RESPONSE TO REVIEWER COMMENTS

Title: Aircraft-derived particle fluxes distinguish entrainment zone and decoupled layer nucleation in marine boundary layers

Journal: Atmospheric Chemistry and Physics

Ref: egusphere-2026-61

Referee Comments in 12-point italicized font

Authors' Response in indented, 12-point normal font.

Changes to the manuscript in quotes, 12-point blue font.

Response to 1st Referee's Comments

Specific Comments

#1. *Was there no measure of coarse-mode concentrations that are expected to be small in the SPE regions? This should be discussed, given that one of the main points of the paper seems to be to dispute the idea that sea-spray surface area inhibits new particle formation. As a proxy, even windspeeds (which aren't discussed except in the framework of flux measurements), might be helpful.*

> **Response:** Coarse-mode concentration data are now included in Figures 3, 6, S8, S11, S14, and S17 of the revised manuscript.

Revised manuscript text: "Figures 3–5 present data collected on January 29, 2018 with an additional example from February 10, 2018, shown in Supplementary Figs. S8–S10. Back-trajectory analysis (not shown) indicates that the sampled airmasses had been circulating around the Azores for the preceding three days and were therefore likely less polluted than North American outflow air masses. Figure 3 shows a multi-panel time series covering approximately 3.5 hours of flight operations. The aircraft initially ascended to ~2,500 m but generally remained below ~1,500 m for most of the flight (Fig. 3a). The flight trajectory (Fig. 3b) reflects predominantly east-west movement across the Azores region, spanning latitudes from approximately 38.7° to 39.4°N and longitudes from −28.4° to −27.4°W. Drizzle number concentration measured by 2DS (purple and blue lines in Fig. 3b) was absent or low during the selected SPE periods. Elevated drizzle number concentration, together with high liquid water content regions (orange in Fig. 3d) mark frequent cloud encounters. Following our quality control procedures, all N_{3-10} concentration data with LWC $\geq 3\text{x}10^{-3}$ g m^{-3} were excluded from analysis to avoid contamination from cloud droplet shattering artifacts. Pink-shaded periods mark the intervals chosen for detailed analysis, which exhibited simultaneous increases in both N_3 and N_{10} concentrations exceeding 10^4 cm^{-3} (indicating an SPE). Supermicron particle concentration (blue in Fig. 3c) as well as total particle surface area (green in Fig. 3c) were also low during the selected SPE periods, indicating the absence of particles such as sea spray aerosols."

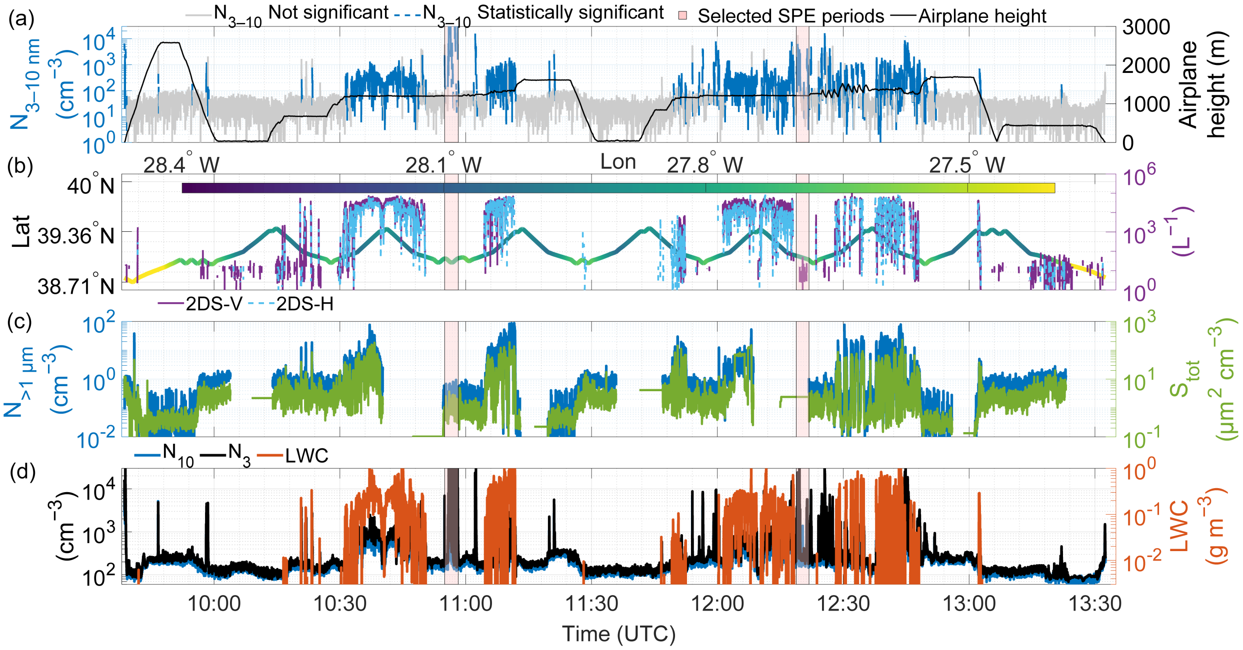

Figure 3. Multi-parameter time series from the January 29, 2018 flight. (a) $N_{3\text{-}10}$ particle concentrations and aircraft altitude; (b) aircraft position (latitude and longitude) and drizzle number concentration; (c) supermicron particle concentration and total particle surface area (S_{tot}) (d) particle number concentrations (N_{10} and N_3) and liquid water content. Gaps in the time series indicate the missing data.

"Figures 6–8 present data from June 21, 2017, with additional examples from July 7, 2017, February 18, 2018 and February 12, 2018, shown in Supplementary Figs. S11–S13, S14–16, and S17–S19. Back-trajectory analysis (not shown) indicates that the sampled airmasses originated from the Arctic and were therefore expected to be relatively clean. Figure 6 covers approximately 4 hours of flight operations, during which the aircraft initially flew at very low altitudes (~30 and 50 m) near 12:00 and 13:30 UTC before gradually ascending to ~1,000 m. Drizzle number concentration by 2DS (purple and blue lines in Fig. 6b) was absent or below the threshold during the selected SPE periods. Multiple events with $N_{3\text{–}10}$ concentrations from 10^2 to 10^4 cm^{-3} were observed throughout the second half of the flight. The flight trajectory (Fig. 6b) reflects predominantly east-west movement, spanning latitudes from approximately 38.6° to 39.3°N and longitudes from −28.4° to −27.4°W. Pink-highlighted intervals show periods with concurrent increases in N_3 and N_{10} concentrations exceeding 10^3 cm^{-3}, indicative occurrences of SPEs. As in Case 1, supermicron particle concentrations (blue in Fig. 6c) and total particle surface area (green in Fig. 6c) remained low during the selected SPE periods, indicating the absence of coarse-mode particles such as sea spray."

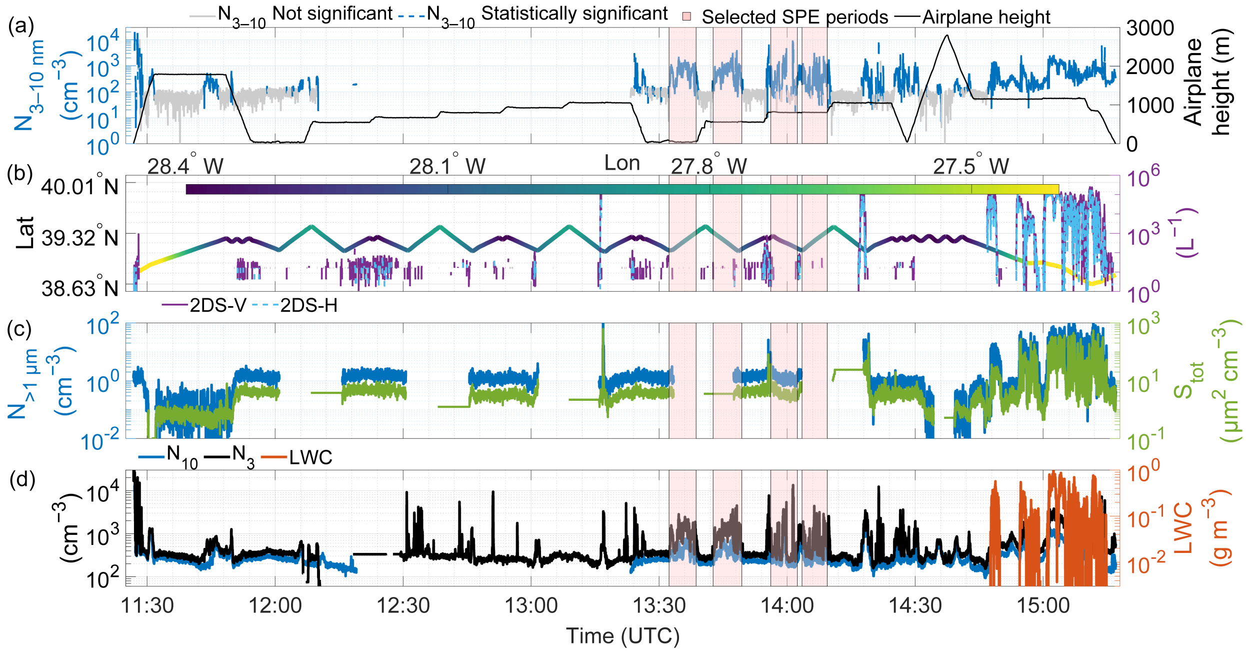

Figure 6. Same as Figure 3 but for June 21, 2017. Gaps in the time series indicate the missing data.

#2. *Also, it would be useful to show time series of aerosol surface area, at least in the sub-micron sizes, as that looks to be a potentially important change that occurs in the SPE regions. Fig 4 indicates almost no particles in the 10-1000 nm range in the second SPE period—is that correct, or is that just missing data? (Stot is shown and briefly discussed in the Fig 4 b-d vertical profiles, but these are not from the actual SPE periods.)*

> **Response:** Total aerosol surface area has been added to Figures 3, 6, S8, S11, S14, and S17. The absence of particle data in the 10–1000 nm range during the second SPE period in Figure 4 reflects a gap in data coverage.

#3. *Lines 428-436: Were these cases when clouds were present above? If so, did you check to rule out drizzle, which would breakup in the inlet and cause artifacts? Below-cloud drizzle would be independent of LWC, which is derived from the CDP (smaller sizes). Also, since droplet number concentrations are never given, it's unknown whether or not these were clean clouds, which are more likely to precipitate. The sometimes high (up to 1 g m^{-3}) LWCs would suggest drizzle as a possibility. Call me skeptical, but I've enough experience with in-situ cloud measurements from aircraft to know that artifacts are common. This point, as well as some discussion of which of the four source regimes (lines 75-78) that the presented measurements represent would be useful.*

> **Response:** Drizzle number concentration derived from the 2DS dataset has been added to the revised Figures 3, 6, S8, S11, S14, and S17. The low supermicron particle concentrations confirm that coarse-mode aerosol loading was minimal during SPE events, while drizzle concentration data confirm that drizzle was not detected outside of high-LWC cloud regions. A discussion of the source regimes represented by the presented measurements has also been added to the revised manuscript. Please also see our reply to Comment #1.

#4. *Fig 3a and discussion on lines 412-413—are these occasional spikes in 3-10 nm particles above the cloud layer--e.g., 11:20-11:25--real? Should be discussed to assure that some of the signals noted below are not from entrainment of particles from above. Unlikely, since the concentrations below are sometimes higher, and these are short spikes, so perhaps they are artifacts of some sort?*

> **Response:** The brief spikes in $N_{3–10}$ concentrations above the cloud layer (e.g., 11:20–11:25 UTC) cannot be conclusively identified as artifacts; however, they do not contribute to any of the analysis presented. A clarifying statement has been added to the revised manuscript

Revised manuscript text: "This interpretation is supported by the near-absence of $N_{3–10}$ at ~1,600 m during 11:14–11:25 and 12:51–13:01 UTC (Fig. 3), with the exception of brief concentration spikes of uncertain origin retained in the record due to insufficient evidence for their removal. The small particle size (3–10 nm) and limited horizontal extent of less than 10 km further argue against a free tropospheric nucleation source, as particles originating in the free troposphere would be expected to have grown substantially and the plume to have diluted during descent to measurement altitude."

#5. *Lines 43-45: "This expectation is based on the relatively high surface area of sea spray aerosols, which act as condensation and coagulation sinks for nucleating vapors and newly formed particles". There are plenty of accumulation-mode sulfate/organic particles in most MBLs that also may act as condensation and coagulation sinks. As do clouds themselves. This should be mentioned.*

Revised manuscript text: "This expectation is based on the high condensation and coagulation sink capacity of the remote MBL, which includes not only sea spray aerosols (Bates et al., 1998; Pirjola et al., 2000) but also accumulation-mode sulfate and organic particles entrained from the free troposphere (Yoon et al., 2001). Clouds further suppress NPF by scavenging Aitken-mode particles (Zheng et al., 2018), accelerating sulfate production on existing droplets through aqueous-phase SO_2 oxidation (Sanchez et al., 2021), and sequestering DMS oxidation products such as HPMTF that would otherwise contribute to sulfuric acid formation (Novak et al., 2021)."

#6. *Perhaps move some of the lengthy data and analysis criteria in Section 2 to the supplement, where some of the associated graphs already are?*

> **Response:** Section 2.3.2 is moved to the supplementary information

#7. *Lines 340-342: "We examine two flight days as case studies of SPEs observed at varying altitudes above the ocean. Additional supporting flights are presented in the Supplementary Information for each case." After this you go immediately into Table 1 that shows six different flight days, which I found confusing. Perhaps discuss Table first and then go into the case studies.*

Revised manuscript text: "We examine two flight days as case studies of SPEs observed at varying altitudes above the ocean. Additional supporting flights are presented in the

Supplementary Information. Table 1 summarizes the N_{3-10} vertical turbulent flux estimates derived from all six flight days analyzed in this study, grouped by the inferred nucleation regime. Flights 1 and 2 (January 29 and February 10, 2018) are classified as entrainment zone nucleation events, where SPEs were detected near the top of the MBL at heights exceeding 1,200 m. Flights 3–6 (June 21 and July 7, 2017; February 18 and 12, 2018) are classified as decoupled layer nucleation events, with SPEs observed across a broader range of altitudes (30–837 m). For all events, the ratio of measured flux to the spectrally-complete flux $\left(\frac{F_m}{F}\right)$ exceeds 0.76, indicating minimal flux loss due to sensor response limitations. The normalized vertical velocity variance ($\sigma_w^2\ w_*^{-2}$) is generally low, consistent with relatively quiescent turbulent conditions during the measurement periods. Negative flux values indicate downward transport of freshly nucleated particles from the entrainment zone toward the surface, while positive values suggest upward transport from a source within the decoupled sub-cloud layer. Two of these flight days, January 29, 2018 (Case 1) and June 21, 2017 (Case 2), are examined in detail as case studies in the following sections, with the remaining four flights presented as supporting examples in the Supplementary Information."

#8. *Figure 4a caption: "The main panel shows size-resolved particle number concentrations (10-600 nm) from FIMS as a function of time and altitude, while N3-10 concentrations in the lower strip." I think "are shown" is missing before "in the lower strip".*

> **Response:** Changed as suggested.

Revised manuscript text: "Figure 4. (a) Size-resolved particle number concentrations (10–600 nm) from FIMS as a function of time and altitude, with N_{3-10} concentrations shown in the lower strip. Pink shading indicates selected SPE periods. (b-d) Vertical profiles of potential temperature (θ), normalized vertical velocity variance ($\sigma_w^2\ w_*^{-2}$), total particle surface area (S_{tot}), and water vapor mixing ratio (MR_{H2O}) for three time periods nearest to the selected SPE periods: (b) 09:51–10:01 UTC, (c) 11:25–11:31 UTC, and (d) 13:01–13:07 UTC. Gaps in the time series indicate the missing data."

#9) *I would suggest making 4a and 4b-d (and 7a and 7b-d) into separate plots, as these are really different from each other and this is too much information to easily assimilate in a single plot.*

> **Response:** We respectfully disagree with this suggestion. Panel (a) provides essential context (the time evolution of size-resolved concentrations and the altitude of the aircraft) that is needed to interpret the vertical profiles shown in panels (b–d). Separating these into distinct figures would obscure the connection between the timing of SPE periods and the concurrent thermodynamic and turbulence structure of the boundary layer. Furthermore, panels (b–d) are intended to be viewed comparatively, as they document the evolution of potential temperature, normalized vertical velocity variance, total aerosol surface area, and water vapor mixing ratio across three successive time windows. We believe the current layout presents this information in the most coherent and space-efficient manner, and have added a brief note to the figure caption to guide the reader through the panel structure.

#10. *Fig 7 has a lot of missing data, so I'd suggest explicitly adding the sentence from Fig 4 caption to Fig 7 caption as well. This confused me until I figured it out. ("Gaps in the time series indicate the missing data.")*

Response: Changed as suggested.

Revised manuscript text: "Figure 7. Same as Figure 4, but for June 21,2017. Gaps in the time series indicate the missing data."

Response to 2nd Referee's Comments

General comment:
This paper describes aircraft-based observations of aerosol particles, primarily in the size range between 3 and 10 nm, which are assumed to be newly nucleated. Turbulent particle fluxes are estimated using additional parameters and wavelet analysis. The technique is discussed as having advantages over other methods, and the detailed description therefore takes up a large portion of the manuscript. Although the determination of particle fluxes remains a major challenge and is therefore of great significance, in my view the classification and interpretation of the results can and should be improved. A key issue here is the imbalance between the details of the flow determination and the interpretation of the results. To take it to an extreme, the conclusions could already be drawn from the sign of the particle fluxes, since the absolute value is not interpreted in the slightest and thus loses its significance. In contrast, in some places the interpretation of observations of particles formation involves a high degree of speculation. Unfortunately, the work is also not comprehensively contextualized with the help of additional literature on particle formation in marine boundary layers; this becomes particularly evident in the last two chapters (see specific comments on this).

For this reason, this manuscript requires a thorough revision ("major revisions") regarding the interpretation and contextualization of the observations and analyses.

As this review proceeds, I will offer both general and specific comments. However, I hope that these comments will be received constructively and will contribute to a successful revision of the manuscript, as I believe the measurements are definitely worth analyzing and publishing.

Abstract :
#1. *It is somewhat misleading to attempt to deduce the location of new particle formation using the flux of particles in the size range between 3 and 10 nm. Firstly, these particles have already aged and grown, and secondly, the flux only indicates the direction and quantity per time in which particles are transported. Perhaps this should be formulated more clearly.*

Revised manuscript text: "3–10 nm sized particles, defined as "small particle event" (SPE), play a critical role in marine boundary layer aerosol budgets and aerosol–cloud interactions, yet the vertical distribution of their sources remains poorly constrained. We identified the vertical location of SPEs by deriving turbulent fluxes of 3–10 nm particles from aircraft measurements during the Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA) campaign. To overcome stationarity limitations of traditional eddy covariance methods, we applied continuous wavelet transform analysis to data collected during June–July 2017 and January–February 2018 flights over the Azores. Our flux-based analysis revealed two distinct SPE scenarios with different vertical structures and spatial extents. The first featured nucleation in the entrainment zone, where free tropospheric air entrains into the boundary layer. The second showed nucleation in the decoupled layer, a stratified region between the well-mixed surface layer and cloud-topped upper boundary layer. In both cases, convergence of air masses from different layers diluted preexisting aerosol surface area to very low levels, creating conditions favorable for nucleation and generating strong downward particle fluxes. SPEs occurred in 15% of flights, challenging prevailing theoretical expectations that new particle formation should rarely occur in marine boundary layers due to high condensation and coagulation sink capacity of sea spray aerosols. Aircraft-derived particle fluxes provide first observational constraints on the vertical location and source strength of likely nucleation regions in the remote marine boundary layer, improving aerosol source representations in climate models and reducing uncertainties in aerosol-cloud interactions."

#2. *It has become common practice to use the following phrasing (or something similar): When ultrafine particles in the range of 3 to 10 nm are measured, we refer to them as "freshly nucleated particles," but never as "new particle formation," since this process definitely cannot be observed with these measurements. As I continue reading the manuscript, I notice that you use a similar term later on; perhaps this should simply be clarified a little earlier.*

> **Response:** We thank the reviewer for this important methodological point. We agree that 3–10 nm particle measurements cannot directly confirm nucleation, as the initial cluster formation at ~1–3 nm is below our detection threshold. Throughout the revised manuscript, we have replaced direct references to 'NPF events' and 'new particle formation' in the abstract and introduction with more precise language that distinguishes between what is measured (freshly nucleated 3–10 nm particles during small particle events, SPEs) and what is inferred (NPF as the most likely source). Specifically, we now refer to 'flux-based observational constraints on the vertical distribution of likely NPF source regions' rather than claiming direct observation of nucleation, and describe our scenarios as 'consistent with NPF' rather than definitive NPF occurrences. This terminology is now introduced in the abstract and introduction, consistent with the SPE framework already established in Section 2.3 following Islam et al. (2022). We note that while direct observation of nucleation is beyond the capability of our measurements, the rapid appearance of 3–10 nm particles with coherent vertical flux structure provides strong indirect evidence for nearby nucleation source regions, and we believe this distinction is now clearly communicated throughout the manuscript.

Revised manuscript text: Abstract: "We identified the vertical location of NPF events by deriving turbulent fluxes of 3–10 nm particles" was changed to "We identified the vertical location of SPEs by deriving turbulent fluxes of 3–10 nm particles..." Also "Our flux-based analysis revealed two distinct NPF scenarios" was changed to "We identified the vertical location of SPEs by deriving turbulent fluxes of 3–10 nm particles." "NPF occurred in 15% of flights" was changed to "SPEs occurred in 15% of flights…"
Introduction: "vertical turbulent flux measurements of freshly nucleated particles have emerged as particularly valuable tools for characterizing the vertical location of particle nucleation" was changed to "vertical turbulent flux measurements of freshly nucleated 3–10 nm particles have emerged as particularly valuable tools for inferring the vertical location of likely nucleation source regions..." "The flux direction provides direct evidence of nucleation location" was changed to "The flux direction provides indirect evidence of the likely nucleation location..." "we provide the first direct observational constraints on the vertical distribution of new particle formation" was changed to "we provide the first flux-based observational constraints on the vertical distribution of likely NPF source regions."

#3. *I always find it a bit difficult when people talk about new particle formation on the one hand and CCN on the other; these are completely different modes, and it remains questionable whether locating NPF will lead to a better understanding of the number concentration of CCN. Why not argue that we want to understand the budget of aerosol particles in general? It is not at all clear whether the small particles will ever grow into CCN. I think a couple more sentences could be added here.*

> **Response:** We appreciate the reviewer's comment and acknowledge that the connection between NPF and CCN is not instantaneous and depends critically on growth rates, coagulation losses, and the timescales involved. However, we respectfully maintain that NPF is a well-established pathway to CCN production in marine environments, provided sufficient time for particle growth. The timescale for freshly nucleated particles to grow to CCN-relevant sizes (>50–80 nm) at typical marine growth rates of 1–3 nm hr^{-1} is on the order of 24–48 hours, which is consistent with the residence times of air masses in the remote marine boundary layer. Short-term measurements that do not capture this full growth cycle may therefore underestimate the NPF contribution to CCN budgets. We have added several sentences to the introduction and discussion acknowledging that the CCN activation potential of freshly nucleated particles depends on growth timescales and loss processes, and that understanding the vertical origin and flux magnitude of 3–10 nm particles is a necessary first step toward constraining their ultimate contribution to the marine aerosol number budget more broadly.

Revised manuscript text: Introduction: "While freshly nucleated particles in the 3–10 nm size range must undergo substantial growth before reaching CCN-relevant sizes (>50–80 nm), this growth pathway is well established in marine environments. At typical marine boundary layer growth rates of 1–3 nm hr^{-1} (Ehn et al., 2010; Nieminen et al., 2018; O'Dowd et al., 2010; Zheng et al., 2018), newly formed particles can reach CCN sizes within 24–48 hours. This timescale is consistent with air mass residence times in the remote marine boundary layer (Zheng et al., 2021; Kulmala et al., 2012). Constraining the vertical location and flux magnitude of freshly nucleated particles therefore represents a critical first step toward understanding the full aerosol number budget in marine environments, including the ultimate contribution of NPF to CCN populations."
Conclusions: "We note that the contribution of freshly nucleated particles to CCN budgets depends on growth rates and loss processes during transport, and cannot be assessed from short-term measurements alone. The flux magnitudes and vertical source locations reported here provide the observational foundation needed to evaluate this contribution quantitatively in future studies combining particle flux measurements with growth rate and CCN closure analyses."

#4. *Incidentally, vertical profiles of aerosols have already been observed above the same island at almost the same time, which also indicate freshly formed aerosol particles. hTps://doi.org/10.1175/BAMS-D-19-0191.1 and references in there (see also my comment about the "discussion" and "conclusion" at the end of my coments).*

> **Response:** We thank the reviewer for pointing out the ACORES campaign observations (Siebert et al., 2021), which we have now cited in the revised manuscript. The ACORES helicopter-borne measurements over Graciosa in July 2017 (overlapping in time and location with our ACE-ENA observations) provide independent corroboration of freshly nucleated particle bursts near the cloud top and in the free troposphere over the eastern

> North Atlantic. Importantly, Siebert et al. (2021) note that the observed NPF events did not influence CCN concentrations within their ~2-hour observation window, which is entirely expected given that newly formed 3–10 nm particles require 24–48 hours to grow to CCN-relevant sizes at typical marine growth rates of 1–3 nm hr^{-1}. Our flux-based measurements complement the ACORES vertical profiles by identifying the vertical source regions of freshly nucleated particles and quantifying their turbulent transport rates, information that is necessary for assessing their eventual contribution to CCN budgets on timescales beyond individual flight periods. The two datasets together strengthen the case that in-situ NPF is a recurring feature of the eastern North Atlantic MBL that warrants inclusion in long-term aerosol budget assessments.

Revised manuscript text: "Independent corroboration of in-situ NPF over the eastern North Atlantic comes from the Azores Stratocumulus Measurements of Radiation, Turbulence and Aerosols (ACORES) campaign, which conducted helicopter-borne observations over Graciosa Island in July 2017, overlapping in time and location with the ACE-ENA campaign analyzed here. Siebert et al. (2021) reported frequent burst-like freshly nucleated particle events near the stratocumulus cloud top and in the free troposphere, while also noting that these particles did not grow to CCN-relevant sizes within the ~2-hour observation window. This outcome is expected given the 24–48 hour growth timescales discussed above. These concentration-based observations, however, could not determine the precise vertical location of the nucleation source regions, motivating the flux-based approach developed in the present study."

#5. *line 106: please specify what you mean with "excellent sensitivity", without numbers or more details such phrases should be avoided*

Revised manuscript text: "The FIMS provides size distribution measurements at 1-second temporal resolution with signal-to-noise characteristics suitable for detecting both remote continental and clean marine aerosol concentrations, as demonstrated in aircraft-based deployments (Kulkarni and Wang, 2006a, b; Olfert et al., 2008)."

#6. *line 116/117: What is the background to this criterion? Why 10%?*

> **Response:** The FIMS measures particles from 10–600 nm, which overlaps almost entirely with the size range of the CPC 3772 (which has a 50% counting efficiency cutoff at 10 nm). Under normal operating conditions, the two instruments should measure comparable total concentrations, since the vast majority of accumulation and Aitken mode particles in the marine boundary layer fall within the FIMS detection range. A CPC concentration that falls below 10% of the simultaneously measured FIMS concentration therefore indicates a severe and physically implausible discrepancy (most likely caused by inlet blockage, flow rate malfunction, or counting electronics failure in the CPC) rather than a genuine atmospheric signal. The 10% threshold is conservative enough to avoid flagging real atmospheric variability (including periods of very low particle concentrations) while still capturing instrument malfunctions.

Revised manuscript text: "FIMS-derived number concentration also served as a quality control flag for the CPC 3772. Since both instruments share an overlapping detection size range (10–600 nm for FIMS; >10 nm for CPC 3772), their total number concentrations should be broadly comparable under normal operating conditions. CPC concentrations falling below 10% of the simultaneously measured FIMS concentration indicate a physically implausible discrepancy inconsistent with real atmospheric variability, and were therefore excluded from analysis as likely instrument malfunctions."

#7. *Section 2.3.2*
At the beginning of the chapter, you describe the problem of temporal synchronization. But what I find missing is a statement about how you specifically correct the problem in your aircraft measurements. It all remains a little vague. In the second paragraph of this section, you then describe statistical collection errors, which has nothing to do with the title. I also think that the comparison with the mast measurements distracts from your specific problem—I'm actually interested in how you deal with your challenges.

> **Response:** We thank the reviewer for this helpful critique. We agree that Section 2.3.2 was poorly organized and insufficiently specific about our actual lag correction procedure. We have restructured the section to focus directly on how lag times were determined for our aircraft measurements, removing the general discussion of tower-based measurement platforms which, as the reviewer correctly notes, distracts from the specific problem at hand. The revised section now clearly states that lag times were determined individually for each flux calculation period using covariance maximization between the vertical wind speed and particle concentration signals, and separately validated using pressure measurements from the isokinetic inlet and the AIMMS-20 probe. The statistical characterization of lag time variability across the campaign has been moved to the Supplementary Information, where it provides context without interrupting the methodological narrative.

Revised manuscript text: "Accurate temporal alignment between the vertical wind speed measured by the AIMMS-20 probe and the particle concentration measured by the CPCs is essential for reliable flux calculations. Because these instruments were located at different positions on the aircraft, a time lag exists between the two signals that must be determined and corrected prior to flux calculation.

To confirm that the two CPCs sampled identical air masses simultaneously, Spearman correlation coefficients were calculated for concentration measurements from both CPCs after removing cloud shattering artifacts and excluding SPE periods. From the complete campaign dataset, 370 randomly selected seconds of data yielded an average Spearman correlation coefficient of 0.97 (Supplementary Figures S1–S2), confirming adequate synchronization between the two concentration records. However, a high correlation coefficient alone does not determine the precise temporal offset between the two signals.

Lag times between the two CPC signals were determined individually for every 20-second interval (representing the time taken for the airplane to traverse 2 km) using covariance maximization, shifting one CPC relative to the other signal to identify the temporal offset that maximizes their covariance. A single fixed lag time across the entire campaign was not appropriate, given the variability in lag times observed across flight segments (Supplementary Figure S3). This approach was independently validated using pressure measurements from the isokinetic inlet and static pressure from the AIMMS-20 probe. The two pressure records yielded a Spearman correlation coefficient of 0.99, confirming that both instruments consistently sampled the same air mass with no systematic offset (Supplementary Figure S5). Similarly, covariance maximization applied to the pressure records confirmed that no single lag time was appropriate across the full campaign (Supplementary Figure S6), consistent with the CPC-based analysis and further supporting the use of individually determined lag times for each flux calculation period."

Revised supplementary text:

"**Time lag correction**

Accurate temporal alignment between the vertical wind speed and particle concentration signals is essential for reliable flux calculations. Since the AIMMS-20 probe and the isokinetic inlet are located at different positions on the aircraft, a time lag exists between the two signals. Lag times were determined individually for each flux calculation period using covariance maximization, as described in Section 2.3.2 of the main text. This section provides the supporting validation analysis.

CPC synchronization validation

Both CPCs draw sample air from the same isokinetic inlet manifold. However, because each CPC detector unit is connected to the manifold via tubing of different length, the air transport time from the inlet to each detector differs (Figure 4, Goldberger, 2020). This difference in transport time means the two CPCs may not sample identical air parcels at exactly the same moment, necessitating validation of their temporal synchronization. Figure S1 shows concentration measurements from both CPCs for a representative flight day. Despite differing absolute values, expected given their different size detection limits (3 nm vs. 10 nm), the temporal patterns closely align. To quantify this agreement, Spearman correlation coefficients were calculated for both CPC records after removing cloud droplet shattering artifacts and excluding SPE periods, since only the ultrafine CPC detects SPEs. From the complete campaign dataset, 370 seconds of data were randomly selected to avoid selection bias, yielding a mean ρ of 0.97 (Figure S2).

The full CPC dataset was segmented into 20-second intervals, each corresponding to approximately 2 km of horizontal flight path, and lag times were determined for each interval using covariance maximization (Figure S3). Lag times of 0 and 1 seconds occurred in 13% and 27% of intervals respectively. The variability in lag times arises primarily from the 1 Hz temporal resolution of the measurements: since the true lag is unlikely to be an exact integer number of seconds, it is effectively rounded to the nearest second when determined by covariance maximization, causing an apparent alternation between 0 and 1 second values. Secondary contributions may arise from minor variations in flow velocity within the inlet tubing due to changes in ambient pressure and temperature with altitude, despite active flow rate regulation. On the spatial scale, lag times of 0 and 1 seconds correspond to spatial displacements of 0–100 m at typical aircraft speeds, which is small relative to the flux-carrying eddy scales resolved in this study. No single lag time was sufficiently dominant to apply uniformly across the campaign; applying a uniform lag would therefore introduce systematic errors in the flux covariance. Sensitivity analysis confirmed that derived fluxes change by an average of less than 16% when the lag is varied by ±1 second, supporting the robustness of the individual lag determination approach.

Pressure-based lag time validation

The lag correction was independently validated by comparing pressure measured at the isokinetic inlet with static pressure measured by the Rosemount 1201F1 sensor on the AIMMS-20 probe. Since both sensors respond to the same ambient pressure field, they should be in close agreement if the instruments are sampling the same air mass. Figure S4 shows pressure time series from both sensors for a representative flight day, confirming that both records track the same pressure variations at 1 Hz resolution. Figure S5 shows the Spearman correlation coefficient between the two pressure records across the full campaign, yielding ρ = 0.99 with values closely following the

1:1 line, confirming that both instruments sampled the same air mass with negligible systematic offset. Covariance maximization applied to the pressure records (Figure S6) again showed that no single lag time was appropriate across the full campaign, consistent with the CPC-based analysis and further supporting the use of individually determined lag times for each flux calculation period."

#8. *For example, in line 156ff, you first describe the vertical distribution of turbulence intensity and then the sampling time required to determine a turbulent flux with a statistical accuracy of 10%. Each part on its own is fine, but the connection between them is not clear to me. In the next sentence, you refer to the spatial resolution of the measurements as a function of the sampling frequency of a sensor, which has nothing to do with the previous statement—a red line of the story is missing here.*

Revised manuscript text: "Flux measurement methods were originally developed for tower-based platforms, and their application to aircraft measurements introduces fundamental differences in sampling characteristics that must be carefully considered. Tower measurements provide continuous observations at fixed heights, capturing the complete turbulent eddy spectrum including low-frequency contributions essential for accurate flux estimates (Helbig et al., 2021; Sakai et al., 2001). Aircraft measurements, by contrast, sample different air masses as the platform moves horizontally, effectively trading temporal for spatial averaging (Desjardins et al., 1989).
Three interconnected challenges arise specifically for aircraft-based flux measurements. First, turbulent intensity in convective boundary layers increases with height above the surface layer before decreasing above 0.3–0.4 z_i (where z_i is the boundary layer height). Maintaining flux variance within 10% therefore requires measurement lengths of 100 to 10^4 times the boundary layer height (Lenschow and Stankov, 1986), a constraint that becomes increasingly difficult to satisfy at the higher altitudes routinely sampled by research aircraft. Second, high aircraft speeds impose strict constraints on sensor response times: for an aircraft traveling at 100 m s^{-1}, a 1 Hz sampling system resolves eddies no smaller than 200 m, approaching the lower limit for capturing the dominant flux-carrying scales. The CPCs used in this study operate at 1 Hz, meaning that contributions from smaller eddies are not resolved and must be accounted for through flux loss corrections (Section 2.8). Third, and most critically for flux calculation, aircraft measurements are inherently non-stationary as the platform continuously moves through different air masses, meteorological conditions, and altitudes. Traditional eddy covariance methods assume stationarity over the averaging period, a condition that is difficult to maintain during aircraft sampling (Gioli et al., 2004).
To address this limitation, this study employs the continuous wavelet transform (CWT) method for flux derivation. The primary advantage of the CWT approach is that it does not require stationarity and eliminates the need for data detrending, thereby avoiding systematic errors that can arise from linear detrending procedures (Rannik and Vesala, 1999). This study follows the CWT flux derivation method of (Torrence and Compo, 1998), described below."

#9. *Line 161: The statement that 30-minute averages based on mast measurements "easily" fulfill stationarity conditions is somewhat harshly worded. I also don't quite understand why you are venturing onto such thin ice in this context.*

> **Response:** We agree with the reviewer on both points. The word 'easily' was poorly chosen and overstates the case as stationarity is a well-recognized challenge even for tower-based

> eddy covariance measurements. More importantly, we have taken the reviewer's broader advice and removed the detailed tower-aircraft comparison from this section entirely, retaining only the statement directly relevant to our methodological choice: that aircraft measurements are inherently non-stationary due to continuous platform movement through varying air masses, and that the CWT method is better suited to handle this than traditional eddy covariance approaches.

Revised manuscript text: "Third, and most critically for flux calculation, aircraft measurements are inherently non-stationary as the platform continuously moves through different air masses, meteorological conditions, and altitudes. Traditional eddy covariance methods assume stationarity over the averaging period, a condition that is difficult to maintain during aircraft sampling (Gioli et al., 2004)."

#10. *Line 167: I don't quite understand the sentence; do the two CPCs now draw their sample air from the same inlet—if so, what is the problem? What do you mean by "same inlet" and "from different positions"—please clarify (maybe a sketch would help here)*

Revised supplementary text: "Both CPCs draw sample air from the same isokinetic inlet manifold. However, because each CPC detector unit is connected to the manifold via tubing of different length, the air transport time from the inlet to each detector differs (Fig. 4, Goldberger, 2020). This difference in transport time means the two CPCs may not sample identical air parcels at exactly the same moment, necessitating validation of their temporal synchronization."

#11. *Why is the time lag for the two CPCs not constant or did I misunderstand this part? Are the flows in the inlet system not regulated? Or is there even a pressure dependency? This should be clarified in this context, because in general I would assume a constant time lag. Under what conditions do you determine a time lag of 0 or even one second? Shouldn't this have a significant influence on particle flow determination?*

> **Response:** The reviewer raises an important point. While the flow rates in the isokinetic inlet system were regulated, the observed variability in lag times arises from two sources. First, at 1 Hz temporal resolution, the true lag (which may be a non-integer number of seconds) is rounded to the nearest integer, meaning that a true constant lag of, say, 0.6 seconds would appear as either 0 or 1 second depending on the specific data segment. Second, minor variations in ambient pressure and temperature with altitude can subtly affect flow velocities within the tubing despite active flow regulation, introducing small but real variability in transport time. The sensitivity analysis confirms that this variability has a modest effect on the derived fluxes (less than 16% when the lag is varied by ±1 second) supporting the robustness of the individual lag determination approach.

Revised supplementary text: "The variability in lag times arises primarily from the 1 Hz temporal resolution of the measurements: since the true lag is unlikely to be an exact integer number of seconds, it is effectively rounded to the nearest second when determined by covariance maximization, causing an apparent alternation between 0 and 1 second values. Secondary contributions may arise from minor variations in flow velocity within the inlet tubing due to changes in ambient pressure and temperature with altitude, despite active flow rate regulation. On the spatial scale, lag times of 0 and 1 seconds correspond to spatial displacements of 0–100 m at typical aircraft speeds, which is small relative to the flux-carrying eddy scales resolved in this study. No single lag time was sufficiently dominant to apply uniformly across the campaign; applying a uniform lag would therefore introduce systematic errors in the flux covariance. Sensitivity analysis confirmed that derived fluxes change by an average of less than 16% when the lag is varied by ±1 second, supporting the robustness of the individual lag determination approach."

#12. *I also think that if you discuss figures in such detail, they should appear in the paper itself and not in the supplement. I generally have a problem with discussing figures from the supplement—I would avoid that entirely. Either a figure contributes to the narrative of your manuscript, in which case it belongs in the manuscript, or you mention it without further discussion.*

> **Response:** We agree that detailed discussion of supplementary figures is better placed in the supplementary information itself rather than in the main text. Accordingly, all detailed discussion of supplementary figures has been moved to the supplementary information section, where it accompanies the figures directly. The main text now references supplementary figures only briefly where necessary, without extended discussion.

13. *Line 187 ff: In line 189 you mention that the time lag is negligible but in the last sentence you state that for each case individual time lags are determined: i) what do you mean with "case" and ii) why do you have to determine a time lag when it is negligible?*
I think the discussion you are having here is necessary, but in some places, it needs to be presented and clarified a little better.

Revised manuscript text: "Lag times between the two CPC signals were determined individually for every 20-second interval (representing the time taken for the airplane to traverse 2 km) using covariance maximization, shifting one CPC relative to the other signal to identify the temporal offset that maximizes their covariance. A single fixed lag time across the entire campaign was not appropriate, given the variability in lag times observed across flight segments (Supplementary Figure S3). This approach was independently validated using pressure measurements from the isokinetic inlet and static pressure from the AIMMS-20 probe. The two pressure records yielded a Spearman correlation coefficient of 0.99, confirming that both instruments consistently sampled the same air mass with no systematic offset (Supplementary Figure S5). Similarly, covariance maximization applied to the pressure records confirmed that no single lag time was appropriate across the full campaign (Supplementary Figure S6), consistent with the CPC-based analysis and further supporting the use of individually determined lag times for each flux calculation period."

#14. *Regarding the spectra in Fig. 1: All spectra are shown up to the Nyquist frequency of 0.5 Hz; while for the vertical wind, the spectrum—especially the one calculated with FFT—certainly flattens out for frequencies above 0.3 Hz and transitions into noise, I am surprised that no flattening can actually be observed for the particle measurements. This somewhat contradicts your statement regarding the CPC behavior – right?*

> **Response:** We thank the reviewer for this careful observation. The new Figure comparing 1 Hz and 10 Hz vertical wind velocity spectra demonstrates that the spectral flattening above 0.3 Hz in the 1 Hz wind spectrum is a sampling artifact. The 10 Hz spectrum

continues to follow the theoretical −5/3 slope well beyond this frequency (as shown in Supplementary Figure S7), confirming that atmospheric turbulent energy exists at these scales but is not resolved at 1 Hz. The absence of visible flattening in the particle concentration spectra above 0.3 Hz is not contradictory but reflects a different manifestation of instrument response limitations. Unlike the wind sensor, the CPC acts as a first-order low-pass filter with a ~3 second response time constant, which attenuates high-frequency concentration fluctuations rather than introducing white noise. This attenuation produces a steeper spectral roll-off rather than a flat noise floor, making the limitation less visually apparent in the concentration spectra while still resulting in flux loss at high frequencies, which we account for through the Horst (1997) correction described in Section 2.8. The revised Figure 1 and the new supplementary figure comparing 1 Hz and 10 Hz vertical wind velocity spectra are shown below.

Revised manuscript text: "Both particle concentration spectra (Figs. 1a,b) and the flux spectrum (Fig. 1d) broadly follow the theoretical −5/3 and −7/3 Kolmogorov scaling across the resolved frequency range. The supplementary figure comparing 1 Hz and 10 Hz vertical wind velocity spectra (Supplementary Fig. S7) confirms that turbulent energy exists at scales beyond the 1 Hz Nyquist frequency. The absence of a visible noise floor in the particle concentration spectra at high frequencies reflects the ~3 second response time of the CPC, which acts as a low-pass filter that attenuates high-frequency concentration fluctuations, producing a steep spectral roll-off rather than a white noise floor. While this results in a cleaner spectrum visually, it still represents real flux loss at high frequencies that is accounted for through the correction described in Section 2.8."

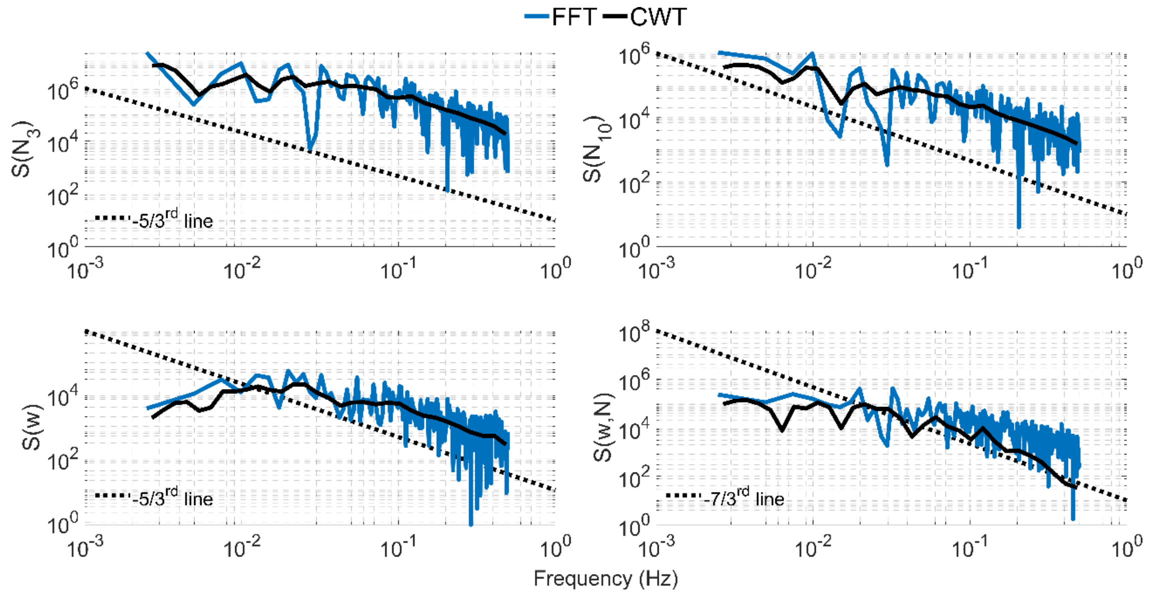

Figure 1. Power spectral density for (a) N_3, (b) N_{10}, (c) vertical wind velocity, and (d) 3-10 nm particle flux.

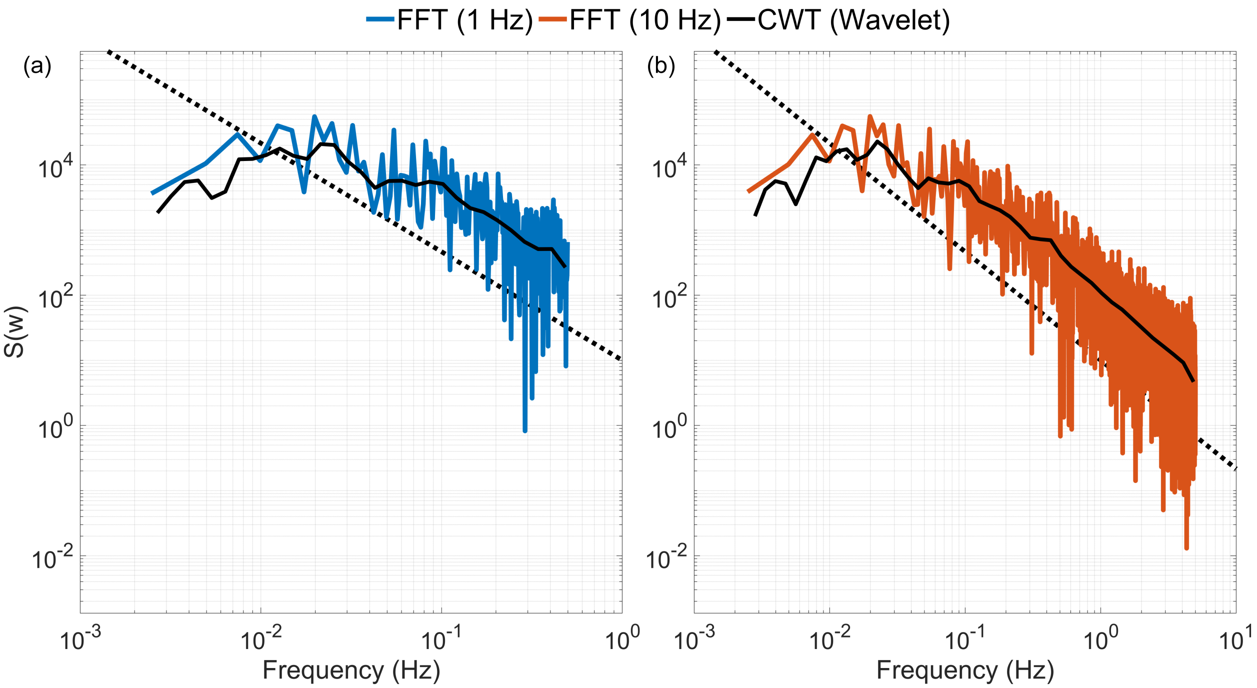

Figure S7. Power spectral density for vertical wind velocity at (a) 1 Hz and (b) 10 Hz.

#15. *Why can't the integral time scale in Sec 2.5 be determined to estimate the sampling error? Shouldn't that be easy to do using the autocorrelation function?*

Revised manuscript text: "The integral time scale was calculated using the cross-correlation function between vertical wind w and 3-10 nm sized particle size concentration, using the method described by (Lenschow et al., 2000; Wulfmeyer et al., 2016). The cross covariance is given as:

$$A_{x,y}(\tau) = \mathrm{cov}(x_t, y_{t+\tau}) \tag{7}$$

Where x_t and y_t are the two signals of interest, with y shifted by the lag time τ. Now the cross covariance is fitted to a model of the form:

$$A_{\mathrm{model}}(\tau) = \nu - k\tau^{\frac{2}{3}} \tag{8}$$

Where ν and k are fitted parameters. The line is fitted till the first zero crossing of the cross covariance. And using this the integral timescale, I is calculated as:

$$I = \frac{2}{5}\left(\frac{\nu}{k}\right)^{3/2} \tag{9}$$

The median value of the integral timescale for the flux events was ~ 11 seconds. The limit of detection (LoD) is defined as 1.96 times (95% confidence interval) the standard deviation of covariance between w and N3-10 when one of the signals is temporarily shifted with respect to the other. To estimate the precision of an individual flux determination, we followed Spirig et al. (2005) and examined fluctuation of the covariance function at time lags far away from the true lag.

The standard deviations of the covariance function were calculated between 4500 to 5700 seconds (corresponding to 45 to 57 km spatial displacement) before and after the peak fluxes. These lag times were selected to be much higher than the integral timescale of turbulence at all measurement heights."

#16. *Line 278: Isn't the response of a CPC also a question of number concentration and therefore of counting statistics? What is the argument for the 3s resolution?*

> **Response:** The reviewer raises a valid point. CPC accuracy is concentration-dependent and governed by counting statistics rather than a fixed response time, as documented in Kuang and Mei (2019). However, following the approach of Islam et al. (2022), we assessed the random uncertainty in particle flux due to counting statistics and found it to be negligible compared to the measured flux magnitudes reported in this study. The original text referencing a fixed 3-second response time has been removed and replaced with a more accurate description of the CPC's concentration-dependent measurement uncertainty.

Revised manuscript text: "Following the approach of Islam et al., (2022), we assessed the random uncertainty in particle flux due to counting statistics. We found it to be 2–3 orders of magnitude lower compared to the measured flux magnitudes reported in this study. "

#17. *Is there an explanation for the flattening of the wind spectrum? I would have expected slightly better performance from the system.*

> **Response**: Please see our reply to comment #14.

#18. *Perhaps it's also a question of style, but you should really avoid repetition in an academic text; here you start mentioning the advantages of wavelets over FFT again – I would avoid that.*

> **Response:** We thank the reviewer for this stylistic observation. The repeated discussion of wavelet advantages over FFT has been removed from the manuscript (lines 281–284 and 286–289).

#19. *Line 285: what is the "true value" of a flux? I thought the eddy covariance is what direct follows the theory – right?*

> **Response:** The reviewer is correct that eddy covariance directly follows from Reynolds decomposition theory, and the term 'true flux' is imprecise and potentially misleading. What we mean is the flux that would be captured by an instrument with infinitely fast frequency response, i.e., one that resolves all eddy scales contributing to the total covariance. Following Li et al. (2023), who use this terminology in the context of wavelet flux uncertainty, we have replaced 'true flux' with 'spectrally complete flux' throughout the manuscript to avoid ambiguity. The Fm/F ratio then represents the fraction of the spectrally complete flux captured by the band-limited CPC measurements.

Revised manuscript text: "…finding that biases can range from 50–100% of the spectrally complete flux."

#20. *Line 287/288: I cannot quite follow your argument regarding the high-frequency part of the spectrum. Of course, there are limits to the resolution and representation of the high-frequency components in the spectrum, but these are mostly technical limitations and not a fundamental problem with FFT. There are airborne atmospheric turbulence measurements with high-resolution hot-wire anemometers down to the cm range, and the spectra look absolutely clean until they drop off steeply due to dissipation.*

Revised manuscript text: "Figure 1 also shows some differences between FFT and CWT flux calculations, especially for fluxes at high frequencies. These differences are attributable primarily to non-stationarity in the aircraft data, as the platform moves through different air masses, meteorological conditions, and altitudes. FFT assumes stationarity over the entire analysis window and can therefore introduce artifacts under such conditions, whereas CWT provides time-localized frequency information that is more robust for non-stationary signals (Schaller et al., 2017). Li et al., (2023) evaluated uncertainties of turbulent flux calculation using both methods, finding that biases can range from 50–100% of the spectrally complete flux. We note that the high-frequency deviations visible in the particle concentration spectra are not a consequence of FFT limitations but reflect the band-limited response of the CPC, which attenuates concentration fluctuations at frequencies above ~0.3 Hz, as discussed in Section 2.8. The CWT's more conservative high-frequency response may better represent the actual resolvable flux contributions (Misztal et al., 2014)."

21. *The figure caption of Fig 2 could be expanded a little bit to better understand what is displayed with the different lines.*

> **Response:** We thank the reviewer for this suggestion. In accordance with ACP journal guidelines, which require that legends appear in the figure itself rather than as verbal descriptions in captions, we have expanded the caption to better describe the content of the figure while keeping line style descriptions within the figure legend. The interpretive discussion of ogive convergence and FFT/CWT agreement is addressed in the main text.

Revised manuscript text: "Figure 2. Normalized ogive function as a function of distance covered by the aircraft for the flight leg shown in Figure 1. The ogive represents the cumulative fractional contribution to total flux from high to low frequencies."

#22. *I think at the end of Sec 2.8, it would be good to mention approximately how large the loss is in determining the flow when you can no longer resolve the integral length scales yourself. This does not anticipate any results from the next chapter, but would simply illustrate how important this correction is.*

> **Response:** We thank the reviewer for this constructive suggestion. We have added a brief summary at the end of Section 2.8 illustrating the magnitude of the flux loss correction across the range of atmospheric conditions and measurement heights encountered during the campaign, without anticipating the specific flux results presented in Section 3.

Revised manuscript text: "To illustrate the practical importance of this correction, the $\frac{F_m}{F}$ ratio varies substantially depending on measurement height and atmospheric stability conditions encountered during the campaign. For measurements conducted near the top of the marine boundary layer (~1,200–1,400 m) under near-neutral to unstable conditions, $\frac{F_m}{F}$ values approach unity (0.93–0.99), indicating that flux losses are modest at these heights where the dominant flux-carrying eddies are large relative to the CPC response limitation. However, for measurements conducted closer to the surface (~30–550 m), $\frac{F_m}{F}$ values range from 0.70 to 0.95, implying that uncorrected fluxes could underestimate the spectrally complete flux by up to 30%. This highlights the importance of applying the flux loss correction, particularly for low-altitude flight legs where eddy sizes are smaller and the CPC response time constant becomes a more significant fraction of the dominant flux-carrying eddy turnover times."

#23. *Starting the results section with a table whose parameters are only described very superficially is not convincing. You begin the chapter by startng that two case studies will be described in more detail, and then start with a table showing six flights—somehow, this doesn't quite fit the picture. Here, too, the table caption is very brief; perhaps I overlooked it, but what does the last column mean?*

Revised manuscript text: "We examine two flight days as case studies of SPEs observed at varying altitudes above the ocean. Additional supporting flights are presented in the Supplementary Information. Table 1 summarizes the N_{3-10} vertical turbulent flux estimates derived from all six flight days analyzed in this study, grouped by the inferred nucleation regime. Flights 1 and 2 (January 29 and February 10, 2018) are classified as entrainment zone nucleation events, where SPEs were detected near the top of the MBL at heights exceeding 1,200 m. Flights 3–6 (June 21 and July 7, 2017; February 18 and 12, 2018) are classified as decoupled layer nucleation events, with SPEs observed across a broader range of altitudes (30–837 m). For all events, the ratio of measured flux to the spectrally-complete flux $\left(\frac{F_m}{F}\right)$ exceeds 0.76, indicating minimal flux loss due to sensor response limitations. The normalized vertical velocity variance ($\sigma_w^2\ w_*^{-2}$) is generally low, consistent with relatively quiescent turbulent conditions during the measurement periods. Negative flux values indicate downward transport of freshly nucleated particles from the entrainment zone toward the surface, while positive values suggest upward transport from a source within the decoupled sub-cloud layer. Two of these flight days, January 29, 2018 (Case 1) and June 21, 2017 (Case 2), are examined in detail as case studies in the following sections, with the remaining four flights presented as supporting examples in the Supplementary Information."

#24 *About Fig 3:*
My first impression of this figure is, why do I need all this detailed information? Somehow, there is no convincing motivation for this figure. For example, why is panel d with the chemical components necessary if the topic is particle fluxes? Panel b describes the flight path and the sampling strategy—okay, that's important, but why don't you show a map with the flight path marked on it? No one is going to visualize the flight path based on the coordinates—so I'm not learning anything from panel b at first glance. If you exclude all cloud passages from the analysis, why are they not already marked in panel a)? It is also somewhat confusing that one first looks at the sections with "significant number concentrations" in panel a (shown in blue) and then looks at panel c and realizes that these sections were in clouds and are therefore excluded from further

analysis. I suggest simplifying this figure significantly and focusing on the things that are evaluated and analyzed—that would make the figure much more convincing. In its current form, it looks like a "quick look plot" in which readers can search for interesting passages instead of the author highlighting the interesting passages and events. The same holds true for Fig. 6.

> **Response:** We thank the reviewer for these detailed and constructive suggestions. The revised Figure 3 addresses several of the reviewer's concerns. First, the chemical composition panel has been removed entirely, as it is not directly relevant to the flux analysis. It has been replaced with supermicron particle concentration and total aerosol surface area, which are directly relevant to establishing the low condensation sink conditions favorable for NPF. Second, drizzle concentration from the 2DS probe has been added to panel (b) to explicitly show that selected SPE periods were free of drizzle. Third, cloud-contaminated periods are now implicitly identifiable through the LWC data in panel (d) and the quality control description in Section 2.3.1.
>
> Regarding the reviewer's suggestion to replace panel (b) with a map showing the flight path: we considered this option but found that the east-west back-and-forth flight pattern, combined with the need to simultaneously convey latitude, longitude, altitude, flight time, and drizzle concentration, was difficult to represent clearly in a single map panel without information loss. The coordinate time series in panel (b) effectively communicates that the aircraft sampled the same horizontal region repeatedly, which is important context for interpreting the flux measurements. We believe the revised figure is now more focused on the quantities directly relevant to the analysis, while retaining the essential sampling context.

Revised manuscript figure:

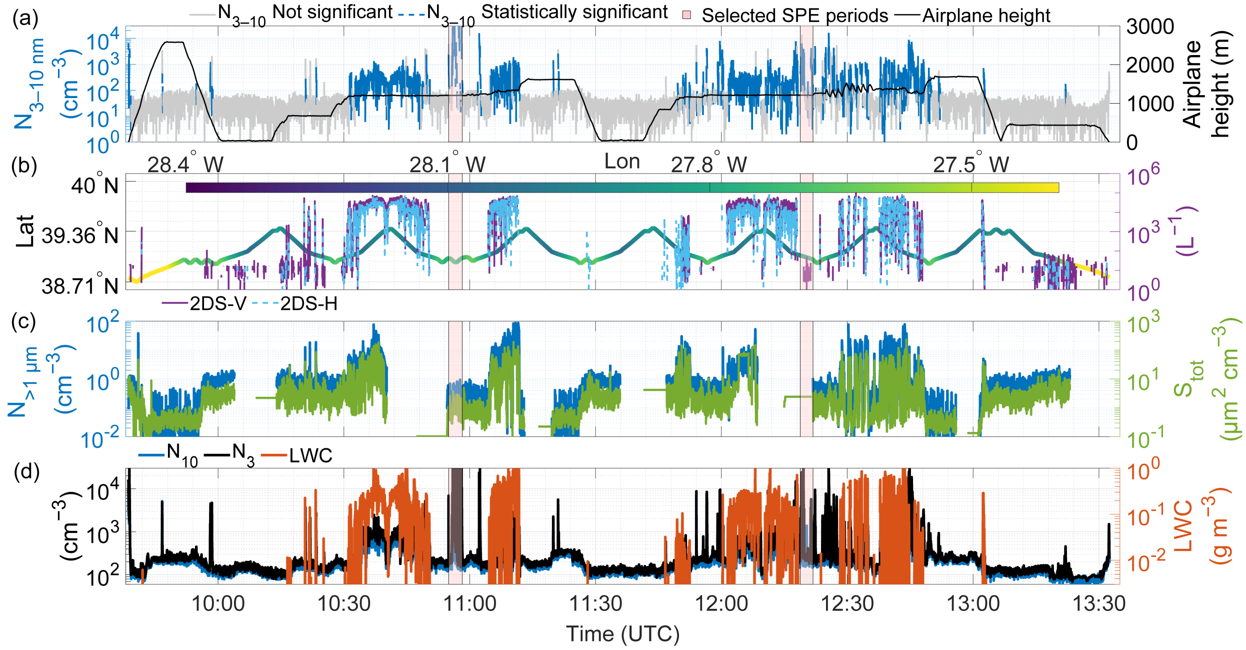

Figure 3. Multi-parameter time series from the January 29, 2018 flight. (a) N_{3-10} particle concentrations and aircraft altitude; (b) aircraft position (latitude and longitude) and drizzle number concentration; (c) supermicron particle concentration and total particle surface area (S_{tot})

(d) particle number concentrations (N_{10} and N_3) and liquid water content. Gaps in the time series indicate the missing data.

#25. *About Fig 4:*
You are mixing a time series with the three excerpts from vertical profiles here; based on the times for the profiles, it is possible to get a rough idea of how everything is connected, but a little help would be useful. I would also consider whether it would make sense to separate the vertical profiles and time series into two figures. The color scale in the upper figure is somehow unconvincing with regard to the tick labels and very close to the legend of the vertical axis—at first, I thought the color scale belonged to the axis label "Airplane height." but that's just a technical note. In the caption the "Stot" should be "S_{tot}" – right?

> **Response:** We thank the reviewer for these detailed comments. We have addressed them as follows:
>
> First, the figure caption has been corrected as suggested.
>
> Second, regarding the suggestion to separate the time series and vertical profiles into two figures: we respectfully maintain that keeping them in a single figure is scientifically preferable, as the vertical profiles are selected specifically to represent the thermodynamic and turbulence structure of the boundary layer during and around the SPE periods shown in panel (a). Separating them would require the reader to cross-reference between two figures to make this connection.

Revised manuscript text: "**Figure 4.** (a) Size-resolved particle number concentrations (10–600 nm) from FIMS as a function of time and altitude, with N_{3-10} concentrations shown in the lower strip. Pink shading indicates selected SPE periods. (b-d) Vertical profiles of potential temperature (θ), normalized vertical velocity variance ($\sigma_w^2\, w_*^{-2}$), total particle surface area (S_{tot}), and water vapor mixing ratio (MR_{H2O}) for three time periods nearest to the selected SPE periods: (b) 09:51–10:01 UTC, (c) 11:25–11:31 UTC, and (d) 13:01–13:07 UTC. Gaps in the time series indicate the missing data."

#26. *Line 373: Regarding the comment on the high particle concentration at 12:18Z: I can't imagine that burst droplets produce particles in the 10 nm size range – but maybe I'm wrong. Could it be possible that you flew into your own exhaust plume? Interestingly, the size distribution directly at the edge of the cloud returns to the more typical distributions before the cloud-free part begins.*

> **Response:** We thank the reviewer for raising this possibility. However, aircraft exhaust contamination can be ruled out in this case, as the nearest prior flight transect over the same region occurred approximately half hour earlier and at an altitude ~60 m lower, more than sufficient temporal and vertical separation for complete plume dispersal. The concurrent LWC and drizzle concentration data strongly support droplet shattering as the most likely explanation, and the data point was excluded from analysis regardless of origin.

Revised manuscript text: "The high particle concentration spike observed ~ 12:18 UTC coincides with elevated LWC and drizzle concentration (Fig. 3d), and was excluded from analysis following our quality control procedures (Section 2.3.1). Weber et al. (1998) documented that cloud droplet shattering can produce artifactual particle concentrations as small as 3 nm in airborne CPC measurements, making this the most plausible explanation for the observed spike. Aircraft exhaust contamination can be ruled out, as the nearest prior flight transect over this region occurred approximately half an hour earlier at an altitude ~60 m lower, more than sufficient time and vertical separation for complete plume dispersal. This data point was therefore excluded from all flux calculations as the concurrent LWC exceeded the quality control threshold of 3×10^{-3} g m^{-3}."

#27. *line 378 I have difficulty with the statement that a strong temperature gradient is an indicator of an entrainment layer. If there is no process that promotes vertical mixing, there is no entrainment even with a strong gradient. A strong gradient initially prevents vertical mixing, and only either shear flow or downward buoyancy due to cloud top cooling for example can cause mixing. But I agree with you, of course, that a strong temperature gradient combined with increased variance in vertical velocity is an indication of entrainment—it's just worded a little awkwardly at the beginning.*

> **Response:** We thank the reviewer for this important clarification. The original wording incorrectly implied that a strong temperature gradient alone is sufficient evidence of an entrainment zone. The revised text now makes clear that it is the co-occurrence of the potential temperature inversion with elevated vertical velocity variance that indicates active entrainment, and briefly notes the thermodynamic mechanisms (cloud-top radiative cooling and wind shear) that drive the mixing process.

Revised manuscript text: "Sharp gradients in the potential temperature (orange) mark the top of the MBL, defining the capping inversion that suppresses vertical mixing between the boundary layer and the free troposphere. The co-occurrence of these temperature gradients with elevated normalized vertical velocity variance ($\sigma_w^2\ w_*^{-2}$) near the inversion level indicates the presence of an entrainment zone, where thermodynamic forcing (including cloud-top radiative cooling and wind shear) drives mixing between the free tropospheric air above and the convective boundary layer below (Boers and Eloranta, 1986)."

#28. *Could you please explain in more detail what exactly you mean by your statement regarding the moisture gradient (lines 386 to 388)? I don't really understand this explanation.*

Revised manuscript text: "The water vapor mixing ratio profiles in Figs. 4b–d reveal evolving boundary layer moisture structure during the flight. The early profile (Fig. 4b) shows a relatively well-mixed moisture distribution below the capping inversion at ~1,200 m, with a sharp decrease into the drier free troposphere above. The later profiles (Figs. 4c–d) exhibit a two-step moisture structure, with a sharper gradient near ~600 m suggesting progressive decoupling of the boundary layer during the course of the flight, separating a moister surface layer from a drier sub-cloud layer above."

#29. *Regarding Figure 5: In principle, I find plots of this kind very illustrative, but unfortunately the flight section with the color-coded particle concentration is so small that it is impossible to relate it to the cloud situation. In the explanations accompanying this figure, the presentation of the observations, explanations, and then interpretations (also in comparison to other studies) are completely mixed up, which does not exactly improve readability. I suggest clearly separating explanations and interpretation.*

> **Response:** We thank the reviewer for these constructive suggestions. The revised Figure 5 addresses the concerns as follows. First, the flight track with color-coded $N_{3–10}$ particle concentrations has been enlarged and repositioned to be more clearly visible relative to the satellite cloud imagery, making it easier to relate the SPE observations to the cloud field context. Second, the flux measurement locations are now explicitly marked with vertical red lines, with the calculated flux values labeled directly on the figure. Third, regarding the mixing of observations, explanations, and interpretations in the accompanying text: we have restructured the relevant paragraph to first describe what is shown in the figure (observations), then explain the physical context (explanations), and finally interpret the results in terms of nucleation source location and comparison with prior studies (interpretations). We believe these revisions significantly improve both the visual clarity of the figure and the readability of the accompanying discussion.

Revised manuscript text: "Figure 5 presents the spatial distribution of N_{3-10} particle concentrations along the flight path at ~1,200 m altitude (dashed lines in Figs. 4b–d), with the calculated vertical turbulent fluxes labeled at their respective measurement locations. Concentrations up to 10,000 cm^{-3} were observed along the flight track, with the highest values concentrated within a horizontal extent of less than 10 km. The substantial downward fluxes of $N_{3–10}$ particles (−41,092 and −2,975 cm^{-2} s^{-1}) at ~1,200 m both exceed their respective limits of detection (Table 1), confirming that the observed downward transport represents a statistically significant atmospheric signal rather than measurement noise. The downward flux direction indicates that the source of freshly nucleated particles was located above the measurement altitude, within the entrainment zone, while the large difference in flux magnitudes between the two events likely reflects spatial heterogeneity in source strength and the proximity of the aircraft to the nucleation zone during each transect. This interpretation is supported by the near-absence of $N_{3–10}$ at ~1,600 m during 11:14–11:25 and 12:51–13:01 UTC (Fig. 3), with the exception of brief concentration spikes of uncertain origin retained in the record due to insufficient evidence for their removal. The small particle size (3–10 nm) and limited horizontal extent of less than 10 km further argue against a free tropospheric nucleation source, as particles originating in the free troposphere would be expected to have grown substantially and the plume to have diluted during descent to measurement altitude.

Several mechanisms could promote nucleation specifically within the entrainment zone: adiabatic cooling in the rising convective plumes, turbulent fluctuation in temperature and vapor concentration generated by entrainment, and dilution of mixed-layer air by the entrained free tropospheric air, causing a sudden reduction in preexisting aerosol surface area (Nilsson et al., 2001). The extremely low S_{tot} values observed in the entrainment zone and free troposphere (Fig.

4), falling well below the campaign averages, are consistent with this interpretation. These conditions are analogous to those identified in previous studies linking entrainment zone nucleation to reduced condensation sink environments (Größ et al., 2018; Meskhidze et al., 2019; Nilsson et al., 2001). Supplementary Figs. S8 and S9 provide additional support, showing a downward flux of N_{3-10} particles ($-1{,}195$ cm^{-2} s^{-1}) at 1,375 m with complete absence of N_{3-10} above ~1,400 meters, consistent with SPE occurrence specifically within the entrainment zone between 1,375–1,400 m.

Figures 3–5 and the flux analysis (Table 1) demonstrate that the entrainment zone nucleation near the MBL top occurred on two days (January 29 and February 10, 2018), representing nearly 5% of flight days. Despite a relatively small horizontal extent (<10 km), these newly formed particles can be entrained in the boundary layer via vertical turbulent processes, potentially playing an important role in marine aerosol number budget and, given sufficient time for growth to CCN-relevant sizes, potentially influencing cloud condensation nuclei concentrations for marine stratocumulus clouds."

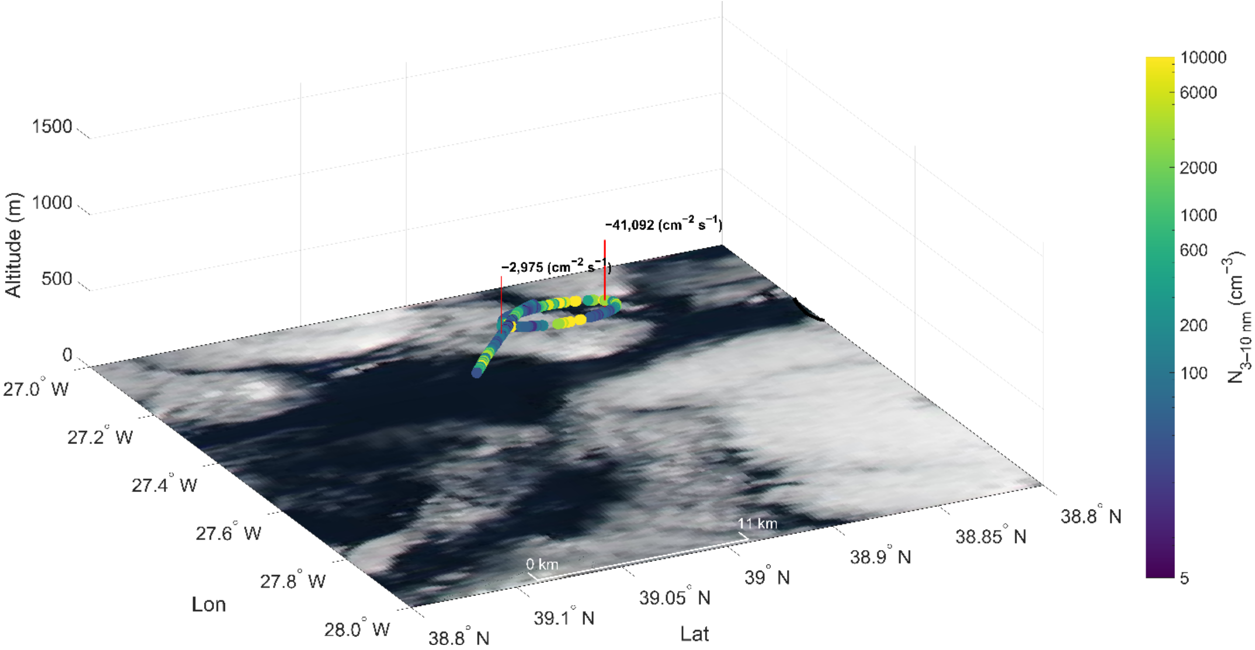

Figure 5. Spatial distribution of N_{3-10} particle concentrations along the flight track at ~1,200 m altitude during the period highlighted in Figures 3 and 4. Calculated vertical turbulent fluxes are labeled at their respective measurement locations. Color scale indicates N_{3-10} particle number concentrations (cm^{-3}). The background shows a true-color satellite-corrected reflectance image from the overpass at 15:15 UTC, with the ocean surface appearing dark and clouds appearing white. Credit: NASA Worldview Snapshots.

#30. *Another important aspect that becomes particularly clear in Fig. 5: Technically speaking, the manuscript deals with the most accurate determination possible of ultrafine particle fluxes. The fluxes are determined for the two SPE periods selected in Figs. 3 & 4, but the amount of the fluxes is not used at all in the further discussion. To put it somewhat provocatively, wouldn't the sign of the particle flux be sufficient for the following discussion? Is the large difference in the calculated* *particle flux for the two events discussed at any point? You are still in the chapter entitled "Results" and are discussing very general possible scenarios in which particle formation can occur. Are there any indications of "cold air outbreaks" over the Azores? Probably not, so in my opinion this discussion belongs in the introduction or discussion section, if anywhere, but not in the results section*

> **Response:** We thank the reviewer for these important and thought-provoking comments. We address each point in turn.
>
> Regarding the suggestion that the sign of the flux alone might be sufficient: while the flux sign is indeed the primary diagnostic for identifying the vertical location of the particle source, the flux magnitude serves an essential and non-negotiable role in the analysis - it must exceed the limit of detection (LoD) for the flux to be considered statistically meaningful and retained for interpretation. As described in Section 2.5, fluxes with magnitudes below the LoD are indistinguishable from measurement noise and are excluded from analysis regardless of their sign. Without quantifying the flux magnitude and comparing it to the LoD, it would be impossible to determine whether any observed flux direction reflects a real atmospheric signal or simply random covariance between the vertical wind and particle concentration signals. Furthermore, the large difference in flux magnitudes between the two SPE periods on January 29, 2018 (−41,092 vs. −2,975 cm^{-2} s^{-1}) likely reflects spatial heterogeneity in the strength of the nucleation source region and the proximity of the aircraft to the nucleation zone, information that flux sign alone cannot provide.
>
> Regarding the general discussion of nucleation scenarios appearing in the Results section: we agree with the reviewer that mechanistic discussion of possible nucleation scenarios is better placed in the Discussion section. The relevant text has been moved accordingly.
>
> Regarding cold air outbreaks over the Azores: we thank the reviewer for raising this point. Zheng et al. (2021) documented the occurrence of cold air outbreaks in the Azores region, providing observational support for the convective roll vortex mechanism we discuss. This reference has been added to the revised manuscript.

Revised manuscript text: "The large difference in flux magnitudes between the two entrainment zone events (−41,092 vs. −2,975 cm^{-2} s^{-1}) provides information that flux sign alone cannot supply: it reflects spatial heterogeneity in source strength and the proximity of the aircraft to the nucleation zone during each transect. These flux magnitudes, integrated over the duration of the events, represent a substantial source of freshly nucleated particles to the marine aerosol number budget, constraints that can be used directly to evaluate nucleation parameterizations in regional and global models. While the ultimate contribution of these particles to CCN populations depends on growth timescales and loss processes during vertical transport (requiring ~24–48 hours at typical marine growth rates of ~1 nm hr^{-1} to reach CCN-relevant sizes; (Zheng et al., 2021)), the flux-based constraints provided here represent a necessary observational foundation for quantifying this contribution in future studies."

"The limited horizontal extent (2–9 km) of these events is consistent with the spatial scales of organized convective structures that develop in the upper decoupled marine boundary layer following cold front passages in the Azores region, where cumulus-associated drizzle reduces the condensation sink to levels favorable for nucleation (Etling and Brown, 1993; Zheng et al., 2021)."

#31. *line 426: The discussion about a possible contribution of nucleation mode particles to CCN is also highly speculative and goes far beyond the observed results, so it does not belong here*

Revised manuscript text: "The revised text now reads: "Despite a relatively small horizontal extent (<10 km), these newly formed particles can be entrained in the boundary layer via vertical turbulent processes, potentially playing an important role in marine aerosol number budget and, given sufficient time for growth to CCN-relevant sizes, potentially influencing cloud condensation nuclei concentrations for marine stratocumulus clouds."

#32. *In the "Discussion" section, you first state that determining particle fluxes is important for a better understanding of particle formation. However, in the sections from lines 490 to 499 and 500 to 511, you describe in much greater detail the conditions prevailing in the two different regions where you measured the NPF. Although the particle flux itself is mentioned, the sign alone would have sufficed for the conclusion you draw from it. As a reader, I wonder why you undertook this significant technical effort to determine the flux, including a very thorough error analysis—which I greatly appreciate—if the magnitude of the flux plays no role whatsoever in your discussion.*

> **Response:** The reviewer is correct that the Discussion section as originally written relied primarily on flux sign to identify nucleation source locations, without fully exploiting the flux magnitude information. We have revised the Discussion section to explicitly address what the flux magnitudes tell us beyond the flux sign.

Revised manuscript text: "Critically, while flux sign alone identifies the vertical location of the particle source, flux magnitude serves two additional essential roles: first, it must exceed the limit of detection to confirm that the observed directional transport represents a statistically significant atmospheric signal rather than measurement noise; and second, it provides quantitative constraints on source strength that cannot be obtained from sign alone."

"Strong downward fluxes (up to −41,092 cm^{-2} s^{-1}) exceeding the limit of detection confirm that nucleation occurs specifically within this ~200 m entrainment layer, while the absence of 3–10 nm particles above the entrainment zone rules out a free tropospheric source. The limited horizontal extent (2–9 km) of these events is consistent with the spatial scales of organized convective structures that develop in the upper decoupled marine boundary layer following cold front passages in the Azores region, where cumulus-associated drizzle reduces the condensation sink to levels favorable for nucleation (Zheng et al., 2021; Etling and Brown, 1993).
The large difference in flux magnitudes between the two entrainment zone events (−41,092 vs. −2,975 cm^{-2} s^{-1}) provides information that flux sign alone cannot supply: it reflects spatial heterogeneity in source strength and the proximity of the aircraft to the nucleation zone during each transect. These flux magnitudes, integrated over the duration of the events, represent a substantial source of freshly nucleated particles to the marine aerosol number budget, constraints that can be used directly to evaluate nucleation parameterizations in regional and global models.

While the ultimate contribution of these particles to CCN populations depends on growth timescales and loss processes during vertical transport (requiring ~24–48 hours at typical marine growth rates of ~1 nm hr^{-1} to reach CCN-relevant sizes; (Zheng et al., 2021)), the flux-based constraints provided here represent a necessary observational foundation for quantifying this contribution in future studies."

"The factor of ~3 difference in flux magnitude between the 550 m and 30 m levels (−2,782 vs. −860 cm^{-2} s^{-1}) is consistent with attenuation of the particle flux during downward transport through dilution with ambient air, as well as losses through coagulation and growth out of the 3–10 nm size range. This vertical divergence in flux magnitude represents a quantitative signature of particle evolution during transport that flux sign alone would be incapable of revealing."

#33. *The interpretation of the event's horizontal extent of a few kilometers in terms of "convective rolls," on the other hand, is highly speculative and rather superfluous in this context—why are you venturing onto such thin ice with this?*

> **Response:** We thank the reviewer for this comment and agree that the interpretation of the limited horizontal extent in terms of convective roll vortices was speculative and not directly supported by our observations. These lines have been removed from the revised manuscript. The limited horizontal extent (<10 km) of the entrainment zone events is now described as being consistent with the spatial scales of organized convective structures that develop in postfrontal conditions in the Azores region (Zheng et al., 2021), without invoking the specific convective roll vortex mechanism that we cannot independently verify from our measurements.

Revised manuscript text: "The limited horizontal extent (2–9 km) of these events is consistent with the spatial scales of organized convective structures that develop in the upper decoupled marine boundary layer following cold front passages in the Azores region, where cumulus-associated drizzle reduces the condensation sink to levels favorable for nucleation (Etling and Brown, 1993; Zheng et al., 2021)."

#34. *In the following "Conclusions" chapter, it sounds a bit as if your measurements have turned conventional wisdom regarding nucleation in marine environments on its head—that goes a bit too far: there are very detailed and well-founded observations of small, newly formed particles even over the ocean.*

Take, for example, the work of Collin O'Dowd, who regularly observed nucleation at low tide along the Irish coast (Mace Head, hTps://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2000JD000206).

Likewise, the pioneering work of Dave Covert (hTps://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/92JD02074) and many others.

And then there was even a study that dealt with particle fluxes over the same area and almost at the same time as your measurements (hTps://acp.copernicus.org/articles/22/10007/2022/), which you completely ignored.

Response: We thank the reviewer for these important references and for pointing out that our conclusions section overstated the novelty of our findings in a way that misrepresented the existing literature. We agree that observations of newly formed particles in marine environments, including over open ocean, have a long and well-documented history. We have revised the conclusions section to position our findings more accurately within this existing body of work, acknowledging the foundational contributions of O'Dowd et al. (2002), Covert et al. (1992), Wiedensohler et al. (1996), and Siebert et al. (2021), among others. We also acknowledge and now cite the particle flux study over the Azores region by Lückerath et al. (2022, ACP), which used helicopter-borne measurements during a period overlapping with ACE-ENA. While that study focused on fluxes of particles above 8.5 nm across the full MBL depth rather than specifically on 3–10 nm particles in the entrainment zone and decoupled layer, it provides complementary observational context for our findings.

Revised manuscript text: "The occurrence of newly formed particles in marine environments has been documented in a number of previous studies. Wiedensohler et al., (1996)observed sub-20 nm particles originating from the free troposphere or cloud tops within the MBL and mixed downwards over the open ocean, suggesting in-situ production based on correlations with absolute humidity. Covert et al., (1992) reported sub-20 nm particle production near precipitating cloud tops within the MBL, where larger particles acting as condensation sinks had been scavenged by precipitation, with sub-20 nm particles dominating 10% of the campaign and indicating episodic rather than continuous production. O'Dowd et al. (2002) documented NPF events at the coastal Mace Head station when marine air masses encountered biogenic emissions from the intertidal zone. The ACORES campaign, conducted over the Azores at the same time as ACE-ENA, reported freshly nucleated particle bursts near the cloud top exceeding background MBL concentrations by more than an order of magnitude (Siebert et al., 2021). Concurrent helicopter-borne particle flux measurements over the Azores during ACORES (Lückerath et al., 2022) also documented particle fluxes in the marine boundary layer, providing complementary observational context to our aircraft-based approach. The prevailing theoretical framework, based on relatively high sea spray aerosol surface area acting as condensation and coagulation sinks (Bates et al., 1998; Pirjola et al., 2000), nonetheless predicted that NPF should rarely occur in remote marine boundary layers over open oceans. Our flux-based observations build on this existing framework by providing the first direct constraints on the vertical source location and strength of freshly nucleated particles in the remote marine boundary layer."

#35. *How could you present a more convincing explanation of your measurements? Of course, the mass balance equation for aerosol particles cannot be solved based solely on the particle flux, but one could at least roughly estimate the contribution of small particles to the total concentration; after all, your original motivation for this topic was to highlight the importance of small particles for CCN development. If you want to keep up this motivation, not address this point and conclude the paper with it?*

Response: We thank the reviewer for this constructive suggestion. We agree that a rough order-of-magnitude estimate of the contribution of the observed particle fluxes to the total aerosol number concentration would significantly strengthen the conclusions and better connect the flux measurements to the broader motivation of the study. We have added such an estimate to the revised conclusions section. Using the measured downward flux magnitudes, the observed horizontal extents of the SPE events, and typical boundary layer mixing timescales, we estimate the contribution of freshly nucleated particles from both modes to the surface mixed layer aerosol number concentration. This provides a quantitative bridge between our flux measurements and their potential significance for the marine aerosol budget.

Revised manuscript text: "For the entrainment zone mode, while the aircraft sampled the SPE for only ~4 minutes during each transect due to its high ground speed, NPF events in marine and continental environments are typically observed to persist for 2–5 hours (Kulmala et al., 2004; Zheng et al., 2021; Islam et al., 2022). Assuming the measured downward flux of −41,092 cm^{-2} s^{-1} is representative of a nucleation event of typical duration of ~3 hours and using a mixed layer depth of ~1,200 m, the estimated increase in vertically integrated particle number concentration is approximately:

$$\Delta N_{3-10} \approx |F| \times \frac{\Delta t}{z_{MBL}} = \frac{41{,}092 \ \times 10{,}800}{1.2 \times 10^5} \approx 3{,}700 \ cm^{-3}$$

where F is the flux in cm^{-2} s^{-1}, Δt is the event duration in seconds, and z_{MBL} is the mixed layer depth in cm. This represents a substantial addition to the total particle number concentration in the surface mixed layer, noting that this estimate refers to freshly nucleated 3–10 nm particles rather than CCN-relevant particles. The fraction surviving to CCN-relevant sizes (>50–80 nm) depends on growth rates and loss processes that cannot be quantified from single aircraft transects alone. However, Zheng et al. (2021) estimated that under favorable conditions at the same site, newly formed particles contributed on average ~50% of total CCN concentrations following cold front passages, suggesting that even accounting for coagulation losses, the contribution of entrainment zone nucleation to the marine CCN budget may be substantial. Entrainment zone nucleation, despite its limited horizontal extent, may contribute significantly to the marine aerosol number budget through sustained downward transport via convective mixing. Though flux magnitudes for the decoupled layer nucleation events (−2,782 cm^{-2} s^{-1}) are less pronounced, their large spatial extent likely results in comparable or larger aggregate contributions to regional aerosol budgets. We note that the contribution of freshly nucleated particles to CCN population depends on growth rates and loss processes during transport and cannot be assessed from short-term measurements alone. The flux magnitudes and vertical source locations reported here provide the observational foundation needed to evaluate this contribution quantitatively in future studies combining particle flux measurements with growth rate and CCN closure analyses."