Delaria et al. report on a series of laboratory experiments to investigate whether chemical processing within the leaf limiting stomatal uptake of NO2 impacts the total NO2 exchange. Their laboratory experiments include testing the effects of nitrogen fertilization and drought. (I think the authors do these experiments to learn more about the in-leaf chemical processing, but it’s not actually very clear.) They also do some multilayer canopy modeling and backhand calculations to test the large-scale impacts of their laboratory results. In general, I find the authors to take too many liberties in moving between spatial scales and in discussing the implications of their work, and the paper to be confusing and disjointed. The discussion about what happens at night is interesting but feels tangential. I think the authors need to more clearly articulate their
experimental design and how it is designed to address current information gaps or conflicting studies, as well as better connect their conclusions with their results and discuss associated uncertainties. While the paper has potential to be an important contribution to the peer reviewed literature, I can’t recommend publication at this stage.

General comments. First, “deposition velocities” measured in the lab are not the same as deposition velocities from a large scale model or estimated from an eddy covariance measurement, which represent the integrated uptake below a certain height, taking into account turbulent transport. I would prefer if the authors chose another term to represent leaf-level uptake, but more importantly, this has implications for the authors’ large scale modeling and backhand calculations. Is it really appropriate to represent true deposition velocities with leaf-level uptake values? What about transport, leaf area, etc.?

The authors make a series of assumptions about resistances to the leaf boundary layer and cuticles in their interpretation of their laboratory results that I think need to be discussed more.

Are the authors maintaining constant temperature, pressure and humidity in the chamber over their forty minute long experiments? How might temporal variations in these quantities, or spatial variations within the chamber, affect measurements?

The canopy scale modeling and discussion in Section 4.1 is confusing. The authors do a fair amount of work in the lab to estimate Rm, and then say an increase from 0.1 s/cm to 0.6 s/cm in Rm doesn’t matter based on canopy scale modeling. The paper could have just been “Rm could be off by an order of magnitude. Does this matter? Let’s see with a model!” I guess I’m asking the authors to more clearly articulate how their setup was designed to build on present knowledge. For example, is the increase much less than they expected based on previous work?

Are there no boundary layer height products for California? I’d like to see at least some discussion of uncertainty in using only one PBL height for all of California for day or
The authors use “significant” to refer to statistical testing and to emphasize the implication of a finding. This is confusing and I ask that they choose another word for the latter.

In some paragraphs multiple verb tenses are used. This is confusing.

Line comments. Line 2 - is it really absorption? Line 11-12 - what do the authors mean by effective? Line 17 - references are needed for this sentence, and the authors should specify what importance is with respect to Line 19 - “after” diffusion rather than “via” Line 28 - are the processes really happening in the mesophyll? Line 35 - a paper from 2000 isn’t exactly recent Line 43-44 - “atmospherically relevant conditions” of what? Line 50 - define compensation point briefly here Line 135 - I think there needs to be a short description of Rb estimation here Line 215/219 - Rb changes with leaf morphology, leaf movement and micrometeorology. I understand Rb is hard to estimate, but I think the authors need to discuss how uncertainty in Rb may play into their results more. For example, how might inferences about stomatal and mesophyll controls be impacted by Rb variations (the authors assume constant Rb)? Line 205: is the only evidence for “believing” this measurement is consistent with a zero compensation point that the concentration is below the limit of quantification? If so, will the authors make this more clear? Line 206: I would be more careful in saying deposition of NO2 âĂŤ perhaps stomatal uptake of NO2 here âĂŤ deposition requires considering Rb,Ra, cuticular deposition Line 207-209: might this be affected by a lack of a diurnal cycle in light in the lab? I know there is evidence for stomatal activity at night generally, but maybe there should be some discussion of uncertainty in moving between the lab and the real world Line 210: It would be helpful if the authors explained what exactly to look for in Table 2 Line 211: the two methods don’t seem that different to me âĂŤ they are relying on the same assumptions âĂŤ seems just like two ways of presenting one method. Line 213: and assuming zero cuticular uptake? Line 230: First, “No significant cuticular resistances” implies cuticular uptake is happening. Second, how do the authors know
that there is no cuticular deposition when the authors are also inferring Rm? How can the authors know that the residual is Rm and not Rc? Also, I think the authors should spell out here what exactly they are suggesting that the Vd/gt ratio means (“attribute to” is a bit vague) and the assumptions involved Line 234 - spelling error Line 242-3: What do the authors mean “behave consistently”? Line 255: It would be helpful if the authors described what is observed as changing in the relationship between gt and vd, instead of just saying that there are changes and referring to a supplemental figure Line 263-6: I’m confused. My interpretation is that there is one slope for every plot in Figure 2. So how are the authors looking at a correlation between gt and the slope for each plot? The description of what the authors are doing on Lines 219-221 could be improved (“slopes were calculated from . . . slopes . . .”). Line 284-6: Not sure what to do with this information. Line 299-300: This seems like a rather broad conclusion based on the limited evidence that the authors have presented. Line 305-6: why is the fertilized group experiencing stress “supported by previous studies [finding] a negligible impact of N fertilization on NO2 uptake”? I think “these” should refer to the sentence before “We did observe . . .” but the writing is unclear. Line 308: uptake can’t ever be bidirectional Line 309: how do the authors know that there is actually accumulation in NO3 and NO2 within the mesophyll after fertilization? Is this from the leaf N measurements? Line 309: “neither . . . nor” (here and elsewhere) Line 310: what does “disproportionation” mean? Line 311: I’m not following why this “further supports . . . atmospheric unimportant” Line 330: I have no idea what the authors mean “atmospherically relevant”. What is/where is this discussed above? Line 340: The authors can’t move like this between lab and model “deposition velocities” Line 345: not true â˘AˇT see 10.1002/2016JD025519 Line 339: instead of saying the models assume this, it would be more appropriate to say the Wesely scheme assumes this. Line 346: Is the box model validated for nighttime chemistry and transport in forests? Line 350: What do the authors mean at such a low degree of stomatal opening? What does “statistically equivalent” mean? That they are similar in magnitude? Line 354: Is this a range in the NOx lifetime to deposition? Or the total lifetime? Also, it doesn’t seem like the authors show anything about lifetime in
Figure 6. Line 358: reference needed for major chemical nighttime sink as PAN Line 360-380: this is a lot of info to take in; please consider a table or a figure. Line 382: what are the significant inconsistencies? Line 390: seems like the authors need to say in June somewhere in the text (it’s only in the figure caption). Also, why June? What years are the authors looking at for LAI and NO2? Line 397: Why do the authors use maximum vd here? It seems like the implications of this need to be emphasized. Line 398: How does one multiply by “land cover”? What are the units of “land cover”? Line 395: How big are the Forest Service plots? Do the authors define forests with less than 50% of the trees measured in the study as “nonforested”? Are they included in white space on the figure? Line 396: clarify what the effective vd is Line 398: can one get midnight measurements of NO2 from OMI? Line 400: what is chaparral? Line 406: what is significant? Line 417: when do the authors look at vapor pressure deficit? Line 419: what does “from an atmospheric perspective” mean? Line 420: I wouldn’t encourage others to overlook the role of transport through turbulence and molecular diffusion at the large scale though Line 424: spelling error Line 421-5: does this really merit discussion in the very short conclusion? The authors look at different species because they have different stomatal conductances. For example, the authors say: “To test this, we measured . . . over a range of stomatal conductances” in the introduction. In other words, I feel like this was the motivation in setting up the study, not a conclusion of it. Line 436: can the authors briefly summarize here their evidence for “large and important”

Figures should be cleaned up to make them more appropriate for publication. The axis labels and tick marks should look better. Figure 2 - what data is included here? No N or drought perturbations right? Figure 3 - specify acronyms used in caption; if the authors briefly described here what we are supposed to take away from helium/zero air differences that would be helpful Figure 4 - if the authors said the meaning of Vd/gt ratio in their last sentence it would be even more helpful. Generally I’m not exactly sure how to interpret this figure â€” what should I be looking at in terms of NH4 and NO3? Figure 5 - spelling error; again helpful to say in plain language what a compensation
point is

Tables could use more context/description in general Table 2 - What does Rm (gt) vs. Rm (gs) mean? Are all compensation points statistically significant or just this one? There are two “e” in the footnotes. Table 3 - Define acronym for IQR