The study conducted by Ren et al. is looking into FIGAERO-I-CIMS thermograms, more precisely the maximum temperatures of the first peaks of the thermograms (Tmax). The study starts with a suite of laboratory experiments with PEG samples of different volatilities which they either inject on the FIGAERO filter using a syringe or nebulize, dry, dilute and collect onto the FIGAERO filter. They attain similar results as Ylisirniö et al. (2021), reproducing a quantitatively similar relationship between saturation vapor pressure (Psat) and Tmax. Ren et al. investigate this relationship further by making mixtures of various PEG, citric acid and erythritol with ammonium sulfate (AS), either as a mixture between one organic component and AS or as one mixture between all organic components and AS. These were then deposited on the FIGAERO filter following the atomizer method. A discrepancy could be observed between the Psat -Tmax relationships derived from these two experiment types. They attribute the mismatch to either AS-derived effects such as organic salt formation, viscosity limitations in evaporation or matrix effects. Finally, the authors utilize the Psat -Tmax relationship derived from the mixture calibration involving all organic compounds compounds to calculate saturation vapor concentrations (C*) for their field data. They further noticed that these Tmax-derived C* displayed as a function of molecular weight showed two groups/clusters with characterized with different O:C-ratios. The authors finally derive two molecular formulae-based C* parameterizations for these groups, respectively. The results are compared to other molecular formulae-based C* parameterizations.

The manuscript could be potentially very useful for the FIGAERO community if the calibration results were investigated further and the reasons behind the mismatch could be narrowed down and the potential influence of inorganic salts or matrix effects on Tmax and therefore C* in the field could be assessed. It would be useful if the authors could carefully evaluate how reliable the C* derived from Tmax are under environments with high mass loading and inorganic salt concentrations. I find this necessary before parameterizations are being derived from these Tmax – C* relationships. I recommend publication after major revisions.

Main comments:
I think it would be crucial to understand/narrow down what actually caused the change in the Psat -Tmax relationship when comparing the line derived from the single PEG+AS mixtures vs one solution. I would suggest you to perform more laboratory experiments such as: 1. A single organic mixture (all organic compounds included) without AS, 1000 μg deposited on filter with the atomizer method; 2. Replicating No.5 experiments and probing the effect of mass loading.

The authors should also provide information of the size distribution and temperature for the different experiments, especially for No.4 and No.5. It should be noted that Ylisirniö et al. (2021) mention their significant role in causing discrepancies in the Psat -Tmax relationship.

The authors should at least provide the 95% credible intervals along the fitted lines in Figure 1 if showing all the data points (replicates) decreases the readability of the graph. It would be useful to see how much scatter there is between replicates. This scatter could even hold some information about possible loading effects.

The authors should think about providing a schematic of their laboratory setup. They mention the possibility of organic salt formation in No.5 mixture that could cause the increase in the observed Tmax values when compared to No.4 experiments. If the sample is dried immediately after the nebulizer there is not much time for any organic salt formation under favorable conditions (high humidity and aerosol liquid water content). Do the authors think there is time for such reactions to actually happen?

After understanding the significance of the matrix effects on the Psat -Tmax relationship in the calibration data – do the authors still recommend deriving Psat from Tmax? Do the authors observe variability in Tmax in the ambient samples that covary with mass loading? How much variability was there in the mass deposited on the 30 filters analyzed and how does it compare to the 1000 μg calibration reference? How would the matrix effects from ambient samples affect the predicted C*?

It is unclear to me whether the FIGAERO was measuring in real time during the field campaign or whether the filters collected (described in Sect. 2.3) were measured with the FIGAERO offline. Could you please clarify.

Minor/technical comments:

- L34: is the Nizkorodov et al. (2011) the best reference for this statement?
- L40-: The descriptions of the past methodologies are incomplete. The description is for example missing classic work with thermodenuders (TD) without CIMS that have been mounted as part of tandem volatility differential mobility analyzers (TDMA) or coupled with an AMS (TD-AMS). The introduction also lacks description of the way the thermograms measured by the TDMA or TD-AMS are being modelled to gain information of C* or VBS (see for example Cappa, 2010, in Atmos. Meas. Tech or Cappa and Jimenez, 2010, in Atmos. Chem. Phys.). In addition, dilution experiments among many others should not be forgotten. The authors seem to be citing more the work for predicting equilibrium partitioning coefficients than actual C* measurements. If the authors wish the provide a list of methods used previously, they should cite more relevant literature and make sure to include a complete description. Alternatively, they could focus on describing just the methods relevant for CIMS.
- L322-L323 Figure 4 does not contain a red or blue dashed circles, maybe ellipse would be a better word. I found this confusing at first.