

We would like to thank the reviewer for their thorough review and helpful suggestions. Each of the comments is listed below in bold text, with our response and corrections following each of them in plain text. We have made a number of small corrections to the manuscript for each of the specific comments as noted. In addition, we have conducted a general reorganization of the results and discussion sections pursuant to the general comments. We have included relevant updated sections in response to general comments, however it may be more effective to reference the marked-up manuscript included at the end of this document to fully address the noted concerns.

Overall.

This paper is clear and well written. Measurements were made in the remote South China Sea/East Sea during the previously unmeasured Southwestern Monsoon and biomass burning season, making the measurements unique. The analysis is described in detail, and the authors took care to provide all information necessary for a clear picture of the project. The classification of the aerosol types is interesting and relevant for this audience. However, the discussion or conclusion could be expanded to include the implications of these results.

General Comments.

More discussion of the back trajectories would be useful. Why was 72-hours back chosen? Was more than one back trajectory per day run? How did the patterns change based on time of day? What time were the daily back trajectories calculated at, or are they representative of a daily average? The height of 500 m is discussed – how do other heights compare?

We have moved the description of use of the HYSPLIT model to the end of section 2.2 and amended the paragraph to address these questions as follows:

“The NOAA Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) Version 4.9 model (Draxler et al., 1999; Draxler and Hess, 1997, 1998) was used to generate daily 72-hour backtrajectories (spawned at 0Z, 8 AM local) from the *Vasco* location with arrival heights of 500m to indicate likely marine boundary layer transport patterns. The GDAS1, 1° x 1° HYSPLIT meteorological dataset was used to drive the model. Trajectory paths were found to be largely influenced by synoptic scale changes in the regional meteorological state of the atmosphere, with no substantial differences due to arrival time of day. Arrival heights between 100 m and 3000 m were examined. Trajectories with arrival heights below 1000 m were generally consistent and representative of boundary layer transport (Atwood et al., 2013; Xian et al., 2013), while higher heights tended to be increasingly influenced by free troposphere transport pathways with a more westerly component. As such, 500 m was selected to be representative of general shifts in synoptic scale boundary layer transport pathways, though more complex vertical interactions and mixing from aloft are a potential influence in the region (Atwood et al., 2013). Trajectory lengths of 72 hours were found to be sufficient to demonstrate general transport path differences between ocean dominated regions of the central portion of the SCS and more terrestrially influenced regions that passed closer to Borneo and Sumatra.”

For the cluster identification in Section 3.3, more details could be added to each group to further explain the identification. While some of the events used to identify the clusters were discussed in an earlier section, they could be added here for emphasis and clarity.

In order to better organize the results in line with several comments, we removed Section 3.2 “Analysis of Daily Observations” and have combined all of the results into the new section 3.2 with relevant descriptions of all results important to identification and description of each cluster. The new section 3.2 is included here, while a marked-up version with changes noted can be found in the full marked-up manuscript at the end of this document:

“3.2 Aerosol Population Type Classification and Properties

The cluster analysis was first conducted to investigate potential aerosol population types in the dataset, followed by physical interpretation of the results against cluster aerosol properties, coincident measurements, and meteorological conditions. The parameter values input to the cluster analysis are shown for each data point and variable in Figure 3, and colored by cluster number for the results of the eight-cluster K-Means analysis. The average value and intra-cluster standard deviation for each cluster parameter and cluster are given in Table 1.

Normalized size distributions for each of these eight aerosol populations are shown in Figure 4; the average CN and CCN number concentrations and hygroscopicities are given in Table 2. Equivalent normalized volume distributions are shown in Supplementary Figure S2. The cluster number associated with each measurement is similarly shown as the background color in Figure 2 and marker color in Figure 5. The aerosol properties, meteorological conditions, and likely transport pathways associated with data points in each cluster were then used to provide a physical interpretation of the results and identify each population type on the basis of its likely sources as discussed below. Clusters 1-4 were the most commonly found (representing 85% of the total observations, Table 2), while clusters 5-8 represented special cases, generally of short duration, that could be identified with specific locations or sampling conditions.

1. Background Marine: Data points associated with this cluster occurred throughout the study, typically following rain in the vicinity of the Vasco or transport from areas further removed from terrestrial regions. In addition, this type was observed following shortly after periods associated with each of the other identified clusters, often appearing as a transition between other types (Figure 2). The measured properties of this population type were similar to the background marine aerosol reported in many prior studies (Hoppel et al., 1994; O’Dowd et al., 1997; Spracklen et al., 2007; Good et al., 2010). The population featured a bimodal size distribution with a Hoppel minimum near 90 nm, thought to be due to cloud or fog processing of marine aerosol (Hoppel et al., 1986). The inner quartile range (IQR: middle 50% of observations between the 25%-75% percentiles) of number concentrations ranged from 382 to 623 cm⁻³, with on average 42% of the total number concentration residing in the accumulation mode as specified by the bimodal fit. Modal hygroscopicities were found to average 0.65 for the accumulation mode and 0.46 for the smaller Aitken mode, while activated fractions were generally moderate across the range of measured supersaturations as compared to other identified population types (Table 1). Each of these findings further reinforced the classification of this cluster as a typical background marine aerosol.

2. Precipitation: This distribution was found during periods immediately following extensive precipitation at or near the Vasco (Figure 2d). Air masses had been substantially scrubbed of particles and accumulation mode particles had been preferentially removed. While the number concentrations of large-mode particles were lower than those in the background marine periods, the number concentrations of smaller particles, particularly those below 40-50 nm, were comparable to the background marine type. The longest contiguous period of this

type occurred on 14 Sept. immediately following the passage of a squall line observed in the satellite visible and IR products (not shown) that left a clean air mass with fewer than 200 cm^{-3} measured in the 17-500 nm range in its wake. Number concentrations tended to be lower than the background marine type with an IQR from 227 to 441 cm^{-3} . Hygroscopicities were similarly lower than the background marine population with κ values of 0.54 and 0.34 for the accumulation and Aitken modes, respectively. Total CCN concentrations across all supersaturations were found to be lower than in background marine air masses due to the combination of fewer total particles, generally smaller particle sizes, and lower hygroscopicities.

3. Smoke: Data points associated with this aerosol type occurred primarily in two events on 14 Sept. and 25-26 Sept., during which backtrajectories were at their furthest south, near burning regions in Borneo (Figure 1a). Normalized size distributions indicated that particles were largely concentrated in a single accumulation mode with a tail of smaller particles. This type was associated with the highest total particle number and estimated submicron mass concentrations observed during the cruise, with the exception of measurements taken in the urban plume of Puerto Princessa. The standard deviations in the normalized size distribution parameters for the dominant accumulation mode in this population (Figure 4, Table 1) were small, even while number concentration varied widely (IQR 1802 to 2780 cm^{-3} ; 81% accumulation modal fraction). Accumulation mode hygroscopicities were lower than either the background marine or precipitation types, with average κ values of 0.40. Aitken mode hygroscopicities showed the opposite behavior from the first two population types with higher κ values of 0.54, though the measured uncertainties (Figure 2 e&f) and standard deviations were considerably higher than for the accumulation mode (0.25 and 0.03, respectively). Interestingly, activated fractions were highest among all population types across the full range of measured supersaturations, owing to the large number fraction of particles in the accumulation mode, while CCN concentrations were the highest of all types (except those measured in port) due primarily to the larger total particle number concentration in these smoke plumes.

Accumulation mode lognormal median diameters around 200 nm with a tail of smaller particles, elevated concentrations of carbon monoxide and benzene, as well as potassium in filter samples during this period (Reid et al., 2016, and Figure 2c) were all consistent with expectations for aged biomass burning smoke (Yokelson et al., 2008; Akagi et al., 2011; Reid et al., 2015; Sakamoto et al., 2015). Additional examination and attribution of this event to biomass burning in Sumatra and Borneo is discussed further in Reid et al., (2016). Finally, while smoke is considered the dominant aerosol source during these periods, anthropogenic pollution may still have been co-emitted along the transport path and contributed to measured results.

4. Mixed Marine: This population was characterized by periods during which the background marine type mixed with other sources of aerosol. Most of the data points associated with this type had transport pathways and biomass burning sources similar to those for the smoke population type, but with number concentrations and size distribution parameters between those of the background marine and smoke types (IQR 782 to 1160 cm^{-3} ; 68% accumulation modal fraction), indicating there was insufficient smoke for it to dominate the properties of the marine background. Accumulation and Aitken mode hygroscopicities of

0.48 and 0.54, along with activated fractions and CCN concentrations, were similarly indicative of mixing between smoke and background marine sources.

While periods of smoke mixing with a background marine air mass appeared to constitute the majority of data points in this cluster, several other periods point to other phenomena of interest being included in this type, perhaps indicating this cluster was relatively more complex than other population types. Short lived intrusions (two to five hours) of accumulation mode particles were regularly observed in both the CCN system and PCASP datasets (e.g. 18-23Z on 22, 23, and 24 Sep) after which the size distributions quickly returned to background marine conditions. These excursions were largely constrained to the pre-dawn hours (sunrise occurs around 22Z) when the boundary layer was thinnest, and when precipitation was occurring in the vicinity of the Vasco. Several prior studies have shown that smoke and anthropogenic pollution aerosol within the wider MC region can be lofted into and transported in the lower free troposphere (Tosca et al., 2011; Robinson et al., 2012; Zender et al., 2012; Campbell et al., 2013; Atwood et al., 2013). The influence of a free tropospheric aerosol layer as a source of MBL aerosol and CCN concentrations has been identified in other remote oceanic regions as well (Clarke et al., 2013). One possible explanation for these events (and possibly for the observed organic and ultrafine events that were characterized by increases in gas phase VOCs as noted in the next clusters) is therefore that aerosol may have been mixed down into the MBL from a layer aloft, perhaps on the edge of rain shafts. Alternatively, they may also be due to intermittent plumes of aerosol that survived stochastic precipitation removal events along a boundary layer transport pathway or human terrestrial activities in the pre-dawn hours. In addition, air masses influenced by anthropogenic pollution may have been included in this cluster as well, but without sufficiently different impacts on aerosol parameters to justify a distinct cluster.

5. Organic Event: An approximately four-hour period starting at 1Z on 23 Sept. had measured particle concentrations between 200 and 325 cm^{-3} , but with significantly ($p < 0.001$) larger median diameters than either the precipitation or background marine types (Figure 3). Both Aitken and accumulation mode particles had among the lowest hygroscopicities measured during the cruise, with κ values around 0.2. During this event measured concentrations of numerous VOCs were much higher than in gas canisters collected approximately 6 hours before and after it, with no associated increase in carbon monoxide (Reid et al., 2016; Figure 2c). The particles had lower hygroscopicities and larger sizes than the background marine particles observed just before this event. While the source of this event is uncertain, Robinson et al. (2012) found occasional organic aerosol above the boundary layer they attributed to biogenic Secondary Organic Aerosol (SOA) formation during an airborne campaign in the outflow regions of Borneo, while Irwin et al. (2011) reported κ values between 0.05 and 0.37 in a terrestrial, biogenically dominated MC environment. Such a source would be consistent with the observed population, perhaps due to growth of a background marine population by condensation of organics, although we lack the ancillary data needed to establish this.

6. Ultrafine Event: This cluster was associated with an approximately 20-hour period on 17-18 Sept. that included that highest concentration of particles below about 30 nm observed throughout the study (Figure 2a), and coincided with a period of elevated VOC measurements at the start of this event. A filter during this period showed very low potassium concentrations, with benzene among the lowest values measured during the study, indicating that biomass

burning was not the likely source for this event. Anthropogenic, shipping, and marine and terrestrial biogenic emissions are known sources of such compounds; isoprene, a common biogenic VOC, was not observed during this event, and a brief period of elevated dimethyl sulfide, associated with marine emissions from phytoplankton, was observed shortly before—but not during—this event (Reid et al., 2016).

A tri-modal best-fit was indicated by the Hussein, et al. (2005) algorithm for a number of these data points (Figure 2a and Supplementary Figure S1). The period had an overall IQR of 482 to 661 cm^{-3} , with generally higher ultrafine number concentrations than other periods with similar total concentrations. The accumulation mode was similar in both size and hygroscopicity ($\kappa = 0.65$) to the accumulation mode of the background marine type, while the smaller Aitken mode showed larger modal fractions and overall number concentrations, and slightly higher hygroscopicities ($\kappa = 0.50$) as compared to the background marine measurements. However, we note that the 0.38% supersaturation hygroscopicity measurement would likely not have been sensitive to these below 30 nm particles, and therefore was likely not representative of this smallest third mode. Additionally, while total number concentration was slightly higher than the background marine population measured CCN concentrations and activated fractions were generally lower, indicating many of the additional particles would not be expected to influence CCN concentrations until higher environmental supersaturations were reached. While not enough information is available to verify the nature of differences between ultrafine particles in these types, the results are consistent with an influx of smaller particles and VOCs into a background marine air mass, and were sufficiently distinct to be identified as a coherent period by the unsupervised K-Means analysis.

7. Transit: This type was associated with measurements taken during a transit away from the port of Puerto Princesa, a city with a population of over 200,000. During this period light, westerly winds advected anthropogenic pollution out over the Sulu Sea and along the path of the Vasco, allowing for sampling of the urban plume as it diluted and mixed with aerosol from other sources. Size distributions were dominated by an Aitken mode with a number median diameter around 80-90 nm, unique in measurements from this study, mixed with an accumulation mode with a smaller modal fraction than other types. The population had an IQR of 738 to 1029 cm^{-3} , while the generally decreasing number concentrations were consistent with an urban plume diluting and mixing with other aerosol populations. Modal hygroscopicity values of 0.58 and 0.62 for the accumulation and Aitken modes, respectively, were closer than those most of the other population types and consistent with high levels of sulfate aerosol in typical urban plumes.

8. Port: This type was assigned to the measurements taken during a short period in the port of Puerto Princesa. Local anthropogenic emissions were dominant during this period, with number concentrations that fluctuated between 4000 and 10,000 cm^{-3} . Ultrafine particles ($D_p < 100$ nm) dominated number concentrations during this period, although large number concentrations of accumulation mode particles with diameters between 100 and 300 nm were also observed. As measurements were fluctuating rapidly and only one CCN scan at each supersaturation setting could be completed before instrumentation was shut down, hygroscopicity results were inconclusive and uncertain. This type is considered separate from the other types as it was not measured in a remote marine area away from the immediate influence of a nearby terrestrial source.

Finally, throughout the study coarse mode particles with diameters larger than about 800 nm were consistently observed in the PCASP volume distributions (Figure 2b). Concentrations of particles in this size range increased with increasing wind speed (Figure 5), consistent with generation of sea spray aerosol due to bubble breaking and wave action (O'Dowd and Leeuw, 2007). In addition, no significant relationship between wind speed and fine mode aerosol population type was noted. Particles in the coarse mode range are not measured or accounted for in our measurements (CCN system range: 17-500 nm), and while the total concentration of coarse particles is small compared to typical CCN concentrations (Figures 2e, f), in the cleanest conditions we measured they represented non-trivial fractions of CCN active at 0.14% and 0.38% supersaturations. The large diameter of these particles makes them likely to activate at very low supersaturations, and they are present in more than sufficient number concentration to impact the microphysical structure and processes in stratocumulus clouds by serving as "giant CCN" (Feingold et al., 1999)."

Correlations with wind speed and other indicators (salt concentrations, etc.) could be useful in identifying the Background Marine cluster. It is discussed later that the coarse mode concentrations increase with increasing wind speed. Was this only for the coarse mode particles?

Wind speed was not found to be correlated to accumulation or Aitken mode particle properties in this analysis. Rather, the noted relationship of wind speed with coarse mode particles (largely thought to be sea spray aerosol, consistent with the finding of similar studies, e.g. O'Dowd and Leeuw, (2007)) was in fact independent of the fine mode particles or population type as can be generally seen in Figure 5.

This has been clarified in the paragraph at the end of the new section 3.2 as follows:

"Finally, throughout the study coarse mode particles with diameters larger than about 800 nm were consistently observed in the PCASP volume distributions (Figure 2b). Concentrations of particles in this size range increased with increasing wind speed (Figure 5), consistent with generation of sea spray aerosol due to bubble breaking and wave action (O'Dowd and Leeuw, 2007). In addition, no significant relationship between wind speed and fine mode aerosol population type was noted."

The Kappa value for background marine is also much lower than that of salt. How does that compare to other marine Kappa values, especially those with high O/C or hydrophilic compounds in the organic fraction?

We have updated a paragraph in the discussion section to discuss both the hygroscopicity of pure sea salt, as well as enhanced organic fractions that have been reported as size decreases in sea spray aerosol, and implications for our population types:

"Fresh sea spray particles, dominated by sodium chloride ($\kappa = 1.28$), are expected to have the highest κ values (Good et al., 2010), although co-emitted organic species and replacement of chlorine by uptake of acidic gases can potentially reduce hygroscopicities. Additionally, increasing organic fractions at smaller sizes have been reported in sea spray aerosol (Prather et al., 2013), leading to decreased hygroscopicities as organic fraction increases. Such findings are

consistent with the lower Aitken mode hygroscopicities found in the background marine populations observed in this dataset, as well as the additional decreases in hygroscopicity noted in the precipitation population that had been further scrubbed of larger accumulation mode particles.”

More discussion of the impact and implications of these results should be included. These clusters of aerosol types are identified, but what does that mean for other aerosol or cloud properties in the area or even globally?

The Conclusions section is somewhat short and vague. A couple more sentences with specific conclusions would be useful to summarize the main points. Include some implications of the results.

We have added more on the implications of the identified aerosol population types, specifically in regards to better understanding cloud development and radiative transfer in the remote marine SCS region. We have amended the second paragraph in the conclusion to read: “Eight aerosol population types were identified in the dataset that were associated with various impacts from background marine particles, smoke, and anthropogenic sources, as well as precipitation impacts and shorter lived events linked to influxes of VOCs or ultrafine particles. Efforts to measure or model the impact of aerosol on cloud development or atmospheric optical properties often rely on proper characterization of aerosol microphysics associated with impacts from various aerosol sources. As such, we provided population type average values and standard deviations for aerosol size distribution and hygroscopicity properties needed to model aerosol hygroscopic growth in humid environments or cloud development. Future work with this dataset will investigate the impact of the identified aerosol population types on CCN properties including supersaturation dependent CCN concentration needed to model development of different types of clouds. Reutter et al., (2009) identified specific regimes of cloud development where aerosol number concentration was important using a cloud parcel model, while Ward et al., (2010) found such results may be further complicated by aerosol size and hygroscopic properties. Inclusion of both population type average properties and the range that they vary across into such a model may help constrain when various properties of the aerosol are relevant to cloud development in the SCS. Additionally, differences in aerosol population type are expected to be relevant to studies of radiative transfer, optical propagation through the atmosphere, and satellite retrievals in sub-saturated marine environments where differences in particle number concentration, size, hygroscopicity, index of refraction, and relative humidity all affect the interaction radiation with particles in complex ways.”

Specific Comments.

Page 3, Line 8: How was sampling shut down? Add a reference to Reid et al., if it is described there.

The ship sampling ports were installed near the front of the ship such that when relative wind was from across the bow, sampled air was clear of self-sampling. During long periods when winds were not across the bow instrumentation was shut down. In short duration cases of incompatible winds, the times were invalidated and removed from the dataset. The text has been updated to reflect this, and further descriptions of these methods included in Reid et al., (2016) have now been cited. The sentence now reads:

“Sampling occurred throughout the cruise, but aerosol measurements were shut down or invalidated and removed from the dataset during periods when representative sampling could not be achieved (i.e., measured relative wind not from over the ship bow, leading to potential self-sampling; see Reid et al., (2016) for more details).”

Page 3, Line 20: Define “DMT PCASP X2”

This has been corrected to “Passive Cavity Aerosol Spectrometer Probe (PCASP)”

Page 4, Line 12: Did the properties of the aerosol remain constant in that 2 hour time period?

Conditions in the SCS were observed to occasionally change rapidly, including during the passage of gust fronts that could result in changes on the order of minutes (Reid et al., 2016). In these cases, the aerosol would in fact not be constant over a full 2 hour scan. However, the nature of the cluster model classification allows for each data point to be handled independently from others that occurred during this nominal 2 hour scan—the caveat being that while each 15 minute data period includes a full size scan, CCN measurements at only one supersaturation are included. In some cases, the aerosol would change on a faster time scale than a single 15 minute scan, however, such situations were rare and would effectively add to noise in the cluster results.

As to the wider question of how a change in aerosol across a 2 hour scan impacts the results, we have added clarification in two locations. First, at the end of the paragraph noted in this comment, we have added some additional explanation on how the 15 minute scans were conducted. Namely, the two lowest supersaturations were run more often than the three higher ones, as more reliable measurements of kappa could be obtained for the relatively larger marine aerosols. As such, the data were better interpreted as 15 minute data points with each point missing some CCN data, rather than a full two hour scan. The added sentence is:

“In addition, rather than continuously running a full two-hour scan across all supersaturation settings, individual scans (approximately 15 minutes) were run more often for the 0.14% and 0.38% settings to take advantage of this outcome.” [“this outcome” referring to valid hygroscopicity measurements at these two supersaturations noted earlier in the paragraph.]

Second, we added some additional explanation to the clustering methodology in the second paragraph of section 2.3 in order to clarify that modal hygroscopicity data was included for data points for which it was available, and treated as missing for data points when it was not.

“As hygroscopicity data was available for at most one mode (during the 0.14% or 0.38% supersaturation scans), Aitken and accumulation mode hygroscopicity was treated as missing for data points without this information. In order to account for missing data and adjust all clustering variables to the same scale, each variable was first standardized to a mean of zero and standard deviation of one, with missing data points imputed to a value a zero (the mean value). As a result, the clustering distance function was insensitive to missing data, but still included information on hygroscopicity when available.”

Page 5, Line 23: Include the parameters input into the clustering here. Did any of the variables dominate the clustering (i.e. distribution variables vs. Kappa)?

We have included a table of clustering variables used as input parameters to the cluster model, along with each cluster's mean and intra-cluster standard deviation to provide some of this information as the new Table 1. Determination of the relative importance of input parameters in a cluster analysis is subject to a number of uncertainties that make a quantitative determination of this value difficult. While analysis of variance methods may allow for comparison of the variance of cluster means against intra-cluster variance using an F-test, such results rely on the assumption of normality of each of these types of variance and produce results based on different sets of data for each variable. There is no guarantee that the cluster means generated by a cluster analysis are normally distributed about the variable mean, and the sometimes small number of cluster members makes it difficult to reject the null hypothesis that cluster values are normally distributed. As a result, we do not include an F-test metric that might provide some indication of relative importance of the variables to the cluster analysis in this table.

Similarly, conducting additional cluster analyses after having removed one of the input variables may provide some indications of relative importance, but evaluation of the similarity of the results to the original clusters suffers from the lack of a specific underlying model that describes the results (i.e. we cannot quantitatively evaluate the difference between two cluster results). As such, quantitative determination of the relative importance of each variable in the clustering is not feasible and we therefore did not speculate as the relative importance of each parameter. Instead we relied on ensuring parameters were relevant to the analysis, and verified that the results maintained a physically realistic interpretation consistent with other measurements.

The new Table 1 is:

Table 1: Aerosol population type parameters used for clustering and the resulting average values (standard deviations in grey parentheses) for each identified population.

Population Type	Total Number		Aitken Mode				Accumulation Mode			
	Concentration (# cm ⁻³)		Median (nm)	Geometric Std Dev	Number Fraction	Kappa	Median (nm)	Geometric Std Dev	Number Fraction	Kappa
1: Back. Marine	510 (181)		50 (7)	1.45 (0.05)	0.57 (0.08)	0.46 (0.17)	162 (18)	1.55 (0.10)	0.42 (0.09)	0.65 (0.11)
2: Precipitation	361 (164)		42 (5)	1.40 (0.10)	0.50 (0.12)	0.34 (0.11)	115 (27)	1.91 (0.20)	0.51 (0.12)	0.54 (0.14)
3: Smoke	2280 (606)		89 (15)	1.53 (0.17)	0.20 (0.06)	0.56 (0.25)	199 (9)	1.55 (0.04)	0.81 (0.06)	0.40 (0.03)
4: Mixed Marine	975 (271)		62 (13)	1.54 (0.18)	0.32 (0.13)	0.54 (0.23)	184 (17)	1.61 (0.08)	0.68 (0.12)	0.48 (0.10)
5: Organic Event	277 (30)		61 (2)	1.45 (0.04)	0.66 (0.01)	0.21 (0.03)	221 (8)	1.48 (0.05)	0.34 (0.01)	0.22 (0.03)
6: Ultrafine Event	577 (158)		50 (6)	1.69 (0.16)	0.82 (0.08)	0.50 (0.10)	174 (19)	1.30 (0.07)	0.19 (0.06)	0.65 (0.09)
7: Transit	976 (384)		80 (6)	1.53 (0.12)	0.73 (0.11)	0.62 (0.16)	209 (42)	1.50 (0.21)	0.26 (0.11)	0.58 (0.08)
8: Port	4890 (2550)		42 (22)	1.62 (0.37)	0.49 (0.18)	0.13 (-)	87 (26)	1.57 (0.22)	0.53 (0.18)	0.49 (-)

Page 8, Line 2: Was there anything besides the bimodal distribution that indicated this was background marine air?

We have addressed this as part of the new results section 3.2 in the background marine type description:

"1. Background Marine: Data points associated with this cluster occurred throughout the study, typically following rain in the vicinity of the Vasco or transport from areas further removed from terrestrial regions. In addition, this type was observed following shortly after periods associated with each of the other identified clusters, often appearing as a transition between other types (Figure 2). The measured properties of this population type were similar

to the background marine aerosol reported in many prior studies (Hoppel et al., 1994; O'Dowd et al., 1997; Spracklen et al., 2007; Good et al., 2010). The population featured a bimodal size distribution with a Hoppel minimum near 90 nm, thought to be due to cloud or fog processing of marine aerosol (Hoppel et al., 1986). The inner quartile range (IQR: middle 50% of observations between the 25%-75% percentiles) of number concentrations ranged from 382 to 623 cm⁻³, with on average 42% of the total number concentration residing in the accumulation mode as specified by the bimodal fit. Modal hygroscopicities were found to average 0.65 for the accumulation mode and 0.46 for the smaller Aitken mode, while activated fractions were generally moderate across the range of measured supersaturations as compared to other identified population types (Table 1). Each of these findings further reinforced the classification of this cluster as a typical background marine aerosol.”

Page 9, Line 14: What was the source of the organic event?

We do not have enough information to determine the source of this event with certainty, however, biogenic SOA formation has been reported in marine regions of the SCS (Robinson et al., 2012). Additional detailed analysis of the gas canister and other sources of data from the cruise forthcoming in other manuscripts may yield more information on this event. We have updated the description of this type in the new section 3.2 to read:

“5. Organic Event: An approximately four-hour period starting at 12Z on 23 Sept. had measured particle concentrations between 200 and 325 cm⁻³, but with significantly ($p < 0.001$) larger median diameters than either the precipitation or background marine types (Figure 3). Both Aitken and accumulation mode particles had among the lowest hygroscopicities measured during the cruise, with κ values around 0.2. During this event measured concentrations of numerous VOCs were much higher than in gas canisters collected approximately 6 hours before and after it, with no associated increase in carbon monoxide (Reid et al., 2016; Figure 2c). The particles had lower hygroscopicities and larger sizes than the background marine particles observed just before this event. While the source of this event is uncertain, Robinson et al. (2012) found occasional organic aerosol above the boundary layer they attributed to biogenic Secondary Organic Aerosol (SOA) formation during an airborne campaign in the outflow regions of Borneo, while Irwin et al. (2011) reported κ values between 0.05 and 0.37 in a terrestrial, biogenically dominated MC environment. Such a source would be consistent with the observed population, perhaps due to growth of a background marine population by condensation of organics, although we lack the ancillary data needed to establish this.”

Page 12, Line 14: If the background marine is comprised of primary marine and anthropogenic/biogenic sources, how is that type different from the Mixed Marine cluster?

We agree that this is an important distinction requiring further explanation. As noted by Shank et al., (2014), a typical background marine aerosol, even in a remote location, still contains some amount of terrestrial or anthropogenic particles. Despite this, it is still useful to have a general description of a “typical” background marine aerosol state. Conceptually, we consider the Mixed Marine type distinct from this background type, occurring when an influx of additional particles from a specific source or plume happens—an impact that is nevertheless insufficient to dominate the background aerosol (i.e. a period when only the properties of a distinct source type are observed). Practically however, we agree that there is not a clear line of

demarcation between a background and a mixed marine aerosol. Rather, it is a spectrum wherein the amount of mixing with a separate source or plume places the population somewhere between the background marine properties and the properties of the mixing population (a feature evident in Figure 3 where the Mixed Marine type generally both overlaps and falls between the smoke and background marine types). In addition, we note at the end of the description of the Mixed Marine population in the new section 3.2 that “In addition, air masses influenced by anthropogenic pollution may have been included in this cluster as well, but without sufficiently different impacts on aerosol parameters to justify a distinct cluster.” We have added additional clarification on this point to the discussion section 4 with the following sentence:

“While the background marine type was earlier noted to be impacted to some extent by background anthropogenic or terrestrial aerosol similar to impacts noted for the mixed marine type, the later was characterized by mixing with a separate, distinct aerosol population, but at levels that were insufficient to dominate the background aerosol properties.”

Further, the discussion section now includes the following paragraph with further treatment of the issue of mixing between the main population types:

“While the cluster analysis assigned each data point to a single cluster, in reality these first four clusters could be better described as a spectrum due to the variable impacts of mixing or meteorological processes, rather than as distinct or mutually exclusive population types. As is evident in Figure 3, overlap between these four clusters occurred in the parameter space for all nine of the measured variables used in the cluster model.”

Figure 1: There appear to be more square markers in the legend than in the figure. Are the square markers representative of the ship location on only the days matching the legend, or is it a range of days? It seems like it is a range of days. If that is the case, a different marker system or extended legend would be useful (i.e. circle for first 3 days, square for others, etc.). Some location labels would also be helpful for the discussion (i.e. Borneo, Sumatra, etc.).

The squares on the figure are indeed indicative of a range of days while at a stationary anchorage. We have included labels to indicate the range of days at the two anchorages, as well as additional labels of the location of Borneo, Sumatra, Palawan, the South China Sea, and the Sulu Sea.

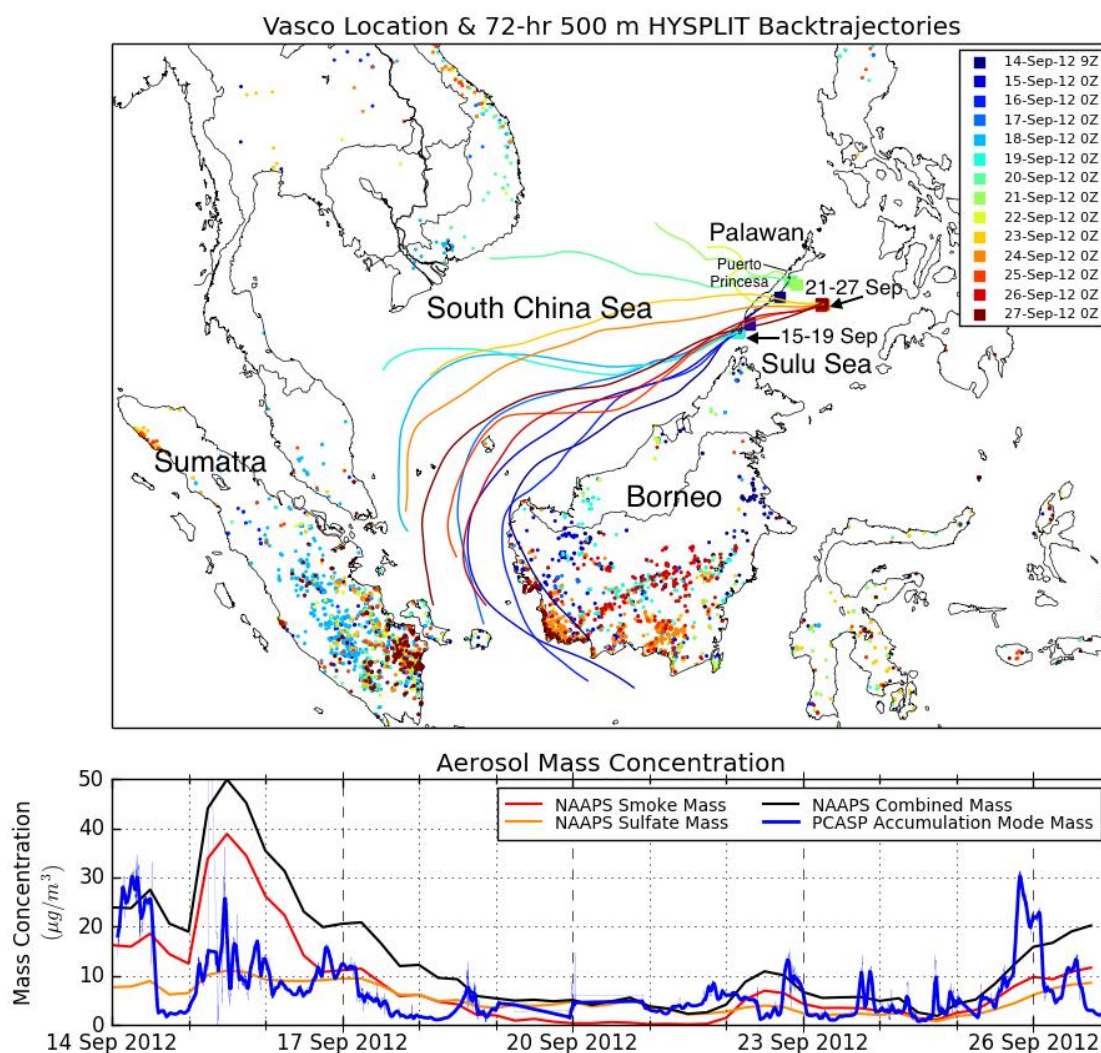


Figure 2: The background colors in the panels should all be the same (a and b are darker than the others). It is hard to associate the variability of the Kappa parameter with the different aerosol types in e and f.

We have corrected the background colors to all have the same colors in all plots, and fixed the Kappa parameter plots to improve contrast. The grey activated fraction bars in the Kappa plots have been replaced with black markers that no longer obstruct the background colors.

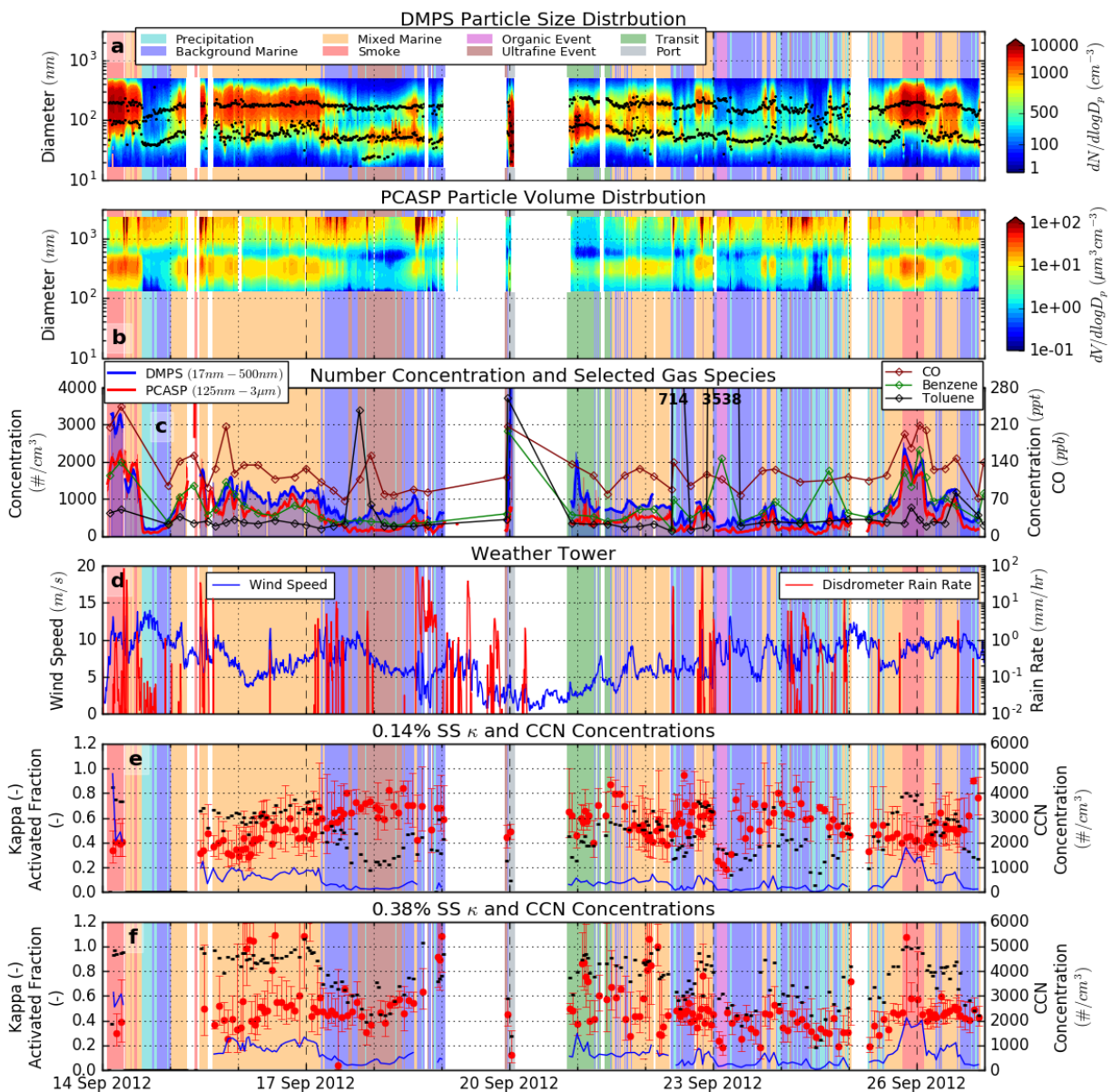


Table 1: The table is really small and should rotated or condensed for publication.

We have included a rotated table in landscape orientation on the page, and will work with the publisher to ensure the new tables are better formatted for publication. If needed, we will split the table into several lines.

Size-resolved aerosol and cloud condensation nuclei (CCN) properties in the remote marine South China Sea, Part 1: Observations and source classification

Samuel A. Atwood¹, Jeffrey S. Reid², Sonia M. Kreidenweis¹, Donald R. Blake³, Haflidi H. Jonsson⁴,
5 Nofel D. Lagrosas⁵, Peng Lynch⁶, Elizabeth A. Reid², Walter R. Sessions^{6,7}, James B. Simpas⁵

¹Department of Atmospheric Science, Colorado State University, Ft. Collins, CO

²Marine Meteorology Division, Naval Research Laboratory, Monterey, CA

³Department of Chemistry, University of California, Irvine, CA

⁴Department of Meteorology, Naval Postgraduate School, Monterey, CA

10 ⁵Manila Observatory, Manila, Philippines

⁶CSC Inc. at Naval Research Laboratory, Monterey, CA

⁷Space Sciences Engineering Center, University of Wisconsin, Madison, WI

Correspondence to: Jeffrey S. Reid (jeffrey.reid@nrlmry.navy.mil)

Abstract. Ship-based measurements of aerosol and cloud condensation nuclei (CCN) properties are presented for two weeks
15 of observations in remote marine regions of the South China Sea/East Sea during the Southwestern Monsoon (SWM) season. Smoke from extensive biomass burning throughout the Maritime Continent advected into this region during the SWM, where it was mixed with anthropogenic continental pollution and emissions from heavy shipping activities. Eight aerosol types were identified using a K-Means cluster analysis with data from a size-resolved CCN characterization system. Interpretation of the clusters was supplemented by additional onboard aerosol and meteorological measurements, ~~and~~
20 satellite, and model products for the region. A typical bimodal marine boundary layer background aerosol population was identified and ~~that was~~ observed mixing with accumulation mode aerosol from other sources, primarily smoke from fires in Borneo and Sumatra. Hygroscopicity was assessed using the κ parameter and was found to average 0.40 for samples dominated by aged, accumulation mode smoke; 0.65 for accumulation mode marine aerosol; 0.60 in an anthropogenic aerosol plume; and 0.22 during a short period that was characterized by elevated levels of volatile organic compounds not
25 associated with biomass burning impacts. As a special subset of the background marine aerosol, clean air masses substantially scrubbed of particles were observed following heavy precipitation or the passage of squall lines, with changes in observed aerosol properties occurring on the order of minutes. Average CN number concentrations, size distributions, and κ values are reported for each population type, along with CCN number concentrations for particles that activated at supersaturations between 0.14% and 0.85%.

30

Keywords. Remote Marine Aerosol, Cloud Condensation Nuclei, Biomass Burning, South China Sea, Source Apportionment, K-Means, Mixing State

1 Introduction

In the Southeast Asian Maritime Continent (MC) and South China Sea/East Sea (SCS) aerosol particles are expected to play an important role modulating cloud development, precipitation, and radiative properties that affect heat transfer through the atmosphere (Reid et al., 2013). Assessment of regional aerosol properties important to understanding such processes in the remote marine segments of this Southeast Asian Maritime Continent (MC) and South China Sea/East Sea (SCS) regions has proven difficult. Extensive cloud cover confounds remote sensing and leads to a clear sky bias in observations (Feng and Christopher, 2013; Reid et al., 2013). Aerosol monitoring has largely been confined to urban centers that are often dominated by local emissions, while in-situ sampling in remote areas has been limited in duration and scope (Irwin et al., 2011; Robinson et al., 2011; Lin et al., 2014; Reid et al., 2015). Airborne measurements have provided some representation of aerosol over wider regions and at various levels (Hewitt et al., 2010; Robinson et al., 2012), but additional questions regarding the ~~impact of various sources on surface aerosol concentrations and the~~ representativeness of such point measurements results across larger time scales remain. Similarly, the impact of various aerosol sources on surface properties and concentrations in remote marine regions, and their relationship to expected transport pathways and the few remotely sensed column measurements that exist is not well understood. Thus, over these remote ocean regions the aerosol optical and physical properties, their variability in time and space, and the processes controlling aerosol lifecycle have not been well constrained. This uncertainty in the aerosol environment itself comes in addition to uncertainty about its impacts on meteorological processes. Aerosol concentration has been found to relate to cloud development, cloud microphysics, and precipitation formation in the region (Yu et al., 2008; Yuan et al., 2011; Wang et al., 2013), while smoke may affect cloud droplet size distributions and the onset of precipitation, similar to processes observed in other tropical regions impacted by biomass burning (Rosenfeld, 1999; Andreae et al., 2004). Improved knowledge of the aerosol environment and aerosol-cloud-climate relationships in the Southeast Asian region has therefore been identified as important regionally, and in regards to links with global climate and large-scale aerosol budgets (Reid et al., 2013).

During the May through October Southwestern Monsoon (SWM) season, burning throughout the MC typically reaches its greatest extent between August and early October as precipitation associated with the ITCZ shifts north into Indochina (Reid et al., 2012). The resulting heavy smoke mixes with urban, industrial, marine, and shipping emissions in an exceedingly complex aerosol mixture (Balasubramanian et al., 2003; Atwood et al., 2013; Reid et al., 2013). During this period, aerosol particles from surface sources are generally advected by low level mean winds throughout the SCS, where they are scavenged by precipitation or eventually removed in the monsoonal trough east of the Philippines (Reid et al., 2012, 2015; Wang et al., 2013; Xian et al., 2013). As a result, the region of the SCS and Sulu Sea to the north and east of Borneo has been predicted to be a receptor for much of these biomass burning and pollution emissions from the greater MC during periods when air masses enter more convective phases of the SWM (Reid et al., 2012; Xian et al., 2013).

Two research cruises in this area were conducted during the 2011 and 2012 SWM to take advantage of these conditions and perform in-situ aerosol and meteorological measurements in the remote marine boundary layer (MBL) in the SCS and Sulu

Sea (Reid et al., 2015, 2016). In this paper, we present observations of aerosol and CCN characteristics during the second cruise, and their relationship to aerosol source type, air mass, and meteorological phenomena. These measurements represent the first in situ observations of size-resolved CCN properties [in the area](#), and fill a gap in knowledge needed to assess aerosol-cloud-precipitation relationships in the in the data poor remote marine SCS region.

5 2 Methods and Cruise Description

The SCS research cruises occurred during the month of September in both 2011 and 2012, and took place aboard the 35m, 186 ton *M/Y Vasco*, operated out of Manila by Cosmix Underwater Research Ltd. A thorough description of the vessel and instrumentation for the 2011 and 2012 cruises can be found in Reid et al. (2015) and Reid et al. (2016), respectively. Here we are concerned with the 2012 cruise, which departed Manila on 4 September from Navotas, Manila Bay, and returned on 29 September. The sampling strategy for these cruises involved moving between various anchorages in the SCS and Sulu Sea around the Palawan Archipelago. Sampling occurred throughout the cruise, but aerosol measurements were shut down [or invalidated and removed from the dataset](#) during periods when representative sampling could not be achieved (i.e., [measured relative wind not from over the ship bow, leading to potential self-sampling; incompatible mean winds and ship orientation](#)~~see self-sampling; see Reid et al., (2016) for more details~~). Remaining periods of self-sampling of ship emissions were identified by anomalous size distributions and high particle number concentrations, and were removed from the data set before analysis. The size-resolved CCN system (Petters et al., 2009) that provided the bulk of the measurements reported here had a computer failure that, once fixed, limited reliable measurements to primarily the last two weeks of the cruise, hence we focus here on data from 14-26 September 2012. This period included several transits along the east side of Palawan Island, stationary measurements at an anchorage between Palawan and Borneo (7.86N, 116.94E), and two anchorages at Tubbataha Reef in the middle of the Sulu Sea (8.80N, 119.26E).

~~The NOAA Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPPLIT) Version 4.9 model (Draxler et al., 1999; Draxler and Hess, 1997, 1998) was used to generate daily 72 hour backtrajectories from the *Vasco* location with arrival heights of 500m to indicate likely boundary layer transport patterns. The GDAS1, 1° x 1° HYSPPLIT meteorological dataset was used to drive the model.~~

25 2.1 Aerosol Measurements

A DMT [Passive Cavity Aerosol Spectrometer Probe \(PCASP\)](#) X2 configured in an aviation pod with heated inlet was mounted at the *Vasco* top mast [approximately 10 m above the water surface](#) to provide optical measurements of dry particle size distributions between approximately 125 nm and 3 μ m. Additional aerosol instrumentation was located in a forward locker, below and slightly behind the aerosol inlet. Sampled air was fed to the locker via a 3 cm diameter, 4 m long inlet located next to the PCASP. Although the inlet was not aspirated, several high flow rate nephelometers sampling from the inlet ensured low residence time in the sample line. Further details and additional instrumentation are discussed in Reid et al.

(2016). A size resolved CCN system sampled air from the inlet just inside the instrument locker with an approximately 2 m, 0.635 cm diameter copper tube. A URG cyclone with an approximate 1 μm 50% size cut was used to remove the largest particles from this CCN sampling line, including coarse mode sea-salt, to minimize wear and corrosion on downstream components. The sample was then dried using a Permapure poly-tube Nafion dryer with low pressure sheath air [to RH values below 30%, as verified by occasional in line checks using a handheld Extech Hygro-Thermometer](#).

Approximately 1.1 liters per minute (lpm) were drawn through the size resolved CCN system, which was comprised of an x-ray neutralizer (TSI Model 3087) and a TSI 3080 Electrostatic Classifier with a long DMA column (TSI Model 3081) for quasi-mono-disperse particle sizing, preceded by a 0.071 cm orifice (0.69 μm 50% cutpoint diameter) impactor. The DMA was operated with a sheath flow rate of 5 lpm (liters per minute) and sample flow rate of 1.1 lpm. The sample flow was then split between a TSI 3782 Water Condensation Particle Counter (CPC) with a flow rate of 0.6 lpm, and a DMT Cloud Condensation Nuclei Counter (CCNc) with a flow rate of 0.5 lpm. Flow rates used to calculate number concentrations were calibrated using a Gilibrator (Models 800285 & 800286) system, [with in-line measurements conducted prior to each CCNc supersaturation calibration session as noted below](#).

The size-resolved CCN system measured CN and CCN (activated particles at [a CCNc set-point supersaturation](#)) concentrations in each of 30 quasi-mono-disperse size bins between 17 nm and 500 nm. The CCNc was operated at five temperature gradient settings and calibrated using ammonium sulfate (following the methods described by Petters et al. (2009)) to measure the corresponding maximum environmental supersaturation within the CCNc column. ~~Calibrated supersaturation set points and their respective standard deviations were $0.14\% \pm 0.01$, $0.38\% \pm 0.01$, $0.52\% \pm 0.01$, $0.71\% \pm 0.02$, and $0.85\% \pm 0.03$.~~ The scan of all 30 size bins at each supersaturation took approximately 15 minutes, while a complete measurement over all 5 supersaturation settings took approximately two hours due to pauses between settings while column temperatures stabilized. The measured CN and CCN particle counts were inverted using the methodology of Petters et al. (2009). The inversion yielded the dry ambient aerosol size distribution over the measured range ($dN/d\log_{10}D_p$ for $17 \leq D_p \leq 500$ nm) and [the equivalent distribution of CCN particles activated](#)~~activation spectra at~~ [at each supersaturation \(\$dCCN/d\log_{10}D_p\$ \). The activated fraction spectrum was then calculated~~expressed~~ as the fraction of total particles \[that formed droplets \\(\\$CCN/CN\\$ \\) at each diameter \\$D_p\\$ that formed droplets\]\(#\). Each activated fraction spectrum was \[then\]\(#\) fit using a three parameter fit similar to the approach of Rose et al. \(2010\). The diameter at which 50% of particles in the fit had activated \(\$D_{50}\$ \) was used to calculate the associated hygroscopicity parameter, \$\kappa\$ \(Petters and Kreidenweis, 2007\). \[Full calibration of the CCN system flow rates and supersaturations took two to five hours to complete, and was therefore conducted four times throughout the study on 15, 20, 27 and 29 September to limit measurement downtime. Each calibration session involved running a calibration scan at each CCNc temperature gradient setting \\(with two full scans conducted at each setting on 27 Sep\\) yielding a total of five calibrations per setting throughout the cruise. Calibrated supersaturation set-points and their respective standard deviations were \\$0.14\% \pm 0.01\\$, \\$0.38\% \pm 0.01\\$, \\$0.52\% \pm 0.01\\$, \\$0.71\% \pm 0.02\\$, and \\$0.85\% \pm 0.03\\$, with no significant trend or calibration drift noted during the cruise. The measured range of 0.14% to 0.85% supersaturation was\]\(#\)](#)

[selected based on values that could both be reliably measured by the CCNc instrument and represented supersaturations expected in the region where aerosol effects may be relevant, ranging from marine stratocumulus with peak supersaturations often below 0.2% to highly convective clouds with supersaturations above 1% \(Reutter et al., 2009; Ward et al., 2010; Tao et al., 2012\).](#)

As the SCS environment tended to have relatively few particles smaller than 50 nm, only the measurements at the 0.14% and 0.38% supersaturation settings had complete activation curves that spanned the measured particle diameter range. For the higher supersaturation settings, the D_{50} diameter tended to occur at small diameters with low CN and CCN concentrations, thereby increasing uncertainty in the associated κ values. As a result, κ values were only reported for the 0.14% and 0.38% supersaturation settings. [In addition, rather than continuously running a full two–2-hour scan across all supersaturation settings, individual scans \(approximately 15 minutes\) were run more often for the 0.14% and 0.38% settings to take advantage of this outcome. The range of \$\kappa\$ values measured for the particles active at supersaturations of 0.14% and 0.38% was typically between about 0.3 and 0.8, although the full range was between 0.2 and 1.1 \(Figures 2e, f\). The \$D_{50}\$ diameters for these hygroscopicity values spanned approximately 96 to 150 nm \(\$\kappa\$ range: 0.8 – 0.2\) for the 0.14% supersaturation setting, and 45 to 80 nm \(\$\kappa\$ range: 1.1 – 0.2\) for the 0.38% setting. The hygroscopicity measurements were therefore more indicative of accumulation mode hygroscopicity during the 0.14% scans and Aitken mode hygroscopicity during those at 0.38%. Additionally, particles outside these size ranges were not well quantified by these measurements.](#)

2.2 Ancillary Measurements and Products

Additional measurements of aerosol composition were used to validate identified source types impacting the measurements throughout the cruise. A series of $PM_{2.5}$ filters were collected by 5 lpm Minivol Tactical Air Samplers [with sampling periods that varied between one and two and half days, and were that were](#) analyzed for elemental concentrations by gravimetric, XRF, and ion chromatography methods, and organic and black carbon concentrations by the thermal-optical methods [\(Reid et al., 2016\)](#). Trace gas concentrations were measured intermittently throughout the cruise by whole-air gas samples for gas chromatography analysis, as discussed further in Reid et al. (2016).

A suite of weather monitoring instruments was located on a 3m bow mast to provide coincident meteorological measurements throughout the study. From this suite, wind speed and wind direction measurements from a Campbell sonic anemometer were used to identify gust front passage. An OTT Parsivel disdrometer was utilized to measure precipitation, from which only the rain rate measurements were used in this analysis.

Several remote sensing and model products were used to characterize the wider SCS atmospheric environment and to identify potential aerosol sources. MODIS visible and IR products were used to identify convection and squall line propagation across the SCS. [The MODIS Collection 6 MOD08 Level 3 daily Aerosol Optical Depth products were utilized for AOD measurements in the region, though cloud cover obscured measurements throughout much of the study.](#) MODIS active fire hotspot analysis and the FLAMBE smoke flux product from Terra and Aqua were used to identify the locations and times during which fires were burning in the MC (Giglio et al., 2003; Reid et al., 2009; Hyer et al., 2013). Simulations

from the Navy NOGAPS model were used to represent surface and 700 hPa winds, interpolated to one-degree spatial resolution and averaged over the study period, to provide an estimate of typical~~and thereby estimate likely~~ aerosol transport pathways ~~throughout the study~~ (Hogan and Rosmond, 1991; Xian et al., 2013)(~~Hogan and Rosmond, 1991~~). Finally, the Navy Aerosol Analysis and Prediction System (NAAPS) was used to predict smoke and sulfate aerosol mass concentrations at the surface along the *Vasco* ship track (Lynch et al., 2016).

The NOAA Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLOT) Version 4.9 model (Draxler et al., 1999; Draxler and Hess, 1997, 1998) was used to generate daily 72-hour backtrajectories (spawned at 0Z, 8 AM local) from the *Vasco* location with arrival heights of 500m to indicate likely marine boundary layer transport patterns. The GDAS1, 1° x 1° HYSPLOT meteorological dataset was used to drive the model. Trajectory paths were found to be largely influenced by synoptic scale changes in the regional meteorological state of the atmosphere, with no substantial differences due to arrival time of day. Arrival heights between 100 m and 3000 m were examined. Trajectories with arrival heights below 1000 m were generally consistent and representative of boundary layer transport (Atwood et al., 2013; Xian et al., 2013), while higher heights tended to be increasingly influenced by free troposphere transport pathways with a more westerly component. As such, 500 m was selected to be representative of general shifts in synoptic scale boundary layer transport pathways, though more complex vertical interactions and mixing from aloft are a potential influence in the region (Atwood et al., 2013). Trajectory lengths of 72 hours were found to be sufficient to demonstrate general transport path differences between ocean dominated regions of the central portion of the SCS and more terrestrially influenced regions that passed closer to Borneo and Sumatra.

2.3 Aerosol Population Type Classification

The $dN/d\log_{10}D_p$ dry particle size distributions obtained every 15 minutes from the CCN system were first normalized by the particle number concentration summed over all bins. Each of these normalized size distributions was then parameterized by fitting a lognormal mixture distribution using an algorithm based on Hussein et al., (2005). This algorithm fit each distribution to one, two, or three lognormal modes, each ~~described~~defined by three lognormal distribution parameters (median diameter, geometric standard deviation, and fractional number concentration)—~~which identified a best fit using t.~~ Two modes were identified as the best fit for 690 of 731 data points in this dataset. In order to have a common point of comparison for classification purposes ~~the a~~ two mode ~~best~~-fit was used for all data points, yielding a parameterized Aitken and accumulation mode for each 15-minute data point. ~~—d~~Data points for which the fitting algorithm selected an additional third mode are noted in the clustering results in Sect. 4.3 and the Supplementary Figure S1.

These normalized size distribution parameters were then combined with hygroscopicity and total number concentration measurements to serve as input variables for an unsupervised cluster analysis (parameters given in Table 1). As hygroscopicity data was available for at most one mode (during the 0.14% or 0.38% supersaturation scans), Aitken and accumulation mode hygroscopicity was treated as missing for data points without this information. In order to account for missing data and adjust all clustering variables to the same scale, each variable was first standardized to a mean of zero and

standard deviation of one, with missing data points given imputed to a value a zero (the mean value). As a result, the clustering distance function was insensitive to missing data, but still included information on hygroscopicity when available.

A hierarchical cluster analysis was first conducted using the *cluster.AgglomerativeClustering* class of the Python scikit-learn package (Pedregosa et al., 2011) using the Ward linkage to help ascertain the number of clusters that can be found in the dataset. Each step in this process involved merging two data points or clusters into a new cluster based on those points with the closest-shortest distance between normalized input measurements (Karl Pearson distance function). A dendrogram and associated measure of the distance between merged clusters for each subsequent clustering step was used to identify potential numbers of clusters appropriate for the dataset. The distance between merged clusters increases at steps that merge substantially different clusters (Wilks, 2011)—in this case indicating 5, 8, 9, and 12 clusters as potentially appropriate for this data set.

A nonhierarchical K-Means cluster analysis was then conducted for each of these four potential cluster numbers using the scikit-learn *cluster.KMeans* class to refine the cluster members. The appropriate number of clusters was selected to classify the dataset according to was identified based on the K-Means result with the least number of clusters that maintained physically distinct and temporally consistent aerosol populations for the associated clusters. In particular, as the timestamp of a data point was not included in the cluster analysis, clusters with smaller numbers of data points were considered distinct if they all occurred during a narrow time frame that could be associated with a transient atmospheric phenomenon. The time frames of all clusters were then compared against other aerosol and meteorological observations to ensure they were physically meaningful. The result was a set of eight identified aerosol population types with associated time periods corresponding to the 15 minute CCN system data points. Finally, size distribution and hygroscopicity measurements were averaged for all time periods associated with each population type.

3 Results

3.1 Overview of Study

The daily positions of the *Vasco* during the two weeks in September 2012 that comprise this study are shown in Figure 1a, together with HYSPLIT 72-hour backtrajectories initiated within the MBL. Several extended periods at the same anchorage are noted by the range of dates spent at these locations. The daily fire hot spots are also indicated in this figure. Average NOGAPS surface winds in the boundary layer were from the southwest, often advecting air parcels from near Borneo, while lower free tropospheric winds, such as those at 700 hPa, were more westerly due to the generally veering structure of the lower atmosphere in the SCS (Reid et al., 2016). Fires detected by MODIS occurred throughout much of Borneo and Southern Sumatra during the study, with surface-level trajectories near the start and near the end of the study period passing close to active fires, whereas those during the middle period remained primarily over open ocean. Results from the NAAPS model, along with limited satellite AOD measurements not obscured by clouds during this period, confirmed this general smoke transport pathway. Accumulation mode aerosol mass concentration estimates (Figure 1b) were initially derived

generated from the PCASP measurements using a density of $1.4 \mu\text{g m}^{-3}$ (Levin et al., 2010), assumed to be representative of a combination of smoldering peat and agricultural fire emissions typical in the MC (Reid et al., 2012) that constituted the largest plumes observed during the study (Reid et al., 2016). Coincident model estimates generated by NAAPS along the ship track indicated generally similar results, with the highest mass concentrations occurring ~~Mass concentrations were highest~~ early and late in the measurement period, in general agreement with ~~times during which the~~ backtrajectories ~~passing passed~~ over terrestrial sources and active fires, ~~and with NAAPS modeled surface aerosol smoke and sulfate concentrations (Figure 1b).~~ Air parcels advecting into the SCS and Sulu Sea during this period that originated from areas further to the north and west were cleaner than those from other sectors, due to fewer emission sources and more precipitation along the trajectories. Changes in the observed particle concentrations also occurred on timescales shorter than these weekly large-scale variations, as shown in the aerosol observation timelines in Figures 2a-c. Many of these higher-frequency fluctuations were associated with squall line passages and heavy local precipitation, as discussed further ~~below in the following section.~~ The timeline of dN/dlogDp size distributions, as measured by the CCN system and shown in Figure 2a, indicates that most of the particle number concentration fell within the 17-500 nm measurement range, except possibly during the highest-concentration periods. A more extensive comparison of model and satellite measurement in situ observations is discussed in Reid et al. (2016).

~~3.2 Analysis of Daily Observations~~

~~At the start of the measurement period on 14 Sept., the Vasco was in transit south from Puerto Princessa, Philippines, in the middle of the island towards Balabac Island at the southern edge of Palawan Archipelago. HYSPLIT 500 m backtrajectories on this day identified MBL transport from the southwest over burning regions in Borneo (Figure 1a). As shown in Figure 2,~~ at the start of the campaign, high number concentrations of accumulation mode particles (1500 to 4000 cm^{-3}) with generally mono-modal size distributions were observed (best fit lognormal median diameters around 200 nm), consistent with expectations for aged smoke. Elevated concentrations of potassium, carbon monoxide, and benzene during this period (Reid et al., 2016, and Figure 2e) all reinforce the classification of this aerosol as smoke (Yokelson et al., 2008; Akagi et al., 2011; Reid et al., 2015). This initial smoke impacted period was followed by a drop in particle number concentrations after 12Z on 14 Sept. as a squall line passed over the Vasco and left a clean air mass scrubbed of many of the particles. Total particle number concentrations were among the lowest values observed during the cruise, with fewer than 200 cm^{-3} measured in the 17-500 nm range. Particle number concentrations did not recover for approximately 10 hours, at which point the aerosol size spectra and concentrations resembled those measured in the air mass before the gust front passage.

The Vasco remained at the same anchorage at Balabac Island on the southern tip of the Palawan Archipelago until 19 Sept. Beginning at about 5Z on 17 Sept., particle number concentrations dropped, but not as low nor as rapidly as during the post-squall line event. A distinct bimodal size distribution was observed that coincided with the onset of precipitation observed at the boat (Figure 2d). Satellite visible and IR imagery (not shown) indicated that no long lived squall line was associated with this 17 Sept. precipitation event. Beginning around 18Z on 17 Sept. number concentrations began to increase and a third

mode of particles with diameters between 20 and 30 nm was observed (Figure 2a). This mode slowly mixed into the Aitken mode as more typical bimodal distributions resumed by the middle of 18 Sept. A filter during this period showed very low potassium concentrations, with benzene among the lowest values measured during the study, indicating that biomass burning was not the likely source for this event. In addition, at the start of this period several volatile organic compound species (VOCs) were elevated as compared to periods before and after the event. Anthropogenic, shipping, and marine and terrestrial biogenic emissions are known sources of such compounds; isoprene, a common biogenic VOC, was not observed during this event, and a brief period of elevated dimethyl sulfide, associated with marine emissions from phytoplankton, was observed shortly before but not during this event (Reid et al., 2016).

The data gap on 19 Sept occurred during the return transit to Puerto Princessa, as the trailing winds caused self-sampling of boat emissions. However, westerly winds allowed for sampling once the *Vasco* turned to head into port, and measurements were continued for several hours to characterize emissions from the port and the city with a population of over 200,000. Number concentrations in port were considerably higher than in the remote marine locations of the SCS, with total number concentrations between 4000 and 10,000 cm⁻³ in the 17–500 nm size range. Ultrafine particles ($D_p < 100$ nm) dominated number concentrations during this period, although large number concentrations of accumulation mode particles with diameters between 100 and 300 nm were also observed.

The *Vasco* departed the port and sailed on an east-southeasterly course late on 20 Sept., during a period when both NOGAPS and the onboard weather tower indicated generally westerly low-level winds. This wind direction allowed for measurements of the Puerto Princessa urban plume as it was transported out over the Sulu Sea. Particle number concentrations between roughly 750 and 2000 cm⁻³ were measured during this period, with size spectra showing a mode with 80–90 nm median diameters. This modal median diameter was unique during the cruise; modal median diameters in the 40–70 nm or 150–225 nm ranges were more commonly observed. As the *Vasco* moved further to the southeast and out of the city plume, the size distribution measurements began to more closely resemble the previously seen bimodal background marine conditions.

The *Vasco* arrived at the remote Tubbataha reef in the middle of the Sulu Sea on 21 Sept and remained there through the end of the measurement period on 27 Sept. Throughout this time, the inflow arm of a nearby tropical cyclone spawned large amounts of intermittent convection and cloud cover over much of the SCS and Borneo. Measured number concentrations and size distributions showed considerable variation during this period as transport of smoke from Borneo was intermixed with cleaner periods associated with precipitation events. A final larger smoke event occurred on 25 and 26 Sept shortly before the end of the measurement period. This event was similar to the early smoke-dominated period with largely mono-modal size distributions and total particle number concentrations above 2000 cm⁻³, and was followed by mixing and a return to background marine conditions.

3.3.2 Aerosol Population Type Classification and Properties

The cluster analysis was first conducted to investigate potential aerosol population types in the dataset, followed by physical interpretation of the results against cluster aerosol properties, coincident measurements, and meteorological conditions. The parameter values input to the cluster analysis are shown for each data point and variable in Figure 3, and colored by cluster number for the results of the eight-cluster K-Means analysis. The average value and intra-cluster standard deviation for each cluster parameter and cluster are given in Table 1. Normalized size distributions for each of these eight aerosol populations are shown in Figure 4; the average CN and CCN number concentrations and hygroscopicities are given in Table 2. Equivalent normalized volume distributions are shown in Supplementary Figure S1S2. The cluster number associated with each measurement is similarly shown as the background color in Figure 2 and marker color in Figure 5. The aerosol properties, meteorological conditions, and likely transport pathways associated with data points in each cluster were then used to provide a physical interpretation of the results and a name identifying the likely source of each population type on the basis of its likely sources as discussed below. A name identifying the likely source of each population type on the basis of its likely sources as discussed below was then assigned as follows, on the basis of the previously identified meteorological and other factors discussed in Sect. 3.2. Clusters 1-4 were the most commonly found (representing 85% of the total observations, Table 2), while clusters 5-8 represented special cases, generally of short duration, that could be identified with by specific locations or sampling conditions.

1. Background Marine: Data points associated with this cluster occurred throughout the study, typically following rain in the vicinity of the Vasco or transport from areas further removed from terrestrial regions. In addition, this type was observed following shortly after periods associated with each of the other identified clusters, often appearing as a transition between other types (Figure 2). The measured properties of this population type were similar to the background marine aerosol reported in many prior studies (e.g. Hoppel et al., 1994; O'Dowd et al., 1997; Spracklen et al., 2007; Good et al., 2010). The population featured a, and consisted of a bimodal size distribution with a Hoppel minimum near 90 nm, thought to be due to cloud or fog processing of marine aerosol (Hoppel et al., 1986) near 90 nm due to cloud or fog processing. The inner quartile range (IQR: middle 50% of observations between the 25%–75% percentiles) of number concentrations ranged from 382 to 623 cm⁻³, with on average 42% of the total number concentration residing in the accumulation mode as specified by the bimodal fit. Modal hygroscopicities were found to average 0.65 for the accumulation mode and 0.46 for the smaller Aitken mode, while activated fractions were generally moderate across the range of measured supersaturations as compared to other identified population types (Table 1). Each of these findings further reinforced the classification of this cluster as a typical background marine aerosol.
2. Precipitation: This distribution was found during periods immediately following extensive precipitation at or near the Vasco (Figure 2d). A in scrubbed air masses had been substantially scrubbed of particles and where accumulation mode particles had been preferentially removed. While the number concentrations of large-mode particles were lower than those in the background marine periods, the number concentrations of smaller particles, particularly

those below 40-50 nm, were comparable to the clean-background marine type. The longest contiguous period of this type occurred on 14 Sept. immediately following the passage of a squall line observed in the satellite visible and IR products (not shown) that left a clean air mass with fewer than 200 cm⁻³ measured in the 17-500 nm range in its wake. Identified periods were nearly always associated with precipitation at or near the Vasco (Figure 2d). ~~and e~~Extended periods of this type occurred in the wake of the squall lines, though not all instances of nearby precipitation lead to this type. Number concentrations tended to be lower than the clean-background marine type with an IQR from 227 to 441 cm⁻³. Hygroscopicities were similarly lower than the background marine population with κ values of 0.54 and 0.34 for the accumulation and Aitken modes, respectively. Total CCN concentrations across all supersaturations were found to be lower than in background marine air masses due to the combination of fewer total particles, generally smaller particle sizes, and lower hygroscopicities.

3. Smoke: Data points associated with this aerosol type, particles occurred primarily in two events on 14 Sept. and 25-26 Sept., during which backtrajectories were at their furthest south, near burning regions in Borneo (Figure 1a). Normalized size distributions indicated that particles were largely concentrated in a single accumulation mode with a tail of smaller particles. This type was associated with the highest total particle number and estimated submicron mass concentrations observed during the cruise, with the exception of measurements taken in the urban plume of Puerto Princessa. The standard deviations in the normalized size distribution parameters for the dominant accumulation mode in this population ~~in~~ (Figure 4, Table 1) were small, even while number concentration varied widely (IQR 1802 to 2780 cm⁻³; 81% accumulation modal fraction). Accumulation mode hygroscopicities were lower than either the background marine or precipitation types, with average κ values of 0.40. Aitken mode hygroscopicities showed the opposite behavior from the first two population types with higher κ values of 0.54, though the measured uncertainties (Figure 2 e&f) and standard deviations were considerably higher than for the accumulation mode (0.25 and 0.03, respectively). Interestingly, activated fractions were highest among all population types across the full range of measured supersaturations, owing to the large number fraction of particles in the accumulation mode, while CCN concentrations were the highest of all types (except those measured in port) due primarily to the larger total particle number concentration in these smoke plumes.

3- Accumulation mode lognormal median diameters around 200 nm with a tail of smaller particles, elevated concentrations of carbon monoxide and benzene, as well as potassium in filter samples during this period (Reid et al., 2016, and Figure 2c) were all consistent with expectations for aged biomass burning smoke (Yokelson et al., 2008; Akagi et al., 2011; Reid et al., 2015; Sakamoto et al., 2015). Additional examination and attribution of this event to biomass burning in Sumatra and Borneo is discussed further in Reid et al., (2016). Finally, ~~W~~while smoke is considered the dominant aerosol source during these periods, anthropogenic pollution may still have been co-emitted along the transport path and contributed to measured results.

4. Mixed Marine: This population was characterized as-by periods ~~when-during which~~ the background marine type was-mixing-mixed with other sources of aerosol. Most of the data points associated with this type had transport

pathways and biomass burning sources similar to those for the smoke population type, but with number concentrations and, size distribution parameters, and hygroscopicities between those of the clean background marine and smoke types (IQR 782 to 1160 cm⁻³; 68% accumulation modal fraction), indicating there was insufficient smoke for it to dominate the properties of the marine background. Accumulation and Aitken mode hygroscopicities of 0.48 and 0.54, along with activated fractions and CCN concentrations, were similarly indicative of mixing between smoke and background marine sources.

4- While periods of smoke mixing with a background marine air mass appeared to constitute the majority of data points in this cluster, several other periods point to other phenomena of interest being included in this type, perhaps indicating this cluster was relatively more complex than other population types. Short lived intrusions (two to five hours) of accumulation mode particles were regularly observed in both the CCN system and PCASP datasets (e.g. 18-23Z on 22, 23, and 24 Sep) after which the size distributions quickly returned to background marine conditions. These excursions were largely constrained to the pre-dawn hours (sunrise occurs around 22Z) when the boundary layer was thinnest, and when precipitation was occurring in the vicinity of the *Vasco*. Several prior studies have shown that smoke and anthropogenic pollution aerosol within the wider MC region can be lofted into and transported in the lower free troposphere (Tosca et al., 2011; Robinson et al., 2012; Zender et al., 2012; Campbell et al., 2013; Atwood et al., 2013). The influence of a free tropospheric aerosol layer as a source of MBL aerosol and CCN has been identified in other remote oceanic regions as well (Clarke et al., 2013). One possible explanation for these events (and possibly for the observed organic and ultrafine events that were characterized by increases in gas phase VOCs as noted in the next clusters) is therefore that aerosol may have been mixed down into the MBL from a layer aloft, perhaps on the edge of rain shafts. Alternatively, they may also be due to intermittent plumes of aerosol that survived stochastic precipitation removal events along a boundary layer transport pathway or human terrestrial activities in the pre-dawn hours. In addition, Air masses influenced by anthropogenic pollution may have been included ~~occur~~ in this cluster as well, but without sufficiently different impacts on aerosol parameters to justify result in a distinct cluster.

5. Organic Event: An approximately ~~four-hour~~four-hour period starting at 1Z on 23 Sept, had measured particle concentrations between 200 and 325 cm⁻³, but with significantly ($p < 0.001$) larger median diameters than either the precipitation or background marine types (Figure 3). Both Aitken and accumulation mode particles had among the lowest hygroscopicities measured during the cruise, with κ values around 0.2. During this event measured concentrations of numerous VOCs were much higher than in gas canisters collected approximately 6 hours before and after it, with no associated increase in carbon monoxide (Reid et al., 2016; Figure 2c). The particles had lower hygroscopicities and larger sizes than the background marine particles ~~which were~~ observed just before this event. While the source of this event is uncertain, Robinson et al. (2012) found occasional organic aerosol above the boundary layer they attributed to biogenic Secondary Organic Aerosol (SOA) formation during an airborne campaign in the outflow regions of Borneo, while Irwin et al. (2011) reported κ values between 0.05 and 0.37 in a

terrestrial, biogenically dominated MC environment. Such a source would be consistent with the observed population, perhaps due to growth of a background marine population by condensation of organics might create aerosol similar to this type, although we lack the ancillary data needed to establish this.

6. Ultrafine Event: This cluster was associated with an approximately 20-hour period on 17-18 Sept. that included the highest concentration of particles below about 30 nm observed throughout the study (Figure 2a), and coincided with a period of elevated VOC measurements as discussed in Sect. 3.2, which coincided with at the start of this event (Figure 2c). A filter during this period showed very low potassium concentrations, while benzene was among the lowest values measured during the study, indicating that biomass burning was not the likely source for this event. Anthropogenic, shipping, and marine and terrestrial biogenic emissions are known sources of such compounds; isoprene, a common biogenic VOC, was not observed during this event, and a brief period of elevated dimethyl sulfide, associated with marine emissions from phytoplankton, was observed shortly before—but not during—this event (Reid et al., 2016).

6. ~~the largest concentrations of particles below about 30 nm during the study~~—A tri-modal best-fit was indicated by the Hussein, et al. (2005) algorithm for a number of these data points (Figure 2a and Supplementary Figure S1). The period had an overall IQR of 482 to 661 cm⁻³, with generally higher ultrafine number concentrations than other periods with similar total concentrations. The accumulation mode was similar in both size and hygroscopicity ($\kappa = 0.65$) to the accumulation mode of the background marine type, while the smaller Aitken mode showed larger modal fractions and overall number concentrations, and slightly higher hygroscopicities ($\kappa = 0.50$) as compared to ~~than~~ the ~~clean background~~ marine measurements. However, we note that the 0.38% supersaturation hygroscopicity measurement would likely not have been sensitive to these below 30 nm particles, and therefore was likely not representative of this smallest third mode. Additionally, while total number concentration was slightly higher than the background marine population, measured CCN concentrations and activated fractions were generally lower, indicating many of the additional particles would not be expected to influence CCN concentrations until higher environmental supersaturations were reached. While not enough information is available to verify the nature of differences between ultrafine particles in these types, the results are consistent with an influx of smaller particles and VOCs into a background marine air mass, and were sufficiently distinct to be identified as a coherent period by the unsupervised K-Means analysis.

7. Transit: This type was ~~assigned associated with to the cluster that identified~~ measurements taken during ~~the a~~ transit away from the port ~~to the east~~ of Puerto Princessa, a city with a population of over 200,000. During this period light, ~~when~~ westerly winds advected anthropogenic pollution out over the Sulu Sea and along the path of the Vasco, allowing for sampling of the urban plume as it diluted and mixed with aerosol from other sources. Size distributions were dominated by an Aitken mode with a number median diameter around 80-90 nm, unique in measurements from this study, mixed with an accumulation mode with a smaller modal fraction than other types. The population had an IQR of 738 to 1029 cm⁻³, while ~~The size distribution measurements~~ the generally decreasing number

concentrations were consistent with an urban plume diluting ~~via and~~ mixing with ~~background marine type other~~ aerosol populations. Modal hygroscopicity values of 0.58 and 0.62 for the accumulation and Aitken modes, respectively, were closer than those most of the other population types and consistent with high levels of sulfate aerosol in typical urban plumes, and had an IQR of 738 to 1029 cm⁻³.

8. Port: This type was assigned to the measurements taken during ~~the a~~ short period in the port of Puerto Princessa. Local anthropogenic emissions were dominant during this period, with number concentrations that fluctuated between 4000 and 10,000 cm⁻³. Ultrafine particles ($D_p < 100$ nm) dominated number concentrations during this period, although large number concentrations of accumulation mode particles with diameters between 100 and 300 nm were also observed. The majority of particles were in the ultrafine size range ($D_p < 100$ nm). As measurements were fluctuating rapidly and only one CCN scan at each supersaturation setting could be completed before instrumentation was shut down, hygroscopicity results were inconclusive and uncertain. This type is considered separate from the other types as it was not measured in a remote marine area away from the immediate influence of a nearby terrestrial source.

Finally, throughout the study coarse mode particles with diameters larger than about 800 nm were consistently observed in the PCASP volume distributions (Figure 2b). Concentrations of particles in this size range increased with increasing wind speed (Figure 5), consistent with generation of sea spray aerosol due to bubble breaking and wave action (O'Dowd and Leeuw, 2007), ~~and with less of a relationship to submicron aerosol population type. In addition, no significant relationship between wind speed and fine mode aerosol population type was noted.~~ Particles in ~~this size~~ the coarse mode range are not measured or accounted for in our measurements (CCN system range: 17-500 nm), ~~and while the total concentration of coarse particles is small. While the number of particles in this size range is small~~ compared to typical CCN concentrations (Figures 2e, f), in the cleanest conditions we measured they represented non-trivial fractions of CCN active at 0.14% and 0.38% supersaturations. The large diameter of these particles makes them likely to activate at very low supersaturations, and they are present in more than sufficient number concentration to impact the microphysical structure and processes in stratocumulus clouds by serving as “giant CCN” (Feingold et al., 1999).

3.4 Aerosol Hygroscopicity

~~The hygroscopicity parameter, κ , can be used to quantify the expected role of particle composition on water uptake and activation to cloud droplets (Petters and Kreidenweis, 2007). Anthropogenic pollution from urban areas often includes highly hygroscopic species such as ammonium sulfate ($\kappa = 0.61$) and ammonium nitrate ($\kappa = 0.67$), although non- or weakly hygroscopic species such as black carbon and nonpolar organic species are also common aerosol components. Fresh sea spray particles, dominated by sodium chloride ($\kappa = 1.28$), are expected to have the highest κ values (Good et al., 2010), although co-emitted organic species and replacement of chlorine by uptake of acidic gases can potentially reduce κ . Aged biomass burning aerosol or organic dominated particle populations have generally been found to have κ values below 0.2, while black carbon ($\kappa \approx 0$) has very low hygroscopicity (Andreae and Rosenfeld, 2008; Petters et al., 2009; Engelhart et al.,~~

2012). Similar size-resolved hygroscopicity measurements were performed in a remote rainforest location in Borneo by Irwin et al. (2011) during a time period with little to no biomass burning. They reported κ values between 0.05 and 0.37 for this terrestrial, biogenically dominated MC environment.

The range of κ values measured for the particles active at supersaturations of 0.14% and 0.38% was typically between about 0.3 and 0.8, although the full range was between 0.2 and 1.1 (Figures 2e, f). Average hygroscopicities and standard deviations for each population type at the 0.14% and 0.38% supersaturations are presented in Table 21, along with the average CN and CCN concentrations across all supersaturation settings. It is important to note that particle sizes (D_{50} : characteristic particle diameter at which 50% of particles in the CCNs have activated) corresponding to these measurements are in the range of 45–150 nm; our measurements did not characterize the hygroscopicities of either the very small particles ($D_{50} < 45$ nm) nor the particles with diameters above 150 nm.

The 0.14% supersaturation scans have D_{50} diameters that span approximately 96 to 150 nm for κ values between the approximate observed range of 0.8 and 0.2, respectively. Hygroscopicity measurements at this lowest supersaturation are therefore more sensitive to particles in the larger accumulation mode—hence our segregation of a subset of observations of κ into accumulation and Aitken parameters for clustering purposes. The averaged properties in Table 21 indicated that such accumulation mode particles had lower average hygroscopicities ($\kappa = 0.40$) in the smoke population type as compared to the precipitation, background marine, ultrafine event, and transit populations ($\kappa = 0.54, 0.65, 0.65, 0.58$, respectively), while the mixed marine population ($\kappa = 0.48$) resided between these.

High concentrations of both SO_2 and sulfate aerosol from numerous sources have been observed in the MC (Robinson et al., 2011; Reid et al., 2013). During this study, multi day filter samples showed average sulfate concentrations between approximately 0.8 and 3 $\mu\text{g}/\text{m}^3$ at the Vasco, potentially increasing during periods of smoke impacts due to burning of sulfur rich peat in the region (Reid et al., 2016). The potential peat source or mixing with other sources of sulfate may explain the higher than typical κ values for aged biomass burning aerosol.

The hygroscopicities derived from measurements at the 0.38% supersaturation set point had different trends. At the 0.38% supersaturation setting, activation occurs for particles sized between roughly 45 and 80 nm for particles in the observed 1.1 to 0.2 κ range, respectively, and thus measurements at this supersaturation were more closely identified with Aitken mode particles. During smoke impacted periods, the Aitken mode particles had κ values of 0.56 as compared with the 0.40 value observed in the accumulation mode. The aged, primary emissions from biomass burning are likely to be confined to particles larger than 100 nm (Figure 4), and it is therefore possible that the Aitken mode particles in this population were largely derived from background sources rather than from biomass burning, leading to the higher observed κ values.

Interestingly, the opposite situation occurred in the background marine and precipitation aerosol populations, where the Aitken mode was less hygroscopic than the accumulation mode (clean marine: $\kappa = 0.46$ and $\kappa = 0.65$, respectively; precipitation: $\kappa = 0.34$ and $\kappa = 0.54$, respectively). Decreasing hygroscopicity with size is consistent with an increasing organic fraction at smaller particle sizes in marine aerosol, as has been observed both in field data and in the laboratory (Cavalli et al., 2004; O'Dowd et al., 2004; Facchini et al., 2008; Prather et al., 2013). These observations are also consistent

with precipitation removal of some of the background sulfate aerosol, leading to lower hygroscopicities in cleaner aerosol populations due to marine organic aerosol becoming more dominant.

Finally, the organic event type had the lowest values ($\kappa \sim 0.2$) in both modes from the entire study. A gas canister grab sample during this period showed elevated levels of a number of organic compounds (Reid et al., 2016; Figure 2e), while the size distributions were similar to those of the background marine population type, but with slightly larger diameters (Figures 3 & 4). These results are consistent with particles dominated by organics across all sizes, perhaps due to growth of a background population by condensation of organics.

4 Discussion

Based on this classification of the SCS remote marine boundary layer aerosol environment, a conceptual picture emerges as to the nature and sources of particles encountered during the *Vasco* 2012 cruise. A bimodal marine aerosol background was present with number concentrations usually between about 300 and 700 cm^{-3} and a Hoppel minimum around 90 nm. Primary emissions via sea spray supply submicron particles consisting of a mixture of sea salt and organic components, with emitted particle diameters as small as 20 nm (Clarke et al., 2006; O'Dowd and Leeuw, 2007; Prather et al., 2013). However, but even in remote marine environments, transported anthropogenic and combustion aerosol may still be an important or even dominant source of small particles (Shank et al., 2012). The background marine population identified in this study is thus therefore considered a background state across the remote SCS that is likely comprised of a mixture of primary marine emissions along with particles derived from anthropogenic and biomass burning sources throughout the region. Departures from the typical range of background marine characteristics and number concentrations occurred under large influxes of aerosol from other sources, such as smoke from biomass burning regions, anthropogenic pollution from population centers or shipping, or when convection and precipitation removed much of the ambient particulate matter and created relatively clean air masses.

During the SWM when large amounts of biomass burning aerosol were being advected into the SCS, a population of aged, accumulation mode smoke particles was periodically injected into the MBL where it mixed with existing particles. When total particle concentrations were above roughly 1500 cm^{-3} , (Figure 3), the smoke particles dominated the background clean marine particles and had characteristic size distribution parameters and hygroscopicities that remained roughly constant regardless of further increases in the concentrations of biomass burning particles. In situations when smoke concentrations were insufficient to dominate the background marine aerosol population, a mixture of smoke and mixed with the background clean marine population types was present, with yielding size distribution and hygroscopicity parameters in between the two types. While the background marine type was earlier noted to be impacted to some extent by background anthropogenic or terrestrial aerosol similar to impacts noted for the mixed marine type, the later was characterized by mixing with a separate, distinct aerosol population, but at levels that were insufficient to dominate the background aerosol properties.

Precipitation ~~removes-removal of~~ particles that ~~have had~~ been advected into the region ~~or ventilation by cleaner air masses~~ when transport pathways changed returned the environment near the surface to its ~~clean-background~~ marine background state. However, when extensive precipitation occurred, accumulation mode particles were removed by wet deposition to a greater extent than Aitken mode particles, leading to lower overall surface number concentrations that were dominated by smaller particles, as evidenced by the emergence of a distinct precipitation population type from the cluster analysis. Based on the two *Vasco* cruises, the cleanest periods were encountered in cold pools following the passage of squall lines with number concentrations as low as 100 to 150 cm⁻³. During these periods, increased number concentrations of coarse mode aerosol were regularly observed (Figure 5; CN_{>800}: 5.5 ± 2.1 cm⁻³), that constituted a potentially important additional source of total CCN not measured by the CCN system (0.14% SS: 44 ± 25 cm⁻³; 0.38% SS: 70 ± 36 cm⁻³), particularly at low supersaturations where they would be expected to activate first.

In addition to these findings, several observed phenomena during the 2012 study were similar to those from the 2011 cruise (Reid et al., 2015). In particular, rapid changes in aerosol properties and source type were noted in the wake of squall lines that left clean air masses in their wake, while longer period fluctuations on the order of days occurred as impacts from anthropogenic and smoke transport mixed with cleaner background marine and precipitation impacted air masses. As both studies were conducted in the remote marine SCS during the biomass burning season and saw similar meteorological phenomena modulating the aerosol populations, the more detailed aerosol property results of the 2012 cruise may be representative of the general nature of changes in SCS remote marine aerosol during the SWM season. Future work in the region to compare surface properties with model results and satellite retrievals will be ultimately required to fully validate these findings.

While the cluster analysis assigned each data point to a single cluster, in reality these first four clusters could be better described as a spectrum due to the variable impacts of mixing or meteorological processes, rather than as distinct or mutually exclusive population types. As is evident in Figure 3, overlap between these four clusters occurred in the parameter space for all nine of the measured variables used in the cluster model.

Deviations from this general picture arose when influxes of other aerosol types occurred. The additional population types each mapped out generally distinct areas in one or more of the parameters, leading to their identification by the cluster model. That such clusters corresponded to temporally distinct periods with physical and meteorological relevance ultimately justified the use of the cluster model to classify aerosol population types and assign rough population boundaries to the parameter space.

While the spectrum of mixing between population types is relevant to the identification of impacts from various sources, additional consideration of these aerosol types against measurements in other regions is also warranted. Fresh sea spray particles, dominated by sodium chloride ($\kappa = 1.28$), are expected to have the highest κ values (Good et al., 2010), although co-emitted organic species and replacement of chlorine by uptake of acidic gases can potentially reduce hygroscopicities. Additionally, increasing organic fractions at smaller sizes have been reported in sea spray aerosol (Prather et al., 2013), leading to decreased hygroscopicities as ~~organic fraction increases~~ particle size decreases. Such findings are consistent with

the lower Aitken mode hygroscopicities found in the background marine populations observed in this dataset, as well as the additional decreases in hygroscopicity noted in the precipitation population that had been further scrubbed of larger accumulation mode particles.

Aged biomass burning aerosol have often been found to have κ values below 0.2 (Andreae and Rosenfeld, 2008; Petters et al., 2009; Engelhart et al., 2012), while below the smoke population type average of approximately 0.4 in the accumulation mode during this study. However, high concentrations of both SO_2 and sulfate aerosol (ammonium sulfate, $\kappa = 0.61$) from numerous sources have been observed in the MC (Robinson et al., 2011; Reid et al., 2013). During this study, multi-day filter samples showed average sulfate concentrations between approximately 0.8 and $3 \mu\text{g}/\text{m}^3$ at the *Vasco*, potentially increasing during periods of smoke impacts due to burning of sulfur rich peat in the region (Reid et al., 2016). The potential peat source or mixing with other sources of sulfate may explain the higher than typical κ values observed for aged biomass burning aerosol in the MC.

Cloud droplet formation updraft velocity and aerosol population parameters

During the transit through an anthropogenic urban plume, size distributions were dominated by an Aitken mode with 80 to 90 nm median diameters that slowly transformed, with distance from the urban center, into a more typical mixed marine type population. After a marked increase in some VOC concentrations at 19Z on 17 Sept (3 AM local) a smaller ultrafine mode of particles was observed several hours later, possibly due to photochemical processes. Larger Aitken modal fractions with higher geometric standard deviations were observed, while the accumulation mode remained similar to the background marine type. A separate event marked by an increase in VOC measured by the gas canister samples occurred on 23 Sept, characterized by a different aerosol distribution that may have mixed from aloft.

In addition to these general classifications, several interesting periods during the 2012 cruise were observed that warrant additional discussion. Short lived intrusions (two to five hours) of accumulation mode particles were regularly observed in both the CCN system and PCASP datasets, and were classified as part of the mixed marine population type (e.g. 18–23Z on 22, 23, and 24 Sep) after which the size distributions quickly returned to clean marine conditions. These excursions were largely constrained to the pre dawn hours (sunrise occurs around 22Z) when the boundary layer was thinnest, and when precipitation was occurring in the vicinity of the *Vasco*. Several prior studies have shown that smoke and anthropogenic pollution aerosol within the wider MC region can be lofted into and transported in the lower free troposphere (Tosea et al., 2011; Robinson et al., 2012; Zender et al., 2012; Campbell et al., 2013; Atwood et al., 2013). The influence of a free tropospheric aerosol layer as a source of MBL aerosol and CCN concentrations has been identified in other remote oceanic regions as well (Clarke et al., 2013). One possible explanation for these events, and possibly for the observed organic and ultrafine events that were characterized by increases in gas phase VOCs, is therefore that aerosol may have been mixed down into the MBL from a layer aloft, perhaps on the edge of rain shafts. However, they may also be due to intermittent plumes of aerosol that survived stochastic precipitation removal events along a boundary layer transport pathway or human terrestrial activities in the pre dawn hours. Future work in the region may provide further insight into these events and the potential importance of an aerosol layer aloft.

5 Conclusion

This study reports ship based measurements of aerosol size distributions and CCN properties conducted as part of the first extensive, in situ aerosol measurement campaign in remote marine regions of the South China Sea/East Sea during the important Southwestern Monsoon and biomass burning season. Analysis of approximately two weeks of measurements found aerosol characteristics consistent with those from a previous pilot study in the region during the same season, indicating that descriptions of aerosol population types and the associated meteorological and transport phenomena that modulate changes and mixing between these populations may be representative of the wider remote marine SCS during the SWM season, the regional and temporal representativeness of these results.

Eight aerosol population types were identified in the dataset that were associated with various impacts from background marine particles, smoke, and anthropogenic sources, as well as precipitation impacts and shorter lived events linked to influxes of VOCs or ultrafine particles. Efforts to measure or model the impact of aerosol on cloud development or atmospheric optical properties often rely on proper characterization of aerosol microphysics associated with impacts from various aerosol sources. As such, we provided population type average values and standard deviations for aerosol size distribution and hygroscopicity properties needed to model aerosol hygroscopic growth in humid environments or cloud development. Future work with this dataset will investigate the impact of the identified aerosol population types on CCN properties including supersaturation dependent CCN concentration needed to model development of different types of clouds. Reutter et al., (2009) identified specific regimes of cloud development where aerosol number concentration was important using a cloud parcel model, while Ward et al., (2010) found such results may be further complicated by aerosol size and hygroscopic properties. Inclusion of both population type average properties and the range that they vary across into such a model may help constrain when various properties of the aerosol are relevant to cloud development in the SCS. Additionally, differences in aerosol population type are expected to be relevant to studies of radiative transfer, optical propagation through the atmosphere, and satellite retrievals in sub-saturated marine environments where differences in particle number concentration, size, hygroscopicity, index of refraction, and relative humidity all affect the interaction radiation with particles in complex ways. ~~Many of the CCN relevant differences between the aerosol population types (particularly for processes involving supersaturations at about 0.3% or higher) occurred in Aitken mode particles. It should be noted that many measurements of aerosol based on either optical or mass properties, including typical aerosol models and satellite retrievals, will be minimally impacted by Aitken mode particles, and therefore may not be especially sensitive to important changes in CCN abundance and properties as detected in with these in situ observations. Future work with this dataset will therefore investigate if coincident optical measurements of both dry and humidified aerosol can be related to CCN properties given the aerosol properties associated with population types found here.~~

Lastly, while specific observed aerosol population types were identified in this dataset, additional open questions remain regarding the relative importance of various sources and transport pathways of aerosol into remote MBL air masses and their impact on aerosol populations. Since the surface-based observations provide only a portion of the observations needed to

construct a true aerosol budget for the MBL, the degree to which MBL aerosol may be impacted by mixing down from a reservoir aloft was not clear. Future airborne aerosol campaigns in the region may be useful to shed light on this important topic.

5 **Acknowledgments.** Funding for this research cruise and analysis was provided from a number of sources. Vasco ship time procurement was provided by the NRL 6.1 Base Program via an ONR Global grant to the Manila Observatory. Core funding for this effort was from Office of Naval Research 322 under Award Number N00014-16-1-2040 and the Naval Research Enterprise Internship Program (NREIP). Funding for NRL scientist participation was provided by the NRL Base Program and ONR 35. This material is based upon research supported by the Office of Naval Research under Award Number
10 N00014-16-1-2040, and by the Colorado State University Center for Geosciences/Atmospheric Research (CG/AR). We are most grateful to the Vasco ship management and crew, operated by Cosmix Underwater Research Ltd, (esp. Luc Heymans and Annabelle du Parc), the Manila Observatory senior management (esp. Antonia Loyzaga and Fr. Daniel McNamara), and the US State Department/ Embassy in Manila (esp. Maria Theresa Villa and Dovas Saulys). [We would also like to thank the two anonymous reviewers for their insightful comments and helpful suggestions.](#)

15 **References**

- [Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D. and Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric models, *Atmos Chem Phys*, 11\(9\), 4039–4072, doi:10.5194/acp-11-4039-2011, 2011.](#)
- [Andreae, M. O. and Rosenfeld, D.: Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, *Earth-Sci. Rev.*, 89\(1–2\), 13–41, doi:10.1016/j.earscirev.2008.03.001, 2008.](#)
- [Andreae, M. O., Rosenfeld, D., Artaxo, P., Costa, A. A., Frank, G. P., Longo, K. M. and Silva-Dias, M. a. F.: Smoking Rain Clouds over the Amazon, *Science*, 303\(5662\), 1337–1342, doi:10.1126/science.1092779, 2004.](#)
- [Atwood, S. A., Reid, J. S., Kreidenweis, S. M., Yu, L. E., Salinas, S. V., Chew, B. N. and Balasubramanian, R.: Analysis of source regions for smoke events in Singapore for the 2009 El Nino burning season, *Atmos. Environ.*, 78, 219–230, doi:10.1016/j.atmosenv.2013.04.047, 2013.](#)
- [Balasubramanian, R., Qian, W.-B., Decesari, S., Facchini, M. C. and Fuzzi, S.: Comprehensive characterization of PM2.5 aerosols in Singapore, *J. Geophys. Res. Atmospheres*, 108\(D16\), 4523, doi:10.1029/2002JD002517, 2003.](#)
- [Campbell, J. R., Reid, J. S., Westphal, D. L., Zhang, J., Tackett, J. L., Chew, B. N., Welton, E. J., Shimizu, A., Sugimoto, N., Aoki, K. and Winker, D. M.: Characterizing the vertical profile of aerosol particle extinction and linear depolarization over Southeast Asia and the Maritime Continent: The 2007–2009 view from CALIOP, *Atmospheric Res.*, 122, 520–543, doi:10.1016/j.atmosres.2012.05.007, 2013.](#)
- [Clarke, A. D., Owens, S. R. and Zhou, J.: An ultrafine sea-salt flux from breaking waves: Implications for cloud condensation nuclei in the remote marine atmosphere, *J. Geophys. Res. Atmospheres*, 111\(D6\), D06202, doi:10.1029/2005JD006565, 2006.](#)

- Clarke, A. D., Freitag, S., Simpson, R. M. C., Hudson, J. G., Howell, S. G., Brekhovskikh, V. L., Campos, T., Kapustin, V. N. and Zhou, J.: Free troposphere as a major source of CCN for the equatorial pacific boundary layer: long-range transport and teleconnections, *Atmos Chem Phys*, 13(15), 7511–7529, doi:10.5194/acp-13-7511-2013, 2013.
- 5 Draxler, R. R. and Hess, G. D.: Description of the HYSPLIT4 modeling system, [online] Available from: <http://warn.arl.noaa.gov/documents/reports/arl-224.pdf> (Accessed 14 April 2015), 1997.
- Draxler, R. R. and Hess, G. D.: An overview of the HYSPLIT_4 modelling system for trajectories, *Aust. Meteorol. Mag.*, 47(4), 295–308, 1998.
- Draxler, R. R., Stunder, B., Rolph, G. and Taylor, A.: HYSPLIT4 user's guide, NOAA Tech. Memo. ERL ARL, 230, 35, 1999.
- 10 Engelhart, G. J., Hennigan, C. J., Miracolo, M. A., Robinson, A. L. and Pandis, S. N.: Cloud condensation nuclei activity of fresh primary and aged biomass burning aerosol, *Atmos Chem Phys*, 12(15), 7285–7293, doi:10.5194/acp-12-7285-2012, 2012.
- Feingold, G., Cotton, W. R., Kreidenweis, S. M. and Davis, J. T.: The Impact of Giant Cloud Condensation Nuclei on Drizzle Formation in Stratocumulus: Implications for Cloud Radiative Properties, *J. Atmospheric Sci.*, 56(24), 4100–4117, doi:10.1175/1520-0469(1999)056<4100:TIOGCC>2.0.CO;2, 1999.
- 15 Feng, N. and Christopher, S. A.: Satellite and surface-based remote sensing of Southeast Asian aerosols and their radiative effects, *Atmospheric Res.*, 122, 544–554, doi:10.1016/j.atmosres.2012.02.018, 2013.
- Giglio, L., Descloitres, J., Justice, C. O. and Kaufman, Y. J.: An Enhanced Contextual Fire Detection Algorithm for MODIS, *Remote Sens. Environ.*, 87(2–3), 273–282, doi:10.1016/S0034-4257(03)00184-6, 2003.
- 20 Good, N., Topping, D. O., Allan, J. D., Flynn, M., Fuentes, E., Irwin, M., Williams, P. I., Coe, H. and McFiggans, G.: Consistency between parameterisations of aerosol hygroscopicity and CCN activity during the RHaMBLe discovery cruise, *Atmos Chem Phys*, 10(7), 3189–3203, doi:10.5194/acp-10-3189-2010, 2010.
- Hewitt, C. N., Lee, J. D., MacKenzie, A. R., Barkley, M. P., Carslaw, N., Carver, G. D., Chappell, N. A., Coe, H., Collier, C., Commane, R., Davies, F., Davison, B., DiCarlo, P., Di Marco, C. F., Dorsey, J. R., Edwards, P. M., Evans, M. J., Fowler, D., Furneaux, K. L., Gallagher, M., Guenther, A., Heard, D. E., Helfter, C., Hopkins, J., Ingham, T., Irwin, M., Jones, C., Karunaharan, A., Langford, B., Lewis, A. C., Lim, S. F., MacDonald, S. M., Mahajan, A. S., Malpass, S., McFiggans, G., Mills, G., Misztal, P., Moller, S., Monks, P. S., Nemitz, E., Nicolas-Perea, V., Oetjen, H., Oram, D. E., Palmer, P. I., Phillips, G. J., Pike, R., Plane, J. M. C., Pugh, T., Pyle, J. A., Reeves, C. E., Robinson, N. H., Stewart, D., Stone, D., Whalley, L. K. and Yin, X.: Overview: oxidant and particle photochemical processes above a south-east Asian tropical rainforest (the OP3 project): introduction, rationale, location characteristics and tools, *Atmos Chem Phys*, 10(1), 169–199, doi:10.5194/acp-10-169-2010, 2010.
- 25 30 Hogan, T. F. and Rosmond, T. E.: The Description of the Navy Operational Global Atmospheric Prediction System's Spectral Forecast Model, *Mon. Weather Rev.*, 119(8), 1786–1815, doi:10.1175/1520-0493(1991)119<1786:TDOTNO>2.0.CO;2, 1991.
- 35 Hoppel, W. A., Frick, G. M. and Larson, R. E.: Effect of nonprecipitating clouds on the aerosol size distribution in the marine boundary layer, *Geophys. Res. Lett.*, 13(2), 125–128, doi:10.1029/GL013i002p00125, 1986.

- Hoppel, W. A., Frick, G. M., Fitzgerald, J. W. and Larson, R. E.: Marine boundary layer measurements of new particle formation and the effects nonprecipitating clouds have on aerosol size distribution, *J. Geophys. Res. Atmospheres*, 99(D7), 14443–14459, doi:10.1029/94JD00797, 1994.
- 5 Hussein, T., Dal Maso, M., Petäjä, T., Koponen, I. K., Paatero, P., Aalto, P. P., Hämeri, K. and Kulmala, M.: Evaluation of an automatic algorithm for fitting the particle number size distributions, *Boreal Environ. Res.*, 10(5), 337–355, 2005.
- Hyer, E. J., Reid, J. S., Prins, E. M., Hoffman, J. P., Schmidt, C. C., Miettinen, J. I. and Giglio, L.: Patterns of fire activity over Indonesia and Malaysia from polar and geostationary satellite observations, *Atmospheric Res.*, 122, 504–519, doi:10.1016/j.atmosres.2012.06.011, 2013.
- 10 Irwin, M., Robinson, N., Allan, J. D., Coe, H. and McFiggans, G.: Size-resolved aerosol water uptake and cloud condensation nuclei measurements as measured above a Southeast Asian rainforest during OP3, *Atmos Chem Phys*, 11(21), 11157–11174, doi:10.5194/acp-11-11157-2011, 2011.
- 15 Levin, E. J. T., McMeeking, G. R., Carrico, C. M., Mack, L. E., Kreidenweis, S. M., Wold, C. E., Moosmüller, H., Arnott, W. P., Hao, W. M., Collett, J. L. and Malm, W. C.: Biomass burning smoke aerosol properties measured during Fire Laboratory at Missoula Experiments (FLAME), *J. Geophys. Res. Atmospheres*, 115(D18), D18210, doi:10.1029/2009JD013601, 2010.
- Lin, N.-H., Sayer, A. M., Wang, S.-H., Loftus, A. M., Hsiao, T.-C., Sheu, G.-R., Hsu, N. C., Tsay, S.-C. and Chantara, S.: Interactions between biomass-burning aerosols and clouds over Southeast Asia: Current status, challenges, and perspectives, *Environ. Pollut.*, 195, 292–307, doi:10.1016/j.envpol.2014.06.036, 2014.
- 20 Lynch, P., Reid, J. S., Westphal, D. L., Zhang, J., Hogan, T. F., Hyer, E. J., Curtis, C. A., Hegg, D. A., Shi, Y., Campbell, J. R., Rubin, J. I., Sessions, W. R., Turk, F. J. and Walker, A. L.: An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for atmospheric and climate sciences, *Geosci Model Dev*, 9(4), 1489–1522, doi:10.5194/gmd-9-1489-2016, 2016.
- O’Dowd, C. D. and Leeuw, G. de: Marine aerosol production: a review of the current knowledge, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, 365(1856), 1753–1774, doi:10.1098/rsta.2007.2043, 2007.
- 25 O’Dowd, C. D., Smith, M. H., Consterdine, I. E. and Lowe, J. A.: Marine aerosol, sea-salt, and the marine sulphur cycle: a short review, *Atmos. Environ.*, 31(1), 73–80, doi:10.1016/S1352-2310(96)00106-9, 1997.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, É.: Scikit-learn: Machine Learning in Python, *J Mach Learn Res*, 12, 2825–2830, 2011.
- 30 Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, *Atmos Chem Phys*, 7(8), 1961–1971, doi:10.5194/acp-7-1961-2007, 2007.
- Petters, M. D., Carrico, C. M., Kreidenweis, S. M., Prenni, A. J., DeMott, P. J., Collett, J. L. and Moosmüller, H.: Cloud condensation nucleation activity of biomass burning aerosol, *J. Geophys. Res. Atmospheres*, 114(D22), D22205, doi:10.1029/2009JD012353, 2009.
- 35 Prather, K. A., Bertram, T. H., Grassian, V. H., Deane, G. B., Stokes, M. D., DeMott, P. J., Aluwihare, L. I., Palenik, B. P., Azam, F., Seinfeld, J. H., Moffet, R. C., Molina, M. J., Cappa, C. D., Geiger, F. M., Roberts, G. C., Russell, L. M., Ault, A. P., Baltrusaitis, J., Collins, D. B., Corrigan, C. E., Cuadra-Rodriguez, L. A., Ebben, C. J., Forestieri, S. D., Guasco, T. L., Hersey, S. P., Kim, M. J., Lambert, W. F., Modini, R. L., Mui, W., Pedler, B. E., Ruppel, M. J., Ryder, O. S., Schoepp, N.

G., Sullivan, R. C. and Zhao, D.: Bringing the ocean into the laboratory to probe the chemical complexity of sea spray aerosol, *Proc. Natl. Acad. Sci.*, 110(19), 7550–7555, doi:10.1073/pnas.1300262110, 2013.

Reid, J. S., Hyer, E. J., Prins, E. M., Westphal, D. L., Zhang, J., Wang, J., Christopher, S. A., Curtis, C. A., Schmidt, C. C., Eleuterio, D. P., Richardson, K. A. and Hoffman, J. P.: Global Monitoring and Forecasting of Biomass-Burning Smoke: Description of and Lessons From the Fire Locating and Modeling of Burning Emissions (FLAMBE) Program, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 2(3), 144–162, doi:10.1109/JSTARS.2009.2027443, 2009.

Reid, J. S., Xian, P., Hyer, E. J., Flatau, M. K., Ramirez, E. M., Turk, F. J., Sampson, C. R., Zhang, C., Fukada, E. M. and Maloney, E. D.: Multi-scale meteorological conceptual analysis of observed active fire hotspot activity and smoke optical depth in the Maritime Continent, *Atmos Chem Phys*, 12(4), 2117–2147, doi:10.5194/acp-12-2117-2012, 2012.

Reid, J. S., Hyer, E. J., Johnson, R. S., Holben, B. N., Yokelson, R. J., Zhang, J., Campbell, J. R., Christopher, S. A., Di Girolamo, L., Giglio, L., Holz, R. E., Kearney, C., Miettinen, J., Reid, E. A., Turk, F. J., Wang, J., Xian, P., Zhao, G., Balasubramanian, R., Chew, B. N., Janjai, S., Lagrosas, N., Lestari, P., Lin, N.-H., Mahmud, M., Nguyen, A. X., Norris, B., Oanh, N. T. K., Oo, M., Salinas, S. V., Welton, E. J. and Liew, S. C.: Observing and understanding the Southeast Asian aerosol system by remote sensing: An initial review and analysis for the Seven Southeast Asian Studies (7SEAS) program, *Atmospheric Res.*, 122, 403–468, doi:10.1016/j.atmosres.2012.06.005, 2013.

Reid, J. S., Lagrosas, N. D., Jonsson, H. H., Reid, E. A., Sessions, W. R., Simpas, J. B., Uy, S. N., Boyd, T. J., Atwood, S. A., Blake, D. R., Campbell, J. R., Cliff, S. S., Holben, B. N., Holz, R. E., Hyer, E. J., Lynch, P., Meinardi, S., Posselt, D. J., Richardson, K. A., Salinas, S. V., Smirnov, A., Wang, Q., Yu, L. and Zhang, J.: Observations of the temporal variability in aerosol properties and their relationships to meteorology in the summer monsoonal South China Sea/East Sea: the scale-dependent role of monsoonal flows, the Madden–Julian Oscillation, tropical cyclones, squall lines and cold pools, *Atmos Chem Phys*, 15(4), 1745–1768, doi:10.5194/acp-15-1745-2015, 2015.

Reid, J. S., Lagrosas, N. D., Jonsson, H. H., Reid, E. A., Atwood, S. A., Boyd, T. J., Ghate, V. P., Lynch, P., Posselt, D. J., Simpas, J. B., Uy, S. N., Zaiger, K., Blake, D. R., Bucholtz, A., Campbell, J. R., Chew, B. N., Cliff, S. S., Holben, B. N., Holz, R. E., Hyer, E. J., Kreidenweis, S. M., Kuciaskas, A. P., Lolli, S., Oo, M., Perry, K. D., Salinas, S. V., Sessions, W. R., Smirnov, A., Walker, A. L., Wang, Q., Yu, L., Zhang, J. and Zhao, Y.: Aerosol meteorology and Philippine receptor observations of Maritime Continent aerosol emissions for the 2012 7SEAS southwest monsoon intensive study, *Atmospheric Chem. Phys. Discuss.*, 1–61, doi:10.5194/acp-2016-214, 2016.

Reutter, P., Su, H., Trentmann, J., Simmel, M., Rose, D., Gunthe, S. S., Wernli, H., Andreae, M. O. and Pöschl, U.: Aerosol- and updraft-limited regimes of cloud droplet formation: influence of particle number, size and hygroscopicity on the activation of cloud condensation nuclei (CCN), *Atmos Chem Phys*, 9(18), 7067–7080, doi:10.5194/acp-9-7067-2009, 2009.

Robinson, N. H., Newton, H. M., Allan, J. D., Irwin, M., Hamilton, J. F., Flynn, M., Bower, K. N., Williams, P. I., Mills, G., Reeves, C. E., McFiggans, G. and Coe, H.: Source attribution of Bornean air masses by back trajectory analysis during the OP3 project, *Atmos Chem Phys*, 11(18), 9605–9630, doi:10.5194/acp-11-9605-2011, 2011.

Robinson, N. H., Allan, J. D., Trembath, J. A., Rosenberg, P. D., Allen, G. and Coe, H.: The lofting of Western Pacific regional aerosol by island thermodynamics as observed around Borneo, *Atmos Chem Phys*, 12(13), 5963–5983, doi:10.5194/acp-12-5963-2012, 2012.

Rose, D., Nowak, A., Achtert, P., Wiedensohler, A., Hu, M., Shao, M., Zhang, Y., Andreae, M. O. and Pöschl, U.: Cloud condensation nuclei in polluted air and biomass burning smoke near the mega-city Guangzhou, China – Part 1: Size-resolved measurements and implications for the modeling of aerosol particle hygroscopicity and CCN activity, *Atmos Chem Phys*, 10(7), 3365–3383, doi:10.5194/acp-10-3365-2010, 2010.

- Rosenfeld, D.: TRMM observed first direct evidence of smoke from forest fires inhibiting rainfall, *Geophys. Res. Lett.*, 26(20), 3105–3108, doi:10.1029/1999GL006066, 1999.
- Sakamoto, K. M., Allan, J. D., Coe, H., Taylor, J. W., Duck, T. J. and Pierce, J. R.: Aged boreal biomass-burning aerosol size distributions from BORTAS 2011, *Atmos Chem Phys*, 15(4), 1633–1646, doi:10.5194/acp-15-1633-2015, 2015.
- 5 Shank, L. M., Howell, S., Clarke, A. D., Freitag, S., Brekhovskikh, V., Kapustin, V., McNaughton, C., Campos, T. and Wood, R.: Organic matter and non-refractory aerosol over the remote Southeast Pacific: oceanic and combustion sources, *Atmos Chem Phys*, 12(1), 557–576, doi:10.5194/acp-12-557-2012, 2012.
- Spracklen, D. V., Pringle, K. J., Carslaw, K. S., Mann, G. W., Manktelow, P. and Heintzenberg, J.: Evaluation of a global aerosol microphysics model against size-resolved particle statistics in the marine atmosphere, *Atmos Chem Phys*, 7(8), 2073–2090, doi:10.5194/acp-7-2073-2007, 2007.
- 10 Tao, W.-K., Chen, J.-P., Li, Z., Wang, C. and Zhang, C.: Impact of aerosols on convective clouds and precipitation, *Rev. Geophys.*, 50(2), RG2001, doi:10.1029/2011RG000369, 2012.
- Tosca, M. G., Randerson, J. T., Zender, C. S., Nelson, D. L., Diner, D. J. and Logan, J. A.: Dynamics of fire plumes and smoke clouds associated with peat and deforestation fires in Indonesia, *J. Geophys. Res. Atmospheres*, 116(D8), D08207, doi:10.1029/2010JD015148, 2011.
- 15 Wang, J., Ge, C., Yang, Z., Hyer, E. J., Reid, J. S., Chew, B.-N., Mahmud, M., Zhang, Y. and Zhang, M.: Mesoscale modeling of smoke transport over the Southeast Asian Maritime Continent: Interplay of sea breeze, trade wind, typhoon, and topography, *Atmospheric Res.*, 122, 486–503, doi:10.1016/j.atmosres.2012.05.009, 2013.
- Ward, D. S., Eidhammer, T., Cotton, W. R. and Kreidenweis, S. M.: The role of the particle size distribution in assessing aerosol composition effects on simulated droplet activation, *Atmos Chem Phys*, 10(12), 5435–5447, doi:10.5194/acp-10-5435-2010, 2010.
- 20 Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, Academic Press., 2011.
- Xian, P., Reid, J. S., Atwood, S. A., Johnson, R. S., Hyer, E. J., Westphal, D. L. and Sessions, W.: Smoke aerosol transport patterns over the Maritime Continent, *Atmospheric Res.*, 122, 469–485, doi:10.1016/j.atmosres.2012.05.006, 2013.
- 25 Yokelson, R. J., Christian, T. J., Karl, T. G. and Guenther, A.: The tropical forest and fire emissions experiment: laboratory fire measurements and synthesis of campaign data, *Atmos Chem Phys*, 8(13), 3509–3527, doi:10.5194/acp-8-3509-2008, 2008.
- Yu, F., Wang, Z., Luo, G. and Turco, R.: Ion-mediated nucleation as an important global source of tropospheric aerosols, *Atmos Chem Phys*, 8(9), 2537–2554, doi:10.5194/acp-8-2537-2008, 2008.
- 30 Yuan, T., Remer, L. A., Pickering, K. E. and Yu, H.: Observational evidence of aerosol enhancement of lightning activity and convective invigoration, *Geophys. Res. Lett.*, 38(4), L04701, doi:10.1029/2010GL046052, 2011.
- Zender, C. S., Krolewski, A. G., Tosca, M. G. and Randerson, J. T.: Tropical biomass burning smoke plume size, shape, reflectance, and age based on 2001–2009 MISR imagery of Borneo, *Atmos Chem Phys*, 12(7), 3437–3454, doi:10.5194/acp-12-3437-2012, 2012.

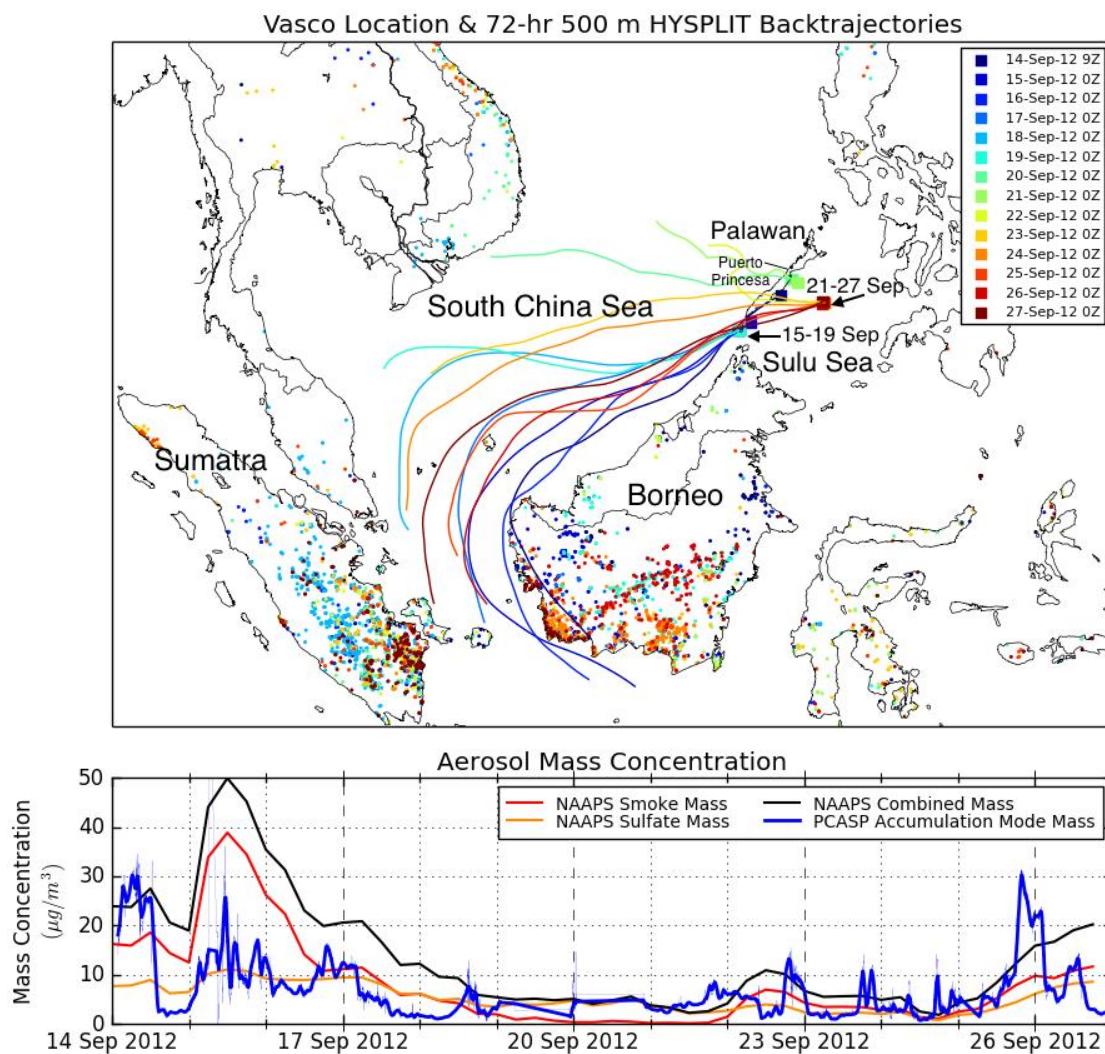


Figure 1: (a) *Vasco* cruise locations (squares) and 72-hour, 500 m HYSPLIT backtrajectories; MODIS fire detections (dots) from Terra and Aqua are included for each day (color coded) during the sampling period. (b) PCASP reconstructed accumulation mode (125nm – 800nm) mass concentration (assumed density $1.4 \mu\text{g m}^{-3}$) and NAAPS estimated smoke and sulfate mass concentration along the *Vasco* ship-track.

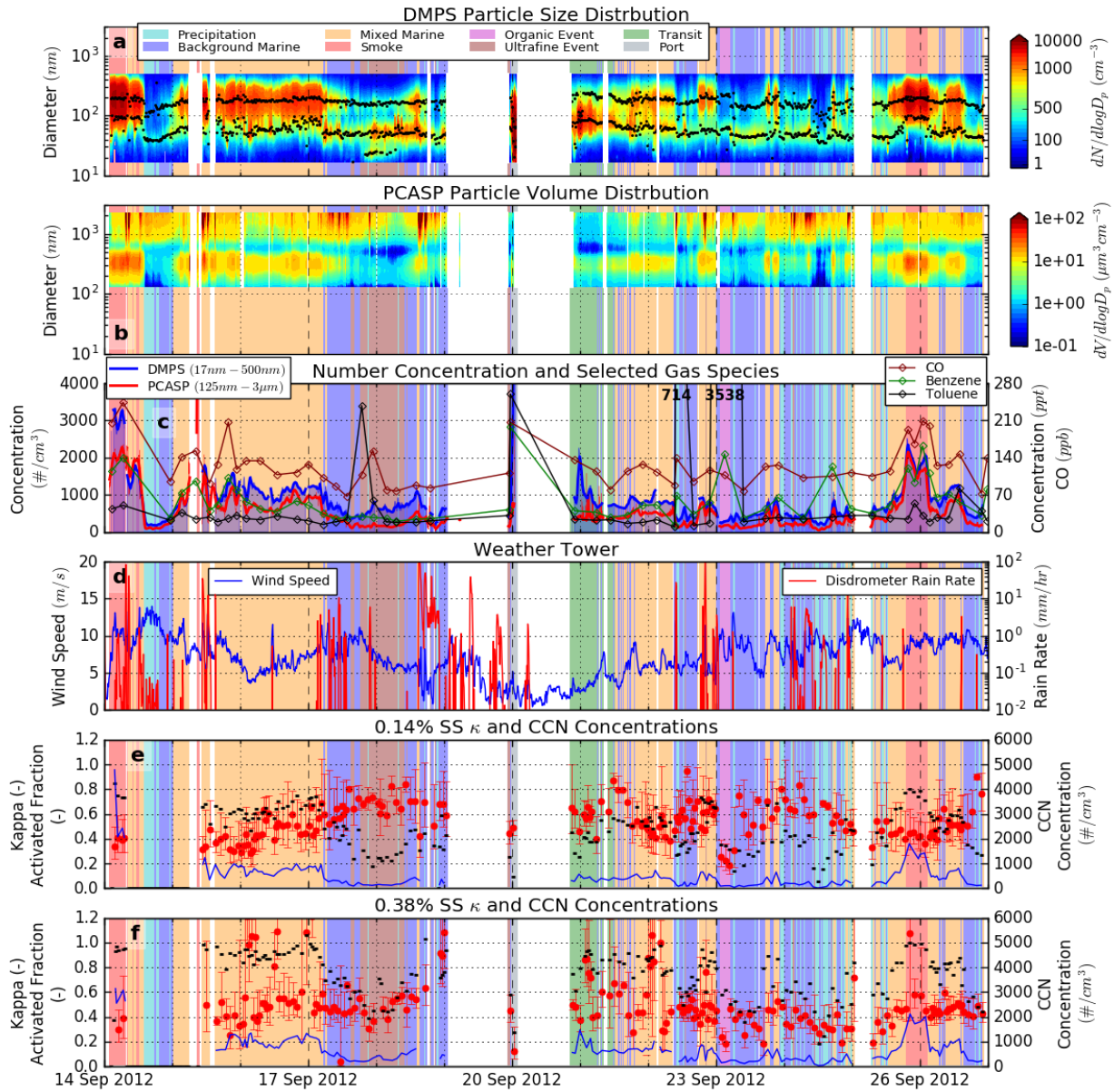


Figure 2: Timelines of measured and derived variables during Vasco 2012 cruise. In all figures, background colors correspond to aerosol type classification from the cluster analysis, as indicated in the legend in panel a. (a) $dN/d\log D_p$ spectra from the CCN system measurements with black dots at best-fit modal median diameters; (b) $dV/d\log D_p$ spectra derived from the PCASP measurements; (c) total number concentrations measured by the CCN system (blue; shaded below for contrast) and the PCASP (red), with 60 minute boxcar average smoothing; gas canister grab sample concentrations for carbon monoxide, benzene, and toluene are shown on the right axis with colored numbers indicating points above the upper scale extent; (d) wind speed and disdrometer rain rate from the Vasco weather tower. (e & f) κ parameter (red) and CCN concentrations (blue) for 0.14% and 0.38% supersaturation settings (corresponding approximately to accumulation and Aitken modes, respectively), with total activated particle number fractions ($CCN_{SS\%} / CN_{Total}$) bars in grey. Error bars on κ data points indicate the κ values associated with 25% / 75% activated fraction curve fits.

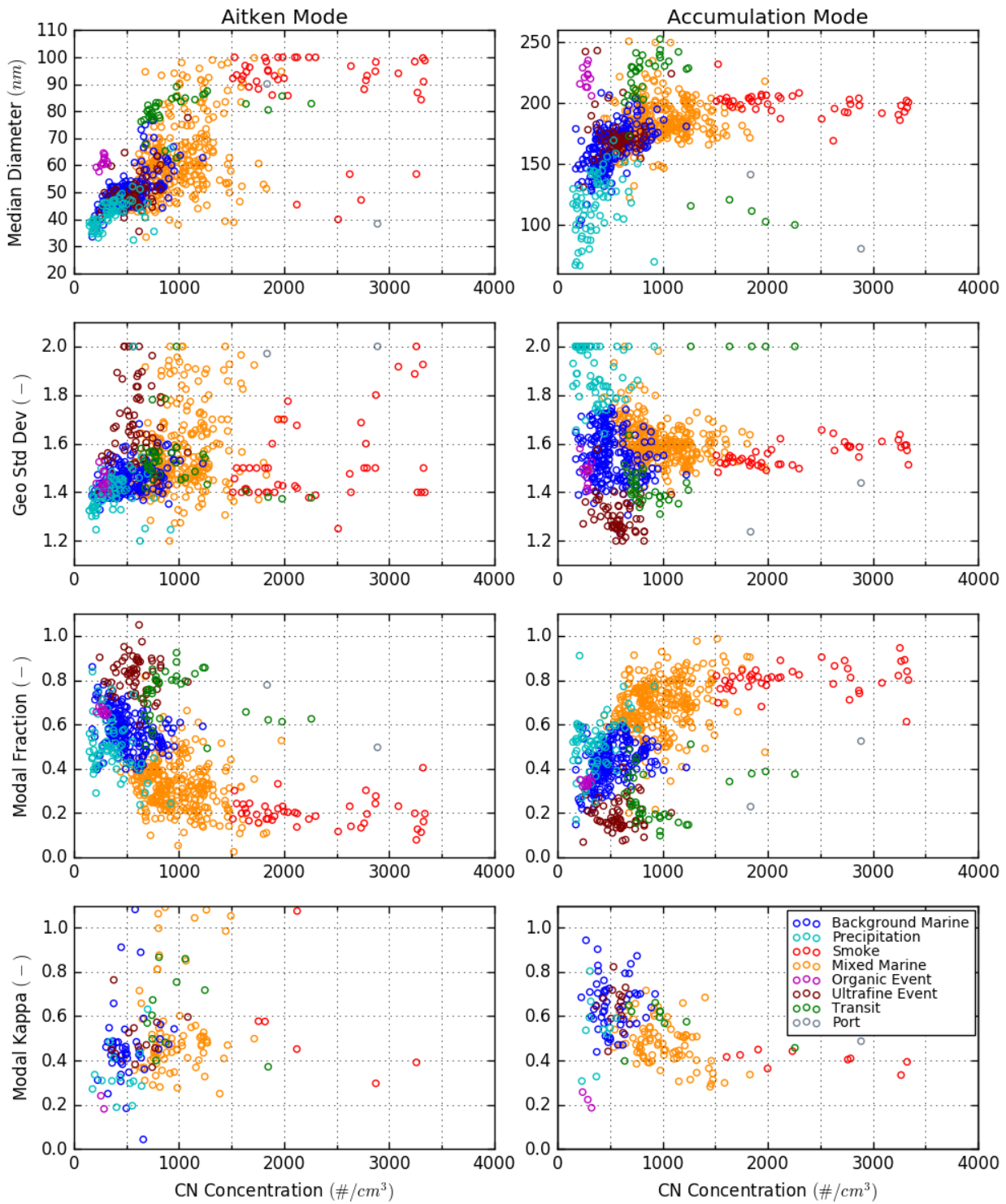


Figure 3: Parameterized variable values for Aitken and accumulation modes (median diameter, geometric standard deviation, modal fraction) at each of the 15 minute data points during the study, along with κ values for data points at CCNc superstation set points of 0.38% (Aitken mode) and 0.14% (accumulation mode). Each data point is colored according to the cluster type to which it was classified.

Normalized dN/dlogDp

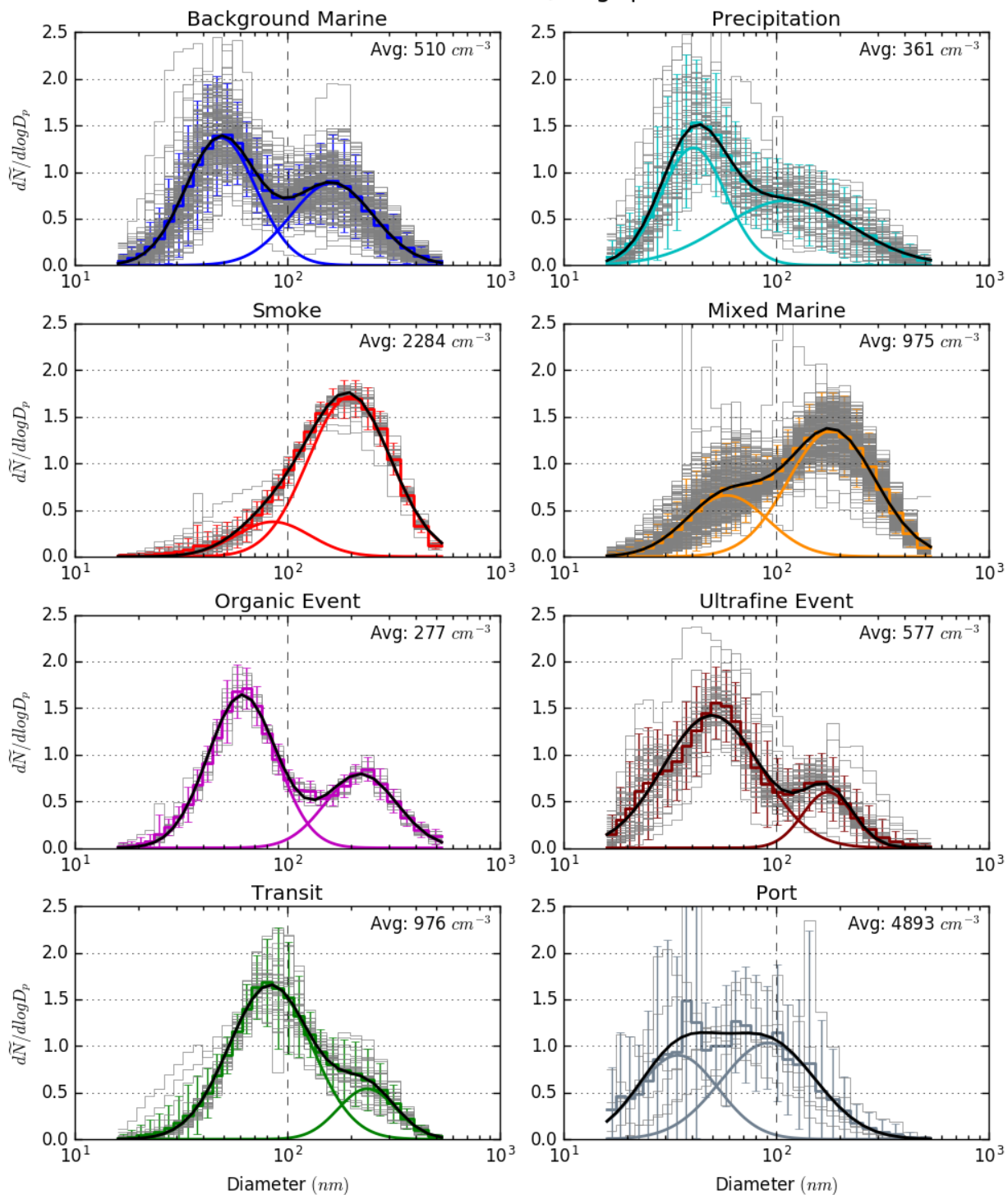


Figure 4: Normalized $dN/d\log D_p$ particle size distributions for each [spectra-spectrum](#) within identified aerosol population types (grey), the associated average bin values with error bars at 95% confidence interval (colored step lines; $1.96 \times$ bin standard deviation), and best fit lognormal modes (colored curves) with bimodal fit (black). The average particle number concentrations of data points within each population type are listed.

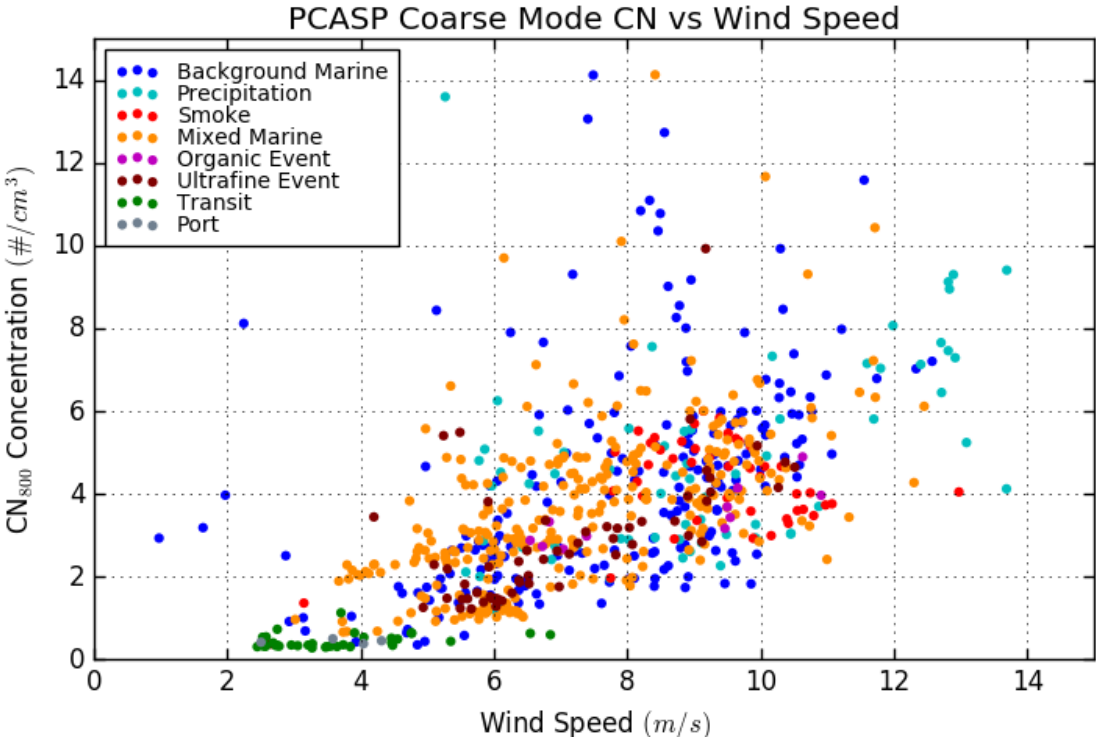


Figure 5: Number concentrations of coarse mode particles ($D_p > 800$ nm) measured by the PCASP as functions of local surface wind speed measured by the onboard *Vasco* weather station. Each point was averaged over the same approximately 15 minute time period as for the CCN system measurements, and is colored by the aerosol type as described in the text.

Table 1: Aerosol population type parameters used for clustering and the resulting average values (standard deviations in grey parentheses) for each identified population.

Population Type	Total Number	Aitken Mode				Accumulation Mode			
		Concentration (# cm ⁻³)	Median (nm)	Geometric Std Dev	Number Fraction	Kappa	Median (nm)	Geometric Std Dev	Number Fraction
1: Back. Marine	510 (181)	50 (7)	1.45 (0.05)	0.57 (0.08)	0.46 (0.17)	162 (18)	1.55 (0.10)	0.42 (0.09)	0.65 (0.11)
2: Precipitation	361 (164)	42 (5)	1.40 (0.10)	0.50 (0.12)	0.34 (0.11)	115 (27)	1.91 (0.20)	0.51 (0.12)	0.54 (0.14)
3: Smoke	2280 (606)	89 (15)	1.53 (0.17)	0.20 (0.06)	0.56 (0.25)	199 (9)	1.55 (0.04)	0.81 (0.06)	0.40 (0.03)
4: Mixed Marine	975 (271)	62 (13)	1.54 (0.18)	0.32 (0.13)	0.54 (0.23)	184 (17)	1.61 (0.08)	0.68 (0.12)	0.48 (0.10)
5: Organic Event	277 (30)	61 (2)	1.45 (0.04)	0.66 (0.01)	0.21 (0.03)	221 (8)	1.48 (0.05)	0.34 (0.01)	0.22 (0.03)
6: Ultrafine Event	577 (158)	50 (6)	1.69 (0.16)	0.82 (0.08)	0.50 (0.10)	174 (19)	1.30 (0.07)	0.19 (0.06)	0.65 (0.09)
7: Transit	976 (384)	80 (6)	1.53 (0.12)	0.73 (0.11)	0.62 (0.16)	209 (42)	1.50 (0.21)	0.26 (0.11)	0.58 (0.08)
8: Port	4890 (2550)	42 (22)	1.62 (0.37)	0.49 (0.18)	0.13 (-)	87 (26)	1.57 (0.22)	0.53 (0.18)	0.49 (-)

Table 21: Average values (standard deviations in grey parentheses) for identified aerosol population types. Shown are number of CCN system data points classified as each type, total number concentrations for the PCASP (125 nm–3 μm) and CCN system (17–500 nm), CCN number concentrations and activated fractions for each CCNc supersaturation set point, and measured κ values for the accumulation mode (0.14% SS) and Aitken mode (0.38% SS) set points.

Population Type	# CCN Meas. (#)	PCASP Number (#/cm ³)	CCN system Number (#/cm ³)	0.14% SS		κ (-)	0.38% SS		κ (-)	0.53% SS		0.71% SS		0.85% SS	
				CCN (#/cm ³)	Act Frac (-)		CCN (#/cm ³)	Act Frac (-)		CCN (#/cm ³)	Act Frac (-)	CCN (#/cm ³)	Act Frac (-)	CCN (#/cm ³)	Act Frac (-)
1: Back. Marine	214	231 (111)	510 (181)	213 (101)	0.38 (0.09)	0.65 (0.11)	320 (148)	0.60 (0.12)	0.46 (0.17)	416 (194)	0.74 (0.11)	444 (239)	0.81 (0.09)	480 (210)	0.87 (0.05)
2: Precipitation	67	142 (79)	361 (164)	96 (58)	0.24 (0.11)	0.54 (0.14)	243 (135)	0.48 (0.15)	0.34 (0.11)	352 (175)	0.65 (0.15)	265 (82)	0.71 (0.09)	228 (100)	0.79 (0.03)
3: Smoke	44	1800 (273)	2280 (606)	1720 (388)	0.72 (0.04)	0.40 (0.03)	2340 (480)	0.93 (0.02)	0.56 (0.25)	1990 (359)	0.97 (0.02)	2080 (396)	0.98 (0.05)	2150 (523)	0.99 (0.02)
4: Mixed Marine	294	689 (295)	975 (271)	591 (201)	0.58 (0.08)	0.48 (0.10)	827 (270)	0.83 (0.07)	0.54 (0.23)	861 (247)	0.89 (0.10)	876 (244)	0.94 (0.06)	893 (271)	0.96 (0.05)
5: Organic Event	11	151 (19)	277 (30)	88 (10)	0.31 (0.02)	0.22 (0.03)	144 (9)	0.53 (0.01)	0.21 (0.03)	182 (26)	0.72 (0.01)	268 (56)	0.89 (0.14)	257 (24)	0.93 (0.07)
6: Ultrafine Event	59	163 (58)	577 (158)	138 (45)	0.25 (0.06)	0.65 (0.09)	361 (172)	0.56 (0.11)	0.50 (0.10)	373 (168)	0.65 (0.12)	439 (163)	0.72 (0.07)	473 (147)	0.79 (0.10)
7: Transit	36	311 (44)	976 (384)	363 (87)	0.37 (0.06)	0.58 (0.08)	772 (263)	0.81 (0.09)	0.62 (0.16)	832 (423)	0.87 (0.05)	877 (370)	0.90 (0.02)	878 (195)	0.95 (0.03)
8: Port	6	671 (210)	4890 (2550)	251 (-)*	0.09 (-)*	0.49 (-)*	1126 (-)*	0.26 (-)*	0.13 (-)*	3936 (-)*	0.40 (-)*	1742 (289)	0.57 (0.22)	2080 (-)*	0.45 (-)*
All Types	731	503 (455)	851 (677)	450 (388)	0.47 (0.16)	0.54 (0.14)	675 (516)	0.72 (0.17)	0.50 (0.21)	698 (555)	0.79 (0.15)	724 (512)	0.85 (0.13)	723 (502)	0.90 (0.10)

* Only one datapoint; Note that Port measurements fluctuated as the Vasco entered port