

Supplement of Atmos. Chem. Phys., 19, 6931–6947, 2019  
<https://doi.org/10.5194/acp-19-6931-2019-supplement>  
© Author(s) 2019. This work is distributed under  
the Creative Commons Attribution 4.0 License.



*Supplement of*

## **Classification of aerosol population type and cloud condensation nuclei properties in a coastal California littoral environment using an unsupervised cluster model**

**Samuel A. Atwood et al.**

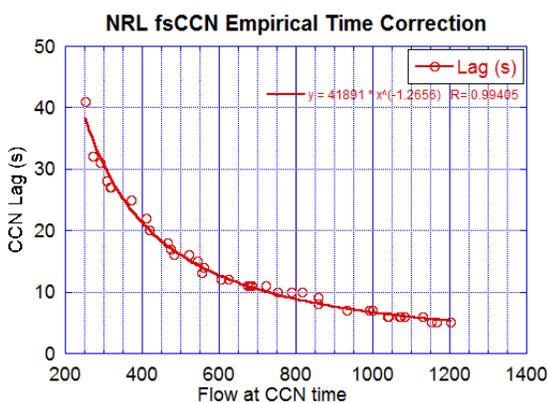
*Correspondence to:* Sonia M. Kreidenweis ([soniak@colostate.edu](mailto:soniak@colostate.edu))

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

## 1 Scanning Flow CCN (sfCCN) System Calibration

2 Calibration of the scanning flow CCN column supersaturation for a given flow rate at each temperature gradient setting  
3 was conducted following the methodology described in Suda et al., (2012). Briefly, a monodisperse ammonium sulfate  
4 aerosol was generated by atomizing an ammonium sulfate solution and size selecting using a TSI 3080 Differential  
5 Mobility Analyzer (DMA) set to a fixed voltage. Concentrations of CN and CCN were measured as flow rate changed.  
6 The flow rate at which 50% of particles had activated ( $Q_{50}$ ) was calculated by fitting a sigmoid curve to the activation  
7 plots for a range of DMA column voltages. Given a  $\kappa$  hygroscopicity parameter of 0.61 (Petters and Kreidenweis,  
8 2007) for the ammonium sulfate particles and the parameters of the DMA, calibration curves for column  
9 supersaturation as a function of flow rate and temperature gradient were calculated. Calibrations were conducted for  
10 column temperature gradients of 4 K and 8 K (referred to as dT4 and dT8, respectively) that were used in these  
11 measurements.

12 As residence time in the sfCCN varied with total flow rate, timestamps were empirically adjusted to align the data  
13 with the parallel CPC. The calibrated CPC timestamp adjustment was made as a function of flow by measuring the  
14 time delay between narrow peaks of particles introduced at specific flow rates. A filter was placed upstream of the  
15 split between instruments and then briefly removed, thereby allowing ambient particles to be sampled by both  
16 instruments. The time lag between measurements was then determined for a number of flow rates for both  
17 measurement and calibration samples. A calibration curve was fit to the aligned time stamps (Fig. S1) and used to  
18 adjust the timestamp of the CPC data points for both ambient and calibration scans.



Supplemental Figure 1 Equivalent CPC time correction function for the sfCCN system at various flow rates (lpm).

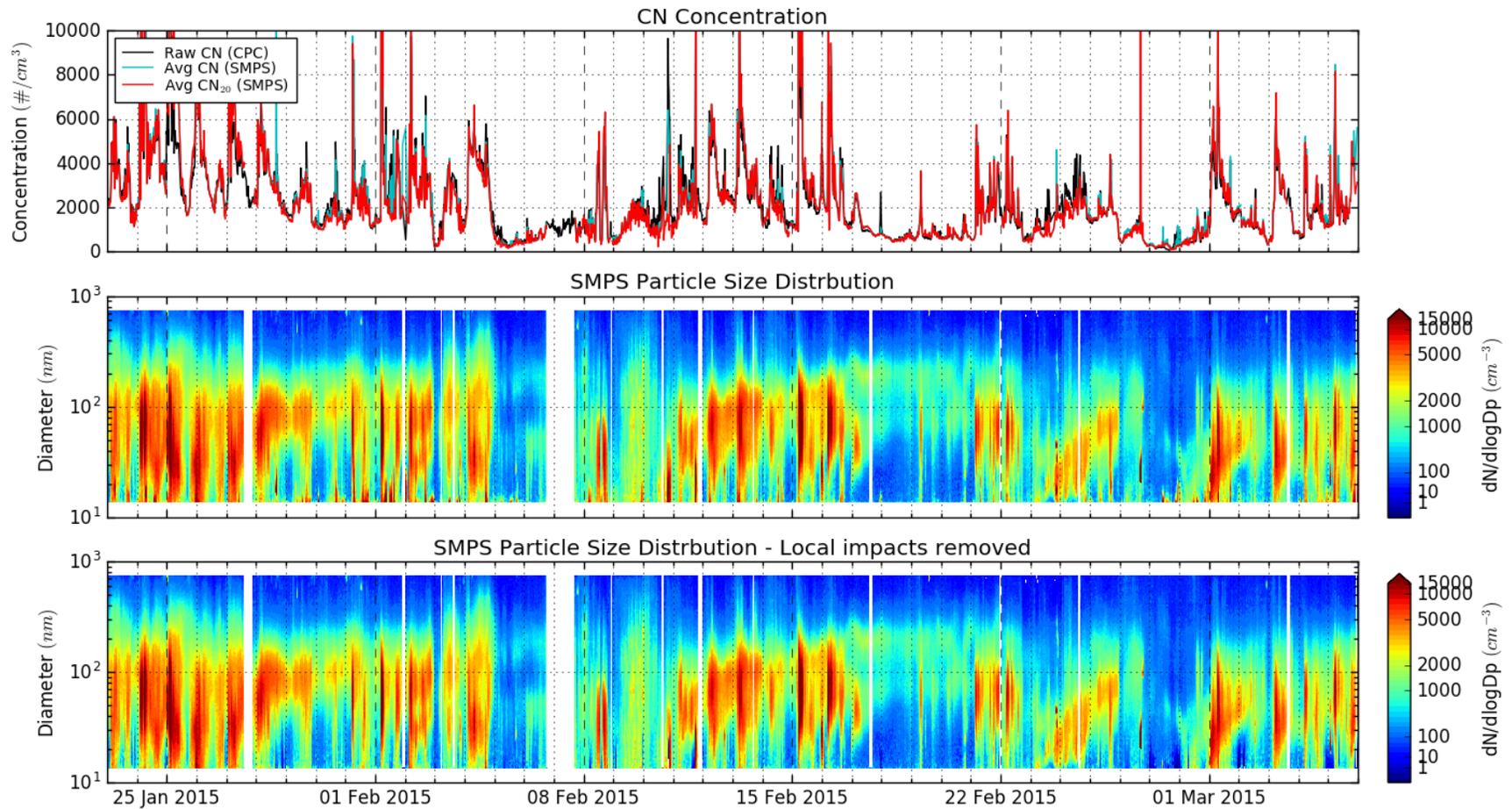
19

20

21 **Aerosol Size Distribution Correction**

22 A raw timeline of particle number concentration and size distribution is shown in Supplemental Figure 2(a&b). The  
23 corrected number concentrations and size distributions after removal of the contamination modes are shown in (a&c).  
24 Both the generally short duration and small particle size of the identified contamination modes are evident in these  
25 figures. Narrow peaks in the light blue line in (a) show that the correction to number concentration occurred over  
26 generally short time periods, while the predominant differences in the corrected size distribution of (c) were for the  
27 smallest particle sizes. Relatively little change in the overall nature of the shape of the overall distributions resulted  
28 from this correction, but it allowed for markedly improved performance of the cluster model.  
29 Further criteria for identification of contamination periods and modes in the number size distribution were considered,  
30 including analysis of signal-to-noise or similar measures of relative smoothness of potential contamination modes.  
31 However, such metrics were not always consistent in their identification of such small contamination modes due to  
32 competing factors from measurement noise and physical variability. Ultimately, the criteria listed in the main text  
33 were found to be sufficiently capable of removing small contamination modes such that the cluster model was able  
34 utilize these data points without skewing the results, and so were used for the purposes of this analysis.

35



36

37 **Supplemental Figure 2 (a) Study timeline with CN concentrations from the bulk CPC (black), reconstructed CN from the SMPS (light blue), and reconstructed CN**  
 38 **corrected for intermittent local contamination modes (red). SMPS particle size distributions are shown (b) before and (c) after correction to remove contamination.**

39

## 40 **Activated Fraction Spectra Fitting**

41 Parameterizations of activated fraction spectra have been presented using various methodologies. The  
42 parameterizations of Paramonov et al., (2015) for various locations in the EUCAARI network, shown in Figure 6 in  
43 the main text, used a two parameter fit of the form  $AF = a \times \ln(S_{eff}) + b$ , where  $S_{eff}$  is the effective environmental  
44 supersaturation. However, this parameterization did not adequately reflect the differences between spectra in the BML  
45 dataset. In order to provide a more accurate parameterization of observed activation spectra, a lognormal cumulative  
46 mixture distribution function was used of the form,

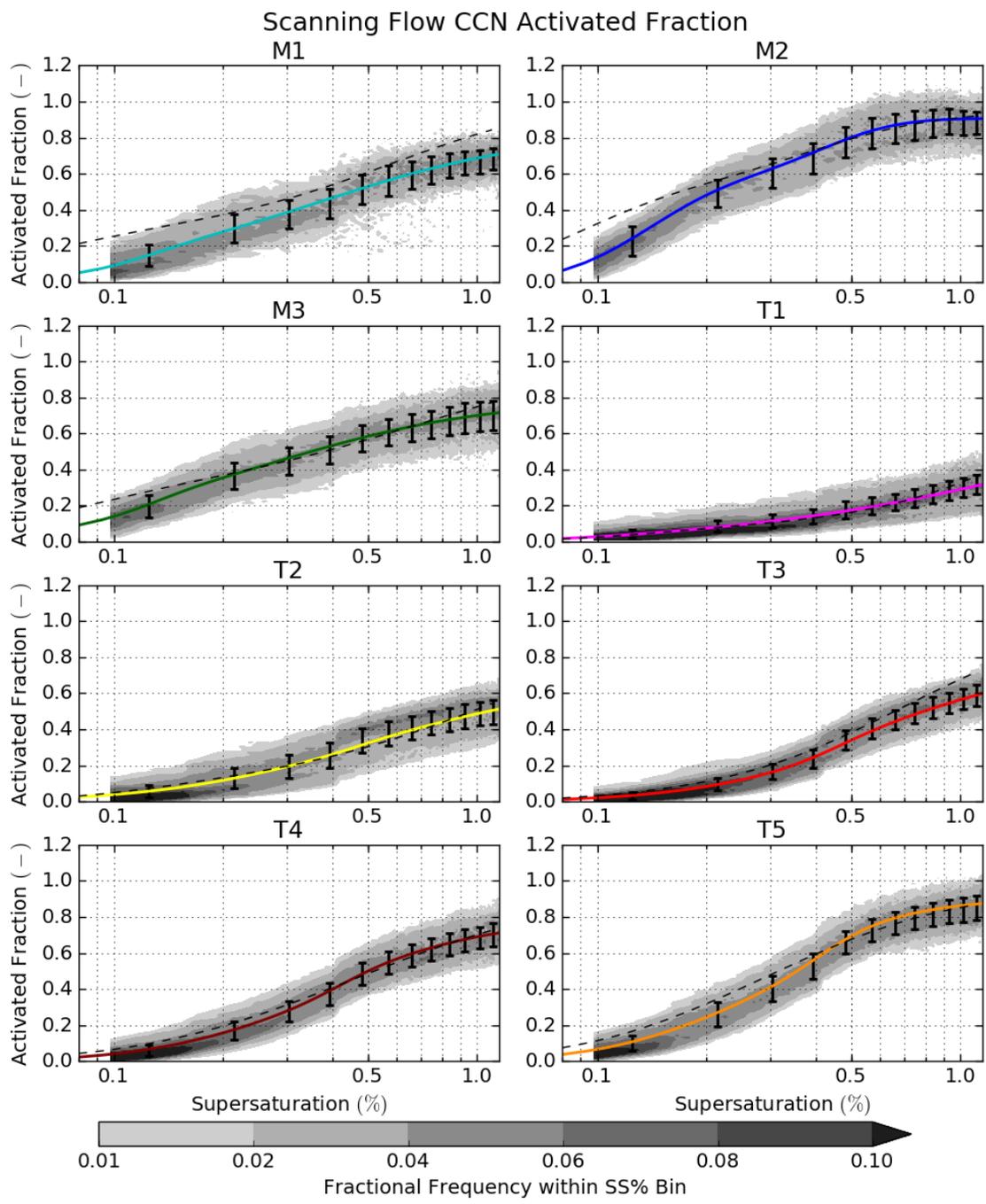
$$47 \quad AF(ss)_{pop} = \sum_i^{FN} \frac{c_i}{2} + \frac{c_i}{2} \operatorname{erf} \left( \frac{\ln(ss/a_i)}{\sqrt{2} \ln b_i} \right)$$

48 where  $AF$  is the parameterized activated fraction for the population,  $ss$  is the environmental supersaturation expressed  
49 as a percentage ( $ss = (S - 1) * 100\%$ ),  $FN$  is the total number of fit modes for each mode  $i$ , and  $a$ ,  $b$ , and  $c$  are the  
50 fit parameters for each lognormal mode. As most cluster average size distributions were bimodal with only a small  
51 fraction of the total number in an additional third identified mode, a bimodal activation spectrum parameterization  
52 was used for all population types here. The associated parameters for each of the activation spectra shown in Figure 6  
53 are given in Table 1.

## 54 **Additional Cluster Figures**

55 Similar to presentation of cluster measurements for size distribution and meteorological data, the activated fraction  
56 spectra for each cluster are shown in Fig. S3. The best-fit distributions are shown as colored lines with activated  
57 fraction data point density contours shown in greyscale around them.

58



59

**Supplemental Figure 3** Activated fraction spectra for each cluster from the 8 cluster K-means analysis. Best-fit bimodal lognormal cumulative distribution functions are shown in colors for each cluster, with black error bars plotted at  $\pm 1$  standard deviation of the mean for data split into 10 linear-spaced supersaturation bins with roughly equal numbers of data points. Grey contours show relative density of valid cluster data points in each supersaturation bin. Dashed black lines show the reconstructed activated fraction spectra based on average measured size distribution parameters and median measured kappa hygroscopicity parameter.

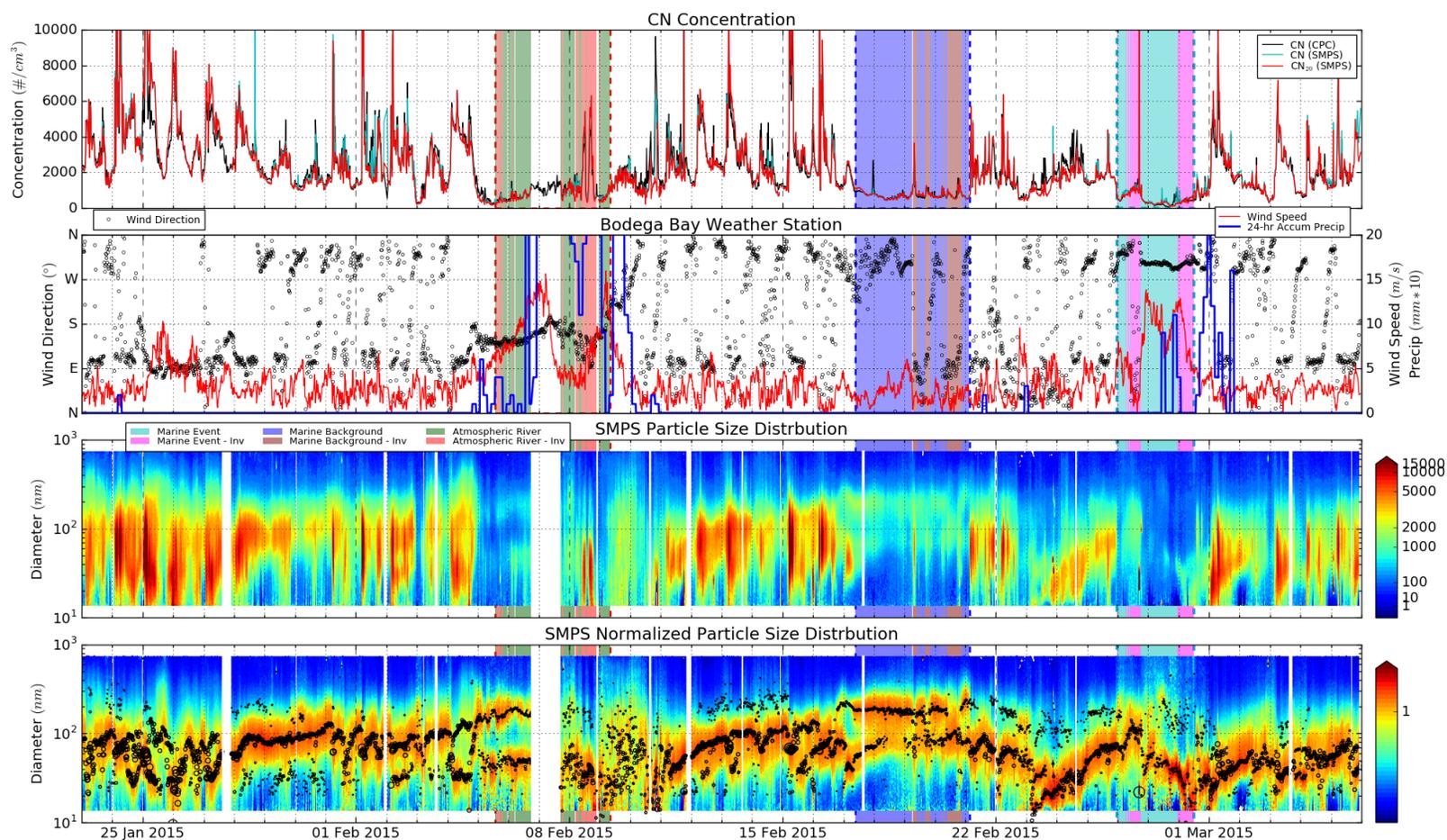
60

## 61 **Validation of clustering methodology during marine aerosol dominated periods**

62 In order to validate that the ability of the clustering model to improve characterization of distinct aerosol population  
63 types, further analysis of the cluster results was performed for three multi-day time periods that were associated with  
64 marine aerosol dominated air masses impacting BML. As noted in the main text, these marine dominated periods were  
65 the precipitation scrubbed clean marine event (25 Feb 23:00 UTC – 28 Feb 12:00 UTC), the cloud-processed  
66 background marine event that involved ventilation of BML from oceanic regions (17 Feb 09:00 UTC – 21 Feb 03:00  
67 UTC), and the atmospheric river (5 Feb 13:00 UTC – 9 Feb 08:00 UTC). These periods are identified in Fig. S4 for  
68 the periods associated with the cluster type associated with the marine event, along with periods of time classified as  
69 other cluster types.

70 Normalized size distributions for all data from each of the event time periods were generated and are shown in the left  
71 column of Fig. S5, all showing typical bimodal size distributions expected of cloud processed marine aerosol  
72 populations. The data points for these time periods were then segregated into only those corresponding to the marine  
73 cluster identified with each event (center column), and all data points from other cluster types that were excluded from  
74 that cluster (right column). In all three cases, noise in the range of distributions associated only with the cluster periods  
75 was reduced and average particle number concentration decreased. The data excluded from the marine clusters  
76 included some distributions that were largely monomodal and similar to the aged terrestrial distributions discussed in  
77 the previous section. Local wind vectors shown in Fig. S6 showed that the marine cluster data points were more  
78 generally associated with northwesterly winds from ocean regions, while the excluded data points occurred during  
79 periods of more easterly winds from inland regions.

80 The consequences to predicted CCN characteristics of removing non-marine data points are obvious in the resulting  
81 activated fraction spectra shown in Fig. S7. For the marine event associated with cluster M1 (top row), averaging all  
82 time points (left column) resulted in two distinct activated fraction curves, with the lower curve resulting from the  
83 inclusion of a period with small particles that had low activated fractions. Excluded data (right column) removed not  
84 only those data points, but also screened time periods with higher activated fractions that belonged to a different  
85 population type from M1. The resulting activated fraction spectrum associated with this event for only cluster periods  
86 (center column) had considerably smaller standard deviations in the result. Similar results were seen in the analysis of  
87 the M3 cluster event, with lower activated fractions associated with the terrestrial or anthropogenic contamination  
88 periods being effectively removed in the clustering activated fraction spectrum. While the M2 marine background  
89 event did not see as large of a change between time-averaging and cluster-averaging, the standard deviations were  
90 smaller in the clustering results, indicating increased precision in the activated fraction parameterization. Additionally,  
91 the excluded data points were themselves generally consistent and had low standard deviations, indicating that that a  
92 single separate population with distinct microphysical properties was likely intermittently observed during this event.  
93 As a result, it was concluded that the clustering methodology using both properties of the aerosol and meteorological  
94 values was able to successfully segregate time periods associated with various aerosol population impacts.  
95 Additionally, it was able to provide improved identification of aerosol population types with less noise in the resulting  
96 parameterizations.

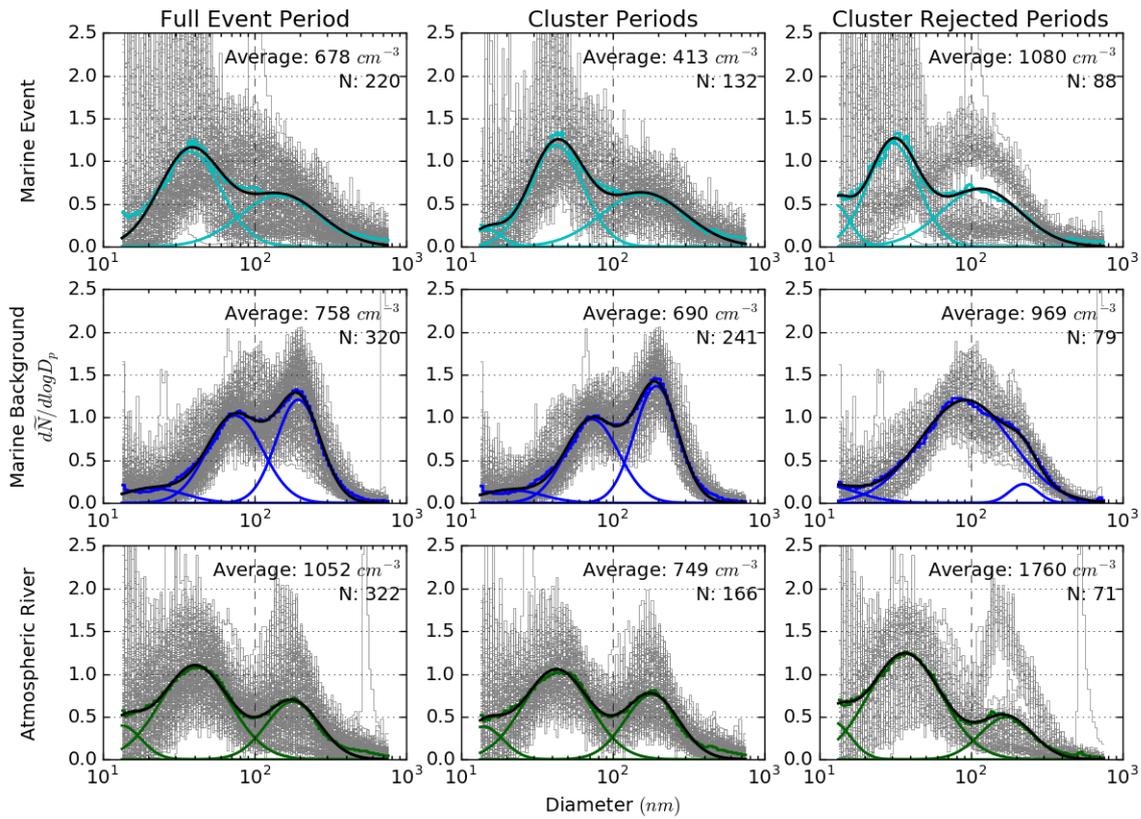


97

**Supplemental Figure 4** Timeline comparison for three marine aerosol events, as identified by the reference classification (periods within dotted lines), and the K-means cluster analysis. Atmospheric river (green), marine background (dark blue), and precipitation impacted clean marine (light blue) events are shown. The invalid periods from the cluster classification are shown in red, brown, and pink, respectively.

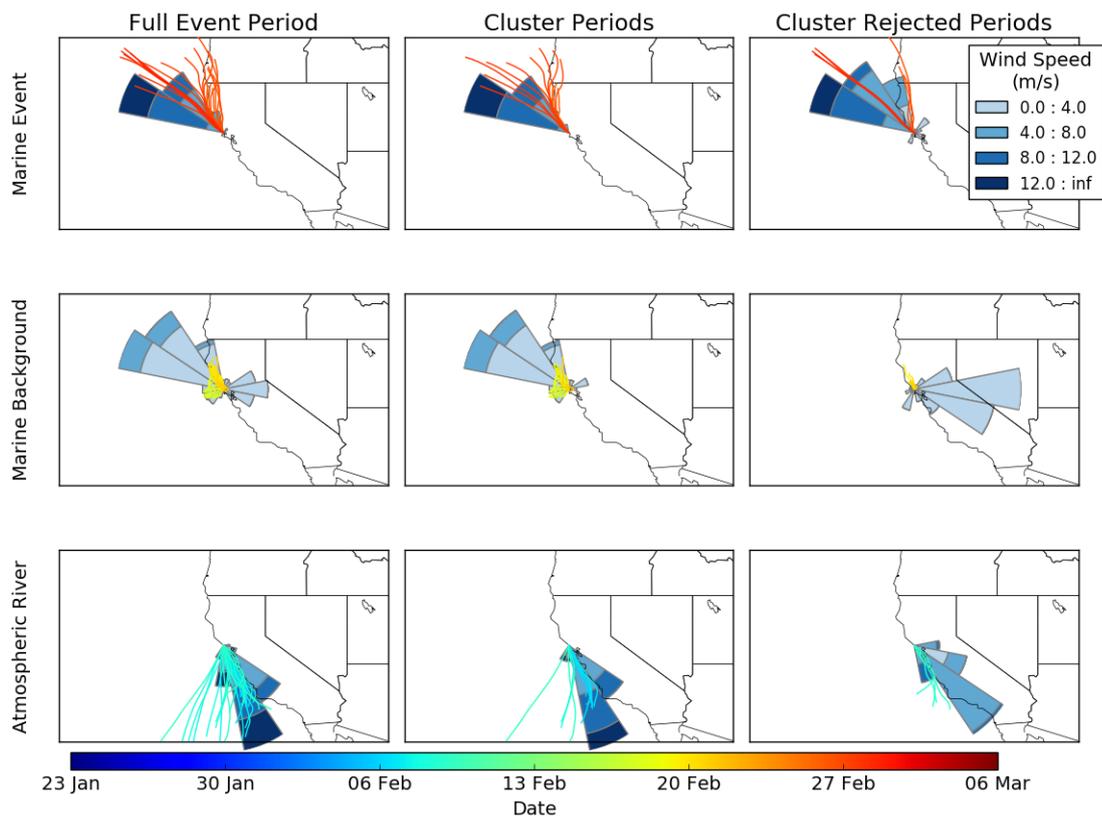
98

## SMPS Normalized $dN/d\log D_p$ - Marine Aerosol Events



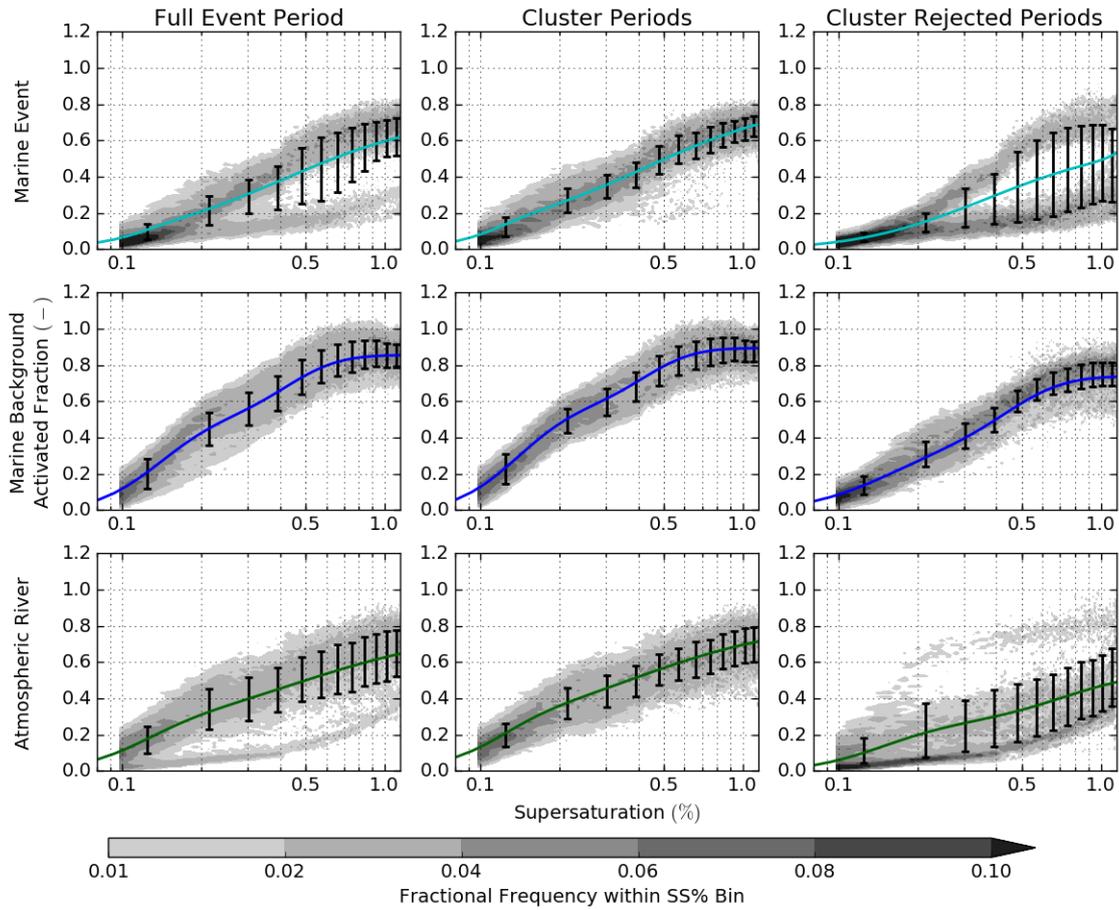
**Supplemental Figure 5** Normalized size distribution comparison for the three marine aerosol events (rows), averaged by the entire event time period (left column), only those times associated with the marine cluster associated with the event (center column), and those times that were excluded from the marine cluster (right column).

Wind Vector and 100m HYSPLIT backtrajectories - Marine Aerosol Events



Supplemental Figure 6 Meteorological comparison for the three marine aerosol events as in Fig. S5.

## Scanning Flow CCN Activated Fraction - Marine Aerosol Events



**Supplemental Figure 7** Activated fraction spectra comparison for the three marine aerosol events as in Fig. S5.