This paper, submitted to the journal, Geoscientific Model Development (GMD), by Q. Li, Serbin, Lamour, Davidson, Ely, and Rogers, entitled “Implementation and evaluation of the unified stomatal optimization approach in the Functionally Assembled Terrestrial Ecosystem Simulator (FATES)”, is a well-written paper that could be accepted after mild revisions. The topic is important for understanding climate and land-surface processes better, and the modeling exhibited here is first rate. I detail my comments below.

I am impressed by the fact that the authors have started the FATES model runs with real-world forest inventory data, as stated on Line 115.

Line 136: “we set the precipitation to 1.47×10^{-5} \text{mm/s}” = 1mm/day? So it is always raining? Is this consistent with the humidity or VPD values of the model experiments? Is it consistent with the PAR values of the model experiments?

Fig 5.: There is not much difference in A_net or g_sw for the 3 days for either BWB_mean or MED_mean, even though the average-peak PAR increases from 700 to 1200 to 1500 mol/m2/s for the 3 successive days. This approximate independence of g_sw on PAR is what can be expected from Fig. 1a, for PAR > 500 mol/m2/s. But from Fig 1a, it might be expected that MED-default_mean and BWB_mean would differ by a factor of 2 in Fig. 5a. Is this Figure 5 actually for MED-B_mean instead of being for MED-default_mean? If it is, then the lack of difference between the modeled values for A_net or g_sw would make more sense.

Or should we be comparing to the ecosystem dependence shown in Fig. 2a, which shows little difference? I would expect the LICOR measurements to be done on a single leaf, instead of measuring over a larger ecosystem.

The case of PAR < 500 mol/m2/s seems to be handled robustly for the date of May 25,
in Fig. 8, where both BWB_mean and MED_mean are lower than the previous 2 days in May, particularly later in the afternoon on May 25. In this case of May 25, BWB_mean does seem to be 2 times higher than MED_mean, even in the morning, which might make a bit more sense if is for MED-default_mean instead of MED-B_mean, this time. On May 23 and on May 24, BWB_mean is 50% greater than MED_mean in the morning, but by mid-day, the models don’t differ much. Maybe the higher VPD that is reached by mid-day on May 23 and May 24 effectively closes the pores, causing the models not to differ? May seems different than (the dry season of) February - April, in that VPD is 0 kPa at night for May.

Lines 376-378: “Our method in keeping VPD in the air constant when studying model response to varying T_air by adjusting specific humidity concurrently is inspiring for other modelers.”
Such future inspiration of other modelers may indeed happen, but the language in this sentence is a bit presumptuous.

Line 619: citation for Pachauri et al. should have 51 authors instead of 10 authors.

Fig S2b: The r^2 value for the MED model is quite a bit lower than for the BWB model. Is this a real effect? Maybe the fit can be improved by removing a single outlier for MED at a value of Modeled g_sw = 0.24? It’s ok sometimes to remove outliers when doing fits. And that outlier seems unusual, too, since it is a MED point that doesn’t have a corresponding nearby BWB point like most of the other points do.