Comment on gmd-2021-154
Martin De Kauwe

In this study, Wang et al. present an interesting analysis where they evaluate one stomatal optimisation and two empirical stomatal models in the new CliMA model. In contrast to many land surface schemes, this model is able to account for greater complexity in the treatment of the canopy space, hence this paper has the potential to offer interesting insights to the state of the knowledge, in particular as we think about more realistic scaling up (from the leaf) of water fluxes. Nevertheless, I currently have some questions about the presentation of the methodology that I think are worth clarifying for the reader, I will outline these below. (Note I've not read beyond the methods...)

- One overall concern I have is about an apparent "conflict" in added complexity in some assumptions with marked simplifications in others. Are these tradeoffs warranted(?) it would be good to add some commentary on this point. For example, a lot is made of the vertical treatment of CliMA but then you assume constant leaf physiology parameters throughout the canopy - so Vcmax does not change with depth through the canopy? I make additional comments below.

- In your parameterisation of the stomatal models (table 1), you used species-level hydraulic parameters to determine the Weibull function in the optimisation model + fit Kmax to site data, but by contrast, you used plant functional type parameters to run the empirical models. Is there any evidence these values are appropriate for the species at these two flux sites? Isn't this akin to "calibrating" the optimisation model but then evaluating its skill improvement relative to an uncalibrated (empirical) model? I also note that the correct citation for the Meldyn g1 values is De Kauwe et al. 2015, GMD not the CLM tech note (as this borrows from that paper and isn't the original source).

- Later when you do fit values (Table 3), are these values sensible? There must be some site estimates of Vcmax that the values could be compared to? I note that the Niwot values are pretty low from a cursory glance. Similarly, a g1 of 16 makes no sense to me, if you look at Lin et al. 2015 NCC, it is above any of the values they derive in their global synthesis. I think these fits are worth double-checking. Similarly, can you also double-check the Kmax values?

- Furthermore, why is Vcmax being fitted differently across schemes? This is a unique
quantity that reflects the canopy, so by varying it across models, aren't we shifting errors between water and carbon fluxes? When g1 is not fitted at the MOFLUX site, Vcmax in the OSM model is half of what it is in the empirical model, this is not a small difference. I note there is an explanation but in the text but this isn't clear to me, "effective Vcmax", I don't see why this would (a) differ across approaches and (b) why it would ever differ by so much.

- By prescribing leaf temperature and soil moisture we are not able to get a good sense of how this would actually work in a LSM scheme when these feedbacks would be important. While I think this approach has value for minimising differences across models, I think it is equally valuable to turn on these feedbacks. I would prefer to see both versions presented.

- Is there a citation to support the division of Kmax (speaks further to my point about apparent complexity vs simplification)?

- Should eqn 7 have an autotrophic respiration loss term?

- Fig 5c, the axis is Pa not KPa.

In summary, I think this will be of great interest to readers as this approach is novel in terms of the role of canopy complexity and stomatal optimisation but there are key methodological points that warrant clarification. I look forward to reading a future revised version.

Martin De Kauwe