Response to the Reviews of ”The potential role of organics in new particle formation and initial growth in the remote tropical upper troposphere”

Response to the Anonymous Referee #1

The authors thank the reviewer for helpful comments that have improved the manuscript. The discussion below contains the original text by the reviewer and our responses (in blue) along with changes made to the revised manuscript. Additional minor changes to the manuscript are indicated at the end of this document.

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 emphasize 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.

These are really good points. The reviewer is right that from Figure 7 both measured and simulated nucleation mode dominated the CS in that specific case, which could lead to a positive correlation between nucleation mode particle number concentration and CS. To clarify that we have remade Figure 5a and 5b to show the CS for particles >12 nm (instead of CS> 7 nm; updated plot is shown below) as a better estimate of the CS prior to nucleation starting. The correlation remained negative and the value of the correlation coefficient, r², increased as compared to the initially presented
CS>7 nm data. When CS is dominated by small particles (CS_3-12nm) the correlation is strongly positive ($r^2$ between 0.97 and 1, Fig. 5c).

One of the assumptions we have made in simulations is that there is no nucleation mode at the start of the trajectory (at the exit of the cloud). Thus, we eliminated all particles with diameters < 12 nm to initiate the model. Figure 5 below is shown in the updated version of the manuscript.

The following text was updated (line 426; changes to the existing text indicated in italic)

“The CS term is calculated for particle diameters >12 nm, and for the diameter range between 3 and 12 nm following Williamson et al. (2019). CS_{>12nm} serves as an estimate of the condensation sink prior nucleation starting, and it is negatively correlated with the number of nucleation mode particles (Fig. 5b), while the nucleation mode is positively correlated with CS_{3-12nm} (Fig. 5c). Over the Atlantic, the maximum concentration of nucleation-mode particles >8 km in altitude averaged from 30° S to 30° N, ~3,000 cm$^{-3}$, is considerably smaller than over the Pacific, but the shape of the profile is similar (Fig. S11).”

Line 436: “Some variability in the strength of NPF and its dependence on CS_{>12nm} can be observed. In general, CS_{>12nm} is weakly negatively correlated ($r^2$ between 0.08 and 0.36 depending on the ATom mission) with the concentration of nucleation mode particles (Fig. 5b), as would be expected if NPF were competing with CS for condensing vapors. When CS is dominated by small particles (CS_{3-12nm}) the correlation is strongly positive ($r^2$ between 0.97-0.1, Fig. 5c).”

We updated the Figure S11 for the Atlantic case in the Supplemental Material (SM) accordingly.
Figure S11: Vertical profile of the median number concentration of nucleation mode particles (3-12 nm) and condensation sink (CS>12 nm) averaged between 30° S - 30° N as a function of altitude for the four ATom deployments. (b-c) One minute average nucleation mode particle concentrations at >8 km in altitude as a function of CS>12nm (b) and CS3-12nm (c) and between 30° S - 30° N over the Pacific Ocean. Pearson correlation coefficient values (r²) are indicated in the legend.

We also checked whether there were instances when CS3-12nm dominated over CS>12nm and found out there were only 7 instances where CS3-12nm exceeded CS>12nm values.

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?

We chose to follow the first of these two alternate suggestions by the reviewer.

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?

Yes, that is correct.

However, since temperature and H2SO4 concentrations vary along the trajectory, this doesn’t need to be true.

Yes, and the reason for this is because the OH is not necessarily at its highest concentration at the start, which, along with the condensation sink, determines the pseudo-steady-state concentrations of H2SO4 and condensable organics.

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?
In all (32) cases, nucleation starts early in the simulation (within the first 20 minutes). The nucleation, however, can have different strengths. In some cases, although nucleation starts early, the maximum nucleation rates are achieved later as the nucleation rate may increase with time. Some nucleation events are very weak, with dN/d\log_{10}D_p never exceeding 10 cm^{-3} in any of the size bins smaller than 12 nm.

In Table S6 below we show a summary of the measured aerosol surface areas for each of the studied cases. The highest surface area values for each case are shaded. These values are split between nucleation, Aitken, and accumulation modes indicating that the case shown in Figure 7, while not representative of all cases used in simulations, is typical. Table S6 is now added to the SM. A few examples of cases where the value of aerosol surface area was either biggest for Aitken or accumulation mode are presented in Fig. S68-70 below and in the SM.

The following text was added to the revised manuscript: Line 531: “In the case presented in Figure 7b, the surface area measured on ATom is dominated by the nucleation mode particles; however, although frequent, this is not a typical pattern among all other cases studied here. In general, in the cases with the highest surface area, values are split between nucleation, Aitken and accumulation modes almost equally (Table S5, Fig. S68-70).”

Table S5. Summary of the measured aerosol surface area for each of the studied cases. The highest surface area values for each case are shaded.

<table>
<thead>
<tr>
<th>ATom2 Case identifier</th>
<th>Surface area, µm² cm⁻³</th>
<th>Nucleation mode 3-12nm</th>
<th>7-12nm</th>
<th>Aitken mode 12-60nm</th>
<th>Accumulation mode &gt;60nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>sd486 (Fig.7 in the manuscript)</td>
<td>4.71</td>
<td>1.90</td>
<td>1.00</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>sd390</td>
<td>2.89</td>
<td>0.89</td>
<td>1.87</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>sd391</td>
<td>2.74</td>
<td>0.64</td>
<td>1.54</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>sd400</td>
<td>1.14</td>
<td>0.25</td>
<td>1.88</td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td>sd404</td>
<td>2.47</td>
<td>0.76</td>
<td>1.98</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>sd446</td>
<td>2.18</td>
<td>0.91</td>
<td>3.44</td>
<td>1.66</td>
<td></td>
</tr>
<tr>
<td>sd448</td>
<td>2.05</td>
<td>0.68</td>
<td>2.95</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td>sd452</td>
<td>2.63</td>
<td>0.90</td>
<td>2.91</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>sd461</td>
<td>1.37</td>
<td>0.36</td>
<td>2.39</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>sd470</td>
<td>0.95</td>
<td>0.47</td>
<td>3.35</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>sd477</td>
<td>3.33</td>
<td>1.21</td>
<td>2.04</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>sd491</td>
<td>2.79</td>
<td>0.48</td>
<td>1.29</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>sd496</td>
<td>4.22</td>
<td>1.32</td>
<td>1.79</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>sd498</td>
<td>3.05</td>
<td>0.82</td>
<td>2.13</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>sd530</td>
<td>0.89</td>
<td>0.37</td>
<td>1.95</td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>sd533</td>
<td>2.13</td>
<td>0.95</td>
<td>1.88</td>
<td>3.33</td>
<td></td>
</tr>
<tr>
<td>sd535</td>
<td>2.23</td>
<td>1.10</td>
<td>1.61</td>
<td>2.49</td>
<td></td>
</tr>
<tr>
<td>sd537</td>
<td>2.25</td>
<td>0.95</td>
<td>1.40</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>sd540</td>
<td>0.53</td>
<td>0.07</td>
<td>0.90</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>sd543</td>
<td>2.71</td>
<td>1.43</td>
<td>1.99</td>
<td>1.62</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATom4 cases</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sd22</td>
<td>0.44</td>
<td>0.33</td>
<td>1.05</td>
<td>1.41</td>
</tr>
<tr>
<td>sd32</td>
<td>0.50</td>
<td>0.37</td>
<td>0.90</td>
<td>1.12</td>
</tr>
<tr>
<td>sd72</td>
<td>0.37</td>
<td>0.31</td>
<td>1.40</td>
<td>1.75</td>
</tr>
<tr>
<td>sd75</td>
<td>0.86</td>
<td>0.62</td>
<td>1.11</td>
<td>1.36</td>
</tr>
<tr>
<td>sd78</td>
<td>0.31</td>
<td>0.22</td>
<td>0.94</td>
<td>1.56</td>
</tr>
<tr>
<td>sd82</td>
<td>0.37</td>
<td>0.31</td>
<td>1.69</td>
<td>2.26</td>
</tr>
<tr>
<td>sd132</td>
<td>0.09</td>
<td>0.06</td>
<td>1.13</td>
<td>1.76</td>
</tr>
<tr>
<td>sd134</td>
<td>0.14</td>
<td>0.09</td>
<td>1.94</td>
<td>2.03</td>
</tr>
<tr>
<td>sd179</td>
<td>0.84</td>
<td>0.56</td>
<td>0.38</td>
<td>0.28</td>
</tr>
<tr>
<td>sd183</td>
<td>0.17</td>
<td>0.08</td>
<td>0.52</td>
<td>0.40</td>
</tr>
<tr>
<td>sd186</td>
<td>0.05</td>
<td>0.02</td>
<td>0.445</td>
<td>0.436</td>
</tr>
<tr>
<td>sd188</td>
<td>0.07</td>
<td>0.02</td>
<td>0.26</td>
<td>0.22</td>
</tr>
</tbody>
</table>

TOTAL number of cases with highest surface area in each size range: 11 0 10 11
Figure S68: as in Fig. 7 but for case 461 (NME=0.08). Measured on ATom aerosol surface area is dominated by the Aitken mode (b). Initial SO$_2$=1.6 pptv, and initial organics=23.4 pptv.

Figure S69: as in Fig. 7 above but for case 470 (NME=0.28). Measured on ATom aerosol surface area is dominated by the Aitken mode (b). Initial SO$_2$=1 pptv, and initial organics=3.4 pptv.

Figure S70: as in Fig. 7 but for case 530 (NME=0.13). Aerosol surface area measured on ATom is dominated by the accumulation mode (b). Initial SO$_2$=1 pptv, and initial organics=7 pptv.
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?

The blue dots in Figure S52 correspond to the VEHK nucleation scheme with organics added for the initial particle growth. The best NME=0.14 here was obtained for SO2=13.9 pptv and organics=7.2 pptv when the measured size distribution (dN/dlog_{10}D_p) >8 nm (instead of standard setting of dN/dlog_{10}D_p >12 nm) was used as initial background aerosol to initiate the simulation.

We have added the following text to the manuscript, line 481:

“Varying the pre-existing initial aerosol or completely removing background particles in the VEHK scheme with added organics may improve the fits for certain initial conditions (Fig.S52), making it plausible for better NME values if CS>12nm was 4.12x10^{-5} s^{-1}, and SO2 and organics were 13.9 pptv and 7.2 pptv, respectively, for this particular case. Out of 32 cases studied here, there is no case with NME <0.2 for VEHK scheme. Once we add organics for the initial growth the NME improves resulting in 6 cases with NME<0.2 (Table S4).”

The caption has been updated and reads now: “Figure S52: The effect of varying the initial background aerosol on the NME in TOMAS for case sd486 (ATom2, 2017-02-03, 03:05:31-03:06:30 UTC) and VEHK nucleation scheme and VEHK with organics added for initial growth. OH at solar zenith angle of 0° was set to 3x10^6 molec cm^{-3}. SD refers to the number size distribution dN/dlog_{10}D_p. Table 1 in the main manuscript describes the abbreviations used here.”

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?

In Section 3.3. Discussion line 647 in the revised manuscript we added (in italic here) to the existing text: “We also have not explored the organic-only nucleation scheme described by Kirbky et al. (2016), an updated version of the Vehkamaki et al. (2002) scheme covering a wider range of temperatures and relative humidities by Maattanen et al. (2017) that has been validated against
What motivates the night-time OH value of 1x10^5? Production from ozonolysis of terpenes? It seems like quite a high value to me.

We have chosen a low non-zero value for night-time OH in TOMAS based on our best estimate at the time. There is no night-time data on OH for the tropical upper troposphere during ATom, and 1x10^5 molec cm^{-3} is the approximate detection limit of the instrument. The night-time OH value of 1x10^5 molec cm^{-3} is an order of 10x smaller than the day-time values and the nucleation and growth of particles simulated in this study come almost entirely from the day-time. Thus, lowering the night-time value further should not change the result.

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.

The Traj3D model does not include an explicit convective parameterization. It uses the vertical velocities from the input reanalysis fields, in this case 0.25 deg resolution NCEP GFS, which parameterizes subgrid convection. Thus, the trajectories may not simulate the rapid transport that a specific parcel might undergo in a convective event, but rather to capture the general upward motion at the reanalysis scale.

We added the following text in line 402: “Similar trajectory analyses, in terms of examining the history of the sampled air masses for interactions with deep convection, have been undertaken by Froyd et al. (2009) and Andreae et al. (2018).”

Line 424 Atlantic is misspelt.

corrected

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)

Yes, that is correct. The RH along the trajectory is represented in our simulations. The shorter the time since convective influence (CI) the smaller the change in RH between t_0 (start of the trajectory, at the point air left the cloud) and t_{fin} (at the point of aircraft location/measurement) (Fig.S67 below). We added Figure S67 to the SM.
Fig. S67. The difference in relative humidity (RH) along the trajectory between the trajectory location at the cloud edge \((t_0)\) and the point of the aircraft location \((t_{fin})\) for 32 simulated cases.

We added the following text in the manuscript line referenced above: “The change in RH along the trajectory between the trajectory location at the cloud edge \((t_0)\) and the point of the aircraft location \((t_{fin})\) for 32 simulated cases is presented in Fig. S67.”

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?

The bin emptying will not go away with a shorter timestep, and it is a common feature of 2-moment microphysics schemes that use a top-hat or moving-center approach for condensational growth. We have added these above sentences to the SM (line 457).

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?

The gap is from zeroing out the nucleation mode in the initial size distribution. We start with no particles below 12 nm in the simulations, and the new particles have not grown past a few nm, creating a gap between a few nm and 12 nm.
The blue line in Figure S25 represents output from the MAIA model, which does not experience the issues with bin emptying. We do not apply any smoothing in MAIA. In MAIA we use surface area of the Aitken and accumulation mode as initial background aerosol (size distribution on the right in the plot). Nucleation mode particles that appear in the simulations have not grown yet to bigger sizes, and that is the reason for the gap between these modes.

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.

We deleted this sentence and updated the sentence in line 713 by additional text: “Also, the Riccobono scheme, one of the least constrained nucleation mechanisms, required a large extrapolation in temperature to simulate UT conditions.”

L1037 several references are merged together.
Corrected.

Response to the Anonymous Referee #2

The authors thank the reviewer for helpful comments that have improved the manuscript. The discussion below contains the original text by the reviewer and our responses (in blue) along with changes made to the revised manuscript. Additional minor changes to the manuscript are indicated at the end of this document.

This manuscript tackles a globally important research topic: formation of new aerosol particles in the tropical upper troposphere. The paper is essentially a sensitivity study, aiming to give new insight into which nucleation mechanisms and aerosol precursors, coupled with the initial growth of newly-formed particles, best explain the observed ultrafine particle number size distribution. The conducted analysis is based on box model simulations and statistical analyses of the simulation results. The paper is scientifically sound and relatively well structure. I do have, however, a few issues that should be addressed before the paper is ready for publication.

My major criticism is related to the treatment of aerosol processes in convective outflow regions. As the authors state, they simulate outflow regions of deep convective clouds using a box model.
This is fine as long as both nucleation and growth occur well beyond the region abound cloud boundaries where most of the mixing between cloud outflow and upper troposphere air take place. This may not be the case, as it is quite possible that nucleation and early particle growth take place in the mixing region, or event inside the convective clouds. This would seriously bias the results obtained in the paper. I understand that including the cloud and its immediate outflow region in a box model is almost impossible, and therefor outside the scope of the current paper. However, the authors should bring up this issue more honestly as done in the present paper (brief mentioning on lines 628-629). Furthermore, there are a number of both modeling and observations studied conducted on new particle formation in cloud outflow regions. The authors should better acknowledge such studies when discussion their results, their implications and the associated uncertainties.

More details have been added in section: 3.3 Discussion, line 661:
“The box model used here simulates NPF in the outflow region of deep convective clouds. Although, active NPF was identified in the vicinity of clouds and in the cloud outflow region in many studies (such as Perry and Hobbs 1994, Clarke et al., 1998, 1999; Ström et al., 1999; Clement et al. 2002; Twohy et al., 2002; Weigelt et al., 2009; Waddicor et al., 2012; de Reus et al., 2001; Clarke and Kapustin 2002; or Andreae et al., 2018), the exact location of NPF with respect to cloud remains uncertain (Kulmala et al., 2006; Waddicor et al., 2012). We assume that nucleation does not occur within the cloud, and that the outflow does not immediately mix with the surrounding air in the highly stratified upper troposphere. If NPF were to occur within the cloud or in a zone of turbulent mixing at the cloud edge, as suggested by some studies (Ström et al., 1999; Lee et al., 2004; Weigel, 2011; Kazil et al., 2007; Kulmala et al., 2006), our results would be biased.”

The references mentioned in the above paragraph were added to the list of references.

Comment: Other, minor issues:
The right parenthesis is missing from line 432.
Corrected.

The text on lines 538-543 is not logical. When discusses nucleation mechanism not involving NH3 at all, it is incorrect to say "regardless of NH3 concentrations", as the outcome of such mechanisms does not depend on NH3 concentrations. Please correct.

Thank you for noticing that. We have corrected this sentence and moved it to lines 545-547. The sentence reads now: “Further, regardless of the available SO2, the results strongly suggest that binary sulfuric acid-water nucleation, whether ion-assisted or neutral, and whether coupled with organic growth or not, generally cannot explain the ATom observations.”

The statement on lines 546-548 sounds a bit strange. Is really so that the concentration of both SO2 and organic precursors need to be smaller than some upper limit values to reproduce the observations? Please check out this statement and modify if needed.

Our results show that in order to reproduce observations, the initial conditions of SO2 were found to be within the range of mixing ratios measured on ATom 4 and in other studies (Rollins et al., 2017, 2018). While the initial conditions of organics, not measured on ATom, were found to be
within the chosen plausible range of initial mixing ratios. We needed to set the lower and upper limit for the initial conditions used in simulations and tried to make sure their values were either based on measurements or literature, and if the species were not measured, that the limits are at least plausible.

The SO$_2$ and NH$_3$ mixing ratios were varied between 1 and 100 pptv to explore a large range of plausible conditions. The evaluated SO$_2$ range exceeds that measured on ATom 4 (Supplemental Material Figs. S4 and S5) and covers the <30 pptv mixing ratios reported in other studies in the UT over the central and western tropical Pacific (Thornton et al., 1997; Rollins et al., 2017, 2018). Organic aerosol precursors are unknown in the UT and were not directly measured; thus, we explored a range of probable mixing ratios between 1 and 100 pptv (of precursors with an SOA yield of 1).

We highlight that the simulated size distributions that show good agreement with observations (NME< 0.2) are characterized by rather low initial mixing ratios of SO$_2$. These low mixing ratios are within the range of SO$_2$ mixing ratios measured during ATom (< 30 pptv) at these locations.

We state that: “We note that we performed no simulations with mixing ratios of SO$_2$ or organics above 100 pptv. While we cannot exclude that for some cases the mixing ratios of these precursors at levels above 100 pptv could improve fits, these levels are outside of prior observations so were not considered in this study.”

We modified the statement as follows: “The findings for the case of organic-mediated NPF are summarized in Fig. 8, where we show the SO$_2$ and organic precursor mixing ratios for all sensitivity simulations using the RIC scheme, highlighting the assumptions that yielded the lowest NME for each case.”

Response to the Anonymous Referee #3

The authors thank the reviewer for helpful comments that have improved the manuscript. The discussion below contains the original text by the reviewer and our responses (in blue) along with changes made to the revised manuscript. Additional minor changes to the manuscript are indicated at the end of this document.

The paper investigates the origin of nucleation mode particles in the upper tropical troposphere and makes use of the extensive ATom dataset in the process. The authors use nucleation and growth box models evaluated along the trajectories of air masses reaching from the outflow of tropical convection until the air was encountered during ATom flights. They find that the nucleation parametrization of Riccobono et al. 2014 can describe the measured datasets best, which indicates organic involvement in nucleation. Also, if growth driven by organics is included in the model, the agreement with observations is enhanced as well, further pointing to the importance of organics in the upper tropical troposphere.
The paper offers very interesting insights into the important, but notoriously hard to observe topic of upper tropospheric new particle formation. The authors corroborate their conclusions with an extensive set of measurement and model data that are nicely presented. The paper is well within the scope of ACP; it is well written and clearly structured. The methods used are described clearly as well as their limitations. I recommend publication in ACP and only have some minor comments to add:

The paper focuses on nucleation rates, which is very insightful, however, in the model also growth, especially driven by organics is incorporated. You state that the inclusion of organic growth enhances the agreement between model and observations in many cases. However, you do not give a range for the growth rates needed to do that. Can you add a figure and/or short discussion that indicate the range of growth rates used in your model?

We calculated growth rates (GR) for all cases and all schemes based on the time of first appearance and the diameter of the leading edge of the nucleation mode. We used a threshold concentration for where the leading edge is when \( dN/d\log_{10}D_p > 10 \text{ cm}^{-3} \) in any of the size bins smaller than 12 nm. The growth rates were not calculated for cases with weak nucleation events i.e. with \( dN/d\log_{10}D_p \) never exceeding 10 cm\(^{-3}\) in any of the size bins smaller than 12 nm.

The growth rates for the RIC scheme are presented in Figure S71 below.

The following text was added to line 513:” Growth rates calculated basing on the diameter of the leading edge and threshold value of \( dN/d\log_{10}D_p > 10 \text{ cm}^{-3} \) in any of the size bins below 12 nm were mostly between 0.1 and 3 nm hr\(^{-1}\) for the RIC scheme (Figure S71). The growth rates for all cases investigated here using RIC and other schemes are presented in Figure S72.”
Figure S71. Growth rates calculated for simulations with SO$_2$ that have the lowest (best) NME for each case (32 cases in total) studied using RIC scheme. Calculation is based on the diameter of the leading edge and threshold value of dN/dlog$_{10}$D$_p > 10$ cm$^3$ in any of the size bins below 12 nm.
Figure S72. Growth rates calculated for the leading-edge diameter and threshold value of $dN/d\log_{10}D_p > 10$ cm$^3$ in any of the size bins below 12 nm for simulations with SO$_2$ that have the lowest (best) NME in (a) VEHK and with added organics for the initial growth of the particles, (b) NAPA and with added organics, (c) NAPAt and with added organics, (d) RIC, (e) DUN NH3=0 and with organics added for initial growth, and (f) DUN and with organics added. Note different y-axis ranges in (a,b,c) and (d,e,f). Blue and red symbols represent simulations without and with organics added for initial growth of the particles, respectively. The RIC scheme is the only one here that includes organics both in the particle nucleation and growth.

Are you in Table 1 discussing the ion concentration as indicated by the given parameter name or the ion pair production rate as indicated by the given unit (cm$^{-3}$s$^{-1}$)? Please add a short description in section 2.4 that describes the origin of this quantity. The value 15 in Table 1 would correspond to the ion pair production rate in cm$^{-3}$s$^{-1}$ given in Dunne et al. 2016 for the upper tropical troposphere.

Thank you for noticing. It has an ion pair production rate of 15 cm$^{-3}$ s$^{-1}$, which is typical for the upper tropical troposphere as given in Dunne et al. (2016). We changed the text from ‘ion concentration’ to ‘ion pair production rate’ in Table 1 and added a footnote under it: “value typical for the tropical upper troposphere (Dunne et al., 2016)”
1. 322: You state “We have undertaken sensitivity studies that vary the pre-existing background aerosol used as input parameter (Table 1)”, but in Table 1 there is no information on how this quantity was varied, as it is only named. Can you give more detail in Table 1 on how you varied the pre-existing aerosol and/or refer to the SI part where you discuss this in more depth?

The pre-existing background aerosol in the table refers to the “size distribution” as listed in the table (row below). We modified the table entry to make it more straightforward and combined the name of this parameter: “Background pre-existing aerosol: initial input size distribution (SD)”. In the table’s footnote it is mentioned how this parameter was varied: “initial background aerosol size distribution was varied: SD>12 nm means background SD as described in the text was used to initiate the model; SD>12nm x2 means background SD multiplied by 2; SD>12nm /2 means SD divided by 2; SD=0 means no background aerosol; SD>12 nm-5 nm means SD was shifted by 5 nm to smaller diameters; SD>8nm means measured background SD >8nm was used as initial SD.”

We added: “where SD refers to the number size distribution dN/dlog_{10}D_p”.

1. 485: You state that “varying the scale factor for organics taking part in nucleation (Forgnuc) did not change results significantly”. Can you add a short discussion on why this is the case? You span one order of magnitude in [BioOxOrg] with the scaling factor values you use, so according to the RIC scheme, this should vary the nucleation rate as well by an order of magnitude. So in Fig S48 I would guess that you would need higher organics in the F = 0.1 case to match observations than in the F = 1 case. However, the blue dots are practically identical in all cases. So what is the compensating effect for that?

Nucleation can have strong self-regulating mechanisms, and these are likely contributing here. Increasing Forgnuc increases the nucleation rate initially in the high- Forgnuc simulations, which increases the condensation and coagulation sinks, which decreases the vapor concentrations (and hence nucleation and growth rates), and increases the coagulational loss rates. Each of these feedbacks lead to a buffering of changes to the nucleation mode. Westervelt et al., (2014) describes these buffering mechanisms in detail. Changing Forgnuc does not change the amount of organic material and sulfuric acid that condenses though, so ultimately the size distributions are relatively insensitive to changing Forgnuc because the buffering mechanism and total condensation remain approximately constant.

We added the text above verbatim to line 488.

Figure 7: It is a bit misleading that you place the black symbols for Aitken and accumulation mode at x = 0, as these originated from the measurement as you write and the measurement was not taken at x = 0. Even if you use them as starting point for the model, I would still place them at the point of actual measurement.

We have moved these to the point of actual measurement.
Figure 1: I would suggest to remove grid lines in between the panels.
Done.

1. 577: “5-20 size range”: You mean nm?
Yes. Corrected now.

Figure 2: Add an altitude axis on the right, as in Fig 2 in Williamson et al. 2019
Done. The updated Figure 2 is presented below. We updated Figure S6 for Atlantic in the SM accordingly.

Figure 2: Ambient pressure as a function of latitude colored by the measured number concentration of particles with $D_p$ from 3 - 60 nm over the Pacific Ocean for a) ATom 1, July-August 2016; b) ATom 2, January-February 2017; c) ATom 3, September-October 2017; and d) ATom 4, April-May 2018). Periods of flight in clouds, over continents and near airports have been removed.
The following references have been added to the manuscript:


Additional minor changes to the manuscript / supplemental material:

Case sd82: typo in Table S4: NME of 0.07 corrected to 0.27 for the RIC scheme. This case has been recalculated for all schemes. All relevant tables and plots for this case were updated. This resulted in a change of the number of cases with best NME for the RIC from 22 to 21. The NAPA+org scheme has now 4 best NME cases rather than 3.
Figure 6a and figures S2, S14-46, S50, S53, subfigure (a) in the SM: y-axis label changed from dNDlogDp to dN/dlog_{10}D_p.

Figure S38: typo in a table: DUN+orggr changed from 1 to 100 pptv.
The potential role of organics in new particle formation and initial growth in the remote tropical upper troposphere

Agnieszka Kupc1,2, Christina J. Williamson1,3, Anna L. Hodshire4, Jan Kazil1,3, Eric Ray1,3, T. Paul, Bui5, Maximilian Dollner2, Karl D. Froyd1,3, Kathryn McKain3,6, Andrew Rollins1, Gregory P. Schill1,3, Alexander Thames7, Bernadett B. Weinzierl2, Jeffrey R. Pierce4 and Charles A. Brock1

1Chemical Sciences Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, U.S.A
2Faculty of Physics, Aerosol Physics and Environmental Physics, University of Vienna, 1090 Vienna, Austria
3Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, U.S.A.
4Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523, USA
5Earth Science Division, NASA Ames Research Center, Moffett Field, California, USA
6Global Monitoring Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, 80305, USA
7Department of Meteorology and Atmospheric Science, Pennsylvania State University, University Park, PA, USA

Correspondence to: Agnieszka Kupc (agnieszka.kupc@univie.ac.at)

Abstract.

Global observations and model studies indicate that new particle formation (NPF) in the upper troposphere (UT) and subsequent particles supply 40-60 % of cloud condensation nuclei (CCN) in the lower troposphere, thus affecting the Earth’s radiative budget. There are several plausible nucleation mechanisms and precursor species in this atmospheric region, which, in the absence of observational constraints, lead to uncertainties in modeled aerosols. In particular, the type of nucleation mechanism and concentrations of nucleation precursors, in part, determine the spatial distribution of new particles and resulting spatial distribution of CCN from this source. Although substantial advances in understanding NPF have been made in recent years, NPF processes in the UT in pristine marine regions are still poorly understood and are inadequately represented in global models.

Here, we evaluate commonly used and state-of-the-art NPF schemes in a Lagrangian box model to assess which schemes and precursor concentrations best reproduce detailed in situ observations. Using measurements of aerosol size distributions \((0.003 < D_p < 4.8 \, \mu m)\) in the remote marine troposphere between \(\sim 0.18\) and 13 km altitude obtained during the NASA Atmospheric Tomography (ATom) mission, we show that high concentrations of newly formed particles in the tropical UT over both the Atlantic and Pacific oceans are associated with outflow regions of deep convective clouds. We focus analysis on observations over the remote Pacific Ocean, which is a region less perturbed by continental emissions than the Atlantic. Comparing aerosol size distribution measurements over the remote Pacific with box-model simulations for 32 cases shows that none of the NPF schemes most commonly used in global models, including binary nucleation of sulfuric acid and water
neutral and ion-assisted) and ternary involving sulfuric acid, water, and ammonia, are consistent with observations, regardless of precursor concentrations. Through sensitivity studies, we find that the nucleation scheme among those tested that is able to explain most consistently (21 of 33 cases) the observed size distributions is that of Riccobono et al. (2014), which involves both organic species and sulfuric acid. The method of Dunne et al. (2016), involving charged sulfuric acid-water-ammonia nucleation, when coupled with organic growth of the nucleated particles, was most consistent with the observations for 5 of 32 cases. Similarly, the neutral sulfuric acid-water-ammonia method of Napari (2002), when scaled with a tuning factor and with organic growth added was most consistent for 6 of 32 cases. We find that to best reproduce both nucleation and growth rates, the mixing ratios of gas-phase organic precursors generally need to be at least twice that of SO$_2$, a proxy for dimethyl sulfide (DMS). Unfortunately, we have no information on the nature of oxidized organic species that participated in NPF in this region. Global models rarely include organic-driven nucleation and growth pathways in UT conditions where globally significant NPF takes place, which may result in poor estimates of NPF and CCN abundance and contribute to uncertainties in aerosol-cloud-radiation effects. Furthermore, our results indicate the organic aerosol precursor vapors may be important in the tropical UT above marine regions, a finding that should guide future observational efforts.

1 Introduction

The majority of particles found in the atmosphere are formed through gas-to-particle conversion (i.e. nucleation) from clustering of low-volatility vapors (Gordon et al., 2017; Pierce, 2017). While the formation of these molecular clusters appears to take place almost everywhere and at all times in the atmosphere (Kerminen et al., 2018), the formation of thermodynamically stable aerosol particles with diameters ($D_p$) $\geq$1.5 nm requires favorable conditions in terms of temperature, availability of condensable vapors, and the background of pre-existing bigger particles that compete for condensing vapors, and so may not occur in every atmospheric environment (Kulmala et al., 2014). Most of these newly formed particles are lost by coagulation with larger particles, and do not contribute to particle number (Westervelt et al., 2014). A subset of the nucleated particles grows by condensation to become larger particles with reduced Brownian motion, and hence lower coagulational loss rates (e.g., Pierce and Adams, 2007). Particles with $D_p$ $\geq$50 nm can serve as CCN at supersaturations found in typical marine cumulus and stratocumulus clouds (Quinn et al., 2008), increasing droplet number concentrations and cloud albedo, and thus indirectly affecting the Earth’s radiative budget (Twomey, 1974; IPCC, 2013).

The tropical UT is known to be a major source region of new particles (e.g. Clarke, 1993; Brock et al., 1995; Clarke and Kapustin, 2002; Weigel et al., 2011; Williamson et al., 2019). This strong aerosol production is believed to be linked with frequent deep convection in this region. The mechanism proposed by Clarke (1992) involves the formation of new particles in the UT from convectively lifted and cloud-processed boundary layer air. At the conditions of cold temperatures, high photolytic fluxes, and low concentrations of pre-existing aerosol particles found in the outflow of deep convection at altitudes $>8$ km, aerosol precursor gases that may survive convective transport and scavenging can oxidize and nucleate new particles which then grow to CCN sizes as they descend in the gradually subsiding air that compensates for the upward convection. Raes
(1995) used a box model to determine that observed concentrations of CCN in the remote marine boundary layer (MBL), and their temporal stability, could not be explained without a source of particles being entrained from the free troposphere (FT). Clarke et al. (2006) estimated that entrainment from the FT provides 35-80% of the CCN flux into the MBL over latitudes between 40° S and 40° N with the rest coming from sea salt aerosol. More recently, Quinn et al. (2017) found that at ~0.5% supersaturation, the accumulation mode aerosol, composed primarily of sulfate compounds rather than sea-spray particles, provides ~70% of the CCN population throughout the MBL of the tropics and midlatitudes, and suggested that these particles originate from the FT.

Despite the climatic importance of NPF in the tropical UT, the chemical mechanisms are poorly understood (e.g., English et al., 2011). This lack of understanding is driven by the fundamental complexity and variability of the atmosphere, the range of potential chemical species and mechanisms that could lead to NPF and subsequent growth of the newly formed particles to CCN, and the difficulty in obtaining observations of processes occurring in remote areas, at high altitudes, and over time scales ranging from minutes (NPF) to weeks (condensational growth during gradual descent). Together, these issues have made it difficult to validate NPF schemes used in global models and have hindered our ability to reduce uncertainty in aerosol-cloud-radiation interactions.

Williamson et al. (2019) showed that three of four global models examined in their study underestimated the magnitude of NPF in the tropical Pacific UT and all failed to accurately simulate the abundance of CCN-sized particles in the lower troposphere of the same region (the fourth model significantly overestimated aerosol loadings throughout the troposphere). None of these models used a NPF scheme involving organics, and the three models may lack sufficient precursor vapors for growth, in addition to other deficiencies. Previous model studies (e.g., Kazil et al., 2010; Yu et al., 2010; Zhang et al., 2010; Zhu et al., 2019) show that the choice of NPF mechanism can drive substantial changes in the predicted abundance and spatial distribution of particles. While Westervelt et al. (2014) suggested that the global-mean boundary layer CCN are not very sensitive to the number of particles formed in the UT due to the dampening effects of coagulation (i.e., more nucleation leads to faster coagulational losses), different choices of NPF mechanisms in models might alter the spatial and temporal pattern of NPF, and thus affect the spatial distribution and magnitude of CCN abundance. It is clear that accurate simulation of NPF and growth processes is essential to adequately represent particle size distributions and their spatial distribution in global models and improve predictions of aerosol-cloud-radiation effects (Hodshire et al., 2018; Williamson et al., 2019).

Field measurements have shown that sulfuric acid is a key component in atmospheric NPF in the continental boundary layer (e.g., Weber et al., 1997; Riipinen et al., 2007; Sihto et al., 2006). Several nucleation schemes involving sulfuric acid have been used in global models as a consequence. These include activation nucleation scheme that depends on sulfuric acid only (Kulmala et al., 2006), binary schemes that involve sulfuric acid and water to form new particles (e.g., Vehkamäki et al., 2002), or ternary schemes in which sulfuric acid, water and ammonia condense to form new particles (e.g., Napari et al. 2002). The activation nucleation scheme, however, is an empirical formulation tuned to mid-latitude continental boundary layer observations so it is appropriate to use only there. Binary NPF has been suggested to be favored in the remote tropical UT due to cold temperatures, high relative humidity (RH), and the availability of supersaturated sulfuric acid (Clarke, 1992; Brock et
al., 1995; Clarke and Kapustin, 2002). Ion-assisted nucleation of sulfuric acid and water clusters has been identified as a potential pathway for binary NPF (Kirkby et al., 2011; Lovejoy et al., 2004; Kazil and Lovejoy, 2007; Raes et al., 1997; Yu 2010). Ions stabilize the molecular clusters so that nucleation can occur at warmer temperatures and lower nucleating-vapor concentrations (Yu, 2010).

Recent observations of the composition of molecular clusters present during NPF have highlighted the role that organics may play (Kulmala et al., 2013; Smith et al., 2004). Murphy et al. (2006) and Froyd et al. (2009) found that larger particles ($D_p > 0.15 \mu m$) in the UT contained significant organic matter that was likely secondary, which suggests that condensable gas-phase organic compounds are present in the UT. Andreae et al. (2018) postulated that biogenic volatile organic compounds carried from the boundary layer to the UT by deep convection and oxidized to form condensable species over the Amazon are responsible for NPF observed in this continental UT region. Weigel et al. (2011) also suggested that organics might contribute to NPF events observed in the UT. Other nucleation processes combining sulfuric acid with ammonia (Kürten et al., 2016; Merikanto et al., 2007), amines (Almeida et al., 2013), di-amines (Jen et al., 2016), or organics (Kulmala et al., 2006; Metzger et al., 2010; Riccobono et al., 2014), or organics alone (Kirkby et al., 2016; Bianchi et al., 2016), have been proposed to explain some field and laboratory observations of NPF, primarily at warmer temperatures and continental locations. In a modeling study, Zhu et al. (2019) found that pure organic nucleation from biogenic volatile organic compounds could be an important source of particles, especially in the UT of modern-day pristine, continental environments and during the pre-industrial period.

Because there have been no in situ observations of the composition of molecular clusters and nano-particles found in convective outflow in the UT, it is difficult to ascertain which of these varied mechanisms, if any, contribute to NPF in the remote FT. In this study, we use unique observations obtained during the Atmospheric Tomography Mission (ATom), a multi-year airborne program to measure gas and aerosol properties of the remote troposphere over both the Atlantic and Pacific oceans across four seasons. Recently formed particles observed in the tropical UT were linked to recent outflow from deep convection. We use box models constrained by trajectory calculations to evaluate how well different NPF formation mechanisms can simulate the observed particle size distributions. We perform extensive model sensitivity studies to determine which nucleation mechanisms and initial precursor mixing ratios allow for the model to match observed size distributions.

2. Methods

To establish a link between convection and NPF, and to explore the processes that govern NPF and initial growth in the tropical and subtropical free troposphere over the Pacific Ocean, we couple measured size distributions between 2.6 nm and 4.8 µm in diameter from the four ATom missions with calculated air mass back trajectories and two box models. The back trajectories identify air masses potentially influenced by recent convection. We compare our simulations with in-situ ATom observations of aerosol size distributions. We examine which nucleation schemes best explain the observations, and evaluate whether observed sulfur precursors ($SO_2$ and dimethyl sulfide (DMS)) can explain the NPF and the particles’ initial growth.
In one model, we simulate particle formation by neutral and charged binary and ternary schemes, and a neutral organic scheme, and we also add organics for initial growth of the particles in all schemes. In an additional model, we form particles using both neutral and charged binary schemes.

### 2.1 The Atmospheric Tomography Mission

The NASA Atmospheric Tomography Mission (ATom) was an airborne global survey that used the NASA DC-8 research aircraft to map for the first time the composition of the remote atmosphere over both the Pacific and Atlantic basins (~82° N to ~86° S; Fig. 1) in nearly continuous ascents and descents between ~0.18 and 13 km in all four seasons (July-August 2016, January-February 2017, September-October 2017, and April-May 2018). The primary objectives of the mission were to examine the composition of the remote atmosphere to improve understanding of photochemical mechanisms for reactive and long-lived gas-phase species and to identify the abundance, distribution, sources of aerosol particles in the remote marine troposphere (Wofsy et al., 2018).

Figure 1: Flight tracks of the NASA DC-8 research aircraft during the ATom 1 (July-August 2016), ATom 2 (January-February 2017), ATom 3 (September-October 2017) and ATom 4 (April-May 2018) missions covering the remote marine atmosphere of the Pacific and Atlantic Oceans between ~83° N and ~86° S.
2.1 Measurements

A suite of fast time response (1 Hz) particle counters and optical particle size spectrometers were used to measure dry size distributions between 2.6 nm and 4.8 µm in diameter (Brock et al., 2019). Two nucleation mode aerosol size spectrometers (NMASS; Williamson et al., 2018), each consisting of five continuous-flow condensation particle counters (CPCs) with different fixed cut-off diameters (i.e. diameters at which each CPC detects 50% of the incoming particles) between 3.2 and 59 nm, measured particle number concentration. Two optical aerosol counters, an ultra-high-resolution sensitivity aerosol spectrometer (UHSAS; Kupc et al., 2018) and a laser aerosol spectrometer (LAS; Froyd et al., 2019), measured particle size distributions from 0.06 < D_p < 1 µm and 0.12 < D_p < 4.8 µm, respectively. All concentrations reported here are corrected to standard temperature and pressure (STP), 1013 hPa and 273.15 K. The NOAA Particle Analysis by Laser Mass Spectrometry (PALMS) instrument measured the composition of individual aerosol particles (Froyd et al., 2009; 2019). For this study the PALMS size range is restricted to 0.125-1.5 µm. Due to inlet sampling artifacts (Weber et al., 1998; Murphy et al., 2006), cloudy periods were removed from the analysis. Clouds were detected using a second-generation cloud, aerosol, and precipitation spectrometer (CAPS) mounted under the wing, which also measured coarse aerosols >0.5 µm (Dollner et al., in preparation) at ambient conditions.

Temperature, pressure, and wind speed with high time resolution (1 Hz) were measured with an accuracy of ±0.3 K, ±0.3 hPa, and ±1.0 m s⁻¹, respectively (Scott et al., 1990). Highly sensitive sulfur dioxide (SO₂) measurements were made during ATom 4 using laser-induced fluorescence (Rollins et al., 2016; Rollins et al., 2017) with a precision (1σ) of 1-2 pptv at 10 s and an overall uncertainty of ±(9 % +2 pptv). Laser-induced fluorescence was used to measure OH and HO₂ simultaneously (Faloona et al., 2004; Brune et al., 2020) with an accuracy of ±35 %. Measurements of carbon monoxide (CO) were made using a Picarro G2401m (Chen et al., 2013) with a precision (1σ) of 2-3 ppb at 10 s and an average uncertainty of 4 ppb. All data used in this analysis can be found in Kupc et al. (2020).

2.2 Air mass back trajectories and convective influence

To identify air in the UT influenced by recent deep convection, we calculated 10-day air mass back-trajectories using the Bowman trajectory model (Bowman, 1993) driven with meteorological fields (3 hourly, 0.25° horizontal resolution) from the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). Trajectories were also run with MERRA2 and ERA5 reanalysis meteorology and the results were similar. Meteorological products from NCEP interpolated to the aircraft flight track agreed best with quantities measured on the aircraft during ATom, so all analyses were done using the trajectories based on the NCEP data.

A cluster of 245 trajectories was initialized within a grid (0.3° x 0.3° x 20 hPa; Fig. S1) centered around the DC-8 flight track location every minute of flight (Fig. 1). The back-trajectory time step was 3 hours, based on the reanalysis data, while a time step of 15 minutes interpolated from the 3 hours reanalysis data was used for box model simulations. Uncertainty in the back-trajectory locations is represented by the 3-D spread in the trajectory cluster. The vertical uncertainty is estimated...
as the standard deviation in pressure (hPa) of the trajectory cluster at each time. The horizontal uncertainty is estimated using a probability grid based on the trajectory cluster in longitude and latitude at each time (Fig. S1), where probability grid refers to the number of trajectories at each time that are within each latitude-longitude grid box (2° x 2°). For instance, if 24 of the trajectories are within a certain grid box at a certain time then the probability for that grid box is ~10% (24/245). The probability that air sampled by the aircraft was influenced by deep convection was calculated based on coincidences of the back-trajectory cluster with satellite derived cloud locations and characteristics such as cloud top and base pressure levels (NASA Langley). To isolate deep convection, only clouds with vertical extent >5 km were considered in the convective influence (CI) calculation. The CI probability is the fraction of the trajectories in each cluster that intersected a convective system within a distance tolerance of 0.15° (~10-15 km), and for which the RH with respect to liquid water (RH\textsubscript{w}) of the trajectory was >50%. If the CI probability determined in this manner was >95%, we assume that the aircraft was sampling air strongly influenced by deep convection.

2.3 Description of models

We use two independent aerosol nucleation and growth box models to test if different nucleation schemes are consistent with observations, following the trajectories from convective outflow to the location of the aircraft. These two models are conceptually similar, but differ in size resolution and their support for different nucleation mechanisms. Our primary model, the TwO-Moment Aerosol Sectional (TOMAS; Adams and Seinfeld, 2002; Pierce and Adams, 2009; Pierce et al., 2011) includes both neutral and charged mechanisms. The neutral mechanisms include sulfuric acid and water (binary scheme; Vehkamäki et al. (2002)), sulfuric acid, water, and ammonia (ternary scheme; Napari et al. (2002)), and sulfuric acid with organics (organics scheme; Riccobono et al., 2014; Yu et al., 2017). The charged mechanism is Dunne et al. (2016), which quantifies NPF in terms of sulfuric acid, ammonia, and ion concentration (also including neutral pathways). In addition to testing the role of organics in nucleation and growth, we also test the influence of organics on aerosol initial growth when added as a condensing species following the nucleation of particles formed by each of the nucleation schemes in TOMAS, as described in Sect. 2.3.1.

We also use the Model of Aerosols and Ions in the Atmosphere (MAIA; Lovejoy et al., 2004; Kazil and Lovejoy, 2007) to test ion-assisted nucleation of sulfuric acid and water. Since ion-assisted nucleation simulations using MAIA did not explain the observed size distributions in our work, we focus on TOMAS model description and results, and the details on the MAIA model are in the Supplemental Material (Section S1). Some features common to both models are described below.

The MAIA and TOMAS box models are constrained to follow the meteorological conditions along the trajectories. They are initialized at the point where the trajectories intersect deep convection, and proceed forward in time until reaching the aircraft sampling location and time. The temperature, pressure, and RH\textsubscript{w} vary based on the trajectory. We vary the initial SO\textsubscript{2}, NH\textsubscript{3}, and organic aerosol precursors in TOMAS and SO\textsubscript{2} in MAIA (Table 1 and Supplemental Material Table S1) to see which initial values of these species allow for the best matches to the observed size distribution. We note that neither model explicitly simulates DMS, which is likely to be an important aerosol precursor through its oxidation to form SO\textsubscript{2} and
subsequently \( \text{H}_2\text{SO}_4 \), as well as through its oxidation to methanesulfonic acid (MSA; Hodshire et al., 2019), which is a condensing species that may also be able to participate in NPF (Bork et al., 2014; Chen et al., 2015; Chen and Finlayson-Pitts, 2017). Previous analyses have shown that most of the observed reactive gas phase sulfur above the boundary layer is in the form of \( \text{SO}_2 \) (Veres et al., 2019). In this work, both models are initialized with a measurement-based, pre-existing background aerosol population that acts as a sink for condensable vapors and small particles (see Section 2.4). Nucleation-mode particles are initialized at zero concentration. We calculate the \( \text{OH} \) diurnal cycle using a prescribed peak noontime value based on observations of \( \text{OH} \) on the DC-8 aircraft (Section 2.4 and Supplemental Material Fig. S3). The \( \text{OH} \) concentration along the trajectory and the resulting production rate of \( \text{H}_2\text{SO}_4 \) from oxidation of \( \text{SO}_2 \) are then calculated. We ignore possible enhanced \( \text{OH} \) due to cloud reflectivity in the vicinity of convective outflow and reduced \( \text{OH} \) from shading by higher clouds.

### 2.3.1 The TOMAS box model

The TOMAS model simulates particle nucleation, condensation, and coagulation in 43 logarithmically spaced particle size bins, which represent dry diameters from 0.7 nm to 10 \( \mu \text{m} \). TOMAS tracks the total aerosol number and mass of each species for each size bin. The simulated aerosol species are sulfate, ammonia, a representative oxidation product of biogenic organics, and water. In these simulations, neutral sulfuric acid-water nucleation is based on Vehkamäki et al. (2002), neutral sulfuric acid-water-ammonia nucleation is from Napari et al. (2002), ion-induced and neutral sulfuric acid-water and sulfuric acid-water-ammonia nucleation is from Dunne et al. (2016), and neutral sulfuric acid-organic nucleation is from Riccobono et al. (2014).

Vehkamäki et al. (2002), referred to here as VEHK, describe a parametrization for neutral sulfuric acid-water particle formation based on a classical nucleation model. They use a model for the hydrate formation relying on \textit{ab initio} calculations of small sulfuric acid clusters and on experimental data for vapor pressures and equilibrium constants for hydrate formation. The parameterized formulas are valid at temperatures between 230.15 K and 305.15 K, \( \text{RH}_w \) from 0.01\%-100\%, and sulfuric acid concentrations from \( 10^4 \)-\( 10^{11} \) \text{cm}^{-3}. Temperatures along the trajectories ranged between 218 and 252 K and thus were below the applicable temperatures of this nucleation scheme in 18 out of 32 simulated cases. In these low-temperature cases, we assume the temperature to be 230.15 K (i.e., we do not extrapolate beyond the bounds of the parameterization). When sulfuric acid concentration was \(< 10^4 \) \text{molecules cm}^{-3}, the model assumes a nucleation rate of zero, and it limits the maximum concentration of gas phase sulfuric acid to \( 10^{11} \) \text{molecules cm}^{-3}.

In the Napari et al. (2002) scheme, referred to here as NAPA, the nucleation rate is parameterized using four variables: temperature, \( \text{RH}_w \), \( \text{H}_2\text{SO}_4 \) concentration, and \( \text{NH}_3 \) mixing ratio. The parameterization is valid for temperatures from 240–300 K, \( \text{RH}_w \) from 5–95\%, sulfuric acid concentrations from \( 10^4 \)-\( 10^9 \) \text{molecules cm}^{-3}, ammonia mixing ratios from 0.1–100 ppt, and nucleation rates from \( 10^{-5} \)-\( 10^6 \) \text{cm}^{-3}\text{s}^{-1}. When temperature is \(< 240 \) K or \( > 300 \) K (25 out of 32 simulated cases), or \( \text{RH}_w \) is outside of the limits stated above, the model assumes the temperature to be 240 or 300 K, and \( \text{RH}_w \) to be 5 or 95\%, respectively. When the sulfuric acid concentration is \(< 10^4 \) \text{molecules cm}^{-3} the model assumes a nucleation rate of zero, and it limits the maximum concentration of gas phase sulfuric acid to \( 10^9 \) \text{molecules cm}^{-3}.

8
This parametrization accounts only for hydrate formation and neglects the formation of ammonium bisulfate and its effect on nucleation rate (Zhang et al., 2010). It overpredicts the effect of ammonia on nucleation when compared with laboratory measurements (Zhang et al., 2010). Merikanto et al. (2007) showed that nucleation rates based on the NAPA scheme were biased high, and Lucas and Akimoto (2006) indicated that this scheme in a global model predicted unrealistically high nucleation rates throughout the troposphere. To address these issues, Westervelt et al. (2013) and Jung et al. (2010) used a nucleation rate tuning factor of $1 \times 10^{-5}$ in the boundary layer and found that the model produced a reasonable agreement with observations. In this study we performed simulations both with (NAPA) and without (NAPA) this tuning factor.

In Dunne et al. (2016), referred to here as DUN, the inorganic nucleation rates determined experimentally in the CLOUD chamber are parametrized in four dimensions: sulfuric acid, ammonia, temperature (208-292 K) and ion formation rates (0-75 cm$^{-3}$ s$^{-1}$). Humidity is not included in this parametrization. The overall nucleation rate is given by the sum of the individual processes

\[
J_{b,n} = k_{b,n}(T)[H_2SO_4]^{Pb,n} \\
J_{t,n} = k_{t,n}(T)f_{n}([NH_3],[H_2SO_4]) \\
J_{b,i} = k_{b,i}(T)n_{-}[H_2SO_4]^{Pb,i} \\
J_{t,i} = k_{t,i}(T)n_{-}f_{i}([NH_3],[H_2SO_4])
\]

where $J_{b,n}$ is the binary neutral rate, $J_{b,i}$ is the binary ion-induced rate, $J_{t,n}$ is the ternary neutral rate, $J_{t,i}$ is the ternary ion-induced rate, $n_{-}$ is the steady state concentration of small negative ions and [H$_2$SO$_4$] and [NH$_3$] are gas concentrations (cm$^{-3}$).

In this paper, we investigated separately ion-induced binary (DUN with NH$_3$ set to 0) scheme as well as the overall nucleation scheme (DUN) given by the sum of the above.

Sulfuric acid-organic nucleation was simulated using the scheme described in Riccobono et al. (2014), referred to here as RIC. While this scheme was developed to represent terrestrial organic species, we use it here as a surrogate for marine organic compounds because there are no specific mechanisms that have been developed for remote marine-sourced precursors. The model includes a secondary organics aerosol precursor (SOAP; MW=200 g mol$^{-1}$) variable, which can oxidize to form a condensable aerosol species. This species can both participate in nucleation in the RIC scheme and condense onto particles in all schemes studied here. We assume a reaction rate constant for the oxidation of biogenic organic species against OH is $\sim 3 \times 10^{-12}$ cm$^3$ s$^{-1}$ molec$^{-1}$, which is roughly an average reaction rate of non-methane alkanes according to Table 1 of Atkinson and Arey (2003). This rate constant gives a SOAP lifetime of $\sim 2$ days for a typical diurnally averaged UT OH concentration of $2 \times 10^6$ cm$^{-3}$. The yield of SOAP to secondary organic aerosol (SOA) is set to 1, which allows us to use SOAP as a simple, tunable variable to determine how much SOA may be necessary to match observed aerosol formation and growth. We use the SOAP oxidation product (i.e. condensable organic) in the RIC scheme, but also use it to explore the effects of organics on new particle growth for each of the nucleation schemes (Riipinen et al., 2011).

In the RIC mechanism, nucleation occurs when only a fraction of the oxidation products of biogenic organic compounds (BioOxOrg in the terminology of the RIC mechanism), formed from SOAP oxidation, are able to form stabilized
clusters. The formation rate dependence on sulfuric acid and BioOxOrg concentration is given by a fit to experimental data in the form

$$J_{\text{ORG}} = k_{\text{NUC}}[H_2SO_4]^p[BioOxOrg]^q,$$

(5)

where $J_{\text{ORG}}$ is the formation rate ($\text{cm}^{-3}\text{s}^{-1}$) of stable particles with diameters ~1.7nm, $k_{\text{NUC}}$ is the nucleation rate constant with a value of $3.27\times10^{-21}$ cm$^6$ s$^{-1}$ at 278 K and RH$w$ at 39 %, BioOxOrg represents concentration of later generation oxidation products of biogenic monoterpenes (cm$^{-3}$), and the exponents $p=2$ and $q=1$ represent the power law dependence of $J_{\text{ORG}}$ upon the concentrations of sulfuric acid and BioOxOrg.

Using the RIC scheme, we test the effect of different fractions of condensable organic formed from SOAP oxidation. This fraction, $F_{\text{orgnuc}}$ represents the fraction of the condensable BioOxOrg that may participate in nucleation by stabilizing the cluster. The value of $F_{\text{orgnuc}}$ does not affect the condensation of organics onto already-nucleated or pre-existing particles. Using $F_{\text{orgnuc}}$ allows us to decouple the possible role of organics in nucleation vs. their role in subsequent condensational growth.

Since RIC scheme does not consider the possible effect of temperature on the nucleation rate, we modify the nucleation rates predicted in equation (5) using the temperature dependence (270-310 K) for this nucleation rate from Yu et al. (2017)

$$J_{\text{ORG}} - T = J_{\text{ORG}}f_T$$

(6)

$$f_T = \exp \left[ \frac{\Delta H}{k} \left( \frac{1}{T} - \frac{1}{T_0} \right) \right],$$

(7)

where $f_T$ is the nucleation rate scale factor accounting for the temperature dependence, and $\Delta H$ is the change in enthalpy of -38.3 kcal mol$^{-1}$ associated with the critical cluster formation. We assume that $\Delta H$ is constant throughout our full temperature range.

One of the limitations of our box modeling effort is that the temperatures along the trajectories ranged between 218 and 252 K, often below the applicable temperatures of the three nucleation schemes: VEHK, NAPA and RIC (Supplemental material Table S2). We would expect faster nucleation rates at the lower trajectory temperatures than are simulated by these schemes (e.g. Yue and Hamill, 1979). Using VEHK and NAPA schemes below their lower temperature limit means forcing them to their lowest rated temperature 230.15 K and 240 K respectively. This in turn may result in underestimating particle concentration and size. This bias for cold cases means that VEHK and NAPA schemes may predict SO$_2$ and organic precursors that would be anomalously high. In the RIC scheme the temperature dependence of Yu et al. (2017) is not experimentally verified down to the tropical UT temperatures. Thus, we tested the impact of changing the $\Delta H$ by $\pm 3$ kcal mol$^{-1}$ (Supplemental material Fig. S2). We also have not explored the organic-only nucleation scheme by Kirkby et al. (2016).

### 2.4 TOMAS input data

Measured and estimated inputs needed to initialize the TOMAS box model (Adams and Seinfeld, 2002; Pierce and Adams, 2009; Pierce et al., 2011) are given in Table 1. TOMAS was configured to use measured size distributions (>12 nm) in discrete bins. Each input in Table 1 represents the initial conditions present at the start of the simulation ($t_0$). Hence,
condensing vapor in the gas phase can contribute both to the formation and growth of new particles and growth of the pre-existing background aerosol.

We expect the output of the TOMAS model to be sensitive to the temperature dependence of nucleation rates, the type and number of organic compounds, SO$_2$, OH, NH$_3$ mixing ratios, and the pre-existing background aerosol into which the convective outflow is mixed. The variability of the simulated aerosol size distribution to various initial conditions was examined by conducting sensitivity simulations (Table 1) on SO$_2$, NH$_3$, OH, background aerosol size distribution, organics added for initial growth (e.g., SOAP), and on the RIC scheme scale factor $F_{orgnuc}$ for organics involved in cluster formation.

The pre-existing aerosol is estimated based on the 1-minute averaged size distributions for $D_p$>12 nm as observed at the aircraft location. The concentration of particles with $D_p$<12 nm is set to zero under the assumption that these particles were produced by the new particle formation being modeled and were not present in the background air at the point of mixing with the air detrained from convection. The box model simulations do not explicitly account for the mixing of highly scavenged air detraining from convective outflow with surrounding UT air containing more aged aerosol (e.g., Weigel et al. 2011). We have undertaken sensitivity studies that vary the pre-existing background aerosol used as initial input parameter (Table 1).

The box model simulations were run forward in time from the moment the parcel exited the convection ($t_0$) to the point of measurement by the aircraft ($t_{fin}$), with temperature, pressure, and RH$_w$ varying as a function of time as determined from the back trajectory. The change in RH$_w$ along the trajectory between the trajectory location at the cloud edge ($t_0$) and the point of the aircraft location ($t_{fin}$) for 32 simulated cases is presented in Fig. S67. The concentration of OH at solar zenith angle of 0° in the simulations was set to 3x10$^6$ molecules cm$^{-3}$; however, we also tested OH concentrations of 1x10$^6$ and 4.3x10$^6$ molecules cm$^{-3}$. These estimates agree well with aircraft-measured concentrations (Supplemental Material Fig. S3) and with values given in Seinfeld and Pandis (2006). In TOMAS, OH is parameterized as a function of the cosine of the solar zenith angle, where the night-time OH is 1x10$^5$ molecules cm$^{-3}$. The solar zenith angle is calculated for the time, altitude, latitude, and longitude of the back trajectories.

The SO$_2$ and NH$_3$ mixing ratios were varied between 1 and 100 pptv to explore a large range of plausible conditions. The evaluated SO$_2$ range exceeds that measured on ATom 4 (Supplemental Material Figs. S4 and S5) and covers the <30 pptv mixing ratios previously reported in the UT over the central and western tropical Pacific (Thornton et al., 1997; Rollins et al., 2017; Rollins et al., 2018). Organic aerosol precursors are unknown in the UT and were not directly measured; thus we explored a range of probable mixing ratios between 1 and 100 pptv.

Table 1. Ranges of parameters used for sensitivity studies in the TOMAS box model. Values varied to match the observed size distribution in italic.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial value used</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abbreviation</strong></td>
<td><strong>Unit</strong></td>
</tr>
<tr>
<td>SO$_2$*</td>
<td>pptv</td>
</tr>
<tr>
<td>NH$_3$</td>
<td>pptv</td>
</tr>
<tr>
<td>Secondary organic aerosol precursors (SOAP)</td>
<td>pptv</td>
</tr>
<tr>
<td>$F_{\text{orgnuc}}$**</td>
<td>%</td>
</tr>
<tr>
<td>-----------------</td>
<td>---</td>
</tr>
<tr>
<td>OH at solar zenith angle of 0°</td>
<td>molecule cm$^{-3}$</td>
</tr>
<tr>
<td>OH at night</td>
<td></td>
</tr>
<tr>
<td>Napari et al. (2002) scheme; nucleation rate tuning factor $1 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>Time since CI</td>
<td>hours</td>
</tr>
<tr>
<td>Ion concentration pair production rate</td>
<td>cm$^{-3}$s$^{-1}$</td>
</tr>
</tbody>
</table>

**Background pre-existing aerosol**

- **Background pre-existing aerosol: initial input size distribution (SD)**
- Varied measured initial input size distribution:
  - ***SD>12nm, SD>12nm x2, SD>12nm /2, SD=0, SD>12nm-5nm, SD>8nm, SD as logarithmic fit**

340 | *SO$_2$ measured on ATom 4 only
341 | **Fraction of SOAP participating in nucleation when using Riccobono et al. (2014) in TOMAS.
342 | *** Value typical for the tropical upper troposphere (Dunne et al., 2016)
343 | **** Initial background aerosol size distribution was varied: SD>12nm means background SD as described in the text was used to initiate the model; SD>12nm x2 means background SD multiplied by 2; SD>12nm /2 means SD divided by 2; SD=0 means no background aerosol; SD>12nm-5nm means SD was shifted by 5 nm to smaller diameters; SD>8nm means measured background SD >8nm was used as initial SD. Where SD refers to the number size distribution $dN/d\log_{10}D_p$.

### 2.5 Evaluating simulated size distributions

To determine which sets of parameters allow the models to reproduce the observed size distributions best, we evaluate every simulation against observations using the normalized mean error (NME) statistic of the first four moments ($0^{\text{th}}$ through $3^{\text{rd}}$) of the size distribution for each model simulation as

$$NME = \frac{\sum_{i=0}^{3} |S_i - O_i|}{4},$$

(8)

where $S_i$ and $O_i$ are $i^{\text{th}}$ moments of the simulated and observed size distributions, respectively (Hodshire et al., 2018).

The $i^{\text{th}}$ moment is defined as

$$M_i = \int_{2.6}^{20} n_N D_p^i dD_p,$$

(9)

where $n_N$ is the number of particles in size interval $dD_p$ and $D_p$ is the diameter. Equation (9) is integrated over the diameter range from 2.6-20 nm, and $M_i$ represents either $S_i$ or $O_i$. The zeroth moment ($i=0$) corresponds to the number of particles, the first moment ($i=1$) to the total diameter of particles (i.e. total aerosol length), the second moment ($i=2$) is proportional to the total surface area of particles, and the third moment ($i=3$) is proportional to the total volume of particles. A NME of 0 is a perfect fit between the simulation and observations; a NME of 1.0 indicates that the average bias of the $0^{\text{th}}$ through $3^{\text{rd}}$ moments between the simulation and observations is 100%. As the $NME$ is given as an absolute value, we do not discriminate between cases in which the model is underpredicting or overpredicting the moments on average. Since these moments are equally weighted, a low value of $NME$ can be achieved only if the modeled size distribution accurately simulates both the shape and magnitude of the observed size distribution over the full range of sizes evaluated.
3. Results

3.1 Observations

Our data show seasonally persistent high nanoparticle concentrations over the remote tropical UT (Fig. 2; Williamson et al., 2019). In this region, the highest concentrations of particles were in the nucleation mode (3-12 nm), which have a short lifetime and are the products of recent NPF. This tropical UT feature was observed in all ATom deployments over all four seasons, over both the Pacific and Atlantic basins. The concentrations of particles observed in the UT over the tropical Atlantic were lower in concentration than observed over the Pacific (Supplemental Material Fig. S6). In this study, we focus on observations over the remote Pacific, which is a region less perturbed by continental emissions than the Atlantic (Fig. 3 and Supplemental Material Fig. S7 and S8).
Figure 2: Ambient pressure as a function of latitude colored by the measured number concentration of particles with $D_p$ from 3 - 60 nm over the Pacific Ocean for (a) ATom 1, July-August 2016; (b) ATom 2, January-February 2017; (c) ATom 3, September-October 2017; and (d) ATom 4, April-May 2018). Periods of flight in clouds, over continents and near airports have been removed.

Previous studies (e.g. Clarke, 1992, 1993; Clarke et al., 1998, 2006; Brock et al., 1995; Weber et al., 1995; Raes et al., 1997; Thornton et al., 1997; Weber et al., 1998; Clarke and Kapustin, 2002; Twohy et al., 2002; Froyd et al., 2009; Borrmann et al., 2010; Weigel et al., 2011) have provided strong evidence of NPF in the tropical UT and its link to convective activity. However, these earlier studies did not provide such extensive, representative, and global-scale coverage of the remote marine troposphere over a wide range of altitudes and latitudes (Williamson et al., 2019). The ATom observations also provide accurate and sensitive, state-of-the-science measurements of the chemical composition of the bulk aerosol and the abundance of hundreds of gas-phase species in all four seasons, making these observations the most comprehensive to date. However, no measurements were made during ATom of NH$_3$, the highly oxygenated organic molecules that are likely aerosol precursors, or molecular cluster composition, and measurements of SO$_2$ took place only during the fourth ATom deployment.

Ten-day back trajectories in the region of NPF in the central Pacific showed transport primarily over the Pacific, with some possible terrestrial influence from the western Pacific region (Fig. 3). However, trajectories coming from the western Pacific generally stayed at high altitudes and did not show recent convective uplift from regions influenced by terrestrial sources. Further, CO and other continental tracers were at background levels over the Pacific, confirming little continental influence in the sampled air masses (Supplemental Material Fig. S7), as opposed to the Atlantic (Supplemental Material Fig. S6 and S7). Thus, the precursors of the recently formed particles are likely mostly marine in origin. The latter is also supported by the measurement of particle phase methanesulfonic acid (MSA) that can be considered as a tracer for maritime influence on the tropical UT (Fig. S9).
Figure 3: Flight track and selected 10-day back trajectories initiated for times in flight at pressures <400 hPa (<~260 K) sampled along the DC-8 flight track during ATom 1, 2 (a, b), 3 and 4 (c, d) during the most tropical flight in each deployment (Hawaii-Samoa on ATom 1 and Hawaii-Fiji on ATom 2-4). Trajectories are colored according to the pressure along their pathway.

The observations and trajectory modeling show that newly formed particles were often associated with deep convection. Similar trajectory analyses, in terms of examining the history of the sampled air masses for interactions with deep convection, have been undertaken by Froyd et al. (2009) and Andreae et al. (2018). Here, using the CI probability criterion of 95% to identify when the aircraft was sampling air recently influenced by convection (Sect. 2.2), and considering the latitude range 30° S - 30° N and ambient temperatures <260 K, for ATom 1 and 2, the shorter the time since convection, the higher the number of small particles (Fig. 4a-d). Such strong trends were not evident for ATom 3 and 4 indicating that factors other than time since CI affect nucleation-mode concentrations. The more recent the convection, the smaller the diameter of the nucleation mode (Fig. 4e-h). These relationships are again strongest for ATom 1 and 2 and also 4. Our hypothesis for these relationships is that with increasing time since CI, particles with diameters <12 nm grew by condensation and coagulation and decreased in concentration by coagulation, leading to the decrease in nucleation-mode concentration and increase in diameter. A similar
A trend was observed over the Atlantic (Supplemental Material Fig. S10). The highest concentrations of nucleation-mode particles occurred during ATom 2 and were associated with the shortest times since CI.

Air detraining from deep convection is likely depleted in pre-existing particles due to in-cloud removal, leading to a reduced condensation sink (CS) that enhances the likelihood of NPF (e.g., Clarke, 1992; Williamson et al., 2019). Figure 5 shows the concentration of measured nucleation-mode particles as a function of altitude for the Pacific basin over four ATom missions. The median concentration of nucleation mode particles averaged from 30° S to 30° N is highest at altitudes >10 km, reaching ~40,000 cm⁻³ (Fig. 5a), coinciding with the lowest values of CS₁₂nm, which competes with NPF for condensing vapors. The CS term is calculated for particle diameters >12 nm, and for the diameter range between 3 and 127 nm following Williamson et al. (2019). CS₁₂nm serves as an estimate of the condensation sink prior nucleation starting, and it is negatively correlated with the number of nucleation mode particles (Fig. 5b), while the nucleation mode is positively correlated with CS₃.
Over the Atlantic, the maximum concentration of nucleation-mode particles >8 km in altitude averaged from 30° S to 30° N, ~3,000 cm⁻³, is considerably smaller than over the Pacific, but the shape of the profile is similar (Fig. S11).

Figure 5. a) Vertical profile of the median number concentration of nucleation mode particles (3-12 nm) and condensation sink (CS) averaged between 30° S - 30° N as a function of altitude for the four ATom deployments. b–c) One minute average nucleation mode particle concentrations at >8 km in altitude as a function of CS₃-12nm (b) and CS₃-12nm (c); CS between 30° S - 30° N over the Pacific Ocean. Pearson correlation coefficient values ($r^2$) are indicated in the legend.
Some variability in the strength of NPF and its dependence on CS_{>12nm} can be observed. In general, CS_{>12nm} is weakly negatively correlated ($r^2$ between 0.083 and 0.3626 depending on the ATom mission) with the concentration of nucleation mode particles (Fig. 5b), as would be expected if NPF were competing with CS_{>12nm} for condensing vapors. When CS is dominated by small particles (CS_{3-12nm}) the correlation is strongly positive ($r^2$ between 0.97-1, Fig. 5c). Factors other than CS are also important in controlling the concentrations of newly formed particles. These factors may include temperature and RH$_w$, actinic flux, and OH that drive photochemical reactions that oxidize precursor species, the abundance of those precursor species in the air lifted by convection and in the background air, and the time since the air parcel exited a convective cloud (Figs. S12-S13).

### 3.2 Box model simulations

Case studies were selected for box model simulations based on specific criteria such as temperature and CI probability. We restrict the analysis to data taken nominally in the tropics and subtropics, between 30° S - 30° N latitude. We consider the case for analysis by box modeling if the CI probability is >95%, temperature at the point of measurement is <260 K, and an aerosol number mode with a modal peak diameter <12 nm is present (Table 2). We performed simulations for 32 cases randomly selected from the ATom 2 (20 out of 47 identified cases) and ATom 4 (12 out of 60 identified cases) datasets over the remote tropical Pacific from the total number of 109 cases with time since convection <1 day (Table 2). Data from ATom 2 and ATom 4 were selected for simulations because these deployments had the most identified cases with time since CI <1 day. During ATom 2, we observed the highest numbers of nucleation mode particles, lowest condensation sink, and shortest time since convection (Fig. 5) among all missions. Measurements of SO$_2$ mixing ratio were made only during the ATom 4 deployment, providing an important constraint for the box model simulations. We did not perform simulations on ATom 1 and ATom 3 data as there were only 2 and 0 identified cases with time since CI <1 day, respectively (Table 2).

The correlation between nucleation-mode particles and time since CI was strongest in ATom 2 (Fig. 4), while CO levels, a proxy for continental influence, were the lowest for trajectory times <1 day (Supplemental Material Fig. S7). Although SO$_2$ was not measured during ATom 2, we expect SO$_2$ in this region in the UT to be <30 pptv based on SO$_2$ levels measured during ATom 4 and other missions in the Pacific (Supplemental Material Figs. S4, S5).

<table>
<thead>
<tr>
<th>ATom mission</th>
<th>Number of cases meeting selection criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trajectory age &lt;1 day</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>
The size distributions simulated by TOMAS were smoothed to avoid the artificial distortion of the distribution caused through size-bin emptying (Hodshire et al., 2019; Stevens et al., 1996). The latter and the smoothing technique are described in Supplemental Material Section S2.

We performed box model simulations on the 32 selected cases using the range of values listed in Table 1. The success of each model simulation was evaluated using the NME described by Eq. 8. As an example using a single case, Fig. 6a shows the observed and simulated aerosol size distribution with the best NME obtained for each of the various nucleation schemes tested, along with the corresponding mixing ratios of SO₂, NH₃, or organics. Organics here refer to the SOAP oxidation product (i.e. condensable organic) that participates in nucleation in the RIC scheme (as BioOxOrg), and in the particle condensational growth in all schemes. The value of NME as a function of the mixing ratios of SO₂, NH₃, and organics for each nucleation scheme is also shown (Fig. 6b-j). The summary of each of the 32 simulated cases is presented in Supplemental Material Table S4 and Figs. S15-S45. The simulations in Fig. 6 used the default OH scheme with a maximum concentration of 3x10⁶ cm⁻³ at a solar zenith angle of 0° (Supplemental Material Fig. S3). Sensitivity studies for maximum OH values of 1x10⁶, 3x10⁶ and 4.3x10⁶ cm⁻³ are presented in Supplemental Material Fig. S46-S50.

TOMAS simulations using VEHK scheme substantially underpredict the observed tropical nucleation-mode number concentration, with resulting poor values of NME (Fig. 6). Sensitivity tests that vary the pre-existing initial (background) aerosol or completely remove background particles do not change the results significantly (Supplemental Material Fig. S51-S53). Further, we find that changing initial input parameters such as SO₂ and OH as indicated in Table 1 do not improve the NME for VEHK scheme (Fig. 6b; Supplemental Material Fig. S46). Adding organics to grow particles nucleated by the VEHK scheme, while reducing NME slightly, does not provide adequate agreement with the observations. Varying the pre-existing initial aerosol or completely removing background particles in VEHK scheme with added organics may improve the fits for certain initial conditions (Fig. S52), making it plausible for better NME values if CS₄≥12nm was 4.12x10⁻⁵ s⁻¹, SO₂ and organics were 13.9 pptv and 7.2 pptv, respectively, for this particular case. Out of 32 cases studied here, there is no case with NME <0.2 for VEHK scheme. Similarly, the ion-assisted binary nucleation scheme of the MAIA box model does not provide good matches with observations (Supplemental Material Table S4).

The NAPA scheme, both with (NAPAt) and without (NAPA) the tuning factor, did not significantly reduce the NME values from the VEHK results for this case (Fig. 6c). However, when organics were added to condense on the particles nucleated by this mechanism, the NME improves, resulting in 6 cases with NME <0.2 was reduced to 0.17 (Table S4). The RIC nucleation scheme, updated by the temperature dependence of Yu et al. (2017), provides the best NME (NME=0.02) for all the schemes investigated (Fig. 6f) for the case shown in Fig. 6. We explored this mechanism with 6 more sensitivity simulations, including various combinations of initial SO₂ and organic mixing ratios, to see how sensitive NME is to small changes of initial precursor vapor mixing ratios. For the example case presented in Fig. 6a, organic mixing ratios <10 pptv and
SO₂ mixing ratios <5 pptv were sufficient to produce size distributions that matched the observations with good fidelity (NME = 0.02). Varying the scale factor of organics taking part in nucleation (F_{orgnuc}) did not change the results significantly (Supplemental Material Fig. S48). Nucleation can have strong self-regulating mechanisms, and these are likely contributing here. Increasing F_{orgnuc} increases the nucleation rate initially in the high-F_{orgnuc} simulations, which increases the condensation and coagulation sinks, which decreases the vapor concentrations (and hence nucleation and growth rates), and increases the coagulational loss rates. Each of these feedbacks lead to a buffering of changes to the nucleation mode. Westervelt et al., (2014) describes these buffering mechanisms in detail. Changing F_{orgnuc} does not change the amount of organic material and sulfuric acid that condenses though, so ultimately the size distributions are relatively insensitive to changing F_{orgnuc} because the buffering mechanism and total condensation remain approximately constant.

The most recently developed NPF mechanism, the ion-induced sulfuric acid-water, referred here as DUN with NH₃ set to 0, and the sulfuric acid-water-ammonia (DUN) nucleation scheme from Dunne et al. (2016), did not provide the lowest NME values among the schemes tested, although adding organics for initial growth of the nucleated particles improved the fits (NME=0.04) (Fig. 6i, Supplemental Material Table S4). The addition of organics resulted in best NME values for DUN in 5 out of the 32 cases simulated.

Overall, a reduction in NME when organics are added for initial particle growth was also observed for other schemes (Supplemental Material Table S4). Out of 32 case studies, we found 6 cases when the NAPAt with organics for growth and the tuning factor applied gave lower NME values than all other schemes. However, 4 out of these 6 cases require SO₂ or NH₃ mixing ratios >50 pptv that exceed ATom 4 SO₂ observations and literature values in the tropical UT for SO₂ of <30 pptv (Fig. S4; Rollins et al., 2017; Rollins et al., 2018; Thornton et al., 1997) and for NH₃ of <10 pptv (Höpfner et al., 2016; Feng and Penner, 2007; Adams et al., 1999).

Growth rates calculated basing on the diameter of the leading edge and threshold value of dN/dlog₁₀Dₚ >10 cm⁻³ in any of the size bins below 12 nm were mostly between 0.1 and 3 nm hr⁻¹ for the RIC scheme (Figure S71). The growth rates for all cases investigated here using RIC and other schemes are presented in Figure S72.
Figure 6: Results of simulations using the TOMAS box model for an example case (ATom 2, 2017-02-04, 03:05:31-03:06:30 UTC; case identifier: sd486) where measurements were made 7.3 hours following convective influence, and temperature along the trajectory varied between 218 and 226 K. (a) Observed (shaded bars) and simulated (lines) aerosol size distributions with best
normalized mean error (NME) calculated for $D_p$ between 2.6 and 20 nm (blue shading) for each of the NPF and growth schemes investigated. Best results from the MAIA box model ion-assisted + neutral binary nucleation scheme shown as a dotted black line.

(b) NME between the modeled and measured size distribution for the VEHK scheme with varying organics mixing ratios for condensational growth. The color of the circle indicates the value of NME corresponding to a particular initial mixing ratio of SO$_2$, NH$_3$, or organics that varied between 0 and 100 pptv. Blue represents the best agreement, red poorer agreement, and grey the worst (NME >1). There were 64 sensitivity tests. (c) As in (b), but for the NAPA scheme. d) As in (c), but for the NAPAt scheme. (e) and (f) as in (c) and (d) respectively, but with NH$_3$ fixed and varying organics for condensation growth. (g) as in (b) but for the RIC scheme, which provides the lowest NME. There were 400 sensitivity tests for this scheme. (h) as in (b) but for the DUN scheme with NH$_3$ set to 0 (DUN$_{NH3=0}$). (i) as in (c) but for the DUN scheme. (j) as in (i) but with varying organics for condensation growth. The table presents the NME results for the corresponding size distributions in panel (a) and associated initial mixing ratios of gas-phase precursors.

Figure 7 shows the time evolution for particle number concentration, surface area, and volume for the nucleation, Aitken, and accumulation modes using the Riccobono et al. (2014) scheme for the same case as shown in Fig. 6 for the simulation with the lowest NME in Fig. 6g (SO$_2$=1.3 pptv, organics=5.5 pptv). There is rapid evolution of the nucleation mode and slower changes of the larger modes, and the model effectively matches the number, surface and volume of the measured nucleation mode. **In the case presented in Figure 7b, the surface area measured on ATom is dominated by the nucleation mode particles; however, although frequent, this is not a typical pattern among all other cases studied here. In general, in the cases with the highest surface area, values are split between nucleation, Aitken and accumulation modes almost equally (Table S5, Fig. S68-70).**
concentrations. (c) as in (a), but for volume concentrations. Conditions for the simulations were diurnally varying OH concentrations with solar zenith angle. Initial SO\textsubscript{2}=1.3 pptv, and initial organics=5.5 pptv.

In 212 out of the 32 cases for which multiple box-model simulations were run, the sulfuric-acid-organic nucleation scheme of (Riccobono et al., 2014) produced lower (better) values of NME than the other parameterizations tested (Supplemental Material Table S4). Two of those 212 best NME cases for RIC were tied with NAPA and NAPAt, both with organics added for initial particle growth. The remaining 12 best NMEs came from two different ternary nucleation schemes with added organics for growth of particles. These schemes were the NAPA or NAPAt, or the DUN with both charged and neutral channels. The majority of these ternary cases, however, required initial conditions of NH\textsubscript{3} of 52 pptv or more, much greater than the mixing ratios expected at these locations in the UT (Höpfner et al., 2016). Regardless of the available NH\textsubscript{3}, together these results strongly suggest that pure binary sulfuric acid-water nucleation, whether ion assisted or neutral, and whether coupled with organic growth or not, generally cannot explain the ATom observations. While we are limited by the lack of direct observations of NH\textsubscript{3}, amines, and condensable organic species, it is plausible that there are enough of these compounds—a few to tens of pptv—to participate in ternary nucleation and subsequent growth to be consistent with the ATom measurements. Further, regardless of the available SO\textsubscript{2}, the results strongly suggest that binary sulfuric acid-water nucleation, whether ion-assisted or neutral, and whether coupled with organic growth or not, generally cannot explain the ATom observations.

The findings for the case of organic-mediated NPF are summarized in Fig. 8, where we show the SO\textsubscript{2} and organic precursor mixing ratios for all sensitivity simulations with NME<0.2 for all the cases analyzed using the RIC scheme, highlighting the assumptions that yielded the lowest NME for each case. The results show that for all of the cases where sulfuric acid-organic nucleation most successfully simulated the observations (212 of 32 cases), initial SO\textsubscript{2} mixing ratios <30 pptv and organic precursors <100 pptv (with an assumed yield of 1) were needed. These SO\textsubscript{2} mixing ratios are consistent with observations during ATom 4 (Figs. S4, S5, Table S65) and earlier results (Rollins et al., 2017, Rollins et al., 2018). Lacking measurements of condensable organic species, we can only speculate that a few to tens of pptv are reasonable for the marine tropical UT. Williamson et al. (2019; Extended Data Fig. 7) suggested that organics dominate the composition of smaller particles at pressure <400 hPa. We note that we performed no simulations with mixing ratios of SO\textsubscript{2} or organics above 100 pptv. While we cannot exclude that for some cases the mixing ratios of these precursors at levels above 100 pptv could improve fits, these levels are outside of prior observations so were not considered in this study.

In the case shown in Fig. 6, mixing ratios of SO\textsubscript{2} and organics of ~1.3 pptv and 5.5 pptv, respectively, were sufficient to nucleate particles and produce a size distribution that matched the observations with an NME of 0.02 using the RIC scheme. In a majority of the cases, the RIC scheme predicted SO\textsubscript{2} <5 pptv that are lower than typical UT SO\textsubscript{2} concentrations, suggesting that our temperature extrapolation may overpredict nucleation rates at the typical SO\textsubscript{2} mixing ratios of ~30 pptv in the UT. Overall, the lowest NME values were obtained when initial SO\textsubscript{2} values were low (<30 pptv), while organics varied over a
range of mixing ratios as shown by triangles in Fig. 8. This suggests that organic matter will often contribute significantly to
the composition of the nucleated and growing particles on a mole basis, and even more so on a mass basis because the assumed
molecular weight of organic precursors and products is 200 g mol$^{-1}$ compared 96 g mole$^{-1}$ for SO$_4$. 
Figure 8: Values of NME (colored symbols) for best fits of the sensitivity studies. TOMAS model simulations were made using the RIC sulfuric acid-organic scheme. Among the sensitivity tests using this scheme, the one with the lowest NME case is shown with a triangle located at the initial conditions of SO$_2$ and organics for that case, while the next best NME case (provided NME < 0.2) is shown as a circle. The shaded region represents the approximate parameter space in which the best agreement between model and measurement is found for all the convective influence cases studied. Note a different NME color scale range (0 - 0.4) than the one presented in Fig. 6 (0 - 1).
3.3 Discussion

Comparing aerosol size distribution measurements with box-model simulations shows that none of the binary neutral or ion-assisted NPF schemes are consistent with observations, regardless of precursor concentrations and the presence or absence of condensing organics for further growth. These schemes predict significant nucleation but do not make enough particles in the 5-20 nm size range (Fig. 6) to match observations. Adding organics for initial growth of particles shifts the size distribution to bigger sizes but only slightly improves the model-to-measurement fits (Table S4).

However, schemes that incorporate organic compounds or NH$_3$ to nucleate particles, plus condensing organics as growth agents, can plausibly replicate the observed size distributions. These results suggest that organic precursor species are likely important in NPF and initial growth in the tropical upper troposphere, even above marine regions remote from continental sources. In general, the RIC scheme provided best model-to-measurement fits; however, the improvement in the fit values for DUN scheme when organics are added for initial growth of particles suggests that organics may be more important for growth than for nucleation (Table S4).

We find that to best reproduce both nucleation and growth rates by the RIC scheme, the mixing ratios of gas-phase organic precursors generally needs to be at least twice that of SO$_2$ (Fig. 8). While an example in Figure 6 shows that the source of condensible organics may be even ~5 times the SO$_2$ mixing ratio in the remote tropical UT (Fig. S66), we do not know whether or not there may be that much more organic precursor available in this region. Although, regions where the oceanic source of SOA may be higher than the DMS source have been reported previously (e.g. Croft et al., 2019).

Unfortunately, we have no information on the nature and mixing ratios of oxidized organic species that participated in NPF and initial growth in this environment. The mixing ratios used in this study do not seem out of the range of possibility. Potential precursors to these condensing species, such as isoprene or monoterpenes (e.g. alpha- or beta-pinene), were found to be below the limit of instrument detection (2 pptv for isoprene, 0.1 pptv for alpha-pinene, and 0.2 pptv for beta-pinene) in the tropical UT during the ATom deployments. The exact identification of these condensing organic species would require instrumentation such as an atmospheric-pressure-interface time-of-flight (API-TOF) mass spectrometer to measure the composition of molecular clusters, which was not a part of the suite of instrumentation during the ATom mission. Other studies also suggest that NPF and growth involving organic species may be common in the remote troposphere. Willis et al. (2016) showed that marine organics contribute to the growth of newly formed particles in the summertime Arctic at low altitude; however, it was unclear if marine organics were involved in nucleation. Burkart et al. (2017) found that particle growth in the remote Arctic was largely due to condensation of unidentified organic compounds, possibly of marine origin, associated with oxidation or photochemistry of the sea-surface micro-layer (Abbatt et al., 2019). Andreae et al. (2018) proposed that oxidized biogenic VOCs were the source of recently formed particles found in the outflows and anvils of convective storms over Amazonia.

Chemistry-climate models rarely include organic-driven nucleation pathways in the UT where globally significant NPF takes place. This may result in poor estimates of NPF and CCN abundance and contribute to uncertainties in aerosol-
cloud-radiation effects. Williamson et al. (2019) showed that the production of newly formed particles and their growth to cloud-active sizes during descent towards the surface is not adequately captured in the global chemical transport models, which tend to underestimate the magnitude of tropical UT NPF and subsequent growth. This underestimate might be related to missing organic precursors, missing chemical mechanisms, or structural errors associated with convective parameterizations. According to Williamson et al. (2019), the combined direct and indirect radiative effect of NPF in the tropical UT is ~0.1 W m\(^{-2}\), globally.

The assumptions in our box model simulations point to the need for further observational and modeling studies. For example, we do not directly simulate in TOMAS the oxidation of DMS to SO\(_2\) and MSA. However, the SO\(_2\) mixing ratios estimated in this study may serve as a proxy for DMS in the modeling in our study, although the timescale for forming H\(_2\)SO\(_4\) from SO\(_2\) will be incorrect. We had measurements of SO\(_2\) only during the fourth ATom deployment, and no measurements of NH\(_3\) or highly oxygenated organic molecules that are likely aerosol precursors. Instead, we have constrained the box model simulations using reasonable lower and upper limits of their mixing ratios based on literature data and in case of SO\(_2\), ATom 4 data. Further, nucleation schemes themselves are simply imperfect parameterizations extrapolated from laboratory observations. One of the limitations of our box modeling effort is that the temperatures along the trajectories were often below the lower range limit of three (out of four) nucleation schemes evaluated (Table S65). In these cases (marked with a “*” in Supplemental Material Figs. 15-45 and Table S65) the best-fit SO\(_2\) and organic concentrations are expected to be biased high. Although experimentally unverified, we incorporated temperature dependence into the Riccobono et al. (2014) scheme after Yu et al. (2017). We would expect faster nucleation rates at the lower trajectory temperatures than are simulated by these schemes (e.g. Yue and Hamill, 1979). Further, we have tested the Napari et al (2002) scheme both with and without a tuning factor of 10\(^{-5}\) that was developed for continental regions (Jung et al., 2010; Westervelt et al., 2013), an obvious source of uncertainty when simulating NPF in the UT over the oceans. These are schemes that many models use and they do not appear to (often) work for this region, possibly due to their limited range of operating temperatures. We also have not explored the organic-only nucleation scheme described by Kirbky et al. (2016), an updated version of the Vehkamaki et al. (2002) scheme covering a wider range of temperatures and relative humidities by Maattanen et al. (2017) that has been validated against CLOUD measurements, or the recently published ternary nucleation look up tables for model implementation (Yu et al., 2020), and these schemes are worth investigating in future studies.

Further, we did not account for mixing with surrounding air on the path between the cloud outflow and the point of measurement when running simulations. The box model used here simulates NPF in the outflow region of deep convective clouds. Although, active NPF was identified in the vicinity of clouds and in the cloud outflow region in many studies (such as Perry and Hobs 1994; Clarke et al., 1998, 1999; Ström et al., 1999; Clement et al., 2002; Twohy et al., 2002; Weigelt et al., 2009; Waddicor et al., 2012; de Reus et al., 2001; Clarke and Kapustin 2002; or Andreea et al., 2018), the exact location of NPF with respect to cloud remains uncertain (Kulmala et al., 2006; Waddicor et al., 2012). We assume that nucleation does not occur within the cloud, and that the outflow does not immediately mix with the surrounding air in the highly stratified upper troposphere. If NPF were to occur in a cloud or in a zone of the turbulent mixing at the cloud edge, as suggested by
some studies (Stroem et al., 1999; Lee et al., 2004; Weigel, 2011; Kazil et al., 2007; Kulmala et al., 2006), our results would be biased.

The limitations described above are important and point out the need to undertake further in situ measurements and modeling studies to confirm the suspected role of organics in UT NPF and subsequent growth in the remote troposphere. Better understanding of NPF in the remote UT, and the growth of these particles to cloud-active sizes, could substantially improve model simulations of the preindustrial atmosphere, would allow better evaluations of the effect of current anthropogenic perturbations, and could allow more confident predictions of the evolution of the climate and its response to future emission scenarios. Modeling efforts should focus on developing new nucleation mechanisms based on chamber studies conducted at temperatures more representative of the UT. Further airborne research should focus on measuring the composition of molecular clusters, sulfuric acid, organics, and NH$_3$ over the oceans and tropical continental areas. The planned Chemistry of the Atmosphere: Field Experiment in Brazil (CAFE-Brazil) study is the first expected to combine airborne measurements of nucleation-mode particle size distributions with API-TOF mass spectrometer measurements of the composition of nucleating clusters.

4. Summary

Airborne observations during the ATom mission show a globally significant source region of newly formed particles in the tropical and subtropical UT that persists over both the Atlantic and Pacific Ocean basins over all seasons. These particles are often associated with the outflow from deep convection. Averaged across the tropics and subtropics over the Pacific, the particle number concentrations were a maximum (reaching as high as ~40,000 cm$^{-3}$) at altitudes above 10 km where the condensation sink from pre-existing aerosol particles was a minimum. Using back-trajectories to identify convectively influenced air parcels, the highest concentrations of recently formed particles were generally found where the CI was most recent, particularly during the first and second ATom deployments. The number concentration of nucleation-mode particles decreased with time since CI due to the effects of coagulation and condensational growth. During ATom 1 and 2, particle size increased with time since CI, showing clear evidence for this growth.

We simulated particle nucleation and growth using two box models constrained to follow the calculated trajectories from the point of convective detrainment to the point of measurement by the aircraft, and we performed sensitivity tests varying the nucleation mechanisms and initial conditions such as precursor (SO$_2$, NH$_3$, organics), OH, and pre-existing particle concentrations.

These simulations indicate that nucleation schemes commonly used in global models, such as binary homogeneous H$_2$SO$_4$ (both neutral by Vehkamäki et al. (2002) or ternary H$_2$SO$_4$+ NH$_3$ (neutral with and without a tuning factor by Napari et al. (2002) and Jung et al. (2010))$^3$, as well as the recently developed neutral and ion-induced binary and ternary nucleation scheme by Dunne et al. (2016), were all inconsistent with observed size distributions in all simulated cases when no organics were included for growth. This result also held for the binary nucleation mechanisms even when organics were added as a
condensing, but not nucleating, species. Adding organics for initial growth of particles formed by either of tested ternary schemes (Napari et al. (2002) or Dunne et al. (2016)) provided the best fits in 12 out of 32 simulated cases (Table S4). However, the majority of these ternary inorganic simulations required initial conditions of NH$_3$ $>$ 50 pptv, which is substantially greater than expected at these locations in the UT (Höpfner et al., 2016).

In contrast, a scheme involving oxidation products of biogenic organics and H$_2$SO$_4$ (Riccobono et al., 2014) gave results that were most consistent among the various models with observations in 212 out of 32 cases, while in 2 cases it was tied for the lowest NME with other schemes. These results strongly suggest that organics are involved in NPF and subsequent initial growth in the remote tropical UT. This supports the finding by Simon et al (2019) that organics, despite their lower oxidation level and yield at low temperatures, may be important for nucleation and growth in the UT. However, the predicted SO$_2$ concentrations were often anomalously low (<5 pptv), suggesting that our temperature extrapolation may overestimate the nucleation rates. While the Riccobono scheme was most consistent, the analysis suggests that multiple nucleation mechanisms may be plausible across the 32 cases.

We have assumed that the Riccobono et al. (2014) scheme, which was developed from laboratory measurements of nucleation from the oxidation products of terrestrial biogenic VOCs, represents processes in the remote marine UT of the tropical Pacific. In fact, there is virtually no information on the nature of oxidized organic species (or ammonia and amines) that may participate in NPF in this environment. Also, the Riccobono scheme, one of the least constrained nucleation mechanisms, required a large extrapolation in temperature to simulate UT conditions. Given that NPF in the tropical UT is a major source of CCN over a large portion of the globe (Williamson et al., 2019), we recommend that future work investigate the species contributing to NPF and growth explicitly, including direct measurements when possible. Additionally, we recommend studies that focus on potential tropical marine sources of aerosol precursor gases, the efficiency of their transport to the UT, the products of their oxidation, and the mechanisms of NPF at temperatures <230 K.

**Data availability**

The full ATom dataset is available as given in Wofsy et al. (2018), and may also be accessed at https://espoarchive.nasa.gov/archive/browse/atom. Data presented in this analysis are available at the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center, (DAAC) Kupe et al. (2020).

**Author contributions**

AK, CW, CB, KF, MD, BW, TB, AR collected data. AK, JP, CB, and AH conceived and designed the study. AK performed the analysis and wrote the manuscript with help from CB and JP, with contributions from all co-authors. JP, AH, JK provided TOMAS and MAIA box models and helped with model upgrades. AK performed all model simulations. MD and BW analysed cloud properties. ER calculated the air parcel back-trajectories.
Competing interests
The authors declare that they have no conflicts of interest.

Disclaimer
The contents do not necessarily represent the official views of the University of Colorado, the University of Vienna, NOAA or of the respective granting agencies. The use or mention of commercial products or services does not represent an endorsement by the authors or by any agency.

Acknowledgements
We thank Ken Aikin for contributions to this analysis, and the ATom science team and NASA DC-8 flight crew for their contributions to the ATom data. We are grateful for the hard work of the ATom leadership and logistics teams. We thank the Whole Air Sampler (WAS; UCI) and the Trace Organic Gas Analyzer (TOGA; NCAR) teams for access to their data.

Financial Statement
This work was funded by NASA’s Earth System Science Pathfinder Program under award NNH15AB12I and by NOAA’s Health of the Atmosphere and Atmospheric Chemistry, Carbon Cycle, and Climate Programs. Agnieszka Kupc was supported by the Austrian Science Fund FWF’s Erwin Schrodinger Fellowship J-3613. Bernadett Weinzierl and Maximilian Dollner were supported by European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation framework program under grant 640458 (A LiFe) and by the University of Vienna. Jeffrey Pierce and Anna Hodshire were supported by the US Department of Energy’s Atmospheric System Research, an Office of Science, Office of Biological and Environmental Research program, under grant DE-SC0019000; and the NOAA, Office of Science, Office of Atmospheric Chemistry, Carbon Cycle, and Climate Program, under cooperative agreement award NA17OAR430001.

References


Jen, C. N., Zhao, J., McMurry, P. H., and Hanson, D. R.: Chemical ionization of clusters formed from sulfuric acid and dimethylamine or diamines, Atmos. Chem. Phys., 16, 12513-12529, 10.5194/acp-12513-2016, 2016.


A. P., Riccobono, F., Rissanen, M. P., Rondo, L., Schnitzhofer, R., Schobesberger, S., Smith, J. N.,
Steiner, G., Stožhkov, Y., Tomé, A., Tröstl, J., Tsagkogeorgas, G., Wagner, P. E., Wimmer, D., Ye,
spanning tropospheric sulfuric acid and ammonia abundances, ion production rates, and temperatures,


Lovejoy, E. R., Curtius, J., and Froyd, K. D.: Atmospheric ion-induced nucleation of sulfuric acid and water,

Merikanto, J., Napari, I., Vehkamäki, H., Anttila, T., and Kulmala, M.: New parameterization of sulfuric acid-
ammonia-water ternary nucleation rates at tropospheric conditions, J. Geophys. Res., 112,


Murphy, D. M., Cziczo, D. J., Froyd, K. D., Hudson, P. K., Matthew, B. M., Middlebrook, A. M., Peltier, R.


NASA Langley https://www-pm.larc.nasa.gov/ last accessed on June 1, 2018.


Pierce, J. R., and Adams, P. J.: Efficiency of cloud condensation nuclei formation from ultrafine particles,


Pierce, J. R., Riipinen, I., Kulmala, M., Ehn, M., Petäjä, T., Junninen, H., Worsnop, D. R., and Donahue, N.

Pierce, J. R.: Cosmic rays, aerosols, clouds, and climate: Recent findings from the CLOUD experiment, J.


