We thank the reviewers for their thoughtful and insightful comments. We have addressed their concerns to the best of our abilities. The comments from the reviewers are in black text, while our responses and adjusted text are included in red text. Upon further reflection and discussion with the co-author team, we have made additional changes to the text. This document reflects our final response to the reviewers and editorial staff. Therefore, the line numbers included here reflect the correct line numbers after all revisions, but may deviate from the line numbers in the reviewer replies posted during the live discussion.

The major changes in response to the reviewers' comments include:

- Section 2.4. (Production Flux Measurements) was heavily edited to better clarify the methodological choices and their impacts (along with additions to the supplement).
- Section 3.7. (Temperature Hysteresis Effects) was added to better characterize this aspect of the temperature-SSA relationship.
- All regression metrics were updated with accompanying statistical significance testing to qualify the regression and account for the limited number of samples when analysis was done over temperature-binned data.

Anonymous Review #1

We thank the reviewer for their thoughtful and thorough comments. We have responded to each comment below. The major changes we have implemented include: adding a discussion of loss rates to the SI, adding section 3.7 *Temperature Hysteresis Effects*, providing additional statistical tests and metrics to qualify every R^2 value we report, and clarifying methodological choices and what is being presented. We believe we have addressed the reviewer's concerns to the best of our abilities while keeping the manuscript focused. Further, we thank the reviewer as this manuscript is significantly stronger after these suggested edits.

In this manuscript, the authors present a dataset collected using the Scripps Ocean–Atmosphere Research Simulator (SOARS), a combined wind–wave facility, to investigate the impact of sea surface temperature on sea spray aerosol production. This is a highly relevant scientific question. Studies on this topic date back to the 1980s, yet there remains significant disagreement in the literature regarding the nature of the SST dependence of SSA production. The topic is also of growing importance given that sea surface temperatures have increased due to global climate change and are expected to continue increasing in the future. Understanding the role of SST in determining sea spray aerosol production is therefore critical for assessing potential feedbacks within the climate system. The SOARS infrastructure used in this study is clearly state of the art, and the dataset presented here will be of broad interest to the marine aerosol community. That said, I feel there are a number of important issues that should be addressed before the manuscript can be recommended for publication. My major comments are given below, followed by a series of more minor points that I believe would further strengthen the manuscript.

Major comments

1) The title currently reads as a relatively universal conclusion, whereas the results themselves appear substantially more nuanced. In particular, the manuscript shows that supermicron responses are non-monotonic, number and mass exhibit different SST dependencies, and the inferred fluxes rely on size-dependent methodological assumptions. In addition, the experiments were conducted under a single wind speed, wave state, and seawater system. While I have no doubt that the advances presented in the manuscript are valuable and represent an important contribution towards resolving the SST–SSA question it is my view that the study is an incremental advance that builds upon an increasing body of literature suggesting that aerosol number and mass may respond differently to changing seawater temperature. I therefore suggest that the title be revised to better reflect the size-dependent and system-specific nature of the reported results.

We thank the reviewer for raising this concern. We agree that the results display a nuanced behavior, which is critical for interpretation. We also agree that the original title does not adequately capture the size-dependent complexity of the results — particularly the non-monotonic SST dependence of supermicron number and mass emissions, which is itself an important finding that deserves recognition rather than risk being lost in the title.

Accordingly, we have revised the title to:

"Sea Surface Warming Suppresses Primary Sea Spray Aerosol Number Production"

"Sea spray aerosol" is the more specific and appropriate term for what SOARS produces and is consistent with the terminology used throughout this and previous manuscripts, while "sea surface warming" more concisely conveys the directional temperature forcing central to this work. We believe this targeted revised title addresses the reviewer's concern while preserving the primary scientific contribution of the work to the modelling community.

We have also made the following edits to clarify our results:

Abstract, lines 22-23: “Importantly, we identify size-resolved differences including a non-monotonic SST dependence of supermicron number and mass emissions.”

Abstract, lines 24-25: “These controlled wind-wave-SST experiments demonstrate that increasing SST overall suppresses SSA production.”

Sect. 3.6., lines 499-501: “While they do not account for the size-dependent differences highlighted in this work, they represent a first step towards integrating SOARS measurements into ESMs to improve the simulation of SSA.”

Sect. 4., lines 592-595: “The size-resolved SST dependencies differed in strength and shape, suggesting the occurrence of a physicochemical or mechanistic shift. Surface area and volume

distributions (Fig. 5) showed pronounced shifts while supermicron SSA fluxes and most SSA mass fluxes show non-monotonic SST dependencies sensitive to the SST value."

Sect. 4., lines 600-602: "The general suppression of subsurface bubble concentrations, total SSA number concentrations, and total SSA number fluxes with increasing SST is common across experiments, independent of SST hysteresis."

Additionally, we acknowledge that our findings are not revolutionary for advancing the understanding of SST dependencies of SSA production. However, since we pay attention to creating the CFs that are applicable beyond the scope of our system, and supplying the probability density functions for comparison to other systems or in situ measurements, we believe altering the title to include the system-specific nature of the results would undercut what we are providing to the community.

2) The introduction currently frames the literature primarily in terms of conflicting positive, negative, or non-monotonic SST dependencies. However, in my view, this treatment substantially oversimplifies what is arguably one of the central challenges facing the SSA research community; that previous studies differ not only in the sign of their reported SST dependencies, but also in other important ways, such as the quantities measured (e.g., number, mass, AOD, sodium burden, flux), the particle size ranges considered, the experimental generation mechanisms employed and the degree to which aerosol losses and transport are constrained. To my mind these distinctions are critical because different SSA metrics and size regimes may respond differently to changing SST. Given this, I think that a more structured discussion of the literature, organised by methodology, size range, and measured quantity, would substantially strengthen the manuscript and provide better context for the contribution of the SOARS experiments.

The reviewer raises an excellent point that many attempts have been made to characterize the influence of sea surface temperature on sea spray aerosol production from different perspectives. An exhaustive review of the literature is beyond the scope of this work and would serve as an excellent stand-alone review article. We have included more detail in the supplemental information S1 (Previous Literature) and updated Table S1 to address some of the reviewer's concerns. The following changes were made to the introduction, and we copy Table S1 here:

Sect. 1, lines 48-52: "The review by Song et al. (2023) and the introductions of Markuszewski et al. (2024) and Salter et al. (2014) provide recent summaries of this literature, and with no significant advances since. A thorough comparison of previous studies (e.g., methodology, size range, and measured quantity) is beyond the scope of this paper; we therefore summarize their main findings in the supplement (Text S1, Table S1).

	Reference	Study Type	Analog or Ocean Basin	Measurement Type	SSA-SST Dependence

1	Mårtensson et al. (2003)	Laboratory	Sintered Glass Filter	Size-resolved SSA number and subsurface bubble concentrations	Number Concentrations: Non-monotonic
2	Zábori et al. (2012)	Laboratory	Plunging Jet	Size-resolved SSA number concentrations	Number Concentrations: Non-monotonic
3	Salter et al. (2014)	Laboratory	Plunging Jet	Size-resolved SSA number and subsurface bubble concentrations, air entrainment	Number Concentrations: Negative
4	Salter et al. (2015)	Laboratory	Plunging Jet	Size-resolved SSA number concentrations and flux	Number Concentrations: Negative (submicron), Positive (supermicron) Number Fluxes: Negative (submicron), Positive (supermicron)
5	Forestieri et al. (2018)	Laboratory	miniMART	Size-resolved SSA number concentrations	Number Concentrations: Positive
6	Christiansen et al. (2019)	Laboratory	Plunging Jet Diffuser	Size-resolved SSA number concentrations, air entrainment	Jet Number Concentrations: Non-monotonic Diffuser Number Concentrations: Negative
7	Nielsen & Bilde (2020)	Laboratory	Pipette Bubbler	SSA number concentrations	Number Concentrations: Negative
8	Sofieva et al. (2022)	Laboratory	Diffuser	Size-resolved SSA number concentrations	Number Concentrations: Non-monotonic
9	Zinke et al. (2022)	Laboratory	Plunging jet	SSA number concentrations	Number Concentrations: Non-monotonic
10	Sellegri et al. (2023)	Laboratory	Plunging Jet	SSA number flux, Size resolved SSA number concentrations	Number Fluxes: Negative Number Concentrations: Non-monotonic
11	Hu et al. (2024)	Laboratory	Plunging Jet Plunging Waterfall	Size-resolved SSA number concentrations, surface bubble concentration	Jet Number Concentrations: Non-monotonic Waterfall Number Concentrations: Positive
12	Jaeglé et al. (2011)	Field	Global	Ship- & Station-based mass concentrations; MODIS & AERONET AOD; GEOS-Chem simulations	Mass Concentrations: Positive
13	Grythe et al. (2014)	Field	Global	Ship- & Station-based mass concentrations; FLEXPART simulations	Mass Concentrations: Positive
14	Lehahn et al. (2014)	Field	North Atlantic	Ship-based number concentrations	Number Concentrations: Negative
15	Ovadnevaite et al. (2014)	Field	North Atlantic	Ship-based number fluxes; station-based number concentrations; model-based wave parameters	Number Fluxes: Positive
16	Saliba et al. (2019)	Field	North Atlantic	Ship-based number concentrations	Number Concentrations: Negative Mass Concentrations: Positive
17	Liu & Liu et al. (2021)	Field	Pacific & Atlantic	Airborne volume concentrations	Volume Concentrations: Positive
18	Markuszewski et al. (2024)	Field	North Atlantic, Baltic Sea	Ship-based number and volume fluxes	Number Fluxes: Negative (Baltic Sea all chlorophyll statistically significant) Volume Fluxes: Positive (Baltic Sea low chlorophyll statistically significant; Atlantic negative)
19	**This study**	**Laboratory**	**Wind-Wave Interactions (SOARS)**	**Number concentrations; number fluxes**	**Number concentrations: Negative (submicron), non-monotonic (supermicron)** **Number fluxes: Negative (submicron), non-monotonic (supermicron)** **Mass fluxes: Non-monotonic**

Copy of Table S1

3) A further issue concerns the role of the theoretical framework introduced in Section 3.1 and the supplementary matieral. The manuscript presents a coupled bubble–SSA framework (Eqs. S2–S8) describing the evolution of subsurface bubbles, surface bubbles, and aerosol concentration over time, and states that the governing concentration changes "can be determined experimentally." However, so far as I can see, the manuscript does not appear to actually solve, constrain, or apply these equations quantitatively in the subsequent analysis. Instead, the results rely primarily on empirical regressions between SST and bulk aerosol metrics derived from binned observations. As currently presented, the framework therefore functions primarily as a conceptual interpretation tool rather than a mechanistic analysis framework. While this is not necessarily problematic and maybe completely intentional, I think that the manuscript should more clearly distinguish between the conceptual theoretical framework and the empirical regression analysis that ultimately underpins the reported SST dependencies and correction factors. Alternatively, if the authors intend the framework to play a mechanistic role in interpreting the data, then a more explicit application of the governing equations to the experimental observations would substantially strengthen the manuscript.

We have revised the introduction of the framework to better identify its purpose (a conceptual interpretation tool). We have also added references back to the framework throughout the text to tie it in with our measurements better.

Sect. 3.1, lines 247-250: "This framework functions as a conceptual interpretation tool, where the equations will not be solved empirically as more measurements or assumptions are necessary for full solutions. This framework also serves to highlight areas for future research where the role of SST in controlling bubble-bursting dynamics is understudied and difficult to constrain across bubble sizes and small-scale processes."

Sect. 3.1., lines 260-262: "Liu and Yang (2022) include a useful schematic relating changes in SST to potential changes in whitecaps, which is related to the ñ term in Eq. 8. Our framework extends that schematic to include the entire OAI system through a simple mathematical lens."

Sect. 3.2., lines 308-311: "The WCFs displayed here relate to the surface bubble concentrations (ñ) in Eqs. 8 and S2, although the ñ term is explicitly the concentration of surface bubbles, which is a function of bubble radius. Whitecaps consist of layers of bubbles near the surface contributing to the high albedo, the measured quantity, whereas the ñ term focuses on estimates of the top-most layer of bubbles directly contributing to SSA production."

4) An important methodological issue concerns the derivation of submicron number fluxes from only the initial 5 minutes of the "Wind & Waves" period. This approach differs substantially from many laboratory sea spray generation studies, where fluxes are commonly inferred from steady-state aerosol concentrations and known "sweep" or carrier gas flows. In the present manuscript, the submicron flux estimate is instead based on the early-time slope of the concentration increase, effectively assuming that aerosol losses are negligible during this initial

period. While this may well be a reasonable approximation, it is certainly not equivalent to a steady-state source-flux measurement. The inferred submicron PNFDs are therefore transient accumulation metrics rather than directly measured steady-state source strengths. Given this, the authors should more clearly explain how this method differs from traditional "flow-through" sea spray chamber approaches, justify the choice of the 5-minute fitting window, and discuss how sensitive the inferred fluxes are to this choice. It would also be useful to show whether the inferred SST dependence changes if alternative fitting windows are used. In addition and more genrally, the authors should discuss the limitations of estimating fluxes under non-steady-state conditions, particularly given that the strongest SST dependence is observed in the submicron regime where this transient method is applied. This distinction is important because the resulting correction factors may otherwise be interpreted as steady-state emission flux constraints for models, whereas they are more directly based on early-time aerosol accumulation rates in SOARS.

And further to this, while the manuscript explains that different flux-estimation methods are applied to different particle sizes, it does not adequately address the inconsistencies this may introduce when interpreting the results. Submicron fluxes, where the strongest SST suppression is reported, are inferred from the initial 5-minute buildup rate under the assumption of minimal losses. In contrast, larger particles are treated with approaches that more explicitly account for the finite residence times of particles, steady-state behaviour, and decay-derived particle loss rates. The authors should discuss how these methodological differences may affect comparisons across size ranges and the interpretation of the resulting SST-dependent correction factors.

We appreciate the reviewer's questions regarding our emission flux techniques and agree it's a critical discussion point that shouldn't be oversimplified. We address the reviewer's concerns in two parts: comments here explaining our rationale in more detail and adjustments to the manuscript to clarify our rationale for all readers.

Our approach differs from the traditional "sweep" or carrier gas flows with steady state concentrations of previous studies because SOARS is not a simple system with clearly defined sweep flows. It is true that SOARS is over pressurized to ensure that leaks are unidirectional out of the system, however, the system-wide losses have not been fully characterized. Further, these losses will change from the decay period (without waves, where we estimate loss rates) and the wind-wave period, where the paddle motion alters airflows within the channel (where we estimate flux).

We concur that the submicron initial rise estimate is not equivalent to a steady-state source-flux measurement. To address this concern, previous work (currently under review for publication in *Nature Climate Change*) evaluates emission flux measurements from SOARS as a function of wind speed and wave breaking intensity. During those experiments, high-temporal resolution flux measurements were made wherein 3 electrostatic classifiers were each set to size-select a

single diameter (0.075, 0.150, and 0.450 μm). The concentrations at these sizes were then recorded at rates of either 1 Hz or 0.1 Hz, enabling our approach to be applied at much smaller time frames when losses are expected to be even smaller than in the first 5 minutes. It was found that estimates within the first 12 seconds (approximately the timescale of a particle completing one pass through the entire SOARS system), 30 seconds, and 300 seconds align well with our fully size-resolved 5-minute approach. Our approach represents an estimate of the lower bound of the expected emission rate based on the higher resolution measurements and considering that losses may be greater than 0, as is assumed.

Additionally, we switch flux estimation approaches for the supermicron sizes due to a failed assumption of zero losses within the first 5 minutes for these large particles. These approaches enable us to use the most information available, however, there are still assumptions about the loss rates in SOARS. Relying on unquantified losses in SOARS then limits the comparability of these estimates to previous supermicron flux estimates. Further, a separate manuscript currently under development seeks to characterize the effect of different flux estimation methods using SOARS data. We firmly believe this is an important discussion, but it generally lies outside of the scope of the current manuscript due to the nuances.

We include the confidential copy of the previous work currently under review for the reviewer to further address their concern. The most pertinent sections are Table S2, Figure S6, and paragraphs 3 & 4 of the first subsection of the Online Methods ("SOARS Measurements").

Additionally, we have revised the text of this manuscript to better clarify our approaches, our rationale, and their limitations. In line with this comment, and comments from reviewer #2, we have included some results in the supplement comparing the methods for "typical" SOARS data, the coldest SST, and the warmest SST:

Sect. 2.4., lines 175-220: "Estimates of particle number flux distributions (PNFDs) utilized three methods to account for size-dependent influences on SSA production. Certain assumptions are made to simplify the uncharacterized loss mechanisms in SOARS (Leibensperger III et al., 2026), with all methods assuming a well-mixed reactor and first-order loss rates. Production flux is estimated using Eq. 4 where N is the number concentration at a given aerosol size and time, P is the production flux (# cm^{-3} min^{-1}), and k is the first-order loss rate (min^{-1}).

First, submicron PNFDs were estimated by applying a linear regression to the SSA number concentration time series during the initial 5 minutes of Wind & Waves, referred to as the "initial rise" method. This method assumes a negligible loss rate (k = 0) based on the sufficiently long residence times of submicron SSA and sufficiently low background aerosol concentrations, maintaining a large signal-to-noise ratio. The efficacy of this method is evaluated in a separate manuscript, but it showed that reducing the 5-minute window to 12 seconds (the travel time of a particle throughout SOARS) yielded changes to submicron emission fluxes within typical daily variability, indicating that the simplified assumptions do not significantly bias the PNFDs.

Second, PNFDs for supermicron SSA between 1 and 4 μm were estimated for each size bin by applying a power law regression fit to the entire time series of SSA number concentrations during Wind & Waves, referred to as the "power-fit" method. This power law fitting allows P and k to each be freely fit according to each size bin. This method relies on the entire Wind & Waves period to estimate production and loss rates but requires a reduction of the number of size bins to maintain high counting statistics during production. Supermicron SSA reach steady-state concentrations quickly, limiting the validity of this approach for larger particles. Estimates from the power-fit method align well with those from the initial rise method but tend to be lower by about 5 – 10 %.

Third, for SSA above 4 μm, steady state was reached within the first minute of sampling due to short residence times. The PNFDs of these particles were estimated assuming that the production flux is equivalent to the product of the steady-state concentration and the loss rate estimated during the Decay period, referred to as the "steady state" method. This is to say the left-hand side of Eq. 4 becomes 0 and N(t) becomes $N_{Steady\ State}$, assumed invariable with time. When loss rates could not be directly quantified due to poor counting statistics, a constant loss rate from the last measured size bin was assumed (k is constant). Size- and SST-dependent loss rates exhibit an asymptotic behaviour above 2 μm (approaching a value of 0.55 min^{-1}) and differences within error between experiments (Fig. S12).

The steady state-derived supermicron fluxes are not inherently equal to the initial rise-derived submicron fluxes, but each method is the best approach for the specified size range given the complexities of the SOARS instrument. The loss rate assumptions are likely underestimates, generally leading to underestimates by each method compared to the true interfacial emission flux. However, when the initial rise and steady state methods are compared to one another, the chosen method in each size range represents an overestimate compared to the other method. That is, the initial rise-derived fluxes are larger than the steady state-derived fluxes for submicron SSA (by up to a factor of about 6) and the steady state-derived fluxes are larger than the initial rise-derived fluxes for supermicron SSA (by up to a factor of about 5) (Fig. S5). This variability is of similar magnitude to daily variability of SSA number concentrations in SOARS (Leibensperger III et al., 2026), reinforcing the validity of these estimates. Critically, this variability is systematic across SST values, such that relative SST-dependent trends are unaffected. A thorough comparison of the flux approximation methods is the focus of a forthcoming manuscript. In the remainder of the text, we combine the three size-differentiated approaches into a single PNFD covering the entire size range (0.01-20 μm) following Eqs. 5-7 where P is converted to units of # m^{-2} s^{-1} using the headspace volume and water surface area of SOARS:"

Text S4, lines 124-135: "The chosen methods for flux estimation in SOARS can introduce additional uncertainties. Figure S5 compares the submicron method (applying a linear regression to the initial rise of particle concentrations within the first five minutes) and the supermicron

method (assuming production flux is equal to the product of steady-state concentrations and the loss rate measured during the Decay period). The black lines indicate typical variability across three days from a separate SOARS experiment using the same wind speed, wave height, and filtering procedures as the current data. In panels (a, b), the solid lines indicate the chosen method, the dashed line indicates the non-chosen method, and the dotted lines indicate the chosen method corrected for by the typical ratio of the two methods. Data is shown for the coldest and warmest SSTs (2 and 23 °C) as bounds of the flux estimate variability. The solid lines are typically above the dotted lines, indicating that the chosen methods lead to a higher estimate of flux. These estimates (panel c) can be as high as a factor of about 20 (± 15) for typical SOARS conditions but remained within a factor of 6 for the SST-dependent results. While this uncertainty should not be overlooked, typical uncertainties in flux can be several orders of magnitude, indicating that our estimates are reasonable given the constraints and complexities of the system."

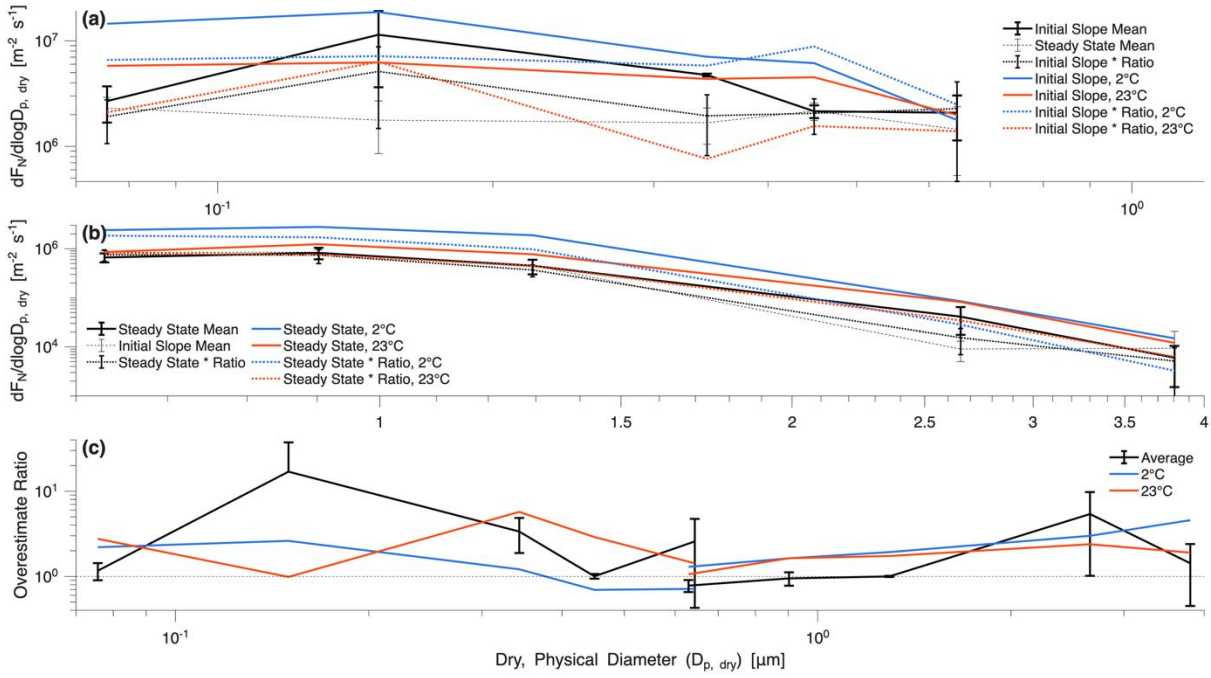


Figure S5

5) Following from this point, another major issue concerns the decay periods. The authors state that decay periods were included and use decay-derived first-order loss rates for parts of the flux calculation, yet these decay data and fitted loss rates are not, as far as I can see, presented or analysed in sufficient detail in the current manuscript. Given that aerosol residence times and loss rates are central to the interpretation of concentration-to-flux conversion in SOARS, these data should be shown explicitly. At minimum, the manuscript should include representative

decay curves, size-resolved loss rates, their SST dependence, and a discussion of how uncertainties in these loss rates propagate into the inferred PNFDs and PMFDs.

We thank the reviewer for this important discussion point. Along with the previous comment, the size-resolved loss rates are included in a separate manuscript currently under review. We agree with the reviewer that including more information about the loss rates is important for the present analysis; however, we believe including the decay curves won't significantly and meaningfully add to the paper. We have edited Figure 1b (copied below) to include a linear vertical axis, making the exponential decay during the decay mode more evident. We believe this is sufficiently representative of the exponential decay nature of the SSA during this operational mode. Additionally, we include Figure S12 showcasing the size-resolved and SST-dependent loss rates used in the flux estimations. Further, we have improved the discussion of the loss rates in Sect. 2.4., Sect. 3.7., and Text S10.

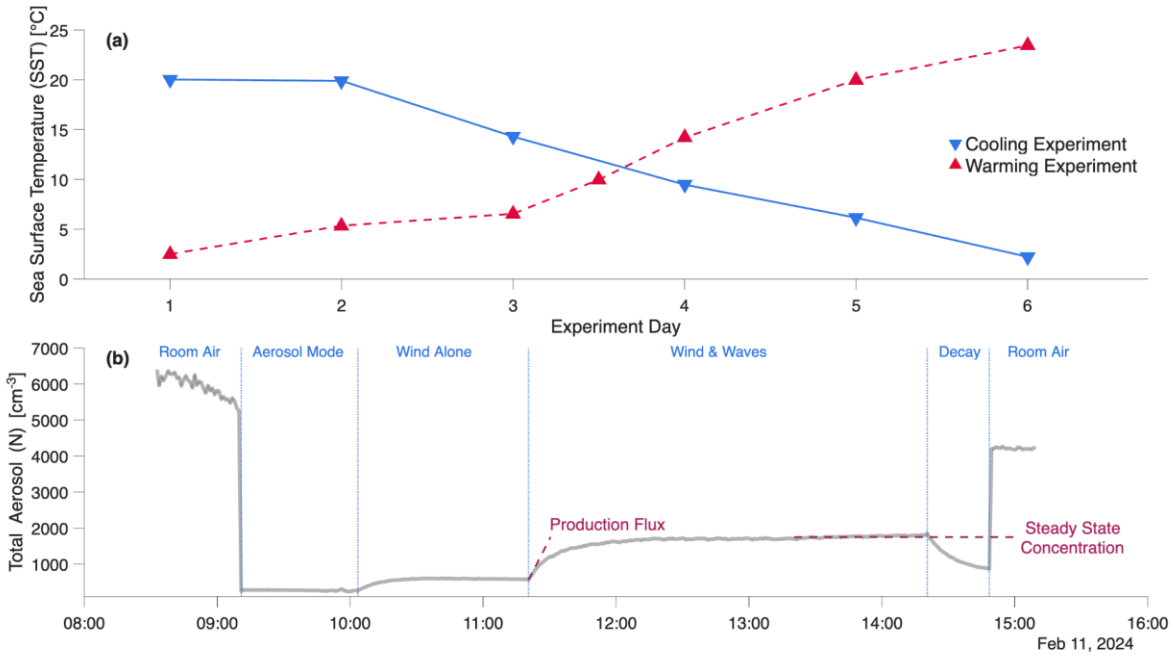


Figure 1

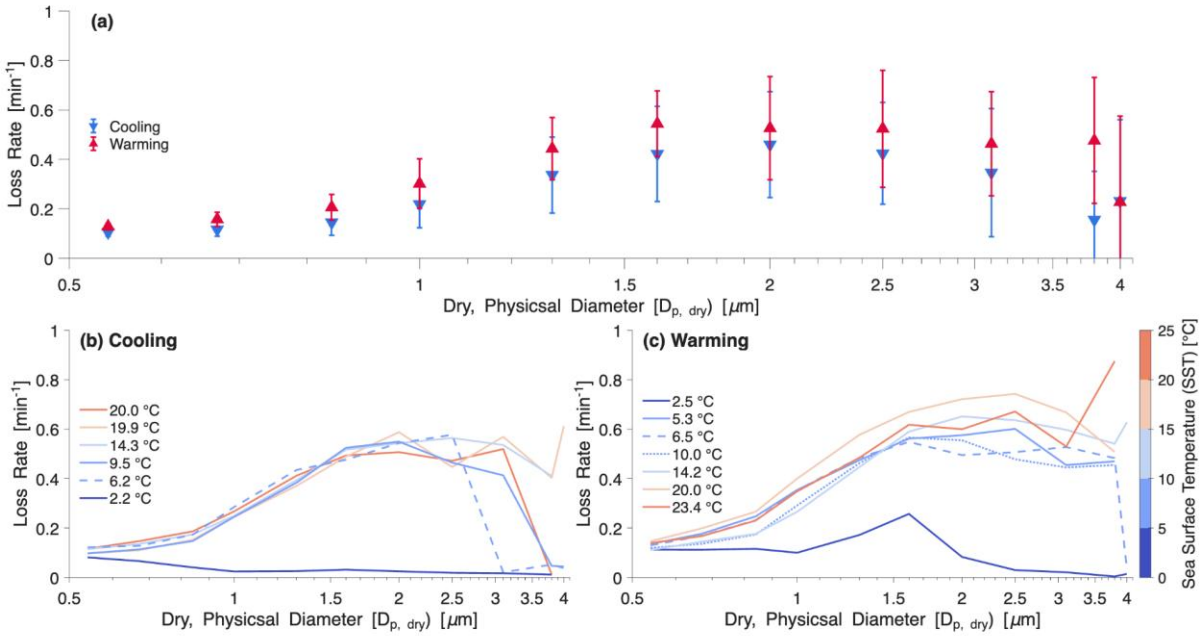


Figure S12

Sect. 2.4., lines 202-203: "Size- and SST-dependent loss rates exhibit an asymptotic behaviour above 2 µm (approaching a value of 0.55 min^{-1}) and differences within error between experiments (Fig. S12)."

Sect. 2.4., lines 206-209: "The loss rate assumptions are likely underestimates, generally leading to underestimates by each method compared to the true interfacial emission flux. However, when the initial rise and steady state methods are compared to one another, the chosen method in each size range represents an overestimate compared to the other method."

Sect. 3.7., lines 545-552: "The loss rates used in the PNFD calculations were similarly separated by each experiment (Fig. S12). To our knowledge, this is the first estimate of the effect of varied SSTs on aerosol loss rates, a process that is best investigated in laboratory studies. The mean loss rate between the experiments and across sizes was found to be 0.30 min^{-1}, with the warming and cooling experiments approaching asymptotes of 0.38 and 0.30 min^{-1}, respectively. The warming experiment exhibits larger loss rates on average (by approximately 20-30%), though the cooling and warming experiments remain within one standard deviation of each other across all sizes. The small variations in loss rates are three to four times smaller than the variations in F_N, and are therefore not expected to drive the larger-scale findings of the PNFD analysis and are not explored further here. Future studies should target SST-dependent loss rate variability to better understand the influence on flux and concentration estimates."

Text S10, lines 245-249: “The loss rates do not appear to be different between the cooling and warming experiments (Figure S12). However, during both experiments the aerosol loss rates measured during the coldest SSTs (SST < 5 °C) are consistently and dramatically less than for the other SSTs across all sizes. This indicates that the loss rates measured here are likely sensitive to subtle, hard-to-detect modifications in the SOARS system that affect turbulence. Future SOARS aerosol studies would benefit from additional loss rate characterization.”

6) While the authors present information on aerosol sampling and transmission efficiencies, several aspects of the sampling system remain insufficiently characterised. In particular, the manuscript does not clearly specify the sampling height within the SOARS headspace or provide the full tubing geometry and residence times. These issues are especially important because supermicron particles are highly sensitive to gravitational settling, inertial deposition, and transport effects, and because surface area and volume moments are disproportionately influenced by the largest particles. The authors should therefore provide a more complete characterisation of the aerosol sampling system and discuss how uncertainties in transmission efficiency may influence the reported SST dependencies, particularly for the supermicron regime and higher-order moments. In addition, Figure S3 appears to show sampling efficiencies exceeding 100% over part of the supermicron size range. This is difficult to interpret as a simple particle transmission efficiency and should therefore be explained more clearly. Again, this issue is particularly important because the supermicron number, surface area, volume, PNFD, and PMFD results are all highly sensitive to size-dependent transmission corrections.

The reviewer raises good points regarding the sensitivity of supermicron SSA to the sampling system. The height of the sample inlets shown in Figure S2 are positioned 0.6 m above the water surface. The particle loss calculator (PLC) described by von der Weiden et al., (2009) was applied to all particle number size distributions reported in this manuscript and follows applications in several publications (Franco et al., 2022; Moore et al., 2022; Schmale et al., 2013; Slater et al., 2014; Teri et al., 2022; Moore et al., 2025). The influence of sample height has been thoroughly investigated in a recently published manuscript (Leibensperger III et al., 2026). Leibensperger III et al. found minimal improvement on sampling efficiency at heights ranging from 0.6 to 1 m above the water surface as well as enhanced sampling efficiency (> 100%) for supermicron SSA, which was strongly dependent on in-channel wind speeds. Regardless of sample inlet positioning, Leibensperger III et al. found that supermicron SSA concentrations in SOARS are low (~1 cm^{-3} at 1 μm). This limitation is incorporated throughout updates to Section 3 Results and Discussion.

The following updates were made to the main text to clarify the use of the PLC for SSA measurements:

Section 2.2, line 116: “at a height of 0.6 m above the water surface”

Section 2.2, lines 121-126: “Loss mechanisms incorporated include diffusion, sedimentation, and inertial deposition in bends and contractions. Sampling efficiency ranged from ~55 to 100% over the size range of 8 to 200 nm. Slight enhancements around 200 nm to 4 μm were observed, consistent with previous SOARS studies (Leibensperger III et al., 2026; Moore et al., 2025). The wind speeds in SOARS produce sub-isokinetic conditions, leading to enhancement in this size range. Beyond 4 μm, sampling efficiency decreases significantly to 20% at 10 μm. These corrections were applied to all measurements prior to subsequent analysis.”

Sect. 3.3., lines 362-364: “Although supermicron concentrations are limited in SOARS (Leibensperger III et al., 2026), supermicron N estimates are dominated by SSA between 1-2 μm, where counting statistics and counting efficiency remain high, providing robust results.”

Sect. 3.3., lines 371-372: “The less pronounced changes in D_g and σ_g again suggest SST primarily influences the number of SSA at steady state rather than the shape of the distribution.”

7) Another issue concerns the effective degree of replication in the SST analysis. Although the experiments include both cooling and warming trajectories with multiple individual operating periods, much of the analysis is ultimately condensed into five 5 °C SST bins. As a result, several of the fitted SST relationships and correction factors appear to rely on a relatively small number of effective temperature intervals. While binning is understandable for reducing variability and combining the warming/cooling datasets, it also obscures variability between individual experimental runs and limits assessment of hysteresis, system drift, and day-to-day reproducibility. I think that this issue is particularly important because laboratory SSA systems commonly exhibit non-negligible run-to-run variability, even under nominally identical operating conditions. The manuscript would therefore benefit from a clearer distinction between the underlying individual experimental measurements and the subsequent binned analyses, including more explicit discussion of the reproducibility of the observed SST trends as well as how the binning procedure influenced the fitted SST dependencies and associated uncertainties. In essence, I wonder how sensitive the reported correction factors would be to repeating the experiments under nominally identical conditions on different days or with a different seawater preparation.

We agree with the reviewer that daily variability and reproducibility are major concerns for laboratory systems, and particularly so for SOARS given its size. As shown in an accompanying manuscript recently published in the Journal of Geophysical Research: Atmospheres (Leibensperger III et al., 2026), we have examined daily drift given nominally identical conditions and across different seawater fills. Further, this concern is addressed in the present analysis through figures S1, S10, S11, S13, and S14, where the data is not binned by temperature and is separated by experiment. While the coefficients of the regressions do vary between experiments, the trends do not. We address this in the new Sect. 3.7.

We have revised Figure S1d to include shading of the standard deviation of the mean distributions across the three days (a-c), copied below. As is clearly shown in this figure, the variability of the SSA (red line) is about 20% across three non-consecutive days compared to the other conditions, exhibiting variability of about 50% across the days. Additionally, Figure S6 shows the PNSDs separated by SST bin. While there is some variability within SST bins, it is relatively low, especially for the SMPS-APS measurements (top row), which are used for the analysis in the main text.

Further, we performed a Chow test for the correction factors. Since the CFs are equivalent to the normalized total number and mass fluxes, we essentially performed the Chow test on panels j, k, and l of figures S13 and S14. The p-values for both the linear and quadratic regressions of the number flux were greater than 0.78, indicating that the regressions were not significantly different between experiments. The p-value for the linear regression of the mass flux was 0.33, also indicating similarity between experiments. To the reviewer's concern, the p-value for the quadratic regression of the mass flux was 0.03, indicating there were statistically significant differences between the experiments. We have addressed this concern in a new section (Sect. 3.7. Temperature Hysteresis Effects) and added a new figure to the supplement (figure S15, copied below).

Finally, we have updated the text, see Sect. 2.3., lines 166-167 to include, "This amounts to 2 (0-5 °C), 5 (5-10 °C), 2 (10-15 °C), 2 (15-20 °C), and 2 (20-25 °C) experimental runs per temperature bin."

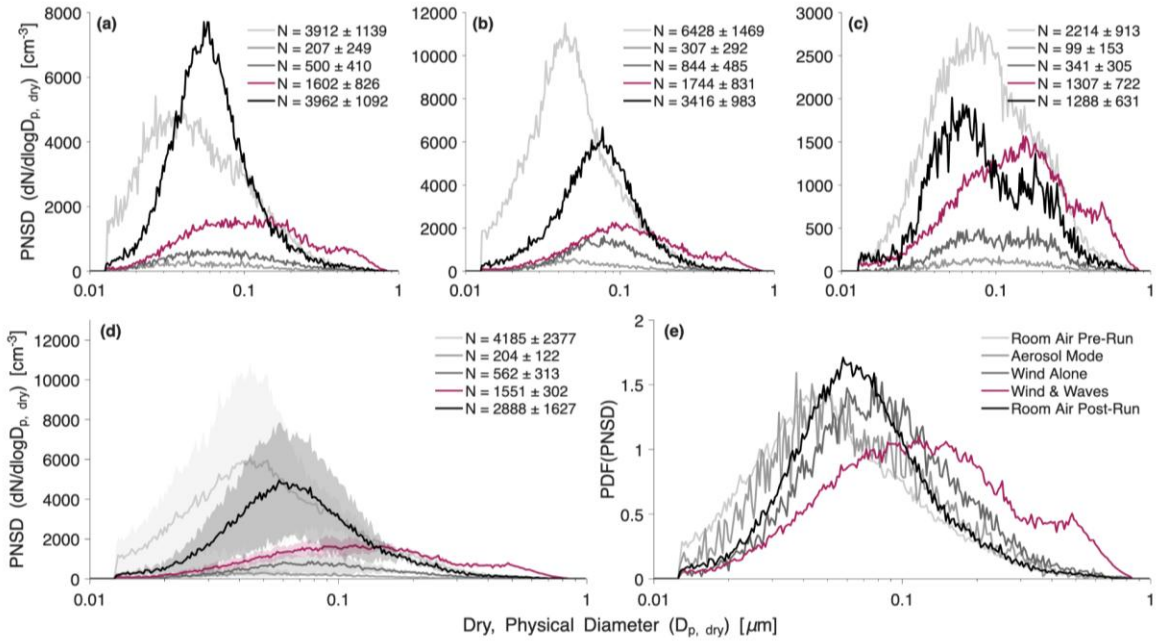


Figure S1

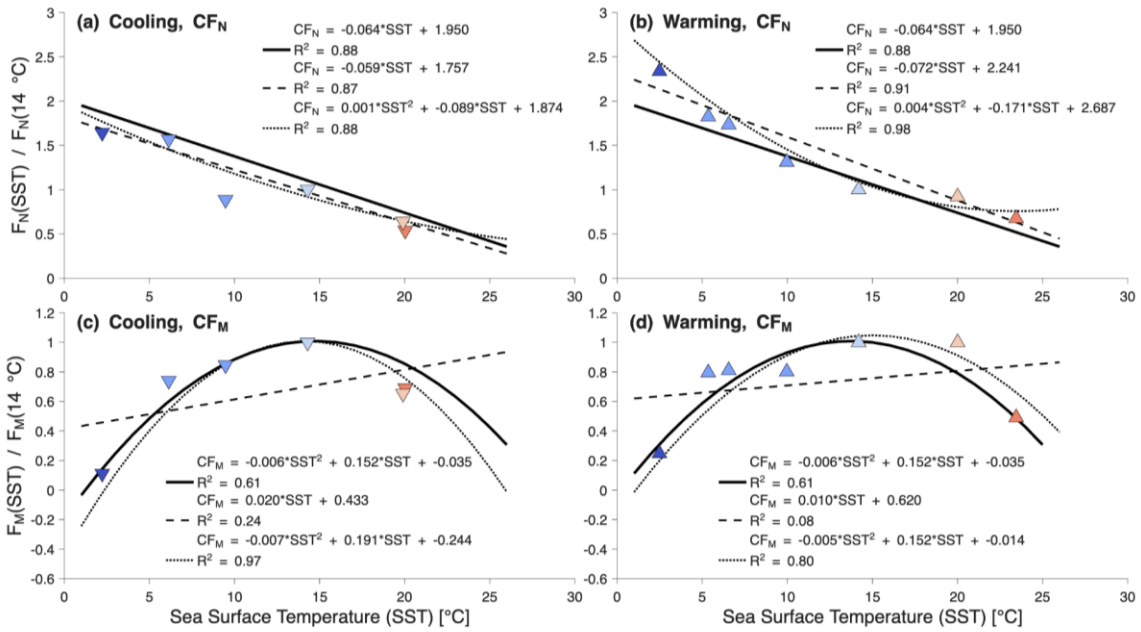


Figure S15

8) The interpretation of the supermicron aerosol response would benefit from substantially more discussion. Throughout the manuscript, supermicron number, surface area, volume, and mass exhibit different, and in some cases non-monotonic, SST dependencies. It is therefore not always clear whether these differences primarily reflect genuine changes in SSA production physics, or shifts in particle size distributions, or size-dependent aerosol losses and transmission efficiencies, or, perhaps most critically, methodological differences in the flux estimation and fitting procedures. The manuscript would therefore benefit from a more explicit discussion of how much confidence can be placed in the supermicron trends and to what extent the observed differences between number, surface area, volume, and mass likely reflect physical versus methodological effects.

Surface area and volume distributions were derived from particle number size distribution measurements, so temperature-dependent shifts are a result of physical changes rather than methodology. This method is limited by the typical assumption that all particles are spherical, and for the APS, particularly, that aerosol density is related to relative humidity (Section 2.3). We have made the following revisions to address this comment:

Sect. 3.3., lines 376-377: “Surface area and volume size distributions were derived from PNSDs, assuming spherical particles, to assess the influence of changing SST on higher-order moments of the size distribution.”

Concerning the sampling efficiency of supermicron SSA, the recently published work by Leibensperger III et al., (2026) demonstrated SOARS' current limitation for sampling large SSA. We have also pointed to this limitation in the text:

Sect. 2.4., lines 226-228: "The PMFDs (along with estimates of S and V) are sensitive to supermicron SSA, which are scarcer than submicron SSA. There are known challenges with sampling supermicron SSA in SOARS, so our estimates likely represent lower limits of these quantities (Leibensperger III et al., 2026)."

The probability density function (PDF) plots shown in Figure S8 and the sampling efficiency shown in Figure S3 indicate that the sampling system achieved efficiencies greater than 80% up to 2-3 μm. Therefore, we believe we have corrected for methodological effects to the best of our ability.

9) The discussion linking the abrupt 2–7 °C modal transition to marine gels and organic coatings currently feels somewhat speculative and would benefit from additional qualification. While this is certainly a plausible hypothesis, similar cold-SST accumulation-mode enhancement has been seen in previous studies using artificial seawater proxies (e.g. Zábori et al., 2012; Nielsen and Bilde, 2020; Salter et al., 2014, 2015). In addition, modal sharpening has also been observed in previous plunging-jet studies, including experiments conducted using artificial seawater systems with substantially reduced biological and organic complexity (e.g., Mårtensson et al., 2003; Salter et al., 2014, 2015). This suggests that at least part of the observed behaviour may arise from more fundamental changes in bubble-bursting physics, such as SST-dependent variations in viscosity, surface tension, bubble fragmentation, film drainage, or film/jet drop production efficiencies. The manuscript would therefore benefit from a more balanced discussion of alternative physical explanations and clearer acknowledgement that the present data do not yet uniquely support a marine-gel interpretation.

We have updated this paragraph to include a physical explanation of the observed results. The following text was updated:

Section 3.3, lines 404-412: "Nielsen and Bilde (2020) reported that removing gel-stabilizing divalent cations (Ca^{2+}, Mg^{2+}) from sea salt mixtures produced a similar suppression of particle concentrations as SSTs increased. Another explanation is a shift in physical production mechanisms modulated by SST. Callaghan et al. (2014) reported that temperature had the greatest effect on subsurface bubbles (n) larger than 1 mm, which are expected to produce submicron film drops (de Leeuw et al., 2011). The rapid decay of larger bubbles suggests that bubble rise times at low SSTs are accelerated. Furthermore, Nielsen and Bilde (2020) support the conclusion that the number of SSA produced per bubble is highest at low SSTs. However, no previous laboratory studies have incorporated wind, which significantly influences SSA concentrations (Leibensperger III et al., 2026). Consequently, these results are expected to be the

most representative of open ocean conditions and SOARS represents an opportunity to test these hypotheses in the future."

10) Some of the reported coefficients of determination appear inconsistent with the visual fit shown in the figures. For example, in Figure 3c the authors report R^2=1.00, yet several plotted points appear visibly offset from the fitted line. The authors should clarify exactly how the R^2 values were calculated, including whether they are based on binned means, weighted means, individual experimental points, or a different dataset from that shown in the figure. If weighted regressions are used, the authors should also clarify whether the plotted line and reported R^2 correspond to the same data shown in the panel. More generally, reporting R^2=1.00 for fits based on only five effective SST bins may give a misleading impression of precision, and the authors should consider reporting additional uncertainty metrics or showing fits to the individual experimental runs.

We thank the reviewer for this critique, and agree that additional metrics will strengthen the results, particularly when R^2 values are so high given the low number of points. We have run additional metric analyses and include p-values with every R^2 value to support our confidence in each metric. We have also clarified the origin of the regressions for each figure and included more explicit details in the methods.

Sect. 2.3., lines 171-172: "All SST-binned analysis in this study uses the uncertainty-weighting unless otherwise specified."

11) The conclusions state that the observed SST dependencies are consistent "regardless of temperature hysteresis, indicating that the magnitude of SST is more important than the direction of change." However, the supplementary figures suggest that the warming and cooling trajectories may differ more substantially than implied by this statement. For example, in Figure S9 the fitted relationships for several modes differ appreciably between the cooling and warming experiments, both in curve shape and goodness-of-fit. The supermicron responses in particular appear substantially different between the two trajectories. While the overall qualitative SST tendencies may be broadly similar, the manuscript does not appear to quantitatively demonstrate that hysteresis effects are negligible. The authors should therefore provide a more explicit statistical comparison between the warming and cooling trajectories, quantify the degree of hysteresis, and, if appropriate, soften the corresponding conclusions in the manuscript. For example, the current statement that "the magnitude of SST is more important than the direction of change" appears somewhat stronger than is fully supported by the presented data.

We thank the reviewer for raising this concern and have edited the text to ensure we do not mislead the readers. Importantly, we added Section 3.7 "Temperature Hysteresis Effects" to directly comment on the role of hysteresis. While SST hysteresis can lead to statistically significant differences in the CF_M coefficients, the quadratic form of the CF_M remains consistent regardless of SST hysteresis. Therefore, we affirm that the magnitude of the SST is more

important than the hysteresis generally. We have adjusted the text in the conclusions to better qualify our assertion and acknowledge the impact of hysteresis on CF_M.

Sect. 4, lines 600-606: "The general suppression of subsurface bubble concentrations, total SSA number concentrations, and total SSA number fluxes with increasing SST is common across experiments, independent of SST hysteresis. Supermicron F_N and F_M for SSA larger than 0.1 μm are suppressed as SSTs diverge from approximately 13 °C. SST hysteresis can lead to meaningful changes in CF_M coefficients, indicating an open area of future research. These results indicate that aerosol number exhibits a stronger dependence on the magnitude of SST than hysteresis, while aerosol mass (and the supermicron mass-containing aerosols) depend on the direction of SST change coupled with the magnitude of the SST."

Minor comments

Abstract, title, conclusions - In several places the manuscript refers to direct "suppression" of SSA production, whereas some of the measurements more directly constrain aerosol concentration trends or transient accumulation rates. Slightly more careful wording in parts of the manuscript would help align the conclusions with the quantities actually measured.

Section 2.4 and the results/discussion surrounding Figure 7 - The manuscript would benefit from more explicit terminology distinguishing concentration tendency, inferred flux, PNFD/PMFD, and steady-state source strength. Given that different size ranges use different inference methods, it is not always clear whether the reported quantities represent transient accumulation rates or steady-state aerosol production.

We have addressed this concern by adding statements to Sect. 2.1. and Sect. 2.4., to orient the reader to what we mean by fluxes throughout the manuscript:

Sect. 2.1., lines 108-111: "During Wind & Waves, a wave field is generated with a peak wave height of 0.42 m, enabling measurements of SSA concentrations and flux estimates (Sect. 2.3–2.4). The features of the aerosol concentration time series explored in the following sections are annotated by the magenta dashed lines in Fig. 1b."

Sect. 2.4., lines 214-217: "In the remainder of the text, we combine the three size-differentiated approaches into a single PNFD covering the entire size range (0.01-20 μm) following Eqs. 5-7 where P is converted to units of # m^{-2} s^{-1} using the headspace volume and water surface area of SOARS:

$$P_{D<1\mu m} = \frac{dN(D)}{dt} \quad (5)$$

$$P_{1<D<4\mu m} = k(D) \cdot N_{Steady\ State}(D) \quad (6)$$

$$P_{4\leq D<20\mu m} = k \cdot N_{Steady\ State}(D) \quad (7)$$

"

Section 2.4 and Figure 1 - While the authors provide a useful overview of the experimental sequence in Figure 1, because different particle size ranges use different flux-inference approaches (i.e. initial buildup, power-law fitting, and steady-state/decay-derived losses), I think that the manuscript would benefit from a more explicit schematic or workflow linking each size regime to its corresponding flux-inference methodology and fitting windows.

We thank the reviewer for this suggestion. In practice, this made the figure more difficult to read, so we have elected to maintain the current figure. However, we agree this visualization would strengthen the discussion of flux estimate approaches, and will add this to the manuscript comparing flux estimation techniques that is currently under development. To address the reviewer's concern, we have modified Figure 1b by making the vertical axis linear to better accentuate the exponential decay feature of the decay mode, and revised the text to better identify how we use the different flux methodologies.

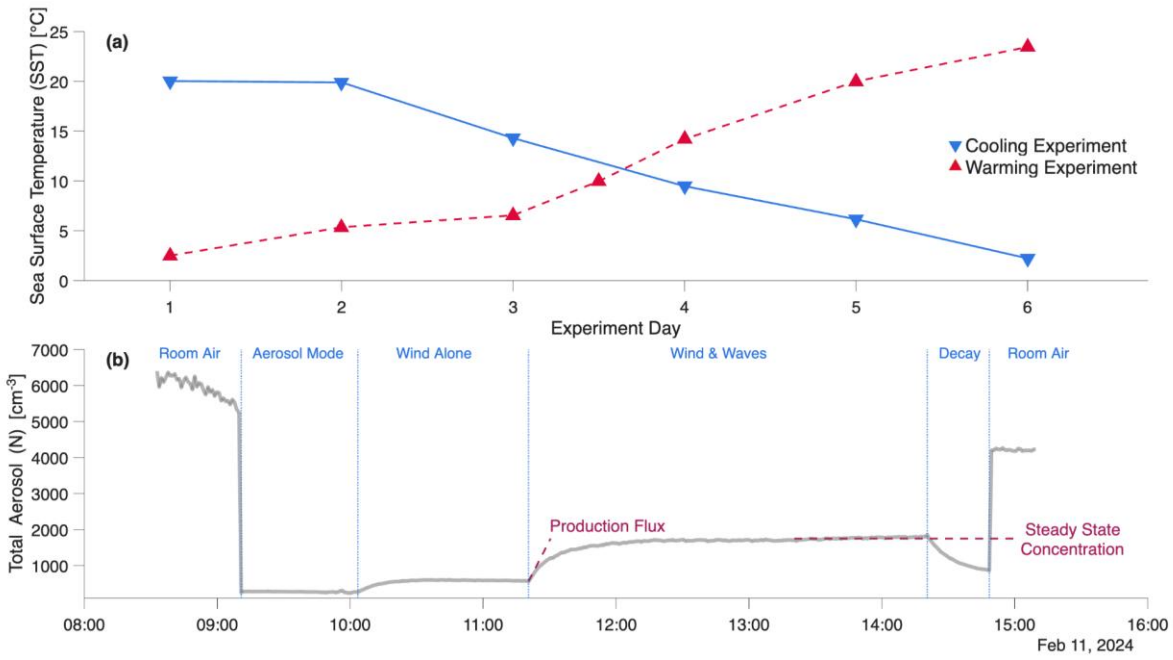


Figure 1

Sect. 2.4., lines 205-206: "The steady state-derived supermicron fluxes are not inherently equal to the initial rise-derived submicron fluxes, but each method is the best approach for the specified size range given the complexities of the SOARS instrument."

Figures 3, 5, 6, and 7 - Several of these figures show steady-state concentrations, while others show inferred fluxes. Because these quantities need not respond identically in SOARS, the manuscript would benefit from more explicit reminders in the figure captions and text regarding which interpretation applies to each figure.

We have updated the figure captions throughout to remind the reader and clarify what quantities are being presented in the figures.

Methods and figure captions throughout - It is not always immediately clear whether plotted points correspond to individual experimental runs, weighted averages, or 5 °C SST-bin means. The manuscript and figure captions would benefit from more explicit explanation of the binning procedure and how the cooling/warming datasets contribute to the combined fits.

In line with the previous comment, we have updated the figure captions throughout to remind the reader and clarify what quantities are being presented in the figures.

Section 3.4 and Figure 8 - I think that the manuscript would benefit from more explicit discussion of how uncertainties associated with SST binning, loss-rate estimation, fitting-window selection, and transmission efficiencies propagate into the final SST correction-factor parameterisations.

We have added more information about the statistics of the CFs and their intentional simplicity:

Sect. 3.6., lines 493-501: "These correction factors are chosen as simple, yet encompassing, representations of the SST dependencies of total F_N and F_M. They rely on all the F_N and F_M measurements made across the two experiments and are not binned by SST. The linear CF_N shows good correlation ($R^2 = 0.88$, $p << 0.01$) and is valid across the SST values ranging from freezing (approximately -1.8 °C) to 30 °C. On the other hand, the quadratic CF_M shows adequate correlation ($R^2 = 0.61$, $p < 0.01$) and is valid for SST values between 0.25 and 25 °C. These validity bounds are determined by when the CFs become negative; however, we advise only adapting them to SSTs between 2 and 25 °C, consistent with the conditions tested here. These CFs mirror the trends of the dominant modes, where F_N is dominated by the accumulation mode, and F_M is dominated by the supermicron mode. While they do not account for the size-dependent differences highlighted in this work, they represent a first step towards integrating SOARS measurements into ESMs to improve the simulation of SSA."

Additionally, following previous comments from the reviewer, we have added Sect. 3.7. "Temperature Hysteresis Effects" to explicitly comment on how temperature hysteresis and the experiment differences contribute to differences in the CFs.

Furthermore, we have added text and a supplemental figure to better compare the flux methods.

Sect. 2.4., lines 208-213: "However, when the initial rise and steady state methods are compared to one another, the chosen method in each size range represents an overestimate compared to the other method. That is, the initial rise-derived fluxes are larger than the steady state-derived fluxes for submicron SSA (by up to a factor of about 6) and the steady state-derived fluxes are larger than the initial rise-derived fluxes for supermicron SSA (by up to a factor of about 5) (Fig. S5).

This variability is of similar magnitude to daily variability of SSA number concentrations in SOARS (Leibensperger III et al., 2026), reinforcing the validity of these estimates."

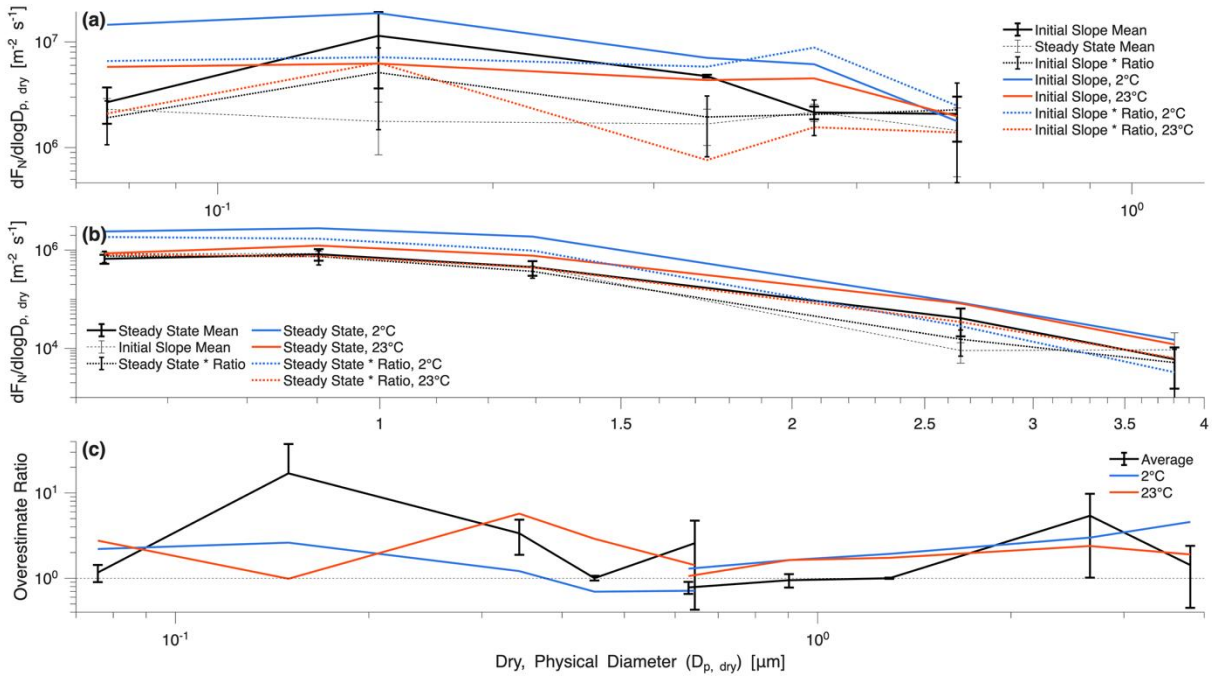


Figure S5

Anonymous Review #2

We thank the reviewer for their insightful comments. Many of which echo comments from reviewer 1, which have been addressed in the updated manuscript. In particular, we have added additional statistical analyses to better qualify the high coefficients of determination, have clarified the text to ensure consistency and not over-simplify the SSA-SST dependence, and improved the discussion of the CFs and their use cases. We believe the manuscript is significantly improved following the reviewer's comments.

This manuscript systematically investigates the impact of sea surface temperature (SST) on sea spray aerosol (SSA) production by a combined wind–wave facility (the Scripps Ocean–Atmosphere Research Simulator, SOARS). SSA is the most prevalent aerosol and is essential to the radiation budget and climate system. The authors characterized the connection between SST, seawater bubble concentrations, SSA number concentrations, and SSA emission fluxes using measurements in the SOARS. This topic is a classic topic in ocean-atmosphere interaction and it is becoming more and more important in the future because of the ongoing global warming. The dataset presented in this study will be of broad interest to the modeling community. Nevertheless, I have several concerns that need to be addressed before the manuscript can be recommended for publication.

Major comments:

1. Section 2.4: Different measurement periods were used for submicron and supermicron PNFDs. Submicron PNFDs were derived from the first 5 minutes of observations, while supermicron PNFDs used the full time series. Would the results change substantially if the full time series were also adopted for submicron PNFDs?

   We thank the reviewer for raising this question, which is in line with some concerns raised by reviewer #1. It's true that we applied different methods to different size ranges, which could introduce uncertainties to the flux estimates. We chose these different methods to use the most information available from each measurement accounting for the limited characterization of aerosol loss mechanisms in the SOARS instrument, see Leibensperger III et al. (2026) for an overview of some of these pertinent, unconstrained processes.

   Additionally, a comparison of applying different methods (the three included here and others) to a common size range and dataset is currently underway for inclusion in a forthcoming manuscript. However, the results indicate that applying the different methods to submicron PNFDs results in variations of about a factor of 6, which is well within the uncertainty of current source functions. Specifically, the chosen method for

each size range yields a higher estimate of flux than the other method, which we now describe in the methods section and supplement (see Text S4):

Sect. 2.4., lines 208-213: "However, when the initial rise and steady state methods are compared to one another, the chosen method in each size range represents an overestimate compared to the other method. That is, the initial rise-derived fluxes are larger than the steady state-derived fluxes for submicron SSA (by up to a factor of about 6) and the steady state-derived fluxes are larger than the initial rise-derived fluxes for supermicron SSA (by up to a factor of about 5) (Fig. S5). This variability is of similar magnitude to daily variability of SSA number concentrations in SOARS (Leibensperger III et al., 2026), reinforcing the validity of these estimates."

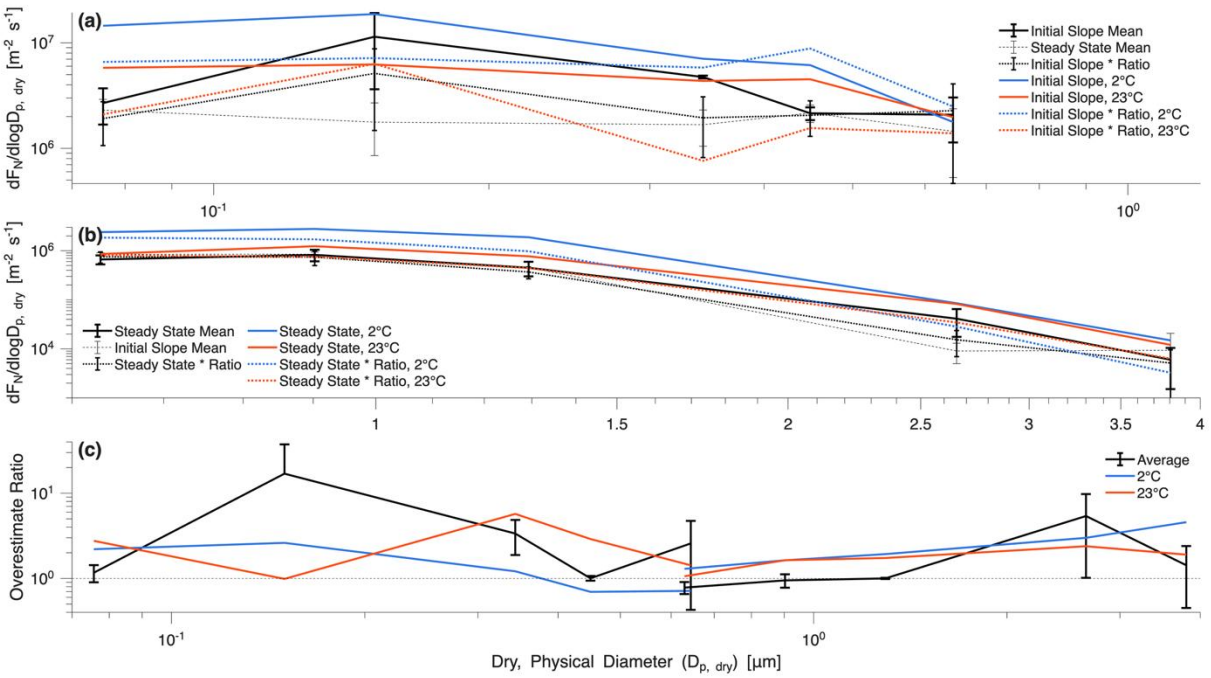


Figure S5.

To address this, we have updated the text in Sect. 2.4. Accordingly:

Sect. 2.4., lines 202-203: "Size- and SST-dependent loss rates exhibit an asymptotic behaviour above 2 µm (approaching a value of 0.55 min^{-1}) and differences within error between experiments (Fig. S12)."

Sect. 2.4., lines 205-208: "The steady state-derived supermicron fluxes are not inherently equal to the initial rise-derived submicron fluxes, but each method is the best approach for the specified size range given the complexities of the SOARS instrument. The loss rate assumptions are likely underestimates, generally leading to underestimates by each method compared to the true interfacial emission flux."

2. In many figures (e.g., figures 3, 5, 6, figures S4, S6, etc.), the coefficients of determination are high (close to 1.0). However, several data points clearly deviate from the fitted lines. Please explain this discrepancy.

   Thank you for pointing out this concern. In part, this stems from our use of the uncertainty weighting (see Text S4 and Figure S4). Accounting for the measurement uncertainty in this way results in stronger regressions despite the limited number of points, but can lead to certain points (especially those between 5-15 °C) seemingly deviating from the fitted lines, due to their higher uncertainty. We have added additional statistical information to every coefficient of determination to indicate whether the regression is statistically significant or not. Some examples to the text additions include:

   Sect. 3.3., lines 346-347: "...the strong linear decrease observed in total number (N, R^2 = 0.99, $p < 0.05$)".

   Sect. 3.3., lines 379-380: "At the SST-binned scale, all three moments exhibit strong, negative correlations with SST, which are well represented by linear regressions ($R^2 \geq$ 0.98, $p < 0.05$)."

   Updated Table 1

Mode	**a_1 ($\times 10^5$) [# m^{-2} s^{-1} $°C^{-1}$]**	**a_2 ($\times 10^6$) [# m^{-2} s^{-1}]**	**R^2**	**p-value**
Aitken	-2.39	6.73	0.99	4.6×10^{-4} **
Accumulation	-2.89	8.72	1.00	6.7×10^{-6} **
Supermicron	-0.0355	0.349	0.06	0.68
Total	-5.08	15.7	1.00	4.7×10^{-5} **

	b_1 ($\times 10^3$) [# m^{-2} s^{-1} $°C^{-2}$]	**b_2 ($\times 10^5$) [# m^{-2} s^{-1} $°C^{-1}$]**	**b_3 ($\times 10^6$) [# m^{-2} s^{-1}]**	**R^2**	**p-value**
Aitken	11.6	-5.00	7.32	1.00	4.5×10^{-3} **
Accumulation	1.81	-3.29	8.81	1.00	5.8×10^{-4} **
Supermicron	-3.95	1.03	-0.109	0.96	4.0×10^{-2} *
Total	-3.26	-4.33	15.6	1.00	2.1×10^{-3} **

Table 1: Least squares regression and coefficients of determination for linear and quadratic fits to total number fluxes. Astricts indicate a statistical significance of p-values exceeding the 95% (*) and 99% confidence levels (**).

3. I'm confused about the influence of SST on SSA production. Lines 250-251 state that 'increasing SST reduces the amount of SSA produced without substantially altering the size distribution', but lines 259-260 state that 'This suggests a mechanistically driven (i.e., film vs. jet drop) size-dependent shift in production and indicates that increasing SSTs have a complex influence on the resulting SSA.' Could the authors clarify whether the impact of SST on SSA production is straightforward or complex?

The impact of SST on SSA is complex, but bulk metrics (i.e., total or modal N and F_N) can be well parameterized by simple first- or second-order regressions. Paragraph 2 of Section 3.3 discusses the results when evaluating the impact on the entire size distribution or from a total concentration approach. From that perspective, submicron SSA dominates number concentrations and the overall number size distribution. So the influence of SST on submicron SSA dominates the results. Shifts in supermicron SSA will not be clearly captured by parameters such as mode diameter (D_{mode}), geometric mean diameter (D_g), or geometric standard deviation (σ_g).

The following paragraph of Section 3.3 approaches the problem from a size-resolved perspective. With that approach, a non-linear SST-SSA dependence is observed for supermicron SSA concentrations (Figure 3d).

To clarify this distinction, we have added text to Sect. 3.3 and Sect. 4 to clarify our conclusions and the complexity of the SSA-SST relationship and its potential drivers.

Sect. 3.3, lines 356-364: "These trends were consistent across cooling and warming experiments (Fig. S10) and suggest a mechanistically driven (i.e., film vs. jet drop) size-dependent shift in production. This indicates that increasing SSTs have a complex influence on the resulting SSA and emphasizes the importance of size-resolved measurements. For example, film and jet drop production may have different SST dependencies, causing size-binned N to vary with SST while the shape of the distribution (a consequence of the ratio of total film to jet drop production) remains constant. More work is needed to isolate the SST dependencies of film versus jet drop production. Subtle shifts in supermicron SSA production are not captured when assessments are made on bulk changes to SSA number concentrations. Although supermicron concentrations are limited in SOARS (Leibensperger III et al., 2026), supermicron N estimates are dominated by SSA between 1-2 μm, where counting statistics and counting efficiency remain high, providing robust results."

Sect. 4, lines 619-621: "As mentioned in Sect. 3.3., single-bubble experiments can also isolate the SST dependencies of film drop production and jet drop production, which are needed to better understand the mechanistic drivers of the SST dependence of bulk SSA measurements."

4. Two SST-dependent emission correction factors are derived in this study: CFN for number flux and CFM for mass flux. These factors represent valuable outcomes. My question is: when applying them separately to the number and mass flux calculations within an SSA emission scheme, do the two corrected fluxes remain mutually consistent?

For instance, if CFN is first applied to the number flux dF/dr, the corrected number flux is written as dF/dr×CFN. The corresponding mass flux is then calculated as dF/dr×CFN×3/4 πr^3×SSA_density.

Alternatively, if the number flux is first converted to mass flux as dF/dr×3/4πr^3×SSA_density, and CFM is applied subsequently, the final mass flux becomes dF/dr×3/4 πr^3×SSA_density×CFM.

Could you clarify whether I am using these correction factors correctly? I tested the two approaches using the Gong (2003) scheme, and the results differ.

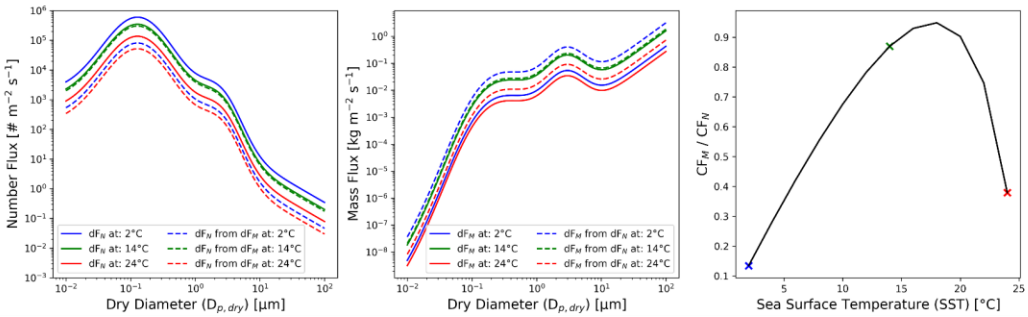
Additionally, could the authors discuss how these new factors differ from previous correction factors adopted in earth system models, such as those presented by Jaeglé et al., Grythe et al., and Sofiev et al.?

The reviewer raises a few excellent points with this critique. Namely, the CF_N and CF_M do not produce mutually consistent results when implemented independently due to the size-resolved weightings. This results in differences when calculating the same end quantity using the different CFs, but these errors are well within an order of magnitude. While this is a concern during implementation, we have clarified this limitation in Sect. 3.6. and repeat your exercise here (figure included below). Further we clarify that we advise the CF_N only be applied to number fluxes, while the CF_M should only be applied to mass or volume fluxes to avoid this issue. Additionally, the CFs are only practical within a given SST range, outside of which they result in negative values. We have addressed this second point and added their constrained SST ranges to the text:

Sect. 3.6., lines 472-477: "The linear CF_N shows good correlation ($R^2 = 0.88$, $p << 0.01$) and is valid across the SST values ranging from freezing (approximately -1.8 °C) to 30 °C. On the other hand, the quadratic CF_M shows adequate correlation ($R^2 = 0.61$, $p < 0.01$) and is valid for SST values between 0.25 and 25 °C. These validity bounds are determined by when the CFs become negative; however, we advise only adapting them to SSTs between 2 and 25 °C, consistent with the conditions here. These CFs mirror the trends of the dominant modes, where F_N is dominated by the accumulation mode, and F_M is dominated by the supermicron mode."

Sect. 3.6., lines 523-534: "These CFs are designed so that the SOARS-derived SST relationships can be adopted to emission schemes regardless of whether they are estimating number or mass (volume) flux. However, one limitation of the presented CFs is that they are not mutually consistent when applied to the same emission scheme. That is to say that applying the CF_N to an SSA number flux emission scheme and then calculating SSA mass produces a different answer than converting the PNFD to a PMFD and then applying the CF_M. This is due to the drastically different relationships of SSA number and SSA mass. These mismatches can be as large as a factor of 8 for SSTs around 2 °C, decrease to a minimum of about 1.05 for SSTs around 18°C,

and increase to about a factor of 3 for the highest SSTs. These differences highlight the variability arising from number versus mass measurements and the importance of retaining quantity consistency (i.e., number, surface area, volume, or mass) between measurements, emission schemes, and correction factors. However, these differences are within typical variability across SSA emission schemes and SST dependencies. Therefore, the CFs improve emission schemes without an SST dependence or with one of the wrong sign; users should apply the CF that matches their emission scheme quantity (number or mass) and not cross-convert."



Your exercise repeated with the Gong (2003) scheme, showing the differences at different temperatures.

We address your additional comment (comparing these CFs to those of Jaeglé et al. (2011), Grythe et al. (2014), and Sofiev et al. (2011)) in the following additions to the text. Of note, the Grythe et al. source function utilizes the SST dependence of Jaeglé et al., so we do not comment on comparisons to Grythe et al. in our additions:

Sect. 3.6., lines 514-521: "Once integrated, these CFs would deviate from the currently utilized SST dependencies of Jaeglé et al. (2011) and Sofiev et al. (2011), referred to here by the abbreviations J11 and S11. The CF_M of J11 is similar to our CF_M for SSTs less than 13 °C. For warmer SSTs, our CF_M would drastically decrease mass production, likely decreasing sea salt aerosol optical depth (SSAOD) over equatorial regions. The S11 CFs mimic our findings in the Aitken mode (increasing F_N with decreasing SST) but deviate in the accumulation and supermicron modes showing an increase in F_N with increasing SSTs. Therefore, replacing the S11 CFs with ours would decrease SSA number concentrations over warm water, likely outside most of the regions they identified for comparison to satellite retrievals of SSAOD."

Minor comments:

1. Lines 64-65: The statement "typically incorporate a positive correlation with SST (Grythe et al., 2014; Jaeglé et al., 2011; Song et al., 2023)" appears inconsistent with the content in Lines 53–54, where non-monotonic relationships are noted with the same reference: *Jaeglé et al. (2011)*.

Thank you for pointing out this inconsistency. We have adjusted the text throughout to ensure Jaeglé et al. is referenced as showing a positive correlation with SST.

2. Line 355, ‘particles whose emission rates increase with SST (Figure 7a, d)’, figure 7d dose not support this statement.

   Thank you for calling attention to this oversight. Panel d has been removed from this figure reference.

3. Supplement section S5, equations S1 and S2 (lines 72 and 81) should be S2 and S3.

   Thank you for catching this oversight we missed while rearranging the supplement. It has been corrected in the updated version.