Response: We thank the two reviewers and Dr. Harry ten Brink for thoughtful suggestions and constructive criticism that have helped us improve our manuscript. Below we provide responses to reviewer concerns and suggestions in blue font. All changes to the manuscript can be identified in the version submitted using Track Changes.

Anonymous Referee #1:

In this very nice paper the authors attack constraints on aerosol-cloud interactions using aircraft data off the California coast over many years of campaigns. Many studies use satellite observations to do this and this study provides an important ground truth evaluation of this that is needed by the field and gives additional information (for instance turbulence) that is not available from space. This study shows the utility of sulfate in predicting variability in Nd, which agrees with other studies. The data set in the study allows the authors to drill down into looking at other species (sea salt, dust, organics) that have more elusive effects on Nd. My corrections are mostly technical in nature.

L53 Adjustments may also include enhanced entrainment at cloud top (Ackerman et al., 2004).

Response: We added this effect and reference:

"For warm marine boundary layer (MBL) clouds at fixed liquid water, higher N_d values result in (i) higher cloud albedo (thus cooling the Earth and counteracting the greenhouse effect) (Twomey, 1977), (ii) delayed and/or reduced precipitation (Albrecht, 1989), and (iii) enhanced entrainment at cloud top (Ackerman et al., 2004)."

L56 This is still the case in more recent reviews (Bellouin et al., 2020).

Response: Reference was added in that line:

"The complex interactions and feedback mechanisms between aerosols, meteorology, and clouds leads to aerosol-cloud interactions as the largest source of uncertainty in climate models (IPCC, 2013; Bellouin et al., 2020)."

L181 In McCoy et al. 2018 the SS and DU was restricted to the submicron size bins from MERRA2 and only hydrophilic BC/OC were used. All mass concentrations were taken at 910 hPa. Not critical to your study, but good to keep in mind to comparing to the better resolved data from aircraft.

Response: A sentence was added at the end of the paragraph, which now reads:

"… A caveat to consider when comparing the findings of McCoy et al. (2018) to other aircraft studies is that McCoy et al. (2018) used mass concentrations retrieved exclusively at the 910 hPa model level (~ 915 m), and only considered mass concentrations pertaining to submicron SS/DU and hydrophilic BC/OC."

L351 While not essential to the analysis being performed here, one interesting possibility is for the authors to train on the NiCE or FASE campaign and test the regression on the other wildfire-affected campaign (reducing the risk of overfitting). One intriguing possibility is that not all fires produce similar aerosol in terms of CCN activitiy and influence on CCN. Were the fires during these campaigns in very different environments?

Response: We thank the reviewer for this insightful suggestion. Even though this study does not make use of training, we do address the reviewer's suggestion by analyzing on the NiCE and FASE campaigns separately. We find that FASE yields similar results to both campaigns combined, but NiCE presents better correlations between all four species analyzed and N_d. Table 8 and a figure in the supplement were modified to include the new results.

Text was added to the end of Section 3.3.2 which reads:

"The NiCE (2015) and FASE (2016) campaigns were influenced by smoke originating from different sources. NiCE was influenced by the Big Windy, Whiskey Complex, and Douglas Complex forest fires near the California-Oregon border, with a transport time of approximately two days to reach the base of aircraft operations in Marina and adjacent areas where most samples were collected (Maudlin et al., 2015). In contrast, FASE was influenced by the Soberanes fire approximately 30 km southwest of aircraft hangar (Braun et al., 2017). Hence, analyzing each campaign separately may provide some insights into the sensitivity of N_d to smoke from both different fuel types and with varying transport trajectories. NiCE fire data were linked to timber, grass and shrub models whereas those from FASE were associated with chaparral, tall grass, and timber (Braun et al., 2017; Mardi et al., 2018). The results are shown in Table 8 and Figure S4. When comparing FASE to both campaigns combined, the prediction of N_d using NSS-SO_4^{2-}, Na, Ox, and Fe is not improved, resulting in a ΔR^2_{adj} of -0.04, -0.04, 0.01, and -0.03, respectively. However, when comparing NiCE to both campaigns combined, the prediction of N_d using NSS-SO_4^{2-}, Na, Ox, and Fe is significantly improved, resulting in a ΔR^2_{adj} of 0.14, 0.29, 0.18, and 0.13, respectively. The difference between NiCE and FASE could be because different forest fires produce aerosols with varying aerosol chemical signatures and size distributions, as studies in the region have shown (Ma et al., 2019; Mardi et al., 2019). Alternatively, the difference could be due to the small sample size of NiCE (31 samples) as compared to FASE (136 samples) (Table 1). Certainly more research, including larger datasets, is warranted to investigate how different fuel types and plume aging times impact aerosol-cloud interactions."

L431 The R2 should always increase with more predictors, but R2_adj won't necessarily?

Response: R^2_{adj} is useful when comparing two regressions that have a different number of predictors. R^2 is corrected to produce R^2_{adj} using the number of predicting variables (P) and the number of data points used in the regression (N) via the equation:

$R^2_{adj} = 1 - (1- R^2)(N-1) / (N-P-1)$.

For large values of N, R^2_{adj} is about equal to R^2. For our data set, R^2 and R^2_{adj} differ by only about 2%. Therefore, the asymptotic behavior in R^2_{adj} is also observed in R^2, i.e., more predictors do not necessarily increase R^2 (or R^2_{adj}). Despite the small difference between R^2 and R^2_{adj}, we decided to use R^2_{adj} throughout the paper for the sake of rigor and consistency.

This issue is addressed by adding some text in Section 2.5, and in Section 3.2. The updated texts now read:

"However, when comparing the performance of correlations between regressions using a different number of predictor variables, it is necessary to use the adjusted coefficient of determination (R^2_{adj}), which is subscripted to distinguish it from the ordinary R^2, and is adjusted by using the number of predictors (P) and the number of data points (N) via the formula $R^2_{adj} = 1 - (1 - R^2)(N - 1)/(N - P - 1)$ (Kahane, 2008). For a large number of data points, $R^2_{adj} \sim R^2$; however, for the sake of rigor and consistency, R^2_{adj} is used instead of the ordinary R^2, except when reporting values from the literature."

"It is also interesting to note how R^2_{adj} increases asymptotically to ~ 0.6; this further makes the point that additional species do not necessarily improve predictability of N_d. The same asymptotic behavior is also exhibited with R^2, as R^2 and R^2_{adj} for these regressions differ by only ~ 2%."

L413 The authors might find it helpful to make a predictor correlation matrix figure for this section: https://seaborn.pydata.org/examples/many_pairwise_correlations.html

Response: We appreciate the suggestion and have added to the supplement a correlation matrix which includes both the 9 predictor variables and the response variable (N_d). This matrix is used to explain the possible multicollinearity causing some coefficients to have negative values. Text was added to Section 3.2, and now reads:

"The physical reason as to why these species have negative coefficients when mixed with NH_4^+ is not clear; perhaps the reason is due to the mathematics of the regression and not physically rooted, as the collinearity among three or more variables (called multicollinearity) can lead to unexpected signs for predictor coefficients (Kahane, 2008). Furthermore, a correlation matrix among the nine predicting species (Figure S2) shows a strong correlation for some pairs of species (NH_4^+-NO_3^-: $R^2_{adj} = 0.48$; NO_3^--V: $R^2_{adj} = 0.49$) and moderate correlation for other pairs (NH_4^+-V: $R^2_{adj} = 0.27$; NO_3^--Fe: $R^2_{adj} = 0.22$)."

L472 See note above regarding use of submicron SS from MERRA2 in the McCoy 2017/18 studies. One potential reason for this discrepancy is that the SS in MERRA2 is partially indicative of dynamical mixing and turbulence, which the present study has information about. Is it possible that the analysis approach in this study has disentangled this? L501 notes the strong dependence of ocean-derived species on turbulence. Would it be possible to make a bivariate plot of Nd as a function of SS and turbulence? This is done in Fig. 5, but going beyond binning into high and low turbulence might be interesting to see.

Response: The reviewer makes an excellent point in suggesting that the discrepancy between the value of the sea salt coefficient between McCoy et al. (2017, 2018) and the present study could be due to the combined effects of turbulence and sea salt, and that the present study offers an opportunity to separate these two effects. The reviewer's suggestions improved the quality of our paper and we are grateful. A new figure was made and added to Section 3.3.1.

Text was added to the end of Section 3.3.1 which reads:

"For Na, there is a better correlation at high turbulent conditions than at smooth conditions (R^2_{adj} = 0.26 and R^2_{adj} = 0.09 for high and low σ_w, respectively). This further strengthens the argument that turbulence plays an important role in the vertical transport of sea salt (and other ocean emissions) from the ocean surface to the cloud base. The present data set allows for deeper analysis into the entangled effects of sea salt and turbulence on N_d. More specifically, aerosol reanalysis products like those from MERRA-2 calculate the mass concentration of sea salt via parameterizations that link wind speed to sea salt emissions (Gong et al., 2003; Randles et al., 2017). Since wind speed affects turbulence, it follows that sea salt concentrations are not independent from turbulence, as turbulence is used to calculate sea salt concentrations. Subsequently, these sea salt concentrations are used to predict N_d (e.g., McCoy et al., 2017, 2018). The present study measured both sea salt (quantified by Na) and turbulence (quantified by σ_w) and thus offers an opportunity to try to isolate the effects of both factors on N_d (Figure 6). Two results emerge. First, more turbulence is correlated to more sea salt, which is consistent with what the models predict (Randles et al., 2017). Second, at a fixed concentration of Na, N_d does not vary significantly with σ_w, as evidenced by a weak change in color. However, at a fixed value of σ_w, N_d does vary significantly with Na, as evidenced by the noticeable change in color. Therefore, the independent measurement of both variables reveals that N_d is more sensitive to changes in Na than to changes in σ_w. We caution that σ_w is not obtained from below the cloud, but from within the cloud during sampling time (Figure S1)."

References:
Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, Nature, 432, 1014-1017, 10.1038/nature03174, 2004.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watsonâ˘AR˘ Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A. L., Dufresne, J. L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Rev Geophys, 58, 10.1029/2019rg000660, 2020.

Anonymous Referee #2:

This paper describes the relationship between cloud droplet number concentration (Nd) and cloud water composition using field measurements by aircraft flights off the California coast over 4 multi-years campaigns. After the chemical analyses of the cloudwater samples, the data were statistically analyzed to find the best correlations between chemical species and Nd. The results highlight the importance of sulfate (both Total and non-sea-salt) in predicting Nd and its variability, confirming findings already reported in previous studies. But the authors investigate also the role of other chemical species (sea-salt, dust, organic matter) as well as of some other factors (i.e., turbulence, cloud height, etc.). This is a very well-written paper that clearly describes measurements, statistical approach and results which are also nicely compared to previous findings. Even the possible drawbacks of the methodology and of the dataset are well discussed by the Authors leaving no space for substantial criticism by my side. The results are of interest for a large community investigating aerosol-cloud interaction from experimental and modelling point of view and so the publication of this work is strongly recommended as it is.

I have only a question/comment (not influencing the final decision on this paper but maybe interesting for future works): have the Authors any measurements/estimations of the acidity of cloud water? pH has an important role in sulfate aerosol formation mechanism (Turnock et al., GRL, 2019), in the gas-particle partitioning of NH4 and NO3 and in solubility of metals (Pye et al., ACP, 2020). Can the Authors comment about the possibility of testing pH as a complementary predictor (maybe partially explaining the negative coefficients of some regressions)?

Pye, H. O. T., Nenes, A., Alexander, B., Ault, A. P., Barth, M. C., Clegg, S. L., Collett Jr., J. L., Fahey, K. M., Hennigan, C. J., Herrmann, H., Kanakidou, M., Kelly, J. T., Ku, I.-T., McNeill, V. F., Riemer, N., Schaefer, T., Shi, G., Tilgner, A., Walker, J. T., Wang, T., Weber, R., Xing, J., Zaveri, R. A., and Zuend, A.: The acidity of atmospheric particles and clouds, Atmos. Chem. Phys., 20, 4809–4888, https://doi.org/10.5194/acp-20-4809-2020, 2020.

Turnock, S. T., Mann, G. W., Woodhouse, M. T., Dalvi, M., O'Connor, F. M., Carslaw, K. S., and Spracklen, D. V.: The Impact of Changes in Cloud-Water pH on Aerosol Radiative Forcing, Geophys. Res. Lett., 46, 4039–4048, https://doi.org/10.1029/2019GL082067, 2019.

Response: We appreciate this thoughtful comment from the reviewer. To address the role of pH on the ability to predict cloud droplet number concentration (N_d), H^+ (as quantified by pH) is now included as a predicting species. Thus, the total number of species is now 80. However, pH is poorly correlated to N_d, thus making it a bad predictor of N_d, and is dropped from the analysis in Step 4 of the filtering algorithm (Figure 2). Therefore, the results of this study were not altered by adding pH as a predicting species. The following parts of the manuscript have been modified to reflect the inclusion of pH:

- Section 2.3 now includes a description of the pH analysis that reads:
  "Cloud water sample acidity was quantified by measuring pH (the aqueous concentration of hydrogen ions, H^+) using a Thermo Scientific Orion 9110DJWP Combination Semi-Micro pH Electrode for E-PEACE, NiCE, and BOAS, and a Thermo Scientific Orion 8103BNUWP

Ross Ultra Semi-Micro pH probe for FASE. […] This study uses air-equivalent concentrations for all species with the exception of H^+ (pH) that uses aqueous concentration."

- Table 2 and Figure 2 now include pH.
- Section 2.4 now includes a sentence highlighting that pH was removed from the analysis that reads:
  "Even though pH plays an important role in the partitioning of gases into particles and droplets, in addition to influencing aqueous reactions in droplets (e.g., Pye et al., 2020), pH was filtered out in Step 4 for being a poor predictor of N_d."

Comment from Dr. Harry ten Brink:

I welcome a study in which the data on aerosol-cloud interaction is generalised. As a surprise I notice that the parameterisation(s) as initiated 25 years ago like B&L still are central in modelling.

Following are comments and questions

-I would have projected that a negative relation of Nd with Na would be seen because the few large seasalt particles favourably compete with the smaller much more numerous sub-submicron CCN composed of nSS (Steve Ghan). While in the remote ocean seasalt could increase CDNC it seems highly unlikely this could occur off the coast in an area with sufficient small CCN as in your case.

Response: We appreciate this insightful comment. Indeed, the effect of giant cloud condensation nuclei (GCCN) like sea salt on cloud droplets and rain drops is of much interest to the aerosol-cloud research community and deserved a better discussion. However, no conclusive results were found in this study. A paragraph was added towards the end of Section 3.2 which reads:

"When considering a multi-species model to predict N_d, it is worthwhile to examine the coefficient of sea salt. Even though it is well established that more CCN leads to more droplets, the effect of giant CCN (GCCN), such as sea salt, is not as clear. Cloud microphysics studies suggest two mechanisms by which more sea salt leads to less N_d: (1) The large size and highly hygroscopic nature of sea salt causes these particles to activate into droplets before other smaller particles. This reduces the amount of available water vapor and creates unfavorable conditions for smaller particles to nucleate into droplets (e.g., Andreae & Rosenfeld, 2008). (2) GCCN nucleate into larger droplets as compared to CCN, which in turn are more likely to collide and coalesce with surrounding droplets. This combination of droplets creates larger but fewer droplets and ultimately leads to the formation of rain drops and precipitation (e.g., Feingold et al., 1999, Jung et al., 2015). Therefore, it is expected that the negative correlation between GCCN and N_d should translate into a negative coefficient for Na (the sea salt tracer) in a multi-predictor regression equation. However, this behavior was not observed in this study. A plausible explanation for this discrepancy is that the effect of GCCN on N_d is highly dependent on conditions like LWC and N_d itself (e.g., Feingold et al., 1999), and that this study did not capture the appropriate conditions to observe this effect. However, McCoy et al. (2017) did observe a negative coefficient for sea salt and ascribed it to a simulation artefact caused by the intimate link between sea salt generation and wind speed (i.e., turbulence). An attempt to isolate the effects of sea salt and turbulence on N_d is provided in Section 3.1.1."

-line 459 e.f. the negative correlation with NO3 in case it is combined with ammonium seems to me of quite some importance given the rather high values of the two as compared with sulphate. What about a combination of sulphate and nitrate or rather nSS and nitrate, both deriving from rather similar sources and possibly similar geographical location.

Response: This is a sensible comment because it is based on the desire to deduce physical meaning from a mathematical result. However, we argue that the methodology used in this study is limited and does not allow us to address such desire satisfactorily. The limitation is not in the method of ordinary least squares (OLS), but rather in the data set fed into the OLS method. More specifically, there are two limitations to the data set: (1) perhaps we did not define a strict enough definition of collinearity when filtering species, and (2) we did not test for multicollinearity in this study. Each limitation is described below.

(1) Say you have one independent (or response) variable, y, that you want to describe in terms of two dependent (or predicting) variables, x_1 and x_2, with a linear model of the form:

$$y = a\,x_1 + b\,x_2 + c$$

The ordinary least squares (OLS) method allows to find the coefficients (a, b, and c) which best describe the data. However, the OLS method rests on the assumption that the predicting variables x_1 and x_2 are not redundant. This redundancy is called "collinearity" and can be assessed by applying the OLS method to x_1 and x_2 with an equation of the form:

$$x_1 = a\,x_2 + b$$

The predicting variables are said to be collinear if the regression yields a large correlation coefficient (R). There is no universal definition of how large R needs to be for two predicting variables to be considered collinear. When collinear predictors are fed into a model, there is no guarantee that the sign or magnitude of the parameters will have any meaning. We decided that two predictors were collinear if $R > 0.6$, but this could very well have been a lenient criterion and could be a possible source of the unexpected sign and magnitude of the NO_3^-, V, and Fe predictors.

(2) Now, say you have one response variable, y, that you want to describe in terms of **three** predicting variables, x_1, x_2, and x_3, with a linear model of the form:

$$y = a\,x_1 + b\,x_2 + c\,x_3 + d$$

The OLS method again allows to find the coefficients (a, b, c, and d), which best describe the data. And again, the OLS method relies on the assumption that the predicting variables x_1, x_2, and x_3 are not redundant. To address this, the concept of "multicollinearity" is introduced, which can be assessed by applying the OLS method to x_1, x_2, and x_3 with an equation of the form:

$$x_1 = a\,x_2 + b\,x_3 + c$$

There is not a single metric to quantify multicollinearity, but for purposes of this rebuttal, we shall use the adjusted correlation coefficient (R_{adj}). Similar to using two collinear predictors, when multicollinear predictors are fed into a model, there is no guarantee that the sign or magnitude of the parameters will have any meaning. **Furthermore, it is critical to point out that just because the pairs x_1-x_2, x_1-x_3, and x_2-x_3 are not collinear does not guarantee**

that the x_1-x_2-x_3 set is not multicollinear. Consider the fictitious data set below. If collinearity and multicollinearity were defined as $R_{adj} > 0.5$, the pairs x_1-x_2, x_1-x_3 and x_2-x_3 are all not collinear, but the x_1-x_2-x_3 set is multicollinear. Thus, it is likely that the coefficients for predictors x_1, x_2, and x_3 might lack meaning.

x_1	x_2	x_3
0.43	0.21	0.51
0.04	-0.18	0.55
0.89	0.49	1.16
0.74	0.52	0.69
0.64	-0.56	2.48
0.65	-0.24	1.87
-0.01	-0.53	1.03
0.24	-0.81	2.16
0.21	0.97	-1.14
0.4	-0.62	2.14
0.75	-0.2	2.12
0.78	0.63	0.53
0.74	-0.6	2.97
0.23	0.28	-0.08
-0.54	-0.88	0.28
0.86	0.99	0.27
-0.44	-0.74	0.14
0.7	0.2	1.28
-0.51	0.29	-1.85
-0.97	-0.66	-1.16

Test for	Linear Equation	R_{adj}	
Collinearity between two variables	$x_1 = a\,x_2 + b$ or $x_2 = a\,x_1 + b$	0.4040	All pairs are not collinear
	$x_1 = a\,x_3 + b$ or $x_3 = a\,x_1 + b$	0.5953	
	$x_2 = a\,x_3 + b$ or $x_3 = a\,x_2 + b$	– 0.3481	
Multi-collinearity between three variables	$x_1 = a\,x_2 + b\,x_3 + c$ or $x_1 = a\,x_3 + b\,x_2 + c$	0.9966	The x_1-x_2-x_3 set is multicollinear
	$x_2 = a\,x_1 + b\,x_3 + c$ or $x_2 = a\,x_3 + b\,x_1 + c$	0.9956	
	$x_3 = a\,x_1 + b\,x_2 + c$ or $x_3 = a\,x_2 + b\,x_1 + c$	0.9968	

The chemical composition of cloud water is a complex system, e.g., not all species can be attributed to their own individual source, and complex chemical reactions take place within droplets. When considering a complex system like the chemical composition of cloud water, it is reasonable to state that the more species are used to predict N_d, the higher the probability that the set of species being considered is multicollinear. We did not test for multicollinearity in this study. Therefore, it is not surprising that unexpected negative coefficients only appear when considering many (five) predictors; recall that at six predictors, all regressions become statistically insignificant. In other words, the unexpected sign and magnitude of the coefficients for NO_3^-, V, and Fe in a regression with five predictors is likely caused by multicollinearity among the predictors. This makes it difficult to gain insight into the physical-chemical processes involved.

It is helpful to keep in mind the intention we had when implementing the multivariable regression method in Section 3.2: qualitatively identify the "ingredient species" that comprise a decent set of N_d predictors, which we found to be a form of sulfate, an ocean emission tracer, and an organic tracer. Not testing for collinearity does not invalidate our finding. However, we appreciate Dr. Harry ten Brink's comment and we added a paragraph at the end of Section 2.5 that reads:

"The correct functioning of the method of ordinary least squares requires that the set of n predicting variables in Equation 3 not be collinear. Multicollinearity is defined by a set of three or more predicting variables being collinear. Using a set of multicollinear predictors can produce unreliable estimates in both magnitude and sign of the coefficients (a_i) (Kahane, 2008). There is no universal marker for multicollinearity. Furthermore, multicollinearity can only be addressed when analyzing all predictors together. For example, for a given set of three predictors (P_1, P_2, and P_3), even though the pairs P_1-P_2, P_1-P_3, and P_2-P_3 are not collinear, there is no guarantee that the P_1-P_2-P_3 set is not multicollinear. When considering a complex system such as the chemical composition of cloud water, it is reasonable to assume that as more species are used to predict N_d, the higher the probability that the set of species is multicollinear. We did not test for multicollinearity in this study; the consequences of not doing so are explored in Section 3.2."

And more discussion is provided in Section 3.2 that now reads:

"In addition, multicollinearity will become more likely as more predictors as considered. Therefore, it is not surprising that unexpected negative coefficients only appear when considering many (five) predictors. Lastly, a correlation matrix among the nine predicting species (Figure S2) shows a strong correlation for some pairs of species (NH_4^+-NO_3^-: R^2_{adj} = 0.48; NO_3^--V: R^2_{adj} = 0.49) and moderate correlation for other pairs (NH_4^+-V: R^2_{adj} = 0.27; NO_3^--Fe: R^2_{adj} = 0.22), thus strengthening the argument that the negative coefficients are due to mathematical multicollinearity and not a physical or chemical reason."

-line 132 sampling was inland in continental clouds

Response: This is a good observation but does not affect the final conclusions of this study. To avoid confusion, the word "continental" was added where appropriate in Section 1, which now reads:

"Leaitch et al. (1986) sampled <u>continental</u> stratiform and cumuliform clouds over Ontario, Canada […]. Leaitch et al. (1992) suggested that […] for both <u>continental</u> stratiform and cumuliform clouds […]."

Furthermore, the word "continental" is also added to the Leaitch et al. (1992) entry in Table 4.

-line 491. "…and it is worth noting that only five of our 385 samples are considered low turbulence according to the criterion of Leaitch et al.". This contradicts the later conclusion that the data can be translated to the NE-coast situation. There should at least some discussion on the absence of stratus-like clouds in your region.

Response: This comment contains several interesting points, please consider the following arguments:

(1) Leaitch et al. (1996) (abbreviated as L96 in this answer) encountered a certain range of turbulence conditions and we encountered a different range of turbulence conditions, as seen in the table below.

	Leaitch et al. (1996)	This study
Number of samples	24	385
Range of turbulence	0.07—0.81 m s^{-1}	0.10—0.51 m s^{-1}
33^{rd} percentile, i.e., “smooth” conditions	0.17 m s^{-1}	0.27 m s^{-1}
66^{th} percentile, i.e., “turbulent” conditions	0.23 m s^{-1}	0.32 m s^{-1}
Number (percentage) of smooth samples	4 (17%)	5 (1%)

We believe that the small overlap between our percentage of “smooth” (i.e., low turbulence) samples (1%) versus L96’s (17%) does not invalidate our statement that the northeast Atlantic region resembles the northwest Pacific region. Rather, we believe that the small overlap can be explained from a statistical point of view, namely: (a) we have 16 times more data points than L96, and (b) we consider four campaigns/summers whereas L96 considers only one.

We found inspiration in L96’s approach to use a distribution of turbulence measurements to statistically define turbulent and smooth conditions in terms of the 33^{rd} and 66^{th} percentile, respectively. Naturally, considering more data points will change the shape of the distribution and consequently also change the statistical definition of “smooth” and “turbulent”. We consider that no edits on the manuscript are required to address this concern.

(2) Even though we think that the critique to our claim that the northeast Atlantic region resembles the northwest Pacific region is not justified based on the overlap of turbulence conditions (argument 1), we do think there is value to this critique because in Section 3.1, we compare our results to Leaitch et al. (1992). As pointed out by Dr. Harry ten Brink in the previous comment, Leaitch et al. (1992) studied continental clouds, whereas, L96 studied marine clouds. To address this valid concern, we adjusted our wording in Section 3.1 from “... suggestive of commonality between two **ocean** regions ...” to “... suggestive of commonality between two **coastal** regions ...”.

(3) We respectfully disagree that stratus-like clouds are absent in our study region, since stratocumulus clouds are a type of stratus clouds. To leave no doubt in the mind of the reader on the abundance of stratocumulus/stratus-like clouds in the study region, Section 2.1 (which was renamed “Aircraft campaigns and study region”) now includes a line of text which reads:

“The persistent summertime stratocumulus cloud deck located off the California coast offers the ideal natural laboratory to study aerosol-cloud-precipitation-meteorology interactions (Russell et al., 2013; Sorooshian et al., 2018).”

1006. first entry in the table: a common error made in citing this reference, though not expected in this paper on cloud-water sulphate: the unit in the Leaitsch et al. paper of 1992 is cw-sulphate in nequivalents/m3.

Response: This is a good point. A footnote on Table 4 mentions the different units of the Leaitch et al. (1992) paper. What is wished to be emphasized when comparing our study to the Leaitch et al. (1992) study is mainly the slope (a_1) for cloud water sulfate air-equivalent concentration. Fortunately, the value of the slope is not affected by the units of sulfate concentration, as shown in the box below. We consider that no edits on the manuscript are required to address this concern.

<table>
<tr><td>Sulfate concentration (x) has units of μg m^{-3}. The slope m is given by:
$$m = \frac{\log(y_2) - \log(y_1)}{\log(x_2) - \log(x_1)}$$
$$m = \frac{log\left(\frac{y_2}{y_1}\right)}{\log\left(\frac{x_2}{x_1}\right)}$$</td><td>Sulfate concentration (x^*) has units of nEq m^{-3}, where x^* is proportional to x, i.e., $x^* = c\ x$. The slope (m^*) is given by:
$$m^* = \frac{\log(y_2) - \log(y_1)}{\log(x_2^*) - \log(x_1^*)} = \frac{log\left(\frac{y_2}{y_1}\right)}{\log\left(\frac{x_2^*}{x_1^*}\right)} = \frac{log\left(\frac{y_2}{y_1}\right)}{\log\left(\frac{c\ x_2}{c\ x_1}\right)}$$
$$m^* = \frac{log\left(\frac{y_2}{y_1}\right)}{\log\left(\frac{x_2}{x_1}\right)}$$</td></tr>
<tr><td colspan="2">$m = m^*$</td></tr>
</table>

However, it is worth mentioning that Leaitch et al. (1992) used a log-log format, whereas Leaitch et al. (1996) used a log-linear format. In Table 4, we show the Leaitch et al. (1992) study, which has the same format (with exception of the units) as the other studies in Table 4; thus, no modifications to the table are required.

Finally I really dearly miss a back-trajectory analysis of at least some typical flights or those with high nSS / NO3 and Na.

Response: We appreciate the observation that including a back-trajectory analysis enriches a paper. However, several previous papers that have analyzed the study region all converge on the same conclusion: the air in the study region is influenced by air mass transport from the north and northwest. To address this concern other readers could also have, a short paragraph was added at the end of Section 2.1 (which was renamed to "Aircraft campaigns and study region"), and reads:

"Previous studies have used back-trajectory analysis to show that air in the MBL in the study region is predominantly influenced by air mass transport from the north and northwest (Schlosser et al., 2020; Wang et al., 2016; Wonaschütz et al., 2013). Thus, the cloud water in this study was influenced by a variety of local and long-range sources such as ship exhaust (Chen et al., 2012; Coggon et al., 2012), biomass burning (Prabhakar et al., 2014; Mardi et al., 2018), ocean emissions (Dadashazar et al., 2017; MacDonald et al., 2018), continental pollution (Ma et al., 2019; Wang et al., 2016), and dust (Mardi et al., 2019; Wang et al., 2014)."