Review of "The potential role of organics in new particle formation and initial growth in the remote tropical upper troposphere" by Kupc et al

The authors present box model studies of new particle formation representative of the remote Pacific upper troposphere sampled during ATom flights. They find the nucleation mechanism of Riccobono et al (2014) represents observations reasonably well. While not well constrained by measurements, organic vapor concentrations had to be at least twice SO2 concentrations to reproduce observed NPF and growth rates. Studies like this will be of interest to the community, and are timely relative to the ATom measurements.

The paper is well structured and well-written and is clearly within the scope of ACP. I recommend publication if the authors can address the following comments.

Minor comments

One very long comment about condensation sink.

Nucleation varies very non-linearly with condensation sink, as highlighted very nicely in several papers by Pierce et al among others. I think it is necessary to emphasise the implications of the variability in CS more explicitly. The measured condensation sink (CS) varies by a lot – more than an order of magnitude. The generally poor correlation of the CS to nucleation-mode particle concentrations is explained in part because the CS probably varies significantly along the trajectories, and the CS at the end of the trajectory is not the same as the CS where nucleation happened. Indeed, if I interpret Figure 7b correctly, it appears that nucleation-mode particles dominate the sink in the case shown. Therefore, it's unclear whether the correlation of observed number concentration to CS should be positive or negative. So I am not sure if Figure 5b is the most helpful way to present the data. Could you try splitting the data in Figure 5b into cases where CS is dominated by 7-12nm particles and cases where it isn't? Maybe when the CS is dominated by small particles it will be positively correlated to number concentration and when not, negatively? I realise that if the CS is always dominated by 7-12nm particles, then this won't be useful.

As another way of digging deeper into this, would it be useful to plot the simulated CS at the time along the trajectory when the simulated nucleation rate is maximum, against the simulated number concentration at the end of the trajectory, just to see if there is a stronger negative correlation? Then maybe compare that to the simulated CS at the end of the trajectory plotted against the simulated number concentration at the end of the trajectory?

I found your Figure 7 very helpful as it shows when nucleation happens along the trajectory. If vapor concentrations and temperature were prescribed to their average values, nucleation would always happen at the start of the trajectory, as nucleation can only ever increase the condensation sink so the most likely time for it to happen is at the beginning of the simulation – correct? However since temperature and H2SO4 concentrations vary along the trajectory, this doesn't need to be true. So is figure 7 typical? Do you sometimes see nucleation only a few hours after the start of the simulation? Is the CS/surface area always dominated by nucleation-mode particles?

At line 470-477, you say that variation in the background aerosol doesn't help the Vehkamaki et al nucleation scheme, which is the scheme that is best constrained by your measurements. But in Figure S52 I see some blue dots. Perhaps (if I'm not wrong) you could say that if the CS were smaller by some factor, the Vehkamaki scheme could explain the observations, or could get reasonably close (NME=0.14), as shown in Figure S52? A NME of

0.14 for Vehkamaki might be just as 'good' as a lower NME for the schemes that include organics and NH3, since these vapors are not constrained by your measurements. Can you consider promoting this figure, and perhaps S51 for the RIC scheme, to the main text, as well as improving the caption (to refer to Table 1 where the various abbreviations are defined) and including a more extensive discussion?

Other minor comments

A mention of the update to the Vehkamaki (2002) nucleation scheme by Maattanen et al (JGR, 2017; https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017JD027429) is warranted. The update is valid over a wider range of temperature and humidity, which is relevant here. The new model is also validated against CLOUD measurements, includes ion-induced nucleation, and updated quantum chemistry calculations. Similarly, the ternary H2SO4-NH3 scheme of Yu et al, GMD 2020, has look-up tables ready for a model implementation: https://gmd.copernicus.org/articles/13/2663/2020/gmd-13-2663-2020.pdf . I think there is no need to update the schemes used for this paper, but perhaps they could be mentioned as worth exploring in future work?

What motivates the night-time OH value of $1x10^5$? Production from ozonolysis of terpenes? It seems like quite a high value to me.

It is stated that trajectories generally stayed at high altitudes (line 387). How does the trajectory model handle deep convection? Presumably it is parameterized? Your Figure 3 shows some evidence that some trajectories do originate in the boundary layer, which seems reasonable, but are there not some big uncertainties here even with the ensemble approach adopted? It would be worth referencing and perhaps discussing Andreae et al (2018) (the paper in ACP, already cited) in this context, as they do a similar trajectory analysis in their paper.

Line 424 Atlantic is misspelt.

How does RH vary with time since convection? I'm guessing it doesn't vary very much and so is not super-important, but would be good to confirm this. I note it is represented in your simulations in any case (line 324)

Supplement S2, line 457, and Figure 6 : Could the rather strange size-bin emptying behaviour be avoided by shortening the timesteps of the box model?
Also, is size bin emptying actually responsible for the rather strange kinks in the size distribution frequently observed around 10nm, and which correspond to a much smaller kink in the observed size distributions? In Figure S25, this is especially apparent. The gaps in the size distribution are far larger than a single bin. Do you have some ideas for what might be happening here?

L618: "Nucleation schemes are simply imperfect parameterizations extrapolated from laboratory observations": this is a bit too much generalization; it's not quite true for Vehkamaki (2002) for example. On the other hand, the RIC scheme you prefer to use is one of the more uncertain parameterizations from the CLOUD experiment, because the organic molecules thought to participate in nucleation were not measured directly in the chamber when that parameterization was developed. For example, in the conclusions, perhaps it's worth pointing out that it is not surprising that one of the least well constrained nucleation mechanisms – the RIC scheme, which is the one you can tune most easily by changing the organic concentration – can be made to agree with observations the best.

L1037 several references are merged together.