## Review of Klovenski et al. 2022

Klovenski et al. present research investigating the impact of drought stress on isoprene emissions and atmospheric composition in the NASA GISS ModelE Earth system model. They implement the drought stress parameterization from the MEGAN3 model, and apply a modelspecific tuning method to best reproduce observed isoprene fluxes at the Missouri. They then compare the results of their simulated drought impacts on atmospheric composition with observed formaldehyde and ozone concentrations. While the work is generally well written and addresses an important research topic, I cannot recommend the manuscript for publication in its current form. Major and minor comments affecting this recommendation are summarized below.

# **Major Comments**

**MOFLUX Point Comparisons** – Much of this work is based around tuning the emissions scheme to observations at one location: the MOFLUX field site. Tuning a global model to one individual site is suboptimal, but necessary in this case given the limited data available on isoprene emissions and drought. However, more work should be done demonstrating that this tuning is not compensating for substantial model errors that lead to differences in prediction. At minimum in order to assess the validity of this tuning factor, it would be useful to see the model performance on other variables necessary for predicting isoprene emissions. Example pertinent questions include:

Does the model properly simulate meteorological drivers of emissions at the MOFLUX site?

Does the land classification in the model match the observed site? Does the model properly represent vegetation properties (e.g., LAI, PFT, etc.)? How do the differences in simulated emissions and observations compare with the substantial uncertainty estimates in the MEGAN model (Guenther et al., 2012)?

The scatterplot in Figure 2 and associated discussion shows a very limited improvement in R<sup>2</sup> (0.03). Since this work is entirely focused on water stress, it would be useful to have those metrics only for the water stressed time periods.

# **Model Tuning**

The authors recommend applying the drought stress when the model grid cell water stress is in below the 4<sup>th</sup> percentile. I understand the logic for this choice, but it should be contextualized further. Is there reason to expect that vegetation respond to relative or absolute water stress? Is this tuning representative of any physical or biological process, or simply statistical?

**Formaldehyde Comparisons** – The analysis of formaldehyde retrievals in this work may be lacking relative to the state-of-the-science and does not support the conclusions in the manuscript.

For apples-to-apples comparisons between satellite observations of formaldehyde and models, the Air Mass Factor (AMF) should be recalculated and applied to the observations. That was not

done in this work. At the very least that point and the associated limitations imposed should be discussed.

The ModelE simulated formaldehyde column disagrees substantially with observations (e.g., Figure 6). It appears as though the column is overestimated by at least a factor of 3. This enormous overestimation is not common across other models of atmospheric chemistry (e.g., GEOS-Chem), and calls into question the validity of ModelE simulated formaldehyde concentrations. While adding the drought stress does improve the simulation, that result alone is not interesting as anything that reduces formaldehyde concentrations would improve the simulation. The authors should make a stronger case as to why the ModelE formaldehyde simulations should be trusted as a useful assessment tool for isoprene emissions changes.

## Statistical significance of results

Many of the results here are lacking detailed statistical treatment to understand if the results are either statistically or practically useful, in particular Figure 2, Figure 4, Figure 5, and Figure 7. All these figures and associated discussion describe noisy results. I am sympathetic to the challenges related to non-drought variability in constraining the process the authors are addressing, but substantially more rigorous assessment is needed before these results can be assessed in depth.

For example, the trends shown in Figure 4 do not appear to be significant in any way. They visually look to be a random sampling of points scattered about y = 0, and do not appear to show "decreasing isoprene emissions" as claimed in the text.

### **Minor Comments**

The introduction includes substantive discussion of drought impacts on SOA, but the analysis does not assess SOA at all, this is confusing.

L548: How was the selection of an alpha value of 100 made? "Best fit" by what metric?

L657: "there is model agreement" is far too strong of a statement given the large scatter in Figure 2.

L695-698: This sentence around HCHO overestimation is confusing. Are there any reasons why the authors suspect these reasons are the culprit for overestimation? If so, why were they not addressed in more detail during model development?

L747: "clear decreases" are not evident given the substantial variability in the figure. See discussion of statistical significance above.

L823: These changes in HCHO and O3 are very small relative to the large biases that still exist, particularly in rows 2 and 3 of Figure 7.