Interactive comment on **"A statistical approach for estimating representative emission rates of biogenic volatile organic compounds and their determination for 192 plant species/genera in China"** *by* **Lingyu Li et al.**

Anonymous Referee #1

I think that having this kind of inventory for VOC emissions from a large region is beneficial for improving air quality models and better constraining other biogenic emissions and earth system models (e.g. CESM, MEGAN). But it appears to me that this work is not sufficient to be much more than qualitative estimates. I know how difficult this work is, and with limited time and resources, it's impractical to get accurate emission rate data from a large number of tree species in a single season. And what they have done, is a reasonable attempt to at least break down a large variety of species into different classes. Their statistical techniques are interesting and are useful for establishing these ranges. Below are a few of concerns regarding the accuracy of the emission

estimates: 1. The authors discuss the volume of the bag and talk about a "semi static" enclosure. They give some dimensions of the bag of 160 cm x 90 cm in the shape of a rectangle and give a volume of 400 L. When I do the calculation using either a cylinder or a rectangular box, I get a volume closer to 1000 L. But if I do take the 400 L volume and their flow rate of 6 L/min into the bag, the residence time is 400/10 = 40 min. For the system to be in steady state, there needs to be ~3-4 turnovers of air, which would take at least two hours. The way they get around this (non-steady state nature of the system) is to determine how much residual air is in the bag by measuring acetylene before and after zero air is added. A typical background level of acetylene is ~500-1000 ppt, so I presume the initial concentration in the bag was approximately this amount. Then the bag was filled with zero air for 6 minutes using 10 L/min and then another 2 L/min for 3 minutes. The total volume displaced would then be 60*6 + 2*3 = 66L, which is 16.5 % of the total 400 L volume. So the concentration of acetylene after this step would be ~(100-16.5)*C0 or ~84% of the original 500-1000 ppt. What is the uncertainty in a measurement of 1000 ppt. vs. 835 ppt? I think the authors really need to describe the precision of their measurements and give some results of this step, as it's important if they are going to use equation 2 instead of waiting for a steady state condition. Uncertainty in this number could result in large uncertainties in emission rates. 2. The samples were taken to be generally in mid-day sun at ~30 C. And then standard algorithms were used to "normalize" the emissions to a set of standard conditions (e.g. 30C). As far as I can tell, there is just one sample per tree or vegetation species, and some of the samples were taken very late or early in the growing season. I think there needs to be more discussion on how seasons, ambient conditions, and number of samples per species influences uncertainty. 3. In their evaluation and conclusion section, the authors discuss other uncertainties, and the recommendations of Niinemets et al. (2011) and readily admit that their measurements do not adhere to these guidelines. They give examples of some emission rates of 838, 707 and 2542 ugC/gdw/hr. Seeing that these are orders of magnitude greater than other reported rates seems like it should have made the authors more skeptical of their other

results. The other thing that seems to be lacking in this paper is biomass distribution. It's not practical to define the whole country, but I think it's necessary to give the reader some context as to how common these different plants/trees are and how much of the total leaf area can be explained (at least in certain regions) by this listing. 4. They discuss the emissions of several other VOC compounds that are not commonly considered in biogenic emission samples. Specifically alkanes and aromatics (e.g. xyxlenes, propane, ethane, isopropylbenzene). I am not familiar of any biogenic emissions of these compounds or biochemical pathways for form them within the leaf structure. It seems like these might be artifacts of the GC system, but without knowing more details and/or seeing results from blank samples, I can't tell. But emissions of these compounds from vegetation are not commonly reported. Since their work focuses primarily on monoterpenes and isoprene, I would suggest to just omit these compounds, as they don't seem to add any valuable information, and their presence is suspicious. My analysis of this work is that it's a nice start, but the uncertainties are very large (and not sufficiently addressed) and these are not sufficient to extrapolate emissions to other times of year or different individuals within the same plants species. Other: The classifying of the emissions into different categories (low, medium, high, etc.) based on statistical distributions of the emissions is valid. And I think this approach is useful for model inputs where the modeler could input a certain mean (+/- range) of emissions based on the species distribution and leaf area index. I just question the accuracy of the emissions for the reasons and examples cited above. Figure 1 looks suspiciously like Figure 1 in Ortega et al. (Chemosphere, 72, p. 365, 2008). Tables 1 and 9 are almost redundant. The authors frequently refer to "pinene", which seems useless to me. I think they should distinguish between this and general Monoterpenes, or specify if they mean alpha pinene, beta pinene, or the sum of the two.