Review of “When does vapor pressure deficit drive or reduce evapotranspiration?” submitted to HESS, by Adam Massmann, Pierre Gentine, and Changjie Lin

The authors address an extremely compelling problem in this manuscript – the question of whether increasing vapor pressure deficit will lead to decrease in evapotranspiration (due to stomatal regulation to reduce water loss) or increase in evapotranspiration (due to increased atmospheric demand). Answering this question will advance our capabilities to understand and predict ecosystem-level differences in water uptake in response to warming, and thus deserve attention and effort.

Methodologically, the authors employ a mix of models from (1) land-atmosphere coupling (Penman-Monteith, based on radiation balance and surface and air conductance of water vapor into atmosphere), (2) leaf-scale optimal stomatal conductance theory (from Medlyn et al. 2011) that relates stomatal conductance Gs to GPP, VPD, and a water use efficiency term g_1, and (3) an empirically-derive relationship between ecosystem scale GPP, ET, VPD, and an “underlying water use efficiency” uWUE:

(1) Penman-Monteith:	ET = f_1(VPD, Gs)
(2) Optimal stomatal conductance (Medlyn et al. 2011):	Gs = f_2(g_1, VPD, GPP)
(3) Underlying water use efficiency (Zhou et al. 2014):	uWUE = f_3(GPP, ET, VPD)

Together, they present a set of three closed equations that can be used to eliminate the dependence on GPP, and relate ET as a function of VPD, g_1, and uWUE only.

All the results follow from the assumptions set out in the derivation of the model, so I will focus my comments mostly on the derivations. I am mainly concerned about two issues:

1. Compounding uncertainties in parameters across each one of the equations used
2. Interpretation and attribution of observed effects to plant physiological responses

1. Compounding uncertainties

A key premise of this work is that g1 and uWUE should exhibit greater variation across ecosystem types than within, thus distinguishing ecosystem types and responses from each other. In trying to see whether this premise is valid, I went back to check on the works of Medlyn et al. (2017) and Zhou et al. (2014), who have already derived values of g1 and uWUE respectively in their own works, and examined the range of previously derived parameter values and compared it with the values fitted using the current model.

The results were unconvincing. The ranking of these parameter values across ecosystems are not preserved. This is shown in Table 2 of the manuscript … whereas Zhou et al. (2014) predicted the crop types (CRO) and evergreen needleleaf forest (ENF) to have the highest uWUE values of all ecosystem types analyzed, the model in the current manuscript predicted both crops and needleleaf forests to only have a moderate, middle-of-the-road uWUEs. The current model instead predicts the highest g1 values for crop types, but those had the lowest mean g1 values as predicted by Medlyn et al. (2017) (Figure 7; C3C and C4C). Thus, the rankings of g1 and uWUE values were found to be inconsistent across all these studies.

How should we interpret this? I'm not sure the authors offer any insights into this question. Certainly, they have acknowledged the existence of uncertainty around the temporal and spatial variations of uWUE and g1 by introducing an uncertainty parameter *sigma*. But given the potential sources of error in working with high resolution ecosystem scale data, and the range of values that both parameters can take on *within* the same ecosystem type, I wonder if it might be more effective to think about the relationship between ET and VPD more probabilistically as a function of pdfs – rather than point estimates – of g1 and uWUE.

The questions still remain, though, that according to Medlyn et al. (2017) Figure 7, the degree of variability in g1 *within* ecosystem types and using different methods of derivation (via leaf, isotope, or flux data), can be as high as variability in g1 *across* ecosystem types. The authors do not seem to have addressed this issue of *within-PFT* range in g1 and uWUE, and without it, I found it very difficult to interpret what these parameter values mean and what they could be useful for. If indeed these discrepancies arise from the selection of sites and/or time periods, then how sensitive would the results be to these choices of analyses?

2. Attribution to physiological responses

I am also concerned about the role of g1 and uWUE in masking potential contribution of soil moisture to ET. Essentially, I don't think it's correct to say that g1 and uWUE are attributes of the PFT *only*, which is another key premise of the authors' interpretation of the results.

There now exist a substantial body of work that suggest that *lambda* (the marginal water cost of leaf carbon used as the Lagrange multiplier in the calculus of variations for optimizing stomatal conductance) varies under water deficient / droughted conditions (see Makala et al. 1996, Annals of Botany; Kirschbaum 1999, Ecological Modeling; and other references within Medlyn et al. 2011). **This means that g1 itself, which is a function of *lambda*, should vary under water-limited conditions.** This functional dependence of g1 on soil moisture is also supported by empirical works from AmeriFlux sites such as those from Novick et al. (2016). Medlyn et al. (2011) itself states that: "*It can be questioned whether the optimization criterion assumed here can still be said to be optimal if drought stress starts to threaten plant survival. It may be that the relationship given by Eqn (11)* (the one used in the current manuscript) *will break down as soil moisture potential is reduced.*"

So, this leads me to conclude that attribution of the derived ET responses to VPD entirely as a result of plant physiology – by using a relationship that derive from an acknowledged limitation in its ability to respond to soil moisture – is inaccurate. This attribution is repeated throughout the manuscript and is illustrated in statements like "plants that are evolved to bred to prioritize primary production over water conservation (e.g., crops) exhibit a higher likelihood of atmospheric demand-driven response (found in the abstract)." An alternative interpretation is that these ecosystem types are responding in this way because they have, on the whole, been subjected to less soil water limitation (due to the non-negligible effects of irrigation?) Without decoupling the effects of VPD from those of soil moisture, I think that the interpretations offered here could be quite misleading.

References

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V., ... & Wingate, L. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134-2144.

Medlyn, B. E., De Kauwe, M. G., Lin, Y. S., Knauer, J., Duursma, R. A., Williams, C. A., ... & Linderson, M. L. (2017). How do leaf and ecosystem measures of water-use efficiency compare? *New Phytologist*, *216*(3), 758-770

MÄKELÄ, A., BERNINGER, F., & HARI, P. (1996). Optimal control of gas exchange during drought: theoretical analysis. *Annals of Botany*, *77*(5), 461-468.

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., ... & Scott, R. L. (2016). The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature Climate Change*, *6*(11), 1023.

Kirschbaum, M. U. (1999). CenW, a forest growth model with linked carbon, energy, nutrient and water cycles. *Ecological Modelling*, *118*(1), 17-59.

Zhou, S., Yu, B., Huang, Y., & Wang, G. (2014). The effect of vapor pressure deficit on water use efficiency at the subdaily time scale. *Geophysical Research Letters*, *41*(14), 5005-5013.