Geophysical Research Letters

Supporting Information for

Aerosol influences on cloud water: Insights from ARM EPCAPE observations with explainable machine learning

Haipeng Zhang[1,2,3], Yunyan Zhang[1], Zhanqing Li[2,3], and Youtong Zheng[4,5]

[1]Lawrence Livermore National Laboratory, Livermore, CA, USA

[2]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA

[3]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA

[4]Department of Earth and Atmospheric Science, University of Houston, Houston, TX, USA

[5]Institute of Climate and Atmospheric Science, University of Houston, Houston, TX, USA

Contents of this file

Text S1:

The *dq* (the most influential meteorological variable as seen in Figure 1) and LTS demonstrate non-linear relationships in EPCAPE. LWP initially increases with *dq* and then declines. To understand this relationship, we first need to note that on hourly time scales *dq* is determined by PBL moisture (q_{1000}) rather than FT moisture (q_{700}) (Figure S5a vs. Figure S5b). Therefore, the positive *dq*-LWP relationship when *dq* is relatively small (<0.007 kg/kg) is mainly due to the fact that increased PBL moisture favors larger LWP. However, once *dq* surpasses this threshold, larger *dq* suggests enhanced entrainment drying that outweighs moistening effects due to increased moisture sources, thereby depleting LWP. Another important meteorological factor, LTS, also exhibits a non-linear relationship with LWP, which increases with LTS first but declines at higher LTS values. The positive relationship due to more moisture be capped in the PBL with stronger cloud-top inversion (Bretherton et al., 2013), but the negative relationship at higher LTS values warrants future investigation. These non-linear relationships highlight the greater complexity of meteorological controls on LWP on hourly time scales compared to daily time scales (Eastman et al., 2021; Zhang et al., 2024).

Text S2:

Previous studies suggested that aerosol-cloud interaction processes take hours to reach equilibrium (~20 hours), i.e., lagged aerosol effects on cloud properties (Glassmeier et al., 2021; Gryspeerdt et al., 2021; Qiu et al., 2024). Our findings above are based on instantaneous relationships between aerosols and LWP. To explore the time dependence of LWP sensitivities, we repeated the analysis from Figure 2b by leading N_d with 5, 10, and 20 hours, with others being equal. The results are summarized in Figure S10, showing that LWP sensitivities are strongest when N_d has no leading time and gradually weaken as leading time increases. However, it is important to note that EPCAPE observations are not Lagrangian-based, making the leading aerosol influence on the current cloud properties less realistic.

Text S3: Difference in slope definitions between this and previous studies

In our study, we use the slope $d\ln LWP_{Nd}/d\ln N_d$, derived from SHAP analysis. Based on Equations (1) and (2) in the main text, this slope relates to the partial derivative of LWP w.r.t. N_d as follow:

$$\frac{\partial \ln LWP}{\partial \ln N_d} = \frac{LWP_{Nd}}{LWP} \cdot \frac{d\ln LWP_{Nd}}{d\ln N_d}, (1)$$

This formula shows that the SHAP-derived slope $d\ln LWP_{Nd}/d\ln N_d$ approximates $\partial \ln LWP_{Nd}/\partial \ln N_d$ or the individual effect of N_d on LWP, when $LWP_{Nd} \approx LWP$ (i.e., when total meteorological contributions are small). In practice, the value of the slope $d\ln LWP_{Nd}/d\ln N_d$ in Figure 2b (-1.1) is found to be insensitive to sampling conditions where $LWP_{Nd} \approx LWP$, supporting its interpretation as a robust estimate of the individual aerosol effect. Additionally, when examining the dependence of the N_d-LWP relationship on meteorological factors (Figure 3), we can also use $dLWP_{Nd}/d\ln N_d$ that is exactly equal to $\partial LWP/\partial \ln N_d$; our conclusion still holds.

By contrast, previous studies (e.g., Hoffmann et al., 2024; Qiu et al., 2024) typically computed the full derivative $d\ln LWP/d\ln N_d$, sometimes stratified by cloud regimes or one to two meteorological factors. Given that LWP depends on both N_d and a set of meteorological variables M_i (including LTS, dq, RH_{700}, ω_{700}, q_{adv}, T_{adv}, LHF, and U_{1000}), and N_d itself is also a function of M_i, the full derivative of LWP w.r.t. N_d can be expanded via the chain rule as:

$$\frac{d\ln LWP}{d\ln N_d} = \frac{\partial \ln LWP}{\partial \ln N_d} + \sum_{i=1}^{8} \frac{\partial \ln LWP}{\partial M_i} \cdot \frac{dM_i}{d\ln N_d}, (2)$$

with

$$\frac{dM_i}{d\ln N_d} = \left(\frac{\partial \ln N_d}{\partial M_i}\right)^{-1}, (3)$$

Substituting Eq. (3) into (2) yields:

$$\frac{d\ln LWP}{d\ln N_d} = \frac{\partial \ln LWP}{\partial \ln N_d} + \sum_{i=1}^{8} \frac{\partial \ln LWP}{\partial M_i} \cdot \left(\frac{\partial \ln N_d}{\partial M_i}\right)^{-1}, (4)$$

This decomposition shows that the full derivative combines the individual/direct effect of N_d and indirect effects mediated through meteorology. These meteorological influences can partially offset the aerosol impacts. For instance, stronger surface winds may simultaneously increase aerosols by enhancing surface aerosol fluxes and enhance LWP via turbulent moistening, contributing a positive term to the second component in Eq. (4) and thereby reducing the overall slope magnitude. As a result, our slopes derived from partial derivatives tend to be more negative than those based on full derivatives, unless their meteorological effects are reduced or controlled.

Figures:

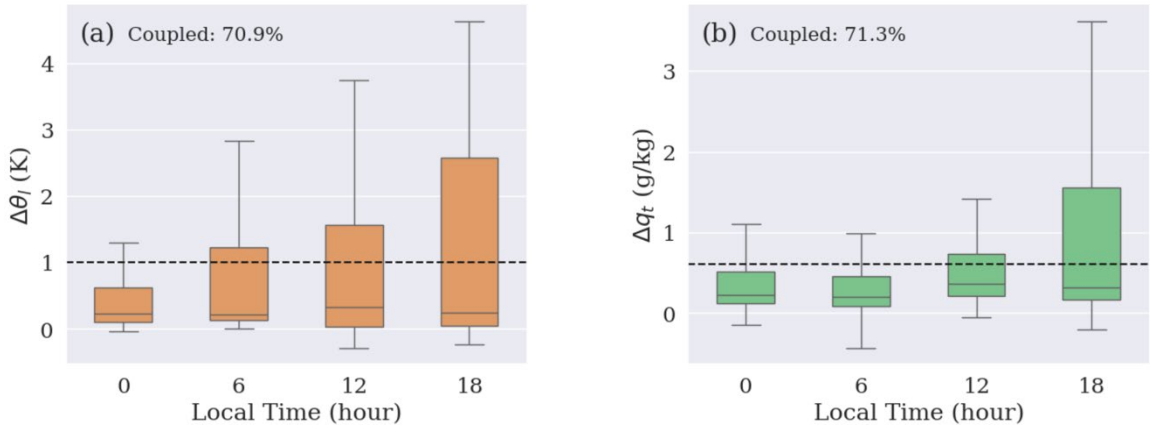

Figure S1. Boxplots showing the differences in average (a) liquid water potential temperature ($\Delta\theta_l$) or (b) total liquid water mixing ratio (Δq_t) between the upper and the lower 25 % of the sub-cloud layer. Both variables are calculated by assuming no liquid water within the sub-cloud layers. Dashed lines in each panel indicate the coupling thresholds, following Wang et al. (2016): the cloud layer is considered coupled with the surface if $\Delta\theta_l$<1 K or Δq_t<0.6 g/kg. The percentage shown in the upper left corner represents the fraction of coupled cases based on temperature or moisture criteria alone. In total, 76 % of the cases are coupled.

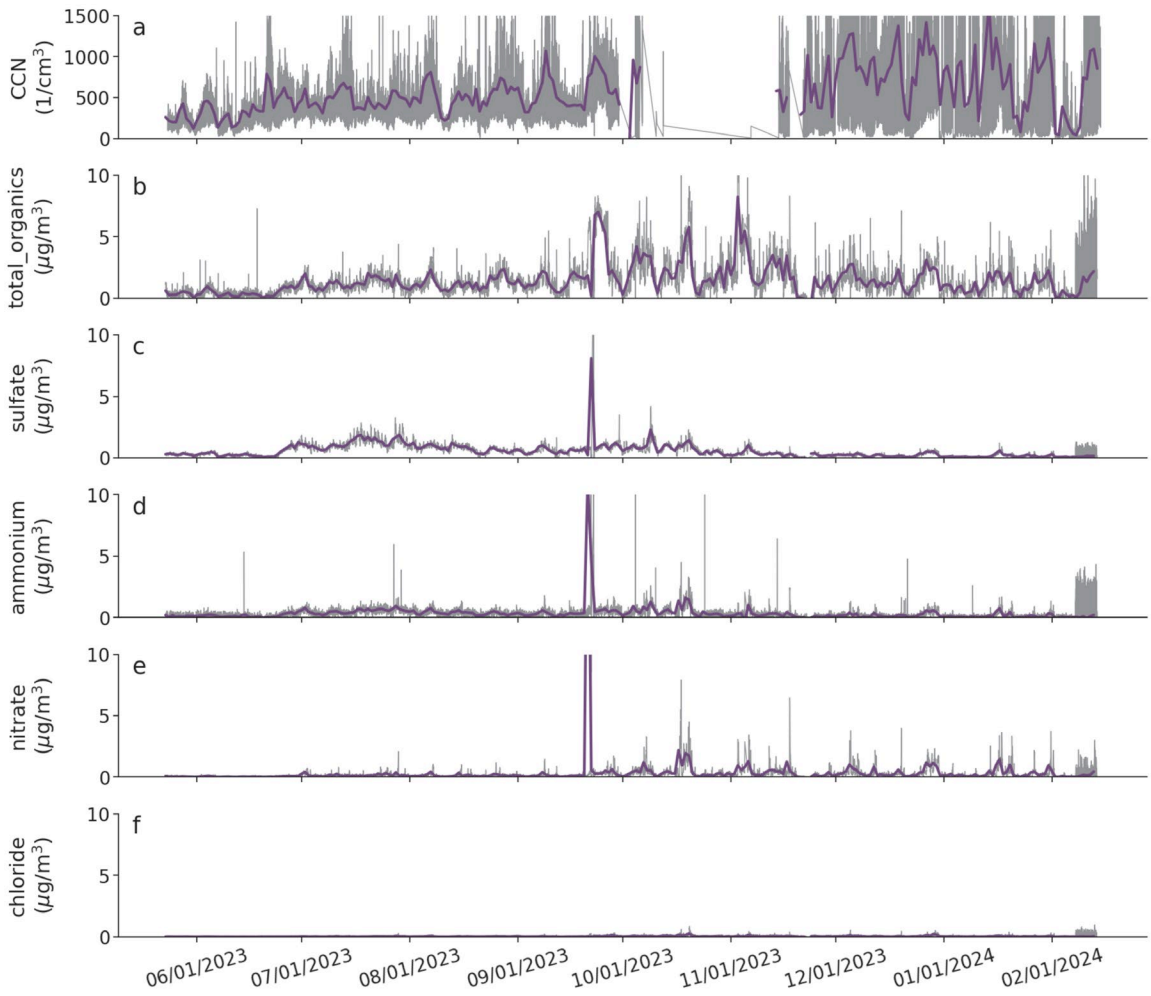

Figure S2. Time series of (a) activated CCN number concentration, (b) total organics mass concentration, (c) sulfate mass concentration, (d) ammonium mass concentration, (e) nitrate mass concentration, and (f) chloride mass concentration, from the EPCAPE main observation site. The gray lines show the hourly values, and the purple lines denote the daily average. The chemical component mass concentrations are obtained from aerosol chemical speciation monitor (ACSM) data.

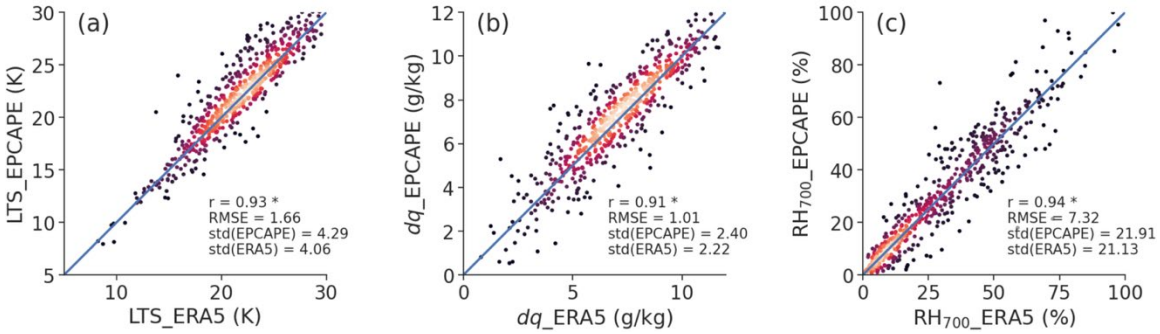

Figure S3. Validation of hourly variables either derived or directly obtained from ERA5 against EPCAPE observations: (a) lower-tropospheric stability (LTS), (b) moisture contrast between the surface and 700 hPa (*dq*), and (c) relative humidity at 700 hPa (RH_{700}). In each panel, the diagonal line represents the one-to-one line. The numbers from top to bottom represent: the correlation coefficient (**r**) between variables **x** and **y**, root mean squared errors (RMSEs) between **x** and **y**, standard deviation from EPCAPE, and standard deviation from ERA5. The asterisk "*" indicates that the correlation coefficient is statistically significant at the 0.01 level. The colors depict scatter density, varying across panels.

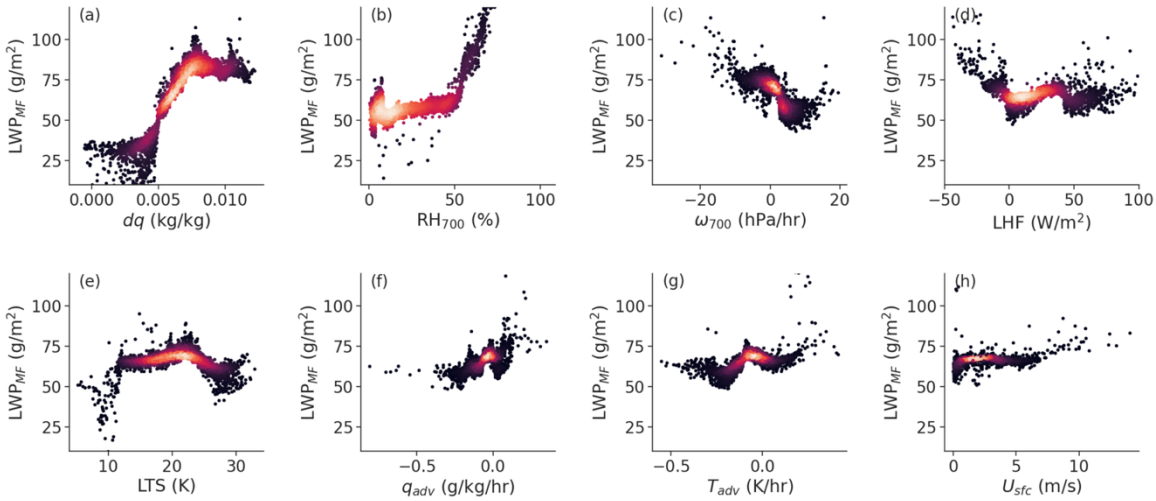

Figure S4. Same as Figure 2b, but showing the relationship between each meteorological factor (MF) and the LWP contributed solely by that MF (LWP_{MF}). The colors indicate scatter density, varying across panels.

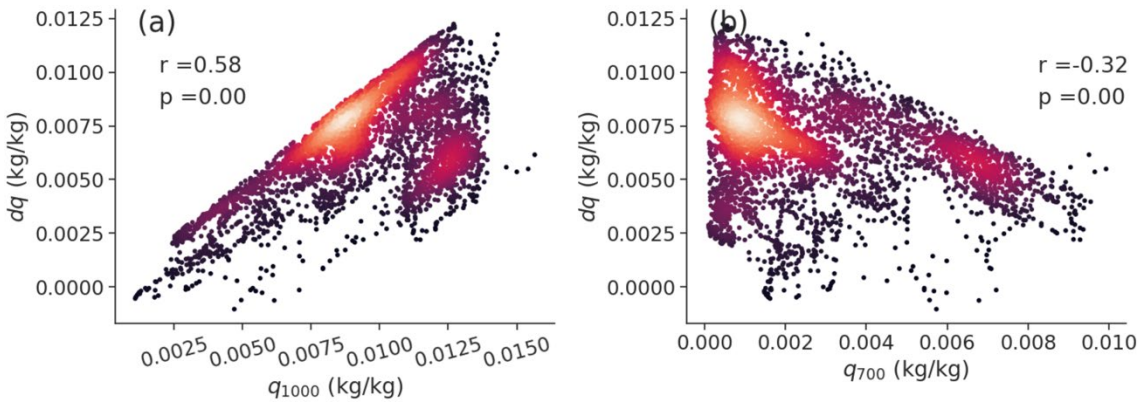

Figure S5. Relationship between hourly *dq* and (a) q_{1000} or (b) q_{700}. Correlation coefficients and p-values are shown in each panel. The colors indicate scatter density, varying across panels.

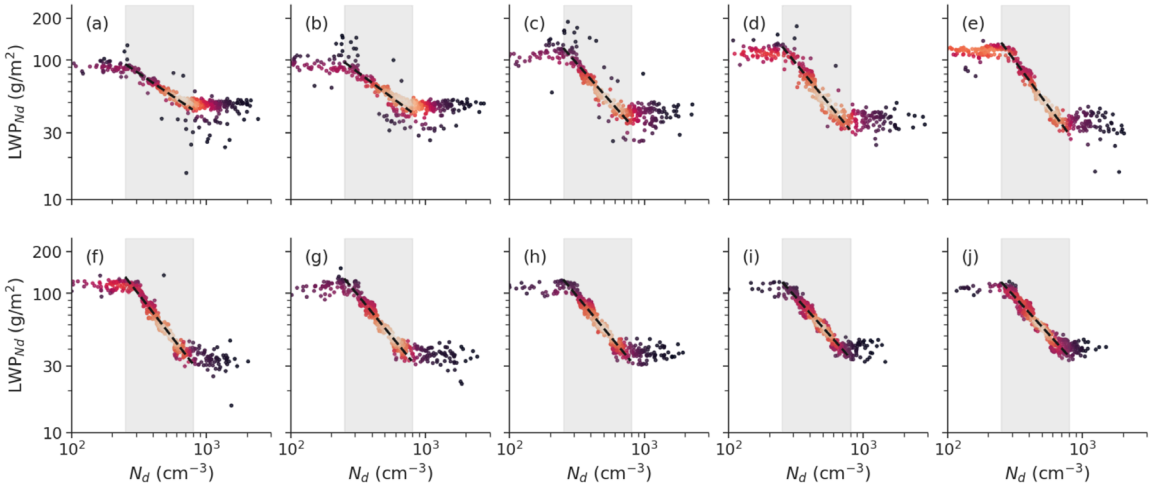

Figure S6. Same as Figure 2b, but with data samples sorted into ten quantile bins based on *dq* values. Panels (a–j) correspond to the first through tenth *dq* quantile bins. The gray box in each panel highlights the N_d range of 250–800 cm^{-3}. The colors represent scatter density, which varies across panels. The dashed line shows the best-fit linear regression for the gray box region.

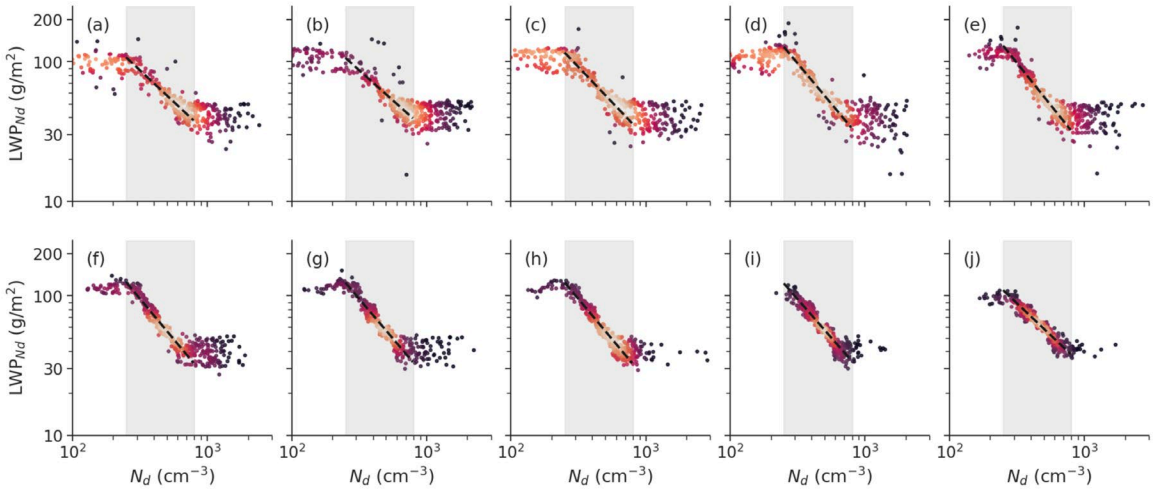

Figure S7. Same as Figure S6, but data sampled are sorted by LTS.

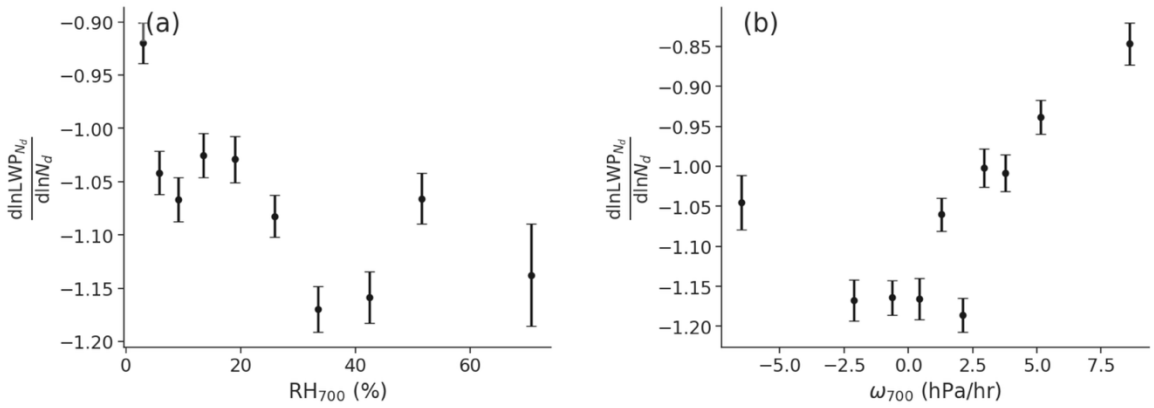

Figure S8. Same as Figure 3a, but for (a) RH_{700} and (b) ω_{700}.

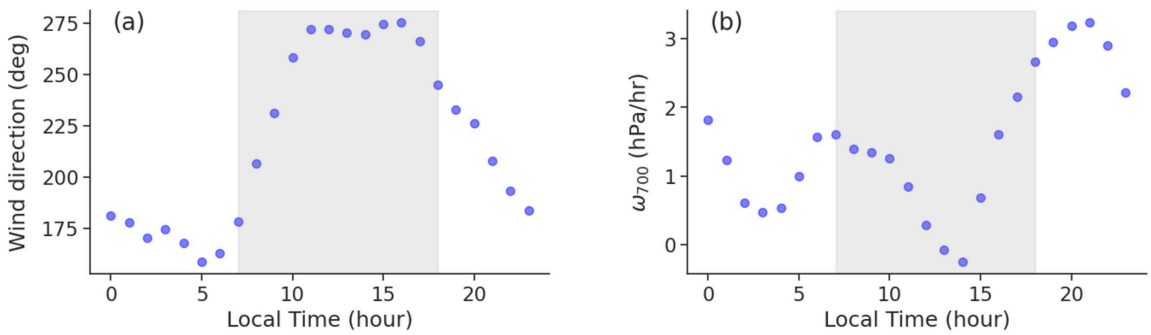

Figure S9. Same as Figure 3d, but for (a) surface wind direction and (b) ω_{700}. In panel (a), zero degrees represent the north direction, and measurements are taken clockwise from this point.

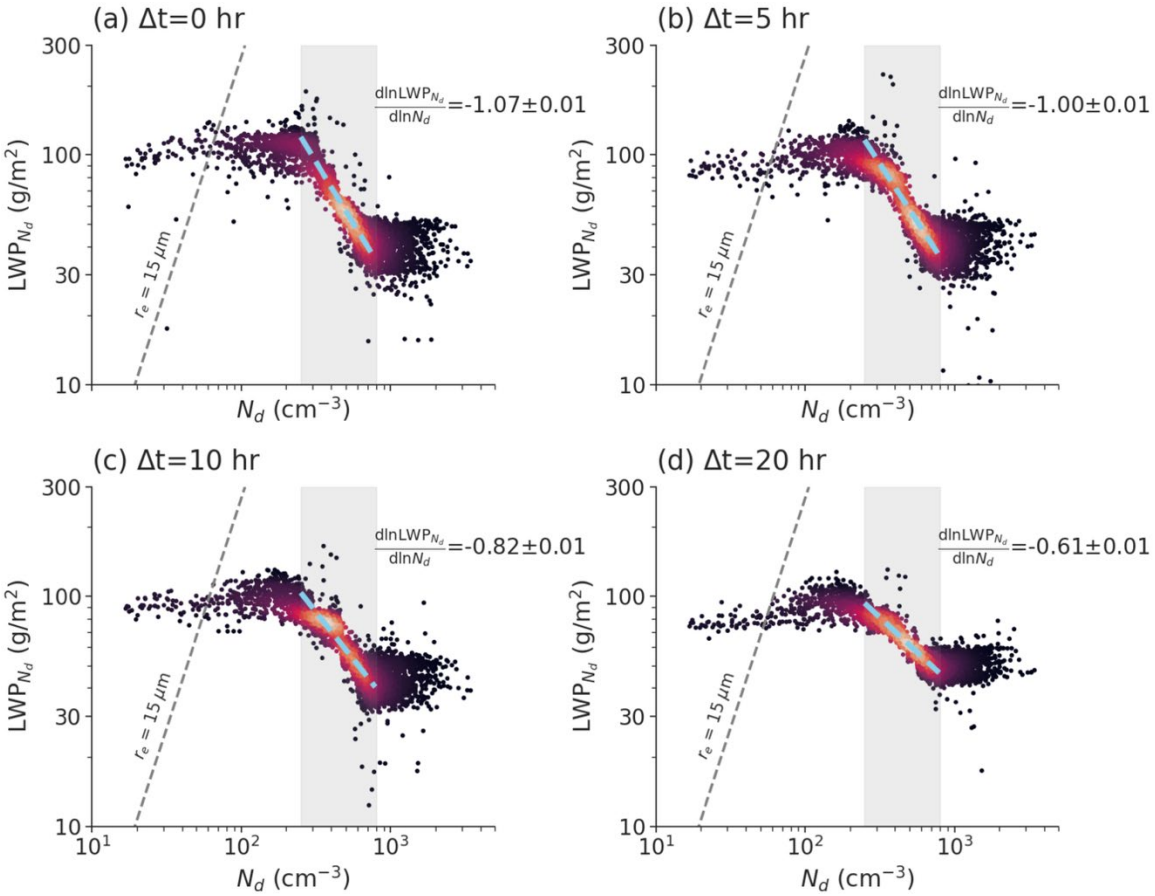

Figure S10. Relationship between hourly LWP_{Nd} and N_d derived from EPCAPE observations, with N_d leading by (a) 0 hours, (b) 5 hours, (c) 10 hours, and (d) 20 hours. The color bar indicates scatter density, varying across panels. The dashed line represents the 15-μm cloud-top effective droplet radius. The gray box highlights the N_d range of 250–800 cm^{-3}. The sky-blue dashed line shows the best-fit linear regression for the gray box region, with the slope and standard deviation provided.

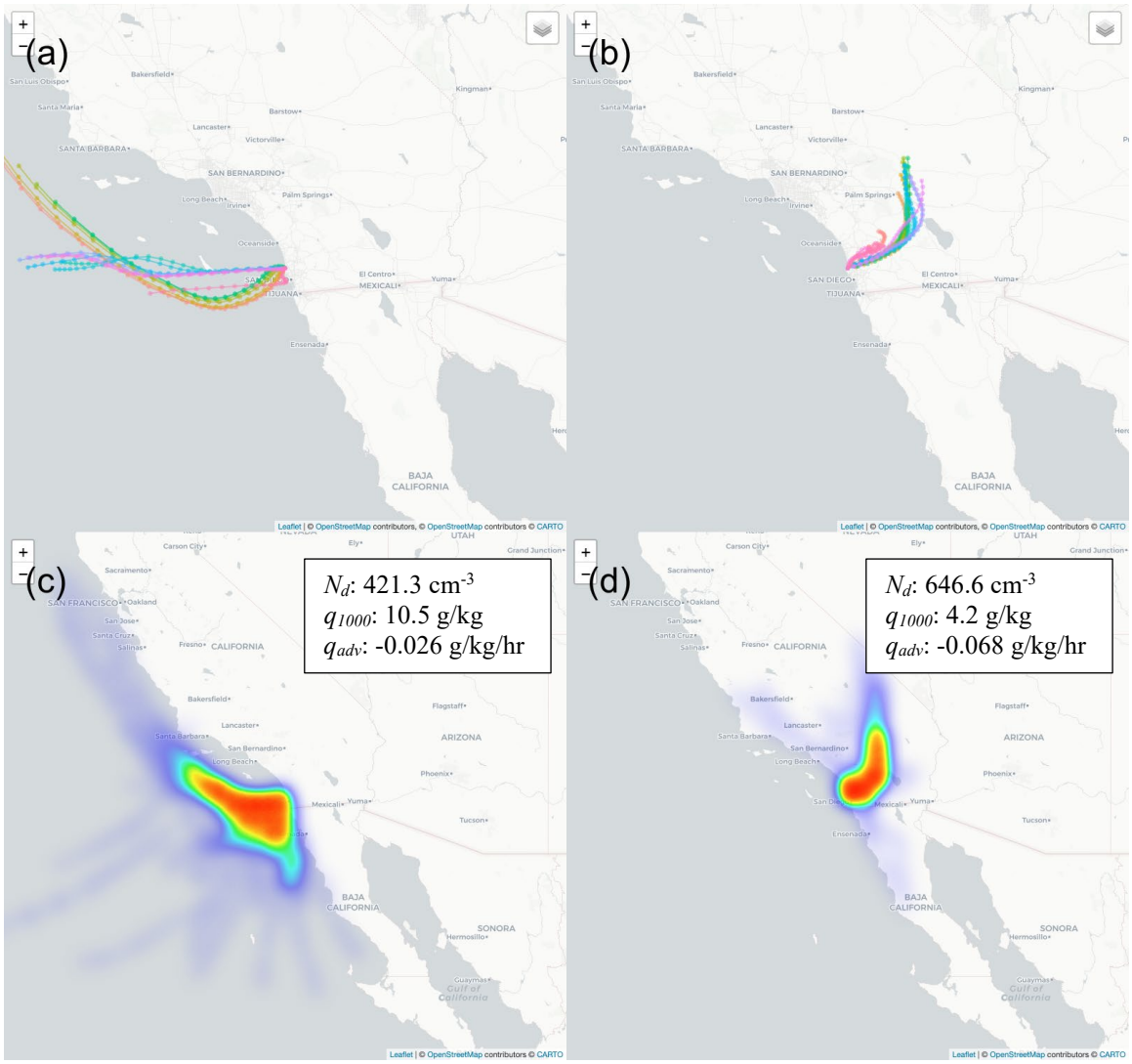

Figure S11. 24-hour backward trajectories that were run each hour from May 2023 to Jan 2024, initiated at 30 meters above the sea level on an isobaric surface. Trajectories were calculated using the HYSPLIT model with wind fields from the global 1-degree GDAS (Global Data Assimilation System) dataset. (a) Example of 15 oceanic trajectories selected from late Nov 2023, originating at least 100 km offshore from Long Beach, CA (to remove the influence from remote Los Angeles/Long Beach pollution sources). (b) Example of 15 continental trajectories from the same period. *Oceanic or continental trajectories are defined as those that remain entirely over the ocean or continent, respectively, throughout the 24-hour period*. (c) and (d) Trajectory densities for all oceanic (960 trajectories) and continental (475 trajectories) cases, respectively. The numbers in panels (c) and (d) indicate the aerosol number concentration (N_d), specific humidity at 1000 hPa (q_{1000}), and horizontal moisture advection (q_{adv}) observed at the main observation site, averaged for airmasses originating from the ocean and continent, respectively.

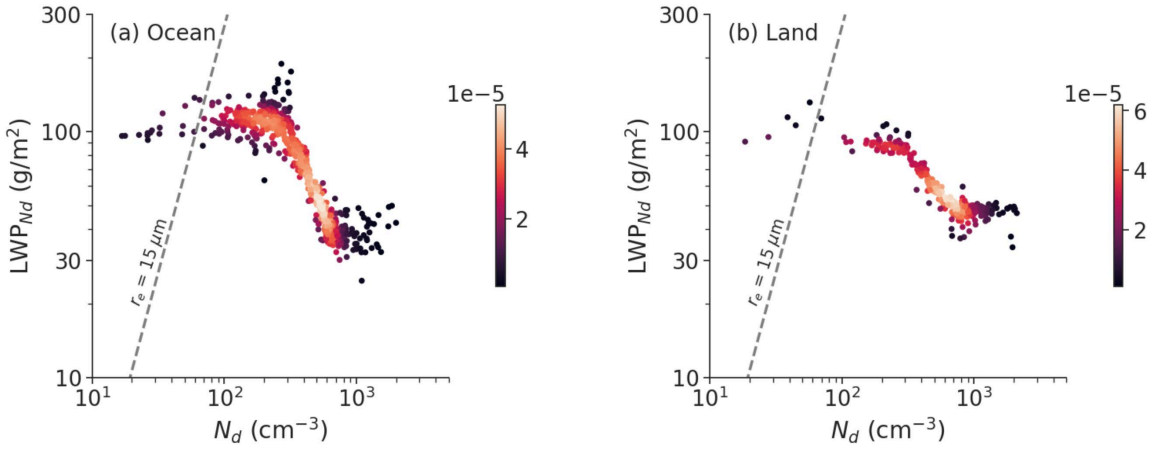

Figure S12. Similar to Figure 2b, but with data samples sorted into (a) oceanic airmasses and (b) continental airmasses, as identified in Figure S11. The color bar represents scatter densities.

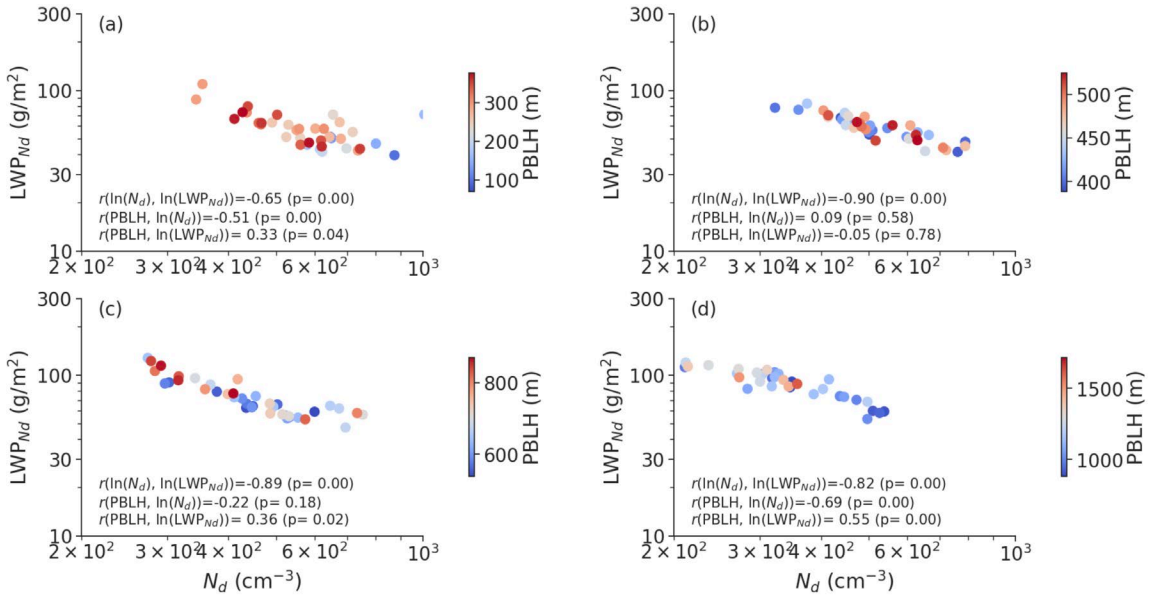

Figure S13. Similar to Figure 2b, but with data samples sorted into (a) the first quantile, (b) the second quantile, (c) the third quantile, and (d) the fourth quantile of planetary boundary later height (PBLH). The color bar represents PBLH values. The values in the lower left of each panel show the correlation between the two selected variables.

Reference:

Bretherton, C., Blossey, P. N., & Jones, C. R. (2013). Mechanisms of marine low cloud sensitivity to idealized climate perturbations: A single-LES exploration extending the CGILS cases. *Journal of Advances in Modeling Earth Systems*, *5*(2), 316–337. https://doi.org/10.1002/jame.20019

Eastman, R., Terai, C. R., Grosvenor, D. P., & Wood, R. (2021). Evaluating the lagrangian evolution of subtropical low clouds in GCMs using observations: Mean evolution, time scales, and responses to predictors. *Journal of the Atmospheric Sciences*, *78*(2), 553–572. https://doi.org/10.1175/JAS-D-20-0178.1

Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., & Feingold, G. (2021). Aerosol-cloud-climate cooling overestimated by ship-track data. *Science*, *371*(6528), 485–489. https://doi.org/10.1126/science.abd3980

Gryspeerdt, E., Goren, T., & Smith, T. W. P. (2021). Observing the timescales of aerosol-cloud interactions in snapshot satellite images. *Atmospheric Chemistry and Physics*, *21*(8), 6093–6109. https://doi.org/10.5194/acp-21-6093-2021

Hoffmann, F., Glassmeier, F., & Feingold, G. (2024). The impact of aerosol on cloud water: a heuristic perspective. *Atmospheric Chemistry and Physics*, *24*(23), 13403–13412. https://doi.org/10.5194/acp-24-13403-2024

Qiu, S., Zheng, X., Painemal, D., Terai, C. R., & Zhou, X. (2024). Daytime variation in the aerosol indirect effect for warm marine boundary layer clouds in the eastern North Atlantic. *Atmospheric Chemistry and Physics*, *24*(5), 2913–2935. https://doi.org/10.5194/acp-24-2913-2024

Wang, Z., Mora Ramirez, M., Dadashazar, H., MacDonald, A. B., Crosbie, E., Bates, K. H., et al. (2016). Contrasting cloud composition between coupled and decoupled marine boundary layer clouds. *Journal of Geophysical Research: Atmospheres*, *121*(19), 238–238. https://doi.org/10.1002/2016JD025695

Zhang, H., Zheng, Y., & Li, Z. (2024). Evaluation of stratocumulus evolution under contrasting temperature advections in CESM2 through a Lagrangian framework. *Geophysical Research Letters*, *51*(4). https://doi.org/10.1029/2023GL106856