valentini, R., Verbeeck, H., Wijaya, A. and Willcock, S.: An integrated pan-tropical biomass map using multiple reference datasets, *Glob. Chang. Biol.*, 22(4), 1406–1420, doi:10.1111/gcb.13139, 2016.
- Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Saatchi, S., Santoro, M., Thurner, M. and Weber, U.: Global covariation of carbon turnover times with climate in terrestrial ecosystems, *Nature*, 514(7521), 213–217, 2014.
- Drüke, M., Forke, M., Bloh, W. Von, Sakschewski, B., Cardoso, M., Bustamante, M., Kurths, J. and Thonicke, K.: Improving the LPJmL4-SPITFIRE vegetation-fire model for South America using satellite data, *Geosci. Model Dev.*, 12(12), 5029–5054, doi:10.5194/gmd-12-5029-2019, 2019.
- Fearnside, P. M.: Brazil's Amazonian forest carbon: the key to Southern Amazonia's significance for global climate, *Reg. Environ. Chang.*, 18(1), 47–61, doi:10.1007/s10113-016-1007-2, 2016.
- Jackson, R. B., Canadell, J., Ehleringer, J., Mooney, H., Sala, O. and Schulze, E.: A global analysis of root distributions for terrestrial biomes, *Oecologia*, 108, 389–411, 1996.
- Johnson, M. O., Galbraith, D., Gloor, M., De Deurwaerder, H., Guimberteau, M., Rammig, A., Thonicke, K., Verbeeck, H., von Randow, C., Monteagudo, A., Phillips, O. L., Brienen, R. J. W., Feldpausch, T. R., Lopez Gonzalez, G., Fauset, S., Quesada, C. A., Christoffersen, B., Ciais, P., Sampaio, G., Kruijt, B., Meir, P., Moorcroft, P., Zhang, K., Alvarez-Davila, E., Alves de Oliveira, A., Amaral, I., Andrade, A., Aragao, L. E. O. C., Araujo-Murakami, A., Arets, E. J. M. M., Arroyo, L., Aymard, G. A., Baraloto, C., Barroso, J., Bonal, D., Boot, R., Camargo, J., Chave, J., Cogollo, A., Cornejo Valverde, F., Lola da Costa, A. C., Di Fiore, A., Ferreira, L., Higuchi, N., Honorio, E. N., Killeen, T. J., Laurance, S. G., Laurance, W. F., Licona, J., Lovejoy, T., Malhi, Y., Marimon, B., Marimon, B. H., Matos, D. C. L., Mendoza, C., Neill, D. A., Pardo, G., Peña-Claros, M., Pitman, N. C. A., Poorter, L., Prieto, A., Ramirez-Angulo, H., Roopsind, A., Rudas, A., Salomao, R. P., Silveira, M., Stropp, J., ter Steege, H., Terborgh, J., Thomas, R., Toledo, M., Torres-Lezama, A., van der Heijden, G. M. F., Vasquez, R., Guimarães Vieira, I. C., Vilanova, E., Vos, V. A. and Baker, T. R.: Variation in stem mortality rates determines patterns of above-ground biomass in Amazonian forests: implications for dynamic global vegetation models, *Glob. Chang. Biol.*, 22(12), 3996–4013, doi:10.1111/gcb.13315, 2016.
- Kleidon, A. and Heimann, M.: A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water and carbon cycle, *Glob. Chang. Biol.*, 4(3), 275–286, doi:10.1046/j.1365-2486.1998.00152.x, 1998.
- Schaphoff, S., von Bloh, W., Rammig, A., Thonicke, K., Biemans, H., Forkel, M., Gerten, D., Heinke, J., Jägermeyr, J., Knauer, J., Langerwisch, F., Lucht, W., Müller, C., Rolinski, S. and Waha, K.: LPJmL4 – a dynamic global vegetation model with managed land – Part 1: Model description, *Geosci. Model Dev.*, 11(4), 1343–1375, doi:10.5194/gmd-11-1343-2018, 2018.
- Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J. and Hutley, L. B.: An optimality-based model of the coupled soil moisture and root dynamics, *Hydrol. Earth Syst. Sci.*, 12(3), 913–932, doi:10.5194/hess-12-913-2008, 2008.
- Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography

and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Glob. Chang. Biol.*, 9(2), 161–185, doi:10.1046/j.1365-2486.2003.00569.x, 2003.

Thonicke, K., Venevsky, S., Sitch, S. and Cramer, W.: The role of fire disturbance for global vegetation dynamics: Coupling fire into a dynamic global vegetation model, *Glob. Ecol. Biogeogr.*, 10(6), 661–677, doi:10.1046/j.1466-822X.2001.00175.x, 2001.

## Response to the comments of referee 2 on Sakschewski et al. 2020: Variable tree rooting strategies improve tropical productivity and evapotranspiration in a dynamic global vegetation model

I like many aspects of this paper. In particular I appreciate the approach of carefully examining the behaviour of the model across a range of climate scenarios, Overall I think this paper and model variant have great potential. The new scheme is clever and I have little to critique in its design. The various analyses do suggest the new model is providing a better match to available data. But I am yet to be convinced this improved match is for the reasons the authors claim (better handling of soil water & rooting).

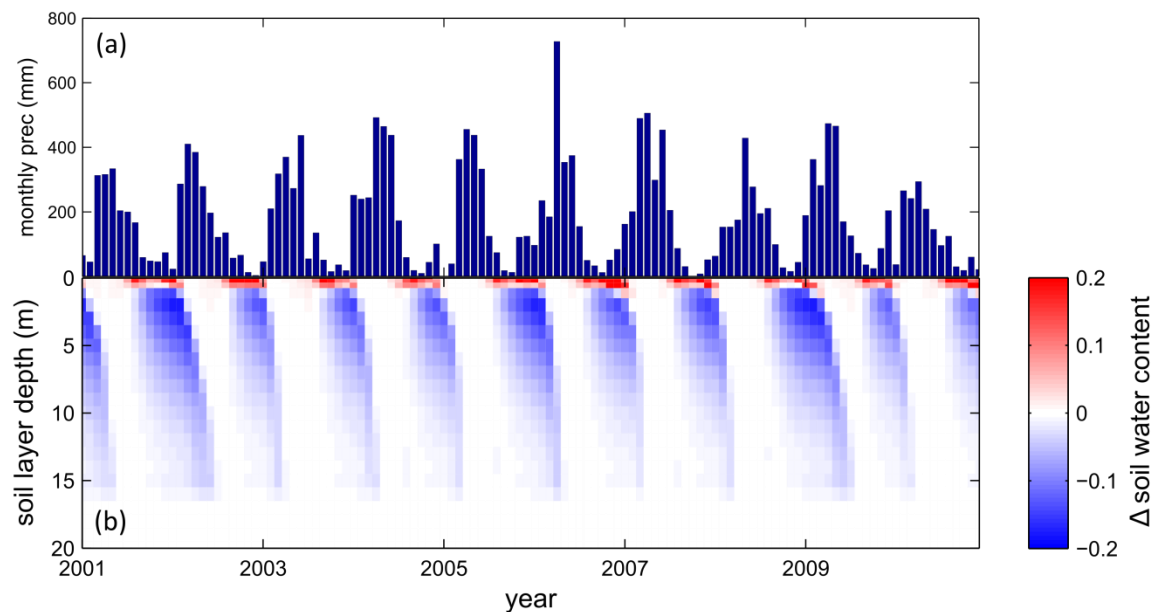
Dear Daniel Falster,

Thank you very much for your positive evaluation of our approach.

The biggest concern I have is that despite all the results presented, these mostly for aggregate outputs at largish spatial scales. Almost no evidence is presented to show the effect of the new scheme on the actual water balance in the soil. Moreover, we don't even know how water is modelled in the different versions. Sure, the soil and root depth is changed, but what does this mean for water balance at different soil depths? Surely this is key to assessing why the model behaves differently. This is important, as the changes in rooting will also change the way carbon is allocated within the model (e.g. deeper roots divert carbon away from leaves). Are the model improvements due to changes in hydraulics, or changes in carbon allocated between leaves and roots?

Thank you very much for raising this point. Indeed we did not explicitly explain how the water percolation scheme and water balance of all 3 LPJmL model versions works. We inserted a new description of this part of the model in Sect. 2.2.2 line 136-151. We agree that we do not present results of how the new scheme affects the water balance in the soil. However, we compare the results of our new scheme (LPJmL4.0-VR) to a baseline model (LPJmL4.0-VR-base). LPJmL4.0-VR-base differs from LPJmL4.0-VR only in the amount of tree rooting strategies present. Both models have the same soil depth information. Hence, soil hydraulics are the same as well and differences in evapotranspiration, net primary productivity, biomass and PFT distribution only arise from the absence/presence of different tree rooting strategies. Deeper roots clearly correlate with higher evapotranspiration and productivity during dry seasons. When shallow roots lead to decreasing evapotranspiration in dry seasons, then the water used for evapotranspiration by deep rooted sub-PFTs must come from deeper soil layers. Regardless of knowing the exact soil water balance we argue that this evidence is presented. The fact that deep rooted sub-PFTs are selected for in regions with deep soils and a dry season, even though deeper roots are more costly than shallow roots, proves that diverting carbon away from the leaves into a deeper root system can be beneficial for the overall carbon balance. Therefore, the differences in the results between LPJmL4.0-VR and LPJmL4.0-VR-base must be related to the new rooting scheme and not to soil hydraulics. Nevertheless, we prepared a new potential Appendix figure (Fig. A7) which we pasted right under this specific response. It shows the difference in monthly relative soil water saturation between LPJmL4.0-VR-base and LPJmL4.0-VR at STM KM67, for 2001-2010, for all 23 soil layers (so LPJmL4.0-VR-base minus LPJmL4.0-VR). Here blue colors depict a lower soil water content in LPJmL4.0-VR compared to LPJmL4.0-VR-base. The figure clearly shows how deep water is extracted by deep roots in the dry

season and how the soil water balance changes over the year. We refer to this new assisting appendix figure now in line 596-598 and hope this presents enough evidence.



**Figure A7: Difference in soil water reaction to seasonal precipitation between LPJmL4.0-VR-base and LPJmL4.0-VR at Fluxnet site STM KM67 a) Mean monthly precipitation input from CRU for 2001-2010. b) Difference in monthly relative soil water content between LPJmL4.0-VR-base and LPJmL4.0-VR forced with CRU climate for 2001-2010. The underlying model output variable “soil water content” of each model version is a number between 0 and 1 depicting the relative water saturation of the soil. Blue colors denote lower soil water content in LPJmL4.0-VR and red colors a lower soil water content in LPJmL4.0-VR-base.**

Second, I feel the authors need to come up with a stronger story and reduced set of results. The paper is currently very long and dense. The authors have made many comparisons using a variety of datasets. Consequently, there is a large number of figures (15) and tables (8). This makes it hard for us to know where to put our attention.

Thank you very much for this suggestion. Indeed the manuscript is very long as variable roots changed a variety of model results for the better. We agree that a clearer story helps to convey the most important messages. As we will now highlighted in the introduction in lines 101-103 we now focus on “the hypothesis that varying tree rooting depth is key to explain major patterns of evapotranspiration, productivity and the geographical distribution of tropical evergreen forests in South America” and structure the manuscript accordingly. We will shorten the manuscript according to this question and transfer several figures and tables to the supplement. Finally, more work is needed to make the different results accessible and easy to interpret. I found that each figure required a fair bit of work to interpret what is going on. Some simple changes could make it much easier for the reader, then we could spend less time deciphering your results and more time thinking about the science!

Thank you for raising this very important point. We will check every figure, insert more information like labels into the figures and simplify the captions and labels.

As examples,

Thank you very much for pointing out all these examples.

- In Fig 1: confusing caption. Simplify labels in legend.

We now simplified the caption and reduced the complexity of the legend labels.

- In Fig 6: label panels with dataset name, so that we don't need to refer to legend as much

We now label each panel as suggested in this figure which is now Fig.5.

- In Fig 9, uses different colours in the map and traces, otherwise these are easily confused. Label each subplot with site name.

We colored the markers more intensively and inserted the site name in each panel as suggested for this figure which is now Fig. 7. We applied the same approach in Fig. 8.

- In Fig 12, put labels on the columns (evergreen, deciduous) and rows (models), so that we can easily see what the different panels are without constantly referring to the caption.

We labelled all panels according to the suggestions for this figure which is now Fig. 10.

Some minor issues

- Some line breaks between paragraphs would make the text much easier to read

Thank you for raising this point. During shortening of the manuscript and strengthening the story line we will insert more line breaks.

- Eq 8: I've looked over this a few times and wonder if the  $n_{est\_tree}$  at the end should be removed?

Thank you for this question. Yes this is true. The 2 times occurring  $n_{est\_tree}$  can be cancelled from this equation. We kept it in for an easy comparison to the original equation (Eq. 7). We think this makes it easier for model developers to copy our approach.

- In 413 – It didn't make much sense to me to compare your results to a modelled product of rooting depth.

Thank you for raising this point. We completely understand this criticism. Since empirical data on rooting depth in this region is very sparse, we wanted at least evaluate our results in the context of a totally different modelling approach. With regard to shortening the manuscript we now moved the comparison to the Appendix and avoid a detailed comparison of our results with Fan et al. (2017).

- I found the talk of offspring, saplings, and "growth" throughout the paper a bit misleading.

My understanding of this version of LPJ is that each patch has a single functional type which has a density 'n' of average sized individuals. When new offspring are recruited, they don't grow from seed to adult, but rather enter fully formed at the average size (this occurs by increasing n, the number of individuals). The only time the individuals seemingly grow from small to large plants, is when starting from bare soil, i.e. during spin up. Yet, often the paper gave the impression that individuals could be born and grow.

Thank you very much for this criticism. We fully agree. Apparently, it is always a thin line between easy wording and correctness of model description. We now clarify that PFTs in LPJmL are average individuals in the methods now in line 237-239 and avoid any misleading formulations in the entire manuscript.