Final response

We would like to thank the editor and the reviewers for assessing the manuscript. We made the necessary changes following the comments of the reviewers. Below, we first address the comments of Referee #2, followed by our response to Referee #3. In the following, the referee comments are written in italics.

Response to Referee #2

I compliment the authors on their thorough revisions to the manuscript and thank them for taking the time to run new simulations (see Fig. 4, S1.41, S1.56, S1.57) to answer my previous comments. They have addressed most of my comments thoroughly; however, I do require more clarification on the differing responses of perennial ET and GPP to soil texture (see comments on Lines 357-367 below). In addition, I have included minor grammatical comments and suggestions for making the text easier to read. I recommend this technical note for publication once these minor changes are addressed and look forward to the final product. Thank you.

We are happy the referee valued our efforts to improve the manuscript, and would like to thank the referee also this time for his thorough review.

Line 8-9: I would replace "original results" with "original analysis of Schymanski et al. (2015)" and adjust the rest of the sentence accordingly.

We changed this to "original results of Schymanski et al. (2015), and we implemented these changes one at a time".

Line 16: Is this an underestimation of ET and GPP compared to Schymanski et al. (2015) or to the observations? Or both? I think this means compared to Schymanski et al. (2015) as mean annual GPP is overestimated for AoB2015 and v0.5 (Lines 20-21).

Indeed, we discuss here in comparison with Schymanski et al. (2015). We changed this accordingly.

Line 20-21: I would specify that GPP is overestimated by 17.8% and 14.7%.

We rephrased this to "whereas the relative errors for the mean annual GPP remained similar with an overestimation that changed from 17.8% to 14.7%."

Line 34: "...photosynthesis, water uptake and storage..."

Changed accordingly.

Line 75-76: ", but an evaluation of the effect of this change on the original simulations was not included in Nijzink et al. (2021)." Were any of the effects mentioned in the above paragraphs included in Nijzink et al. (2021)? If not, I would remove this line.

We removed this line.

Line 87: change "works" to "work"

Changed accordingly.

Line 111: Remove "also" from "see also Figure 2"

Changed accordingly.

Line 111-112: Run-on sentence. Add comma after "(grasses)" or rephrase.

We added a comma.

Line 119: Replace "mol fraction" with "mole fraction"

Changed accordingly.

Line 125: I would define what c_Rl is rather than saying it is a constant. I am guessing it is the cost factor for leaf respiration?

This factor was defined by Schymanski et al. (2007) and comes from the results of Givnish (1988). This study showed that for a range of species the leaf respiration equals 7% of photosynthetic capacity. We clarified this in the manuscript.

Line 127: I would define h_a and h_d rather than saying they are parameters.

We added a description of these parameters, that were originally taken from Medlyn et al. (2002).

Line 135: I would include the variable name and units after "Root water uptake"

Changed accordingly.

Line 166: The equations for root surface area distribution are not defined anywhere. It is easy to imagine how the short term optimization works with M_a,s and J_max,25, but not with S_A,d. Can the authors either add an equation or discuss how roots are distributed in each layer?

We added the equations and described how the root surface areas are distributed over the layers. Briefly, it is first assessed how far tissue water content gets depleted in the course of the day. This decides whether root surface area will be increased or reduced in the overall profile. In a next step, the effectiveness per layer is determined, i.e. the root water uptake per unit root surface area in each layer. The relative effectiveness of each layer is eventually used to distribute the increases/decreases over the layers, with the biggest increase for the most efficient layer and the largest reduction for the most inefficient layer.

Line 178: Where does the value of 0.22 comes from?

This was based on an analysis by Schymanski et al. (2007) of the Glopnet dataset (Wright et al., 2004). We clarified this.

Line 192: Remove "also" from "see also Figure 2"

Changed accordingly.

Line 197: Here and throughout the paper replace "Van Genuchten" with "van Genuchten".

Changed accordingly.

Line 207-208: I am okay with not delving into the full soil evaporation model, but would recommend a *parenthetical comment referring to the section of the previous papers that explain the model.*

We added a reference to Schymanski et al (2009), including the equation numbers.

Line 223: I appreciate the authors addressing my previous time-stepping comments. However, this conversion from daily to hourly brought up another question. Is a diurnal variation imposed when converting fluxes like temperature and radiation from daily to hourly? Maybe this is covered in Schymanski et al. (2009), but it may be good to briefly mention here. Otherwise the hourly Cowan and Farquhar calculations would not be very meaningful if the atmospheric conditions were constant over the day.

We now clarify in Section 2.2.7 that diurnal variation was imposed for global radiation and temperature, and consequently for atmospheric vapour deficit, referring the reader to Appendix A in Schymanski et al. (2009) for details.

Line 263: Remove "also" from "see also Supplement S2"

Changed accordingly.

Table 3: Is there a more specific name for the "water use parameters"? The exponent and multiplicative factor have the same description in Table 3.

They do not have more specific names, besides defining the non-linearity of the relationship through the exponential factor and the multiplication through the other factor. We changed the names in the table to "exponential water use parameter" and "multiplicative water use parameter" for clarity.

Sect 3.1: I really like that the authors added further explanation of alterations to vegetation properties under each case as well as Figure 4. I have two (hopefully minor) recommendations for this section to help make it easier to follow:
1) I would indicate the case being discussed in each paragraph to help the reader to attach the text to Table 4. For example, in Line 334 it would be helpful to write "In contrast, changing the fixed atmospheric CO2-levels (350 ppm) in the VOM-AoB2015 to variable atmospheric CO2-levels (Case 2)..."
2) The authors have chosen to discuss the modifications to GPP and ET for each case separately from the modifications to vegetation properties. This can be a little tedious for the reader and may obfuscate the findings. For example, in Lines 335-339, the authors discuss how variable CO2 led to increased GPP, while 10 paragraphs later (Lines 378-382), they discuss how the variable CO2 yields larger perennial vegetation cover. The increase in GPP in the first paragraph is influenced by the vegetation modification in the second paragraph, so it makes sense to combine the two. I would recommend the authors assimilate Lines 375-402 into their respective case paragraphs in Lines 331-374. This will ease the reading by discussing each case completely once as well as make the connection between the modifications of GPP/ET and vegetation properties clearer.

Thank you, these are very good suggestions and we made changes accordingly.

Lines 357-367: I am still struggling with this explanation, which implies that increased vegetation cover and soil storage capacity benefit perennial GPP, while, simultaneously, higher suction head (lower matric potential) and decreased hydraulic conductivity reduce perennial ET. This seems *contradictory as stomata are controlling both perennial GPP and ET, which means both fluxes should have a similar response (in sign at least) to changes in cover, soil storage, soil water potential and hydraulic conductivity. The only way different GPP and ET responses make sense, are if water use efficiency and/or photosynthetic capacity per leaf area changes. The authors have illustrated with new analysis (Fig. S1.56-1.57) that the new water use parameters in VOM-v0.5 do not have a large effect on this result. However, I think the answer may lie in the effect of the new soil texture on overall λp through the higher suction heads. The new soil suctions seem to create more efficient water use in the wet season compared to VOM-AoB2015 (Fig. S1.53d) and could lead to the different sensitivity of GPP and ET to changes in soil texture. To summarize, can the authors explain how perennial GPP and ET can have opposite responses to increased vegetation cover/soil water capacity and decreased water potential/conductivity when they are both controlled by stomata?*

Thank you for your constructive comments on this matter, we fully agree that it is counter-intuitive and we want to have a good explanation. Our definition of GPP and ET per unit ground area instead of per leaf area may cause some confusion. On a per leaf area basis, it is correct to assume that changes in ET and GPP should have the same sign. However, if leaf area per ground area increases, the increased light attenuation can result in much greater GPP per ground area for the same or even lower ET, similarly to increasing photosynthetic capacity, as the reviewer noted. In this particular case, the increase in perennial cover is mainly due to improved access to soil moisture during the dry season, but it results in largely increased WUE in the wet season, as light attenuation is increased at the same time as soil moisture access is slightly reduced during the wet season due to higher soil suction and reduced hydraulic conductivity. This results in the different *λp*-values, as noted by the referee. We clarified this in the revised manuscript.

Line 361-363: How does a larger soil moisture capacity and carry-over lead to higher suction heads? Is my understanding that higher suction head means more negative matric potential correct? If so, then wouldn't higher suction heads compound (and not be compensated by) the effects of lower hydraulic conductivity mentioned in Line 364? It appears from Figure S1.40c-e that although there is more water in the soil under the new texture, it will be harder for the roots to extract it due to lower water potentials and less conductivity. This intuitively explains the reduction in perennial ET even though there is greater vegetative cover, but does not address how there can be greater perennial GPP (see previous comment). Can the author's elaborate?

We agree that our formulation was confusing here, and thank the reviewer for pointing it out. Contrary to what our statement implied, it is not that higher storage capacity leads to higher suction heads, but finer texture results in reduced hydraulic conductivity and higher suction heads at relatively high soil water contents, and therefore stronger water holding capacity. Stronger water retention is beneficial for dry season conditions, therefore permitting greater perennial cover, but higher suction heads and reduced conductivity make it harder to suck water out during the wet season, as pointed out by the referee. As explained above, improved dry season conditions increase perennial cover, which increases perennial wet season WUE, resulting in the simulated increase in wet season perennial GPP at slightly reduced ET.

Lines 372-374: I am not sure the authors meant to make this a new paragraph, but it seems to be part of the previous paragraph.

This should be part of the previous paragraph indeed, we corrected this.

Sect. 3.2: I still find the purpose of this section unclear. Currently, the section compares overall performance of the VOM-v0.5 (Case 12) and VOM-AoB2015 (Case 1) at predicting Howard Springs ET and GPP (Fig. 6). Next, the section explores the mechanisms for the model differences driven primarily by subsurface changes (Fig. 7). This section could be helped by clearly stating that you are comparing VOM-v0.5 (Case 12) used in the companion paper to the VOM-AoB2015 (Case 1) to the Howard Springs data. Then, let the reader know you are diving further into the model differences by exploring the compensating effects of the two most important factors from Sect. 3.1, soil texture and free drainage. Lastly, the primary focus in this section is on ET, but it would be helpful to say a bit more about GPP effects.

We added two introductory sentences here, and elaborated also on GPP effects. However, the strongest influences are by the hydrology and the soil, which is why we focused on these here.

Line 405: I would write "...mean annual GPP changed from 17.8% to 14.7% overestimation." to be consistent with the ET description.

Changed accordingly.

Line 414-415: What is the significance of this difference? Is it due to the fact you are comparing soil moisture at 5 cm to integrated soil moisture at 20 cm? I would expect the model to be wetter since deeper soil layers tend to be wetter than near the surface, where transpiration, soil evaporation and loss to deeper layers via gravity and suction occur.

Thank you for this comment, this is a good point. We also did not expect that the observed and simulated soil moisture would exactly match, but looked also more at the similarity in dynamics. We rephrased this sentence, and added also the point brought forward by the referee, that the soil moisture is expected to be wetter for the model.

Lines 420-425: This portion will need to be updated based on the response to Lines 361-367.

We updated this part.

Line 429: Remove extra parentheses after "(Nijzink et al., 2021)"

Changed accordingly.

Line 437 – 444: It would be helpful to reference the figures that illustrate these conclusions.

We referenced the figures.

Line 447: Can the authors elaborate on why this effect is more important on highly permeable soils. I could see tighter soils having a larger capillary fringe and interacting with the root zone significantly.

We had here the rather permeable soils at Howard Springs in mind, in comparison with the other sites along the North Australian Tropical Transect. However, we believe this statement is not fully supported by what we present in this manuscript, so we decided to remove it.

Response to Referee #3

I am not familiar with the VOM approach. This paper (as well as the accompanying paper in HESS) is not complete enough for understanding this approach. Complete, open-source land surface models able to work at all spatial scales, making use of all available observations (including satellite-derived products) are now available. Why do we need this new approach? The Authors claim that VOM does not need calibration of model parameters but in the end they find that soil water transfer processes are key (not a surprise to me!) and that soil properties need to be described. This means that parameter values have to be prescribed at some stage. Tuning rooting depth is a good example of model parameter tuning. This sounds like a contradiction.

We are sorry that the referee feels that the manuscript, as well as the accompanying paper in HESS, is not complete enough for understanding the approach. We made changes in the Methods section to improve the clarity, by moving the section about the water balance and the carbon costs before the sections about the optimization. At the same time, we added an extra equation defining the Net Carbon Profit in order to clarify the definition of this objective (Equation 10), which is independent from observations.

The referee argues that existing open-source land surface models work at all spatial scales and make use of all available observations. However, this is exactly the problem that we are pointing out. Using observations of vegetation properties and behaviour as input in these models means that we do not understand how the ecosystem functions, and that making predictions in future scenarios remains highly uncertain. We added some extra sentences in the introduction to underline this.

The VOM reduces the need for observed vegetation properties or behaviour as model input or for model calibration by predicting them based on optimality theory. Consequently, contrary to the reviewer's interpretation, we do not tune rooting depths, but we optimize these for maximizing the Net Carbon Profit, and then compare the simulations with observations and with simulations based on prescribed rooting depths.

I am extremely concerned by the lack of clarity on how leaf area index (LAI, in m2m-2) is represented. LAI is a key driver of all surface fluxes, including plant transpiration, soil evaporation, rainwater interception. LAI is also related to other quantities like surface albedo. How are LAI and surface albedo represented? How is interception represented? Etc. A Figure showing how these variables compare to observations would be useful. Why not plotting LAI time series in Figures 5 and 6? LAI is mentionned on Line 200 (Eq. 8). But in Eq. 8, a constant "clumped LAI" is used. What is the definition of "clumped LAI"? Why using a constant value of 2.5? Does "clumped LAI" mean "effective LAI"? How do you calculate and validate the clumping index relating true LAI to effective LAI? How and why is Rf from Eq. 8 used in the model?

The VOM uses a big leaf approach, where LAI is only used to connect the absorbed fraction of PAR with foliage turnover costs. For this purpose, the VOM assumes that a LAI of 2.5 is needed to absorb all the PAR, and that this LAI is reached within the vegetated fraction of a catchment. Since the VOM dynamically predicts the fraction of area covered by vegetation and distinguishes between vegetated and bare soil area fractions, we refer to the LAI within the vegetated area fraction as "clumped LAI", whereas the site-averaged LAI would be that multiplied by the vegetated area fraction. Due to this simplistic represenation of LAI, we do not present a detailed analysis of the LAI dynamics and per-leaf-area fluxes. This is also one of our discussion points in the accompanying HESS paper. We clarified in sect 2.2.2 that the LAI is not modelled explicitly and in 2.2.3 we now mention what is meant by clumped LAI and refer to Schymanski et al., 2007 for more details.

Rf is the carbon cost related to the maintenance of leaf area, we defined this now also in the text. It is used to calculate the NCP, which we defined with the new Equation 10.

Interception was assumed to be negligible at these sites, as there is a strong seasonality with heavy rainfalls. The VOM does not calculate a surface energy balance and hence does not consider the surface albedo.

Finally, simulations at the site level are presented. Is the VOM able to make 2D simulations? If yes, at which spatial resolution? If not, what would be needed to acquire this capability?

This is an interesting question. Currently, the VOM only works at the point scale, but a gridded version could be considered in the future. To do so, adjustments to especially the optimization algorithm are necessary, as the vegetation parameters adjust to the local climate and could be different per grid cell. In addition, adjustments to the water balance part may be needed as well, in order to account for the routing of water through the system. However, the focus of this paper is on the processes governing vegetation response to the environment and if/how improvements in process representation propagate into improved model predictions, so we did not expand on the issue of spatially explicit modelling.