

Atmos. Chem. Phys., 20, 7645–7665, 2020

<https://doi.org/10.5194/acp-20-7645-2020>

© Author(s) 2020. This work is distributed under the Creative Commons Attribution 4.0 License.



# On the relationship between cloud water composition and cloud droplet number concentration

Alexander B. MacDonald<sup>1</sup>, Ali Hossein Mardi<sup>1</sup>, Hossein Dadashazar<sup>1</sup>, Mojtaba Azadi Aghdam<sup>1</sup>, Ewan Crosbie<sup>2,3</sup>, Haffidi H. Jonsson<sup>4</sup>, Richard C. Flagan<sup>5</sup>, John H. Seinfeld<sup>5</sup>, and Armin Sorooshian<sup>1,6</sup><sup>1</sup>Department of Chemical and Environmental Engineering, University of Arizona, Tucson, AZ, USA<sup>2</sup>Science Systems and Applications, Inc., Hampton, VA, USA<sup>3</sup>NASA Langley Research Center, Hampton, VA, USA<sup>4</sup>Naval Postgraduate School, Monterey, CA, USA<sup>5</sup>Department of Chemical Engineering, California Institute of Technology, Pasadena, CA, USA<sup>6</sup>Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA**Correspondence:** Armin Sorooshian ([armin@email.arizona.edu](mailto:armin@email.arizona.edu))

Received: 13 March 2020 – Discussion started: 19 March 2020

Revised: 28 May 2020 – Accepted: 7 June 2020 – Published: 2 July 2020

**Abstract.** Aerosol–cloud interactions are the largest source of uncertainty in quantifying anthropogenic radiative forcing. The large uncertainty is, in part, due to the difficulty of predicting cloud microphysical parameters, such as the cloud droplet number concentration ( $N_d$ ). Even though rigorous first-principle approaches exist to calculate  $N_d$ , the cloud and aerosol research community also relies on empirical approaches such as relating  $N_d$  to aerosol mass concentration. Here we analyze relationships between  $N_d$  and cloud water chemical composition, in addition to the effect of environmental factors on the degree of the relationships. Warm, marine, stratocumulus clouds off the California coast were sampled throughout four summer campaigns between 2011 and 2016. A total of 385 cloud water samples were collected and analyzed for 80 chemical species. Single- and multispecies log–log linear regressions were performed to predict  $N_d$  using chemical composition. Single-species regressions reveal that the species that best predicts  $N_d$  is total sulfate ( $R_{\text{adj}}^2 = 0.40$ ). Multispecies regressions reveal that adding more species does not necessarily produce a better model, as six or more species yield regressions that are statistically insignificant. A commonality among the multispecies regressions that produce the highest correlation with  $N_d$  was that most included sulfate (either total or non-sea-salt), an ocean emissions tracer (such as sodium), and an organic tracer (such as oxalate). Binning the data according to turbulence, smoke influence, and in-cloud height allowed for

examination of the effect of these environmental factors on the composition– $N_d$  correlation. Accounting for turbulence, quantified as the standard deviation of vertical wind speed, showed that the correlation between  $N_d$  with both total sulfate and sodium increased at higher turbulence conditions, consistent with turbulence promoting the mixing between ocean surface and cloud base. Considering the influence of smoke significantly improved the correlation with  $N_d$  for two biomass burning tracer species in the study region, specifically oxalate and iron. When binning by in-cloud height, non-sea-salt sulfate and sodium correlated best with  $N_d$  at cloud top, whereas iron and oxalate correlated best with  $N_d$  at cloud base.

## 1 Introduction

To assess the degree to which humans have altered Earth's climate, it is necessary to quantify the effect that particles in the air (i.e., aerosols) have on clouds. Some fraction of aerosols (called cloud condensation nuclei, CCN) activate into cloud droplets, thus impacting the cloud droplet number concentration ( $N_d$ ). For warm marine boundary layer (MBL) clouds at fixed liquid water, higher  $N_d$  values result in (i) higher cloud albedo (thus cooling the Earth and counteracting the greenhouse effect) (Twomey, 1977), (ii) delayed and/or reduced precipitation (Albrecht, 1989), and (iii) en-

hanced entrainment at cloud top (Ackerman et al., 2004). The complex interactions and feedback mechanisms between aerosols, meteorology, and clouds leads to aerosol–cloud interactions as the largest source of uncertainty in climate models (IPCC, 2013; Bellouin et al., 2020).

It is indispensable to know the value of  $N_d$ , but this is a difficult parameter to accurately simulate and retrieve (Fountoukis and Nenes, 2005). There is a need to improve  $N_d$  retrievals from satellite remote sensors, which provide broad spatial and temporal coverage in contrast to surface sites and airborne research flights. Currently,  $N_d$  retrievals are limited to inferred values based on values of cloud optical depth, cloud droplet effective radius, and temperature, along with assumptions such as vertical homogeneity of  $N_d$  and monotonic increases in liquid water content at a constant fraction of its adiabatic value (Grosvenor et al., 2018). Ultimately, measurements are needed to better inform climate models about the cloud droplet activation process and better constrain  $N_d$  values. Current general circulation models (GCMs) calculate  $N_d$  using the properties of aerosol particles in one of two ways (Ghan et al., 1997; Menon et al., 2002). First, there is a rigorous approach that is based on physical principles that predicts  $N_d$  based on aerosol properties and meteorological conditions (Abdul-Razzak and Ghan, 2000). Second, there is an empirical approach that parameterizes  $N_d$  using either the number concentration of aerosols,  $N_a$  ( $\text{cm}^{-3}$ ); the number concentration of CCN,  $N_{\text{CCN}}$  ( $\text{cm}^{-3}$ ); or the mass concentration of chemical species that comprise the aerosols (Ghan et al., 1997).

The rigorous approach predicts  $N_d$  by considering aerosol properties (e.g., size distribution and chemical composition), microphysical processes (e.g., the seeding of cloud droplets by particles, droplet growth, and droplet evaporation), and meteorological parameters (e.g., relative humidity and the vertical updraft velocity transporting aerosols to cloud base) (e.g., Chuang et al., 1992; Chuang and Penner, 1995; Nenes and Seinfeld, 2003; Partridge et al., 2012). This method is based on the physical principle that an aerosol particle needs to be a cloud condensation nucleus in order to seed a cloud droplet; consequently, the input for this approach is  $N_a$ , from which to calculate  $N_{\text{CCN}}$ , and subsequently  $N_d$ . The requisite information for these calculations may not be readily available for GCMs. A limitation is that the spatial resolution of a GCM may be too coarse to capture the small-scale spatial variation of updraft velocity (Ghan et al., 2011; West et al., 2014).

The empirical parameterization approach of interest in the present study uses the mass concentration of one or several chemical species and correlates it directly to  $N_{\text{CCN}}$  or  $N_d$ . Aerosols containing the sulfate ion ( $\text{SO}_4^{2-}$ ) have long been known to serve as effective CCN (Andreae and Rosenfeld, 2008; Charlson et al., 1992; Lance et al., 2009; Medina et al., 2007). Sulfate is both contained in sea salt and is a product of the oxidation of gaseous sulfur dioxide ( $\text{SO}_2$ ) (Hegg et al., 1981; Quinn et al., 2017), so it is customary to isolate

the anthropogenic contribution to total  $\text{SO}_4^{2-}$  by considering its non-sea-salt fraction (NSS- $\text{SO}_4^{2-}$ ). Therefore, most studies choose either total  $\text{SO}_4^{2-}$  (denoted hereafter as Tot- $\text{SO}_4^{2-}$ ) or NSS- $\text{SO}_4^{2-}$  to predict  $N_{\text{CCN}}$  and  $N_d$  (e.g., Leaitch et al., 1992; Novakov et al., 1994; Saxena and Menon, 1999). Using the mass concentration of  $\text{SO}_4^{2-}$  or any other chemical species to predict  $N_d$  (i) circumvents the complex intermediate microphysical steps to go from an aerosol particle to a cloud droplet and implicitly accounts for such meteorological variables like updraft velocity, (ii) is based on actual measurements, and (iii) can be compared directly to the mass concentration of different species produced by aerosol transport models (e.g., Boucher and Lohmann, 1995; Chen and Penner, 2005). The limitations of using an empirical parameterization are (i) assuming a mass size distribution of the aerosols, (ii) assuming that one or a few chemical species are responsible for all CCN, and (iii) uncertainty in generalizing field data from one region (or a few regions) under specific conditions to the entire globe for all conditions (Pringle et al., 2009). Despite these drawbacks, empirical correlations of  $N_d$  and the mass concentration of different species are valuable. For example, of the 20 studies addressing the cloud albedo effect considered in the IPCC Fourth Assessment Report (IPCC, 2007), half relied on empirical relationships to calculate  $N_d$  (Pringle et al., 2009).

Several studies have developed empirical correlations between  $N_{\text{CCN}}$  and the mass concentration of  $\text{SO}_4^{2-}$  (e.g., Adams and Seinfeld, 2003; Hegg et al., 1993; Matsumoto et al., 1997). However, the present objective is to focus on improving the prediction of  $N_d$ , not  $N_{\text{CCN}}$ , using the mass concentration of  $\text{SO}_4^{2-}$  in addition to other species. A log–log relation is often used to correlate the mass concentration of  $\text{SO}_4^{2-}$  to  $N_d$  with an equation of the following form (e.g., Lowenthal et al., 2004):

$$\log(N_d) = a_0 + a_1 \log\left(\left[\text{SO}_4^{2-}\right]\right), \quad (1)$$

where  $\text{SO}_4^{2-}$  is the mass concentration in air ( $\mu\text{g m}^{-3}$ ), and  $a_0$  and  $a_1$  are fitting parameters. A log–log relation is chosen to accommodate large ranges in  $N_d$  and  $\text{SO}_4^{2-}$  and to reduce sensitivity of results to the measurement accuracy of each individual parameter (Boucher and Lohmann, 1995). The mass concentration of  $\text{SO}_4^{2-}$  can be obtained by analyzing either aerosol particles or cloud water. When analyzing cloud water, the mass concentration of  $\text{SO}_4^{2-}$  dissolved in droplets ( $\text{mg L}^{-1}$ ) is converted to the air-equivalent mass concentration ( $\mu\text{g m}^{-3}$ ) by multiplying by the liquid water content, LWC ( $\text{g m}^{-3}$ ), in a cloud. The data used to create  $N_d$ – $\text{SO}_4^{2-}$  empirical parameterizations are typically derived from field campaigns, which differ in the region of analysis, sampling platforms (aircraft or ground-based), measurement approach (e.g., in particle form or dissolved in cloud water), and number of species analyzed. While the literature evaluating relationships between cloud water composition and  $N_d$

is limited and largely from aircraft studies from more than a decade ago, there is a growing number of data sets characterizing  $N_d$  and cloud water composition that are of interest to continue this line of research. Examples include the recently completed Cloud, Aerosol, and Monsoon Processes Philippines Experiment (CAMP<sup>2</sup>Ex), the North Atlantic Aerosols and Marine Ecosystems Study (NAAMES) (Behrenfeld et al., 2019), and the current multiyear Aerosol Cloud Meteorology Interactions over the western Atlantic Experiment (ACTIVATE) (Sorooshian et al., 2020). A summary of relevant past field work follows.

Leaitch et al. (1986) sampled continental stratiform and cumuliform clouds over Ontario, Canada, and showed a roughly linear relationship between  $N_d$  and  $\text{SO}_4^{2-}$  at low  $\text{SO}_4^{2-}$  concentrations (below  $5 \mu\text{g m}^{-3}$ ), and they showed that the relationship leveled out at higher concentrations (Novakov et al., 1994). Leaitch et al. (1992) suggested that the low  $R^2$  values for the linear regression between  $N_d$  and  $\text{SO}_4^{2-}$  for both continental stratiform and cumuliform clouds (0.30 and 0.49, respectively) stemmed from factors such as (i) other chemical species besides  $\text{SO}_4^{2-}$  and variability in both (ii) updraft wind speed and (iii) temperature. Pueschel et al. (1986) sampled clouds originating from marine and continental air masses at a ground-based observatory at Whiteface Mountain, New York. They found that emissions contributed strongly to  $\text{SO}_4^{2-}$  and that a significant portion of  $\text{SO}_4^{2-}$ -containing particles acted as CCN and thus likely impacted  $N_d$ . Novakov et al. (1994) sampled marine cumulus and stratocumulus clouds by El Yunque peak in Puerto Rico. Although they showed that  $N_{\text{CCN}}$  and  $\text{SO}_4^{2-}$  were highly correlated in both cumulus and stratocumulus clouds, they also found that  $N_d$  and  $\text{SO}_4^{2-}$  were weakly correlated for stratocumulus clouds and not correlated for cumulus clouds. They attributed this difference to the effect of entrainment and mixing on cloud microphysics. Leaitch et al. (1996) sampled marine stratus clouds over the Gulf of Maine and the Bay of Fundy during the North Atlantic Regional Experiment (NARE) and showed that  $\text{SO}_4^{2-}$  was better correlated with  $N_d$  than nitrate ( $\text{NO}_3^-$ ) (with  $R^2$  values of 0.30 and 0.12, respectively). The  $R^2$  between  $N_d$  and  $\text{SO}_4^{2-}$  increased when the data were stratified into bins of low and high turbulence, which was quantified as the standard deviation of vertical wind speed. They found that in situations with lower supersaturations,  $N_d$  was more influenced by turbulence than by either  $\text{SO}_4^{2-}$  or  $N_a$ . Menon and Saxena (1998) and Saxena and Menon (1999) sampled orographic clouds at a ground-based station at Mt. Mitchell, North Carolina. They found that  $\text{SO}_4^{2-}$  was the main contributor to cloud water acidity and a reliable tracer for anthropogenic pollution. Log–log regressions of  $\text{SO}_4^{2-}$ – $N_d$  were binned according to the level of  $\text{SO}_4^{2-}$ , with not much difference observed between the different levels of pollution. Borys et al. (1998) and Lowenthal and Borys (2000) sampled warm marine stratiform clouds on the island of Tenerife in the Canary Islands. They found that

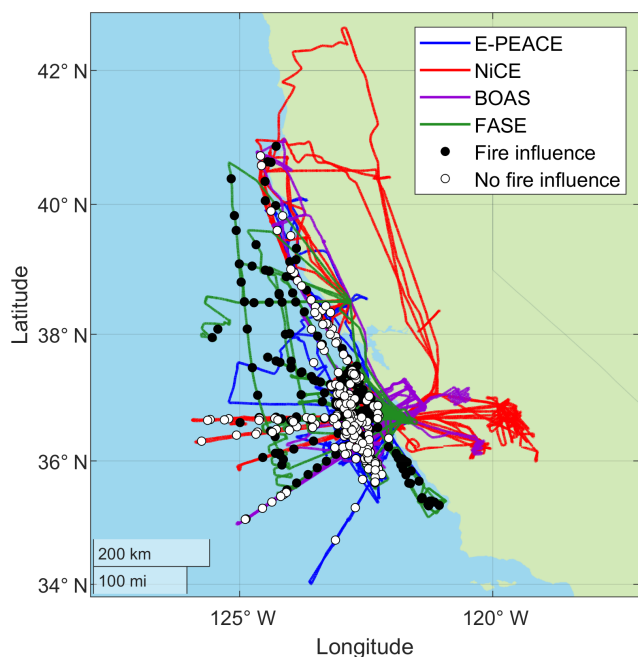
$N_d$  was influenced by  $\text{NSS-SO}_4^{2-}$ ,  $\text{NO}_3^-$ , pollution-derived trace elements, and elemental carbon (EC), signifying that species other than  $\text{SO}_4^{2-}$  influenced  $N_d$ . Despite the sampling site proximity to African deserts, the mass concentration of crustal elements contained in dust was found to have little correlation with  $N_d$ . Also, the sea salt tracer sodium ( $\text{Na}^+$ ) was found to have little correlation with  $N_d$ . Several studies (e.g., Boucher and Lohmann, 1995; Lowenthal et al., 2004; Menon et al., 2002; Van Dingenen et al., 1995) have combined field data, such as those mentioned above, in addition to other data sets, with the intention of producing a robust empirical prediction of  $N_d$ . Menon et al. (2002) provided a log–log multispecies prediction of  $N_d$  using  $\text{SO}_4^{2-}$ , organic matter, and sea salt. Organic carbon has been shown to increase  $N_d$ , as it affects the surface tension of cloud droplets (e.g., Facchini et al., 1999; Nenes et al., 2002). Additionally, nitric acid ( $\text{HNO}_3$ ) has been linked with increased CCN activity and  $N_d$  based on modeling studies (Hegg, 2000; Kulmala et al., 1993; Xue and Feingold, 2004).

McCoy et al. (2017) used  $N_d$  data from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite instead of in situ measurements. Second, aerosol mass concentration data were obtained from the Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2; Gelaro et al., 2017) reanalysis product and various aerosol transport models instead of in situ measurements. Third, the study region was more global in nature (albeit focusing on marine stratocumulus clouds) instead of a specific region. Fourth, since reanalysis data were used, a multi-species, multi-variable linear regression was performed:

$$\log(N_d) = a_0 + a_1 \log(\text{SO}_4^{2-}) + a_2 \log(\text{SS}) + a_3 \log(\text{BC}) + a_4 \log(\text{OC}) + a_5 \log(\text{DU}), \quad (2)$$

where SS is sea salt, BC is black carbon, OC is organic carbon, and DU is dust. McCoy et al. (2017) found that  $\text{SO}_4^{2-}$  was predominantly correlated with  $N_d$ , with sea salt, black carbon, organic carbon, and dust accounting for smaller contributions. A caveat to consider when comparing the findings of McCoy et al. (2018) to other aircraft studies is that McCoy et al. (2018) used mass concentrations retrieved exclusively at the 910 hPa model level ( $\sim 915$  m) and only considered mass concentrations pertaining to submicron SS/DU and hydrophilic BC/OC.

The field studies cited above still leave a series of unanswered questions that the current study aims to address there: (i) how is the  $\text{SO}_4^{2-}$ – $N_d$  relationship affected by vertical wind speed (Leaitch et al., 1992), turbulence (Leaitch et al., 1996), and entrainment (Novakov et al., 1994); (ii) why do species such as sea salt and dust play such a minor role in influencing  $N_d$ , even when located over the ocean and near a desert (Borys et al., 1998; McCoy et al., 2017, 2018); (iii) what is the relationship between organic matter and  $N_d$  (McCoy et al., 2018; Nenes et al., 2002); and (iv) can the  $\text{SO}_4^{2-}$ – $N_d$  correlation be improved by considering other chemical species (e.g.,



**Figure 1.** Flight paths for each of the four campaigns used in this study. Markers indicate the average location at which the cloud water samples were collected. Smoke- and non-smoke-influenced samples are indicated with filled and open markers, respectively.

Hegg et al., 1993; Leaitch et al., 1992; Novakov and Penner, 1993). The present study will examine these questions using a data set comprised of in situ aircraft measurements collected off the California coast during four field campaigns. In addition to meteorological and aerosol and cloud microphysical measurements, a total of 385 cloud water samples were collected and analyzed for 80 chemical species (ions and elements). Even though measurements were collected in only one localized region, it is expected that the variety of conditions encountered over four summers, together with the large number of chemical species analyzed, will help address the questions noted above. The results of this work have implications for simulations and retrievals of  $N_d$ , in addition to studies examining relationships between atmospheric chemistry and cloud microphysics.

## 2 Methodology

### 2.1 Aircraft campaigns and study region

This work reports results relevant to warm marine stratocumulus clouds off the California coast based on field measurements from four field campaigns between 2011 and 2016, each during the months of July and August. The persistent summertime stratocumulus cloud deck located off the California coast offers the ideal natural laboratory to study aerosol–cloud–precipitation–meteorology interactions (Rus-

sell et al., 2013; Sorooshian et al., 2018). For all field campaigns, the Center for Interdisciplinary Remotely-Piloted Aircraft Studies (CIRPAS) Twin Otter was deployed out of Marina, California, with an almost identical instrumentation payload. The four campaigns addressed in this study are the Eastern Pacific Emitted Aerosol Cloud Experiment (E-PEACE) (Russell et al., 2013; Wonaschütz et al., 2013), the Nucleation in California Experiment (NiCE) (Crosbie et al., 2016; Maudlin et al., 2015), the Biological and Oceanic Atmospheric Study (BOAS) (Wang et al., 2016), and the Fog and Stratocumulus Evolution (FASE) experiment (Dadas-hazar et al., 2017; MacDonald et al., 2018). Research flight information and tracks are shown in Table 1 and Fig. 1, respectively.

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

### 2.2 Aircraft instrumentation

Aircraft instrumentation used in each campaign is described in detail in Sorooshian et al. (2018). The relevant instrumentation used in the present study is as follows: aerosol size distribution was measured using a passive cavity aerosol spectrometer probe (PCASP; particle diameter ( $D_p$ ) of  $\sim 0.1$ – $2.6 \mu\text{m}$ ; Strapp et al., 1992); cloud droplet size distribution was measured using a forward scattering spectrometer probe (FSSP;  $D_p$  of  $\sim 2$ – $45 \mu\text{m}$ ; Gerber et al., 1999) and a cloud and aerosol spectrometer – forward scattering (CASF;  $D_p$  of  $\sim 1$ – $61 \mu\text{m}$ ; Baumgardner et al., 2001); rain drop size distribution was measured using a cloud imaging probe (CIP;  $D_p$  of  $\sim 25$ – $1600 \mu\text{m}$ ; Baumgardner et al., 2001); cloud liquid water content (LWC) was measured using a particulate volume monitor (PVM-100A;  $D_p$  of  $\sim 3$ – $50 \mu\text{m}$ ; Gerber, 1994); and three-dimensional wind speeds were calculated by combining the pressure measurements from a five-hole Radome gust probe plumbed into the aircraft nose together with the aircraft velocity and altitude measurements provided by the aircraft's Global Positioning System and inertial navigation system (GPS/INS).

Since LWC played a critical role in converting aqueous concentration to air-equivalent concentration, the size range used to calculate  $N_d$  was bracketed to resemble the size range of the PVM-100A. Therefore,  $N_d$  was defined in this study to be equivalent to the integration of the cloud droplet size distribution between  $D_p$  of  $\sim 3$ – $50 \mu\text{m}$  and was calculated using CASF (for E-PEACE) and FSSP (NiCE, BOAS, and FASE).

**Table 1.** Summary of field campaign data sets used in this study and statistics related to cloud water sample collection. Smoke-influenced research flights (RFs) were NiCE RFs 16–23 and FASE RFs 3–11 and 13–15.

Field campaign	Dates (mm/dd/yyyy)	No. of RFs	No. of samples	No. of fire- impacted samples
Eastern Pacific Emitted Aerosol Cloud Experiment (E-PEACE)	07/08/2011–08/18/2011	30	82	0
Nucleation in California Experiment (NiCE)	07/08/2013–08/07/2013	23	119	31
Biological and Oceanic Atmospheric Study (BOAS)	07/02/2015–07/24/2015	15	29	0
Fog and Stratocumulus Evolution experiment (FASE)	07/18/2016–08/12/2016	16	155	136

For the NiCE campaign, LWC measurements from the PVM-100A instrument were unreliable; therefore, the LWC for NiCE was calculated instead using FSSP data between  $D_p$  of  $\sim 3$ – $50 \mu\text{m}$ .

### 2.3 Cloud water collection and chemical analysis

A total of 385 cloud water samples were collected throughout the four campaigns using a modified Mohnen slotted-rod collector, reported to collect droplets with  $D_p$  of  $\sim 5$ – $35 \mu\text{m}$  (Hegg and Hobbs, 1986). The cloud water was collected in polyethylene bottles and stored at  $\sim 5^\circ\text{C}$  for subsequent off-line chemical analysis. The spatially averaged location of each cloud water sample is shown in Fig. 1. Cloud water samples were chemically analyzed post-flight for ions using ion chromatography (IC; Dionex ICS-2100) and for elements using inductively coupled plasma mass spectrometry (ICP-MS; Agilent 7700 Series) for E-PEACE, BOAS, and NiCE or triple quadrupole inductively coupled plasma mass spectrometry (ICP-QQQ; Agilent 8800 Series) for FASE. The limit of detection (LOD) for each ion and element measured is shown in Table S1 in the Supplement. The concentration of non-sea-salt (NSS) species was calculated using the relative abundance of a NSS species to  $\text{Na}^+$  in natural sea salt (Seinfeld and Pandis, 2016). Cloud water sample acidity was quantified by measuring pH (the aqueous concentration of hydrogen ions,  $\text{H}^+$ ) using a Thermo Scientific Orion 9110DJWP Combination Semi-Micro pH electrode for E-PEACE, NiCE, and BOAS and a Thermo Scientific Orion 8103BNUWP Ross Ultra Semi-Micro pH probe for FASE. Aqueous concentrations (i.e., mass concentrations in the droplets,  $\text{mg L}^{-1}$ ) were converted to air-equivalent concentrations (i.e., mass concentrations in the air,  $\mu\text{g m}_{\text{air}}^{-3}$ ) by multiplying aqueous concentrations by the LWC and dividing by the mass density of water. This study uses air-equivalent concentrations for all species with the exception of  $\text{H}^+$  (pH) that uses aqueous concentration.

A total of 80 species (29 measured ionic species, 46 measured elemental species, measured pH, and 4 NSS calculated species; Table 2) were considered in this study as an initial pool of candidate species that could potentially be used to predict  $N_d$ . To facilitate the statistical analysis in this study, the amount of chemical species were filtered from 80 to only

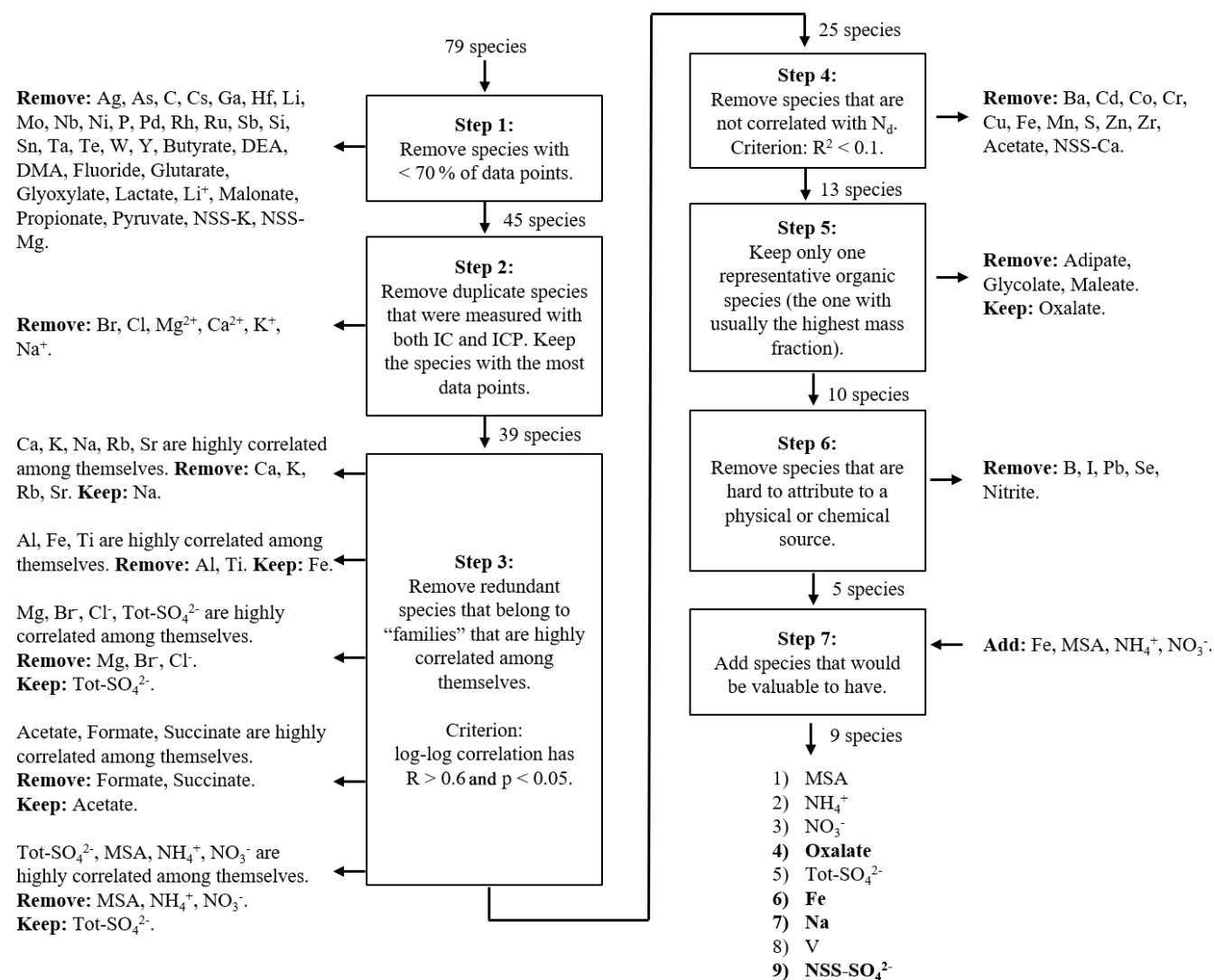
9. The steps used in this filtering process are summarized in the next section.

### 2.4 Filtering of chemical species

A focus in this study is to identify appropriate chemical species to use as predictors in a linear regression model (addressed in Sect. 2.5). Good statistical practice (e.g., Freund et al., 2010) recommends that two conditions must be met to produce a meaningful multivariable regression: (1) the independent/predictor variables must not be redundant, i.e., they must not be highly correlated among themselves (the property of high correlation is called collinearity), and (2) each independent/predictor variable must have some correlation with the dependent/response variable. There is no universal rule to define what is “highly” correlated; rather, it depends on the nature of the data and the user’s judgment.

As using all 80 species is impractical in terms of providing results that could be tested and/or used by others, a filtering method was used to reduce the number of species. The filtering method consisted of seven steps (Fig. 2), the objective of which was to trim the total number of species by an order of magnitude, leaving just a few that exhibited the following conditions: (1) the highest data quality and quantity, (2) the least redundancy among themselves, (3) the highest correlation with  $N_d$ , and (4) the most physical meaning. The decision to remove a species becomes less objective and quantifiable towards the last steps in Fig. 2. Each step is described below.

Step 1 removed species with less than 70 % of data points. A species could have a low amount of points because it was not analyzed in a field campaign or because the data quality from the IC or ICP (ICP-MS or ICP-QQQ) was inadequate. Step 2 removed duplicate species that were measured by both IC and ICP. Step 3 addressed condition (2) by removing species that were collinear (i.e., correlated among themselves). The criterion for a “high” correlation was to have a correlation coefficient ( $R$ )  $> 0.6$  and a  $p$  value  $< 0.05$ . For example, if a fixed number of five species were all highly correlated between each other, then only one of the five species was kept and the rest were removed. This procedure is to consolidate “families” of three or more highly correlated species to a single species and does not apply to pairs of highly con-



**Figure 2.** Algorithm used to filter the number of species from 80 to 9. The four species in bold font are the ones used in Sect. 3.3. ICP represents ICP-MS + ICP-QQ.

solidated species. Step 4 addressed condition (3) by removing species that were not correlated to  $N_d$ . The criterion for a "low" correlation was to have a coefficient of determination ( $R^2$ )  $< 0.1$ . Notice that step 3 uses  $R$ , whereas step 4 uses  $R^2$ ; this is because collinearity is determined not only by the value of  $R$  but also the sign of  $R$ . Step 5 removes all but one organic species, oxalate (Ox), since this species generally had the highest mass concentration of all the organic species and was considered to be representative of all other organic species. Step 6 removed species that could not easily be attributed to a physical process or chemical source. Step 7 added back into the analysis four species that had been removed. This was done for the sake of having species that are known to have relevant sources in the study region. Even though pH plays an important role in the partitioning of gases into particles and droplets, in addition to influencing aqueous

reactions in droplets (e.g., Pye et al., 2020), pH was filtered out in step 4 for being a poor predictor of  $N_d$ .

The nine species that survived the filtering scheme in Fig. 2 are methanesulfonic acid (MSA), ammonium (NH<sub>4</sub><sup>+</sup>), NO<sub>3</sub><sup>-</sup>, Ox, Tot-SO<sub>4</sub><sup>2-</sup>, NSS-SO<sub>4</sub><sup>2-</sup>, Fe, Na, and vanadium (V). These species have known sources as follows. MSA: ocean biogenic (Sorooshian et al., 2009); NH<sub>4</sub><sup>+</sup>: agriculture (Bauer et al., 2016), marine emissions (Bouwman et al., 1997), and wildfires (Reid et al., 1998); NO<sub>3</sub><sup>-</sup> and Ox: fire (Prabhakar et al., 2014; Maudlin et al., 2015); Tot-SO<sub>4</sub><sup>2-</sup>: sea salt (Seinfeld and Pandis, 2016), ocean biogenic (Charlson et al., 1987), and shipping (Coggon et al., 2012), with NSS-SO<sub>4</sub><sup>2-</sup> missing the sea salt contribution; Fe: dust (Jickells et al., 2005) and fire (Maudlin et al., 2015); Na: sea salt (Seinfeld and Pandis, 2016); and V: shipping (Wang et al., 2014). Note that we retained both Tot-SO<sub>4</sub><sup>2-</sup> and NSS-SO<sub>4</sub><sup>2-</sup>; this is to eval-

**Table 2.** Summary of chemical species analyzed in this study. IC = ion chromatography; ICP represents ICP-MS or ICP-QQQ. Note: NSS species, with the exception of  $\text{NSS-SO}_4^{2-}$ , were calculated using elements, not ions, hence they have no superscript charge.

Elements (ICP)	
1	Ag
2	Al
3	As
4	B
5	Ba
6	Br
7	C
8	Ca
9	Cd
10	Cl
11	Co
12	Cr
13	Cs
14	Cu
15	Fe
16	Ga
17	Hf
18	I
19	K
20	Li
21	Mg
22	Mn
23	Mo
24	Na
25	Nb
26	Ni
27	P
28	Pb
29	Pd
30	Rb
31	Rh
32	Ru
33	S
34	Sb
35	Se
36	Si
37	Sn
38	Sr
39	Ta
40	Te
41	Ti
42	V
43	W
44	Y
45	Zn
46	Zr

**Table 2.** Continued.

Inorganic ions (IC)	
47	Ammonium ( $\text{NH}_4^+$ )
48	Bromide ( $\text{Br}^-$ )
49	Calcium ( $\text{Ca}^{2+}$ )
50	Chloride ( $\text{Cl}^-$ )
51	Fluoride ( $\text{F}^-$ )
52	Lithium ( $\text{Li}^+$ )
53	Magnesium ( $\text{Mg}^{2+}$ )
54	Methanesulfonic acid (MSA)
55	Nitrate ( $\text{NO}_3^-$ )
56	Nitrite ( $\text{NO}_2^-$ )
57	Potassium ( $\text{K}^+$ )
58	Sodium ( $\text{Na}^+$ )
59	Sulfate ( $\text{SO}_4^{2-}$ )
Amines (IC)	
60	Diethylamine (DEA)
61	Dimethylamine (DMA)
NSS species (calculated)	
62	NSS calcium (NSS-Ca)
63	NSS potassium (NSS-K)
64	NSS magnesium (NSS-Mg)
65	NSS sulfate (NSS- $\text{SO}_4^{2-}$ )
Organic ions (IC)	
66	Acetate
67	Adipate
68	Butyrate
69	Formate
70	Glutarate
71	Glycolate
72	Glyoxylate
73	Lactate
74	Maleate
75	Malonate
76	Oxalate
77	Propionate
78	Pyruvate
79	Succinate
Acidity (pH)	
80	Hydrogen ion ( $\text{H}^+$ )

1994; Boucher and Lohmann, 1995). Section 3.1 and 3.2 will discuss these nine species, and the rest of Sect. 3 will focus on only four species to be explained later. These species were analyzed by a multivariable regression model, which is described in the next section.

uate which correlates more with  $N_d$ , as some studies have used  $\text{Tot-SO}_4^{2-}$  (e.g., Leaitch et al., 1992; Saxena and Menon, 1999), whereas others have used  $\text{NSS-SO}_4^{2-}$  (Novakov et al.,

## 2.5 Mathematical model

This study examines the relationship between cloud water mass concentration and  $N_d$  with a multivariable linear model similar to that of McCoy et al. (2017, 2018):

$$\log(N_d) = a_0 + a_1 \log(M_1) + a_2 \log(M_2) + \dots + a_n \log(M_n), \quad (3)$$

where  $M_i$  is the air-equivalent mass concentration of species  $i$  ( $\mu\text{g m}^{-3}$ ),  $a_i$  represent fitting parameters, and  $n$  is the number of species being considered.  $N_d$  is the dependent (or response) variable and  $M_1, M_2, \dots, M_n$  are the independent (or predictor) variables. The logarithmic forms of  $N_d$  and  $M_i$  were correlated to account for a numerically large range of several orders of magnitude and because a log–log model is commonly used to correlate chemical composition to  $N_d$  (e.g., Boucher and Lohmann, 1995; Menon et al., 2002; McCoy et al., 2017).

The MATLAB software package was used to obtain multivariable linear regressions of the form of Eq. (3) using the method of ordinary least squares. The performance of a regression was quantified using the coefficient of determination ( $R^2$ ). However, when comparing the performance of correlations between regressions using a different number of predictor variables, it is necessary to use the adjusted coefficient of determination ( $R_{\text{adj}}^2$ ), which is subscripted to distinguish it from the ordinary  $R^2$ , and is adjusted by using the number of predictors ( $P$ ) and the number of data points ( $N$ ) via the formula  $R_{\text{adj}}^2 = 1 - (1 - R^2)(N - 1)/(N - P - 1)$  (Kahane, 2008). For a large number of data points,  $R_{\text{adj}}^2 \approx R^2$ ; however, for the sake of rigor and consistency,  $R_{\text{adj}}^2$  is used instead of the ordinary  $R^2$ , except when reporting values from the literature. The statistical significance of correlations was quantified using the  $p$  value obtained by doing a two-tailed Student's  $t$  test. Both  $R_{\text{adj}}^2$  and  $p$  values were given by the MATLAB software after regression.  $p$  values were obtained for both the overall regression and each individual coefficient in the regression, e.g., if a regression has three predictors, there are a total of five  $p$  values: one for the overall regression, three for the slope of each individual predicting variable, and one for the intercept. In this study, a regression was considered to be statistically significant if all the  $p$  values were  $< 0.05$ .

The correct functioning of the method of ordinary least squares requires that the set of  $n$  predicting variables in Eq. (3) not be collinear. Multicollinearity is defined by a set of three or more predicting variables being collinear. Using a set of multicollinear predictors can produce unreliable estimates in both magnitude and sign of the coefficients ( $a_i$ ) (Kahane, 2008). There is no universal marker for multicollinearity. Furthermore, multicollinearity can only be addressed when analyzing all predictors together. For example, for a given set of three predictors ( $P_1, P_2$ , and  $P_3$ ),

even though the pairs  $P_1 - P_2$ ,  $P_1 - P_3$ , and  $P_2 - P_3$  are not collinear, there is no guarantee that the  $P_1 - P_2 - P_3$  set is not multicollinear. When considering a complex system such as the chemical composition of cloud water, it is reasonable to assume that as more species are used to predict  $N_d$ , the higher the probability that the set of species is multicollinear. We did not test for multicollinearity in this study; the consequences of not doing so are explored in Sect. 3.2.

## 2.6 Calculation of turbulence

Similar to Leaitch et al. (1992) and Feingold et al. (1999), this study analyzes the effect of turbulence on the ability to predict  $N_d$ . Turbulence was considered to be represented by the standard deviation of the vertical wind speed ( $w$ ) and is represented as  $\sigma_w$ . Also similar to Leaitch et al. (1992), this study classified conditions into turbulent and smooth regimes by considering the upper and lower 33rd percentile of  $\sigma_w$ , respectively. Although the rigorous approach to calculate  $\sigma_w$  uses  $w$  from below the cloud (Twomey, 1959), this study used vertical wind speed data collected throughout the sampling time (i.e., mostly inside the cloud but also outside the cloud). This was mainly because not all cloud water samples had an accompanying measurement of  $w$  below the cloud. To justify using  $\sigma_w$  from the sampling time instead of below-cloud  $\sigma_w$ , consider Fig. S1 in the Supplement, which shows a representative time series of altitude,  $w$ , and  $\sigma_w$  for a cloud water sample that was collected minutes before a below-cloud leg, which collected measurements of  $w$ . It can be seen that the plots of  $w$  and  $\sigma_w$  are similar and that an average  $\sigma_w$  calculated either way is still in the bottom 33rd percentile. Therefore, for the purposes of this study, we consider in-cloud turbulence to reasonably approximate below-cloud turbulence.

## 2.7 Determination of smoke influence

One of the objectives of this study is to analyze the extent to which the presence of smoke from wildfires affects the correlation between  $N_d$  and cloud water chemical composition. Thus, it was important to identify cloud water samples that were influenced by smoke. Only the NiCE and FASE campaigns were affected by wildfires. Mardi et al. (2018) identified vertical soundings in the NiCE and FASE campaigns that were influenced by smoke by establishing smoke influence to have a total aerosol number concentration ( $N_a$ )  $\geq 1000 \text{ cm}^{-3}$ , as measured by the PCASP, in addition to visual and olfactory detection of smoke by flight scientists. In this study, a cloud water sample was considered to be influenced by smoke if it was collected during a research flight (RF) that contains a vertical sounding identified by Mardi et al. (2018) to be influenced by smoke, even if the cloud water sample was not necessarily collected near the sounding labeled as smoke influenced; this is a valid assumption based on the work of Mardi et al. (2019). The RFs considered to



**Table 3.** Summary of one-predictor models for  $N_d$  based on using any of nine of the final chemical species that were identified after applying the filtering scheme shown in Fig. 2. The coefficients correspond to a linear model of the form  $\log(N_d) = a_0 + a_1 \log(M_i)$ , where  $M_i$  is the mass concentration of species  $i$ .

Species	$R_{\text{adj}}^2$	Coefficients	
		$a_0$	$a_1$
Tot-SO <sub>4</sub> <sup>2-</sup>	0.40	2.05	0.32
NH <sub>4</sub> <sup>+</sup>	0.34	2.33	0.25
NSS-SO <sub>4</sub> <sup>2-</sup>	0.29	2.13	0.28
MSA	0.26	2.37	0.31
NO <sub>3</sub> <sup>-</sup>	0.24	2.12	0.25
Na	0.19	2.03	0.13
Ox	0.15	2.26	0.18
V	0.14	2.61	0.15
Fe	0.05	2.26	0.09

be smoke influenced in this study were NiCE RFs 16–23 and FASE RFs 3–11 and 13–15.

### 3 Results and discussion

With the refined list of nine physically meaningful species from Sect. 2.4, we now proceed to address the following questions: (1) what single species best predicts  $N_d$ ; (2) how many species are sufficient to predict  $N_d$ ; (3) what is an effective combination of species to predict  $N_d$ ; and (4) how do several factors (i.e., turbulence, smoke-influence, and location along cloud depth) affect the ability to reliably predict  $N_d$ . These questions are addressed in order in Sect. 3.1–3.4.

#### 3.1 Single-variable prediction of $N_d$

In this section, we analyze which of the nine species filtered out in Sect. 2.4 best predicts  $N_d$  by itself without binning by external factors. These single-predictor regressions with no binning are important, as they provide a baseline for subsequent sections in which multi-predictor regressions and binning are used. Table 3 and Fig. 3 display the ability of each of the nine species to predict  $N_d$ . To have consistency with subsequent sections,  $R_{\text{adj}}^2$  is used instead of the ordinary  $R^2$ . The regression and the individual coefficients were all statistically significant.

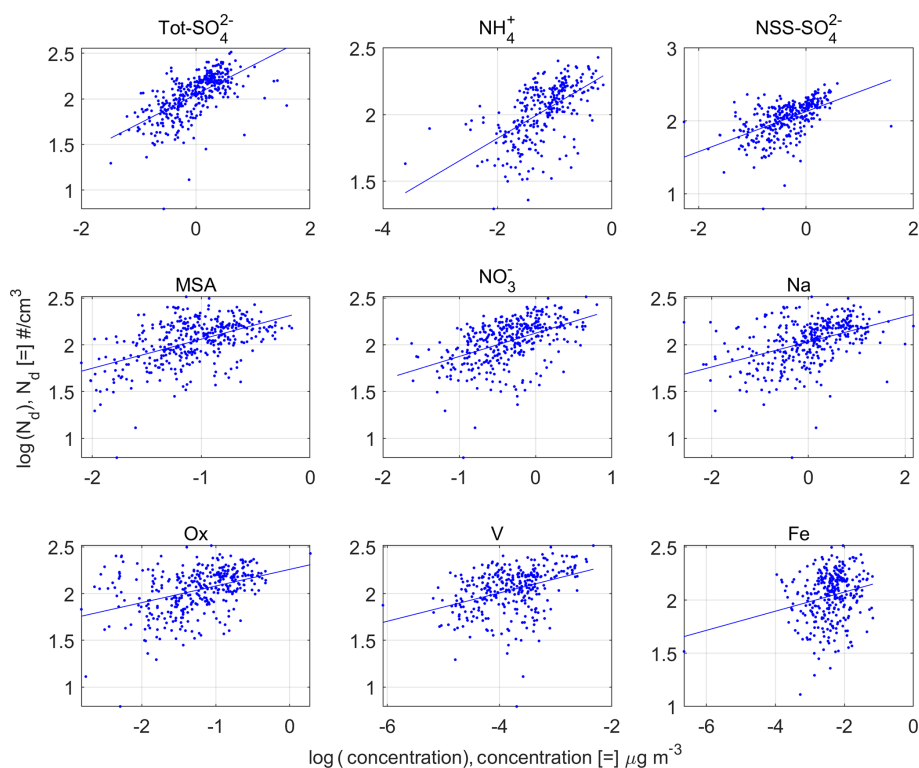
Some previous studies predicted  $N_d$  using Tot-SO<sub>4</sub><sup>2-</sup> (e.g., Leaitch et al., 1992; Saxena and Menon, 1999), whereas other studies used NSS-SO<sub>4</sub><sup>2-</sup> (e.g., Novakov et al., 1994; Lowenthal et al., 2004). We find that Tot-SO<sub>4</sub><sup>2-</sup> is the best predictor and that it is better correlated to  $N_d$  ( $R_{\text{adj}}^2 = 0.40$ ) than NSS-SO<sub>4</sub><sup>2-</sup> ( $R_{\text{adj}}^2 = 0.29$ ). This is likely because Tot-SO<sub>4</sub><sup>2-</sup> encompasses both sea salt particles and non-sea-salt particles and thus gives a better approximation to the to-

tal number concentration of CCN. In addition, Tot-SO<sub>4</sub><sup>2-</sup> also had the largest slope ( $a_1 = 0.32$ ), suggesting that  $N_d$  is more sensitive to changes in Tot-SO<sub>4</sub><sup>2-</sup> than other chemical species. Although HNO<sub>3</sub> has been observed to increase  $N_d$  (e.g., Xue and Feingold, 2004), NO<sub>3</sub><sup>-</sup> was found to be only moderately correlated with  $N_d$  ( $R_{\text{adj}}^2 = 0.24$ ). The species with the lowest correlation was Fe ( $R_{\text{adj}}^2 = 0.05$ ). This low correlation with  $N_d$  was also presented by other crustal metals like Al ( $R_{\text{adj}}^2 = 0.01$ ) and Ti ( $R_{\text{adj}}^2 \sim 0$ ) (not shown in Table 3). The low influence of crustal metals on  $N_d$  is consistent with the findings of Lowenthal and Borys (2000). Some physical meaning can be extracted from the intercept of the regression ( $a_0$ ). If  $N_d$  is insensitive to the mass concentration of a species, then the slope ( $a_1$ ) should be zero, and  $N_d$  would be constant with a value of  $N_d = 10^{a_0}$ . These intercepts yield a range of  $N_d$  of 108–412 cm<sup>-3</sup>. These values are not unrealistic in clouds in this study region (e.g., Chen et al., 2012; Lu et al., 2009; Wang et al., 2016).

To contrast with results of this work, Table 4 shows the regression parameters from other studies when correlating  $N_d$  and SO<sub>4</sub><sup>2-</sup>. For the sake of completeness, Table 4 shows regressions that analyzed non-marine stratocumulus clouds, but in this comparison, we focus only on those regressions that analyzed stratocumulus clouds. Our results (i.e.,  $a_i$  coefficients and  $R^2$ ) for Tot-SO<sub>4</sub><sup>2-</sup> reasonably match the results of Leaitch et al. (1992), suggestive of commonality between two coastal regions with differing meteorological conditions (i.e., northeast Pacific vs northwest Atlantic) (Sorooshian et al., 2019). Our results for NSS-SO<sub>4</sub><sup>2-</sup> also reasonably match those of McCoy et al. (2017), which is noteworthy as McCoy et al. (2017) used satellite retrievals and model aerosol concentrations for several stratocumulus decks around the world, whereas our analysis used in situ data from a relatively small region. However, our NSS-SO<sub>4</sub><sup>2-</sup> results differ significantly from those of Novakov et al. (1994), which is understandable since the regression presented by Novakov et al. (1994) has a  $p$  value > 0.05. Our data set does not achieve the degree of correlation achieved by Lowenthal et al. (2004), who report the highest correlation for marine clouds ( $R^2 = 0.82$ ). The studies that analyzed stratocumulus clouds all report intercept values ( $a_0$ ) of  $\sim 2.0$ , which is consistent with our data.

#### 3.2 Multi-variable prediction of $N_d$

When previous studies correlated  $N_d$  (or  $N_{\text{CCN}}$ ) and the air-equivalent concentration of chemical species and obtained a poor correlation, it was suggested that taking more chemical species into consideration would improve the correlation (e.g., Leaitch et al., 1992; Novakov et al., 1994). In this section we address the following issue: “How many chemical species are necessary to adequately predict  $N_d$ ”. To answer this question, we use the nine filtered species from Sect. 2.4. Regressions of the form of Eq. (3) are performed for every combination of species. The number of predictors in the regressions are varied from one up to eight. The number of

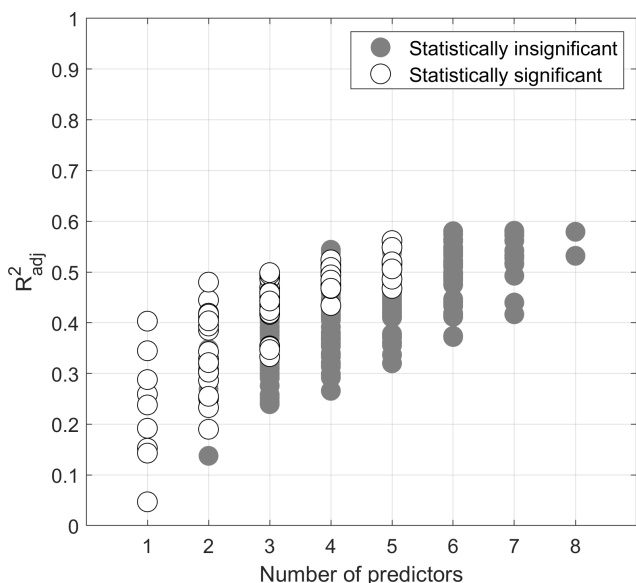


**Figure 3.** Scatter plot for the nine filtered species from Fig. 2. The lines are linear regression models of the form  $\log(N_d) = a_0 + a_1 \log(M_i)$ , where  $M_i$  is the mass concentration of species  $i$ .

**Table 4.** Comparison of coefficient values for studies that correlate  $N_d$  to  $\text{SO}_4^{2-}$  (total or non-sea-salt). The coefficients correspond to a linear model of the form  $\log(N_d) = a_0 + a_1 \log(\text{SO}_4^{2-})$ .

Reference	$a_0$	$a_1$	$\text{SO}_4^{2-}$	$R^2$	Cloud type
Leaitch et al. (1992) <sup>a</sup>	1.95	0.257	Tot	0.3	Continental stratocumulus
	2.33	0.186	Tot	0.49	Continental cumulus
Novakov et al. (1994)	2.323	0.091	NSS	0.50 <sup>b</sup>	Marine stratocumulus
	2.43	-0.056	NSS	0.03	Marine cumulus
Van Dingenen et al. (1995) <sup>c</sup>	2.33	0.4	NSS	0.42	All cloud types combined
Boucher and Lohmann (1995) <sup>c</sup>	2.24	0.257	NSS	<sup>d</sup>	Continental stratus
	2.54	0.186	NSS	<sup>d</sup>	Continental cumulus
	2.06	0.48	NSS	<sup>d</sup>	Marine
	2.21	0.41	NSS	<sup>d</sup>	All cloud types combined
Saxena and Menon (1999)	0.67	0.66	Tot	<sup>d</sup>	Continental orographic clouds
Lowenthal et al. (2004)	2.32	0.74	NSS	0.82	Marine
	2.38	0.49	NSS	0.66	Continental
	2.39	0.5	NSS	0.81	Combined
McCoy et al. (2017)	2.11	0.41	NSS	0.36	Marine stratocumulus

<sup>a</sup> The units of  $\text{SO}_4^{2-}$  for this regression are nanoequivalents per cubic meter ( $\text{nEq m}^{-3}$ ). All other studies report  $\text{SO}_4^{2-}$  in units of micrograms per cubic meter ( $\mu\text{g m}^{-3}$ ). However, the value of the slope ( $a_1$ ) is not affected by the units of concentration. <sup>b</sup> The  $R^2$  has a  $p > 0.05$  due to having few data points. <sup>c</sup> These regressions were made using data compiled from several studies and assume that  $N_{\text{CCN}} \approx N_d$ . <sup>d</sup> Study does not report  $R^2$ .



**Figure 4.** Plot showing which of the 383 regressions are statistically significant. This plot ignores the regressions that use both  $\text{NSS-SO}_4^{2-}$  and  $\text{Tot-SO}_4^{2-}$  simultaneously.

combinations ( $C$ ) that can be made with  $P$  predictors selected from  $S$  species is  $C = S! / (P!(S - P)!)$ . Combinations that include  $\text{Tot-SO}_4^{2-}$  and  $\text{NSS-SO}_4^{2-}$  together are not considered, thus leaving a total of 383 regressions.

Of the total 383 regression, only 67 were considered statistically significant. Figure 4 shows the  $R_{\text{adj}}^2$  as a function of the number of predictors for both statistically significant and insignificant regressions; the percentage of regressions that were statistically significant is shown in Table S2. These results show that adding more predictors does not necessarily improve the correlation, as all correlations that use six or more predictors are statistically insignificant. This behavior is perhaps because the new species being added are redundant with respect to the species that are already in the model (i.e., the new species is mathematically collinear with the old species). It is also interesting to note how  $R_{\text{adj}}^2$  increases asymptotically to  $\sim 0.6$ ; this further makes the point that additional species do not necessarily improve predictability of  $N_d$ . The same asymptotic behavior is also exhibited with  $R^2$ , as  $R^2$  and  $R_{\text{adj}}^2$  for these regressions differ by only  $\sim 2\%$ .

We examined the best regressions produced by a given number of predictors to explore the factors that contribute to a respectable multivariable regression. Table 5 shows the three statistically significant regressions that had the highest  $R_{\text{adj}}^2$  for a given number of predictors (one to five). The predictors are ordered horizontally according to the value of their coefficient in order to show qualitatively which species is more dominant in a regression. Eight of the nine chemical species considered appear at least once in a regression, with the most common species being  $\text{NH}_4^+$ , a form of  $\text{SO}_4^{2-}$  (total or non-sea-salt), Na, Ox, and MSA. Sulfate (total or non-sea-

salt) appears in 12 of the 15 regressions, and in eight regressions it has the largest coefficient; this speaks to the importance of  $\text{SO}_4^{2-}$  in predicting  $N_d$ . However, the appearance of Na and Ox and their non-negligible slope also highlights the importance of considering them as well in a correlation; this is clearly observed in the increase of  $R_{\text{adj}}^2$  when Na and Ox are added to a regression that contains only  $\text{NSS-SO}_4^{2-}$  (Table 6). We believe that the ingredients that yield the higher  $R_{\text{adj}}^2$  in Table 5 are (1) a form of  $\text{SO}_4^{2-}$  (such  $\text{Tot-SO}_4^{2-}$  or  $\text{NSS-SO}_4^{2-}$ ), (2) a sea emissions tracer (such as Na), and (3) an organic tracer (such as Ox).  $\text{NH}_4^+$  was present in all the regressions; however, given that it comes from diverse sources such as agriculture (ApSimon et al., 1987; Bauer et al., 2016), marine emissions (Bouwman et al., 1997; Paulot et al., 2015), and wildfires (Maudlin et al., 2015; Reid et al., 1998), it is difficult to assess if it contributes to the CCN budget or simply accompanies all types of CCN. In other words, we suspect that  $\text{NH}_4^+$  appears in all correlations because it generally accompanies the three ingredients we propose make a good correlation: a form of  $\text{SO}_4^{2-}$ , a marine emissions tracer, and an organic tracer.

It is of interest to note that combining a sea salt tracer (such as Na) with  $\text{NSS-SO}_4^{2-}$  in a two-predictor model has about the same performance ( $R_{\text{adj}}^2 = 0.41$ ; Table 6) as a one-predictor model using  $\text{Tot-SO}_4^{2-}$  ( $R_{\text{adj}}^2 = 0.40$ ; Table 3). We believe this is because  $\text{Tot-SO}_4^{2-}$  encompasses the sea salt and the non-sea-salt contribution to CCN about the same as the artificial mathematical separation of the two. Also of interest is that when only looking at the statistically significant regressions, only 17 regressions have species with negative coefficients (i.e., negative slopes). The species with negative coefficients are  $\text{NO}_3^-$ , Fe, and V (not shown); more specifically,  $\text{NO}_3^-$ , Fe, and V have negative coefficients when they are accompanied by  $\text{NH}_4^+$  in the same regression. The physical reason as to why these species have negative coefficients when mixed with  $\text{NH}_4^+$  is not clear; perhaps the reason is due to the mathematics of the regression and not physically rooted, as multicollinearity can lead to unexpected magnitudes and signs for predictor coefficients (Kahane, 2008). In addition, multicollinearity will become more likely as more predictors as considered. Therefore, it is not surprising that unexpected negative coefficients only appear when considering many (five) predictors. Lastly, a correlation matrix among the nine predicting species (Fig. S2) shows a strong correlation for some pairs of species ( $\text{NH}_4^+ - \text{NO}_3^-$ :  $R_{\text{adj}}^2 = 0.48$ ;  $\text{NO}_3^- - \text{V}$ :  $R_{\text{adj}}^2 = 0.49$ ) and moderate correlation for other pairs ( $\text{NH}_4^+ - \text{V}$ :  $R_{\text{adj}}^2 = 0.27$ ;  $\text{NO}_3^- - \text{Fe}$ :  $R_{\text{adj}}^2 = 0.22$ ), thus strengthening the argument that the negative coefficients are due to mathematical multicollinearity and not a physical or chemical reason.

When considering a multispecies model to predict  $N_d$ , it is worthwhile to examine the coefficient of sea salt. Even though it is well established that more CCN leads to more

**Table 5.** The top three statistically significant regressions with the highest  $R_{\text{adj}}^2$  for a given number of predictors. The coefficients correspond to a linear model of the form  $\log(N_d) = a_0 + \sum a_i \log(P_i)$ .

No. of Predictors	Predictors ( $P_i$ ) and their respective coefficients ( $a_i$ )											$R_{\text{adj}}^2$
	$a_0$	$a_1$	$P_1$	$a_2$	$P_2$	$a_3$	$P_3$	$a_4$	$P_4$	$a_5$	$P_5$	
1	2.05	0.32	Tot-SO <sub>4</sub> <sup>2-</sup>									0.40
	2.33	0.25	NH <sub>4</sub> <sup>+</sup>									0.34
	2.13	0.28	NSS-SO <sub>4</sub> <sup>2-</sup>									0.29
2	2.18	0.22	Tot-SO <sub>4</sub> <sup>2-</sup>	0.12	NH <sub>4</sub> <sup>+</sup>							0.48
	2.43	0.21	MSA	0.15	NH <sub>4</sub> <sup>+</sup>							0.44
	2.25	0.19	NH <sub>4</sub> <sup>+</sup>	0.09	Na							0.42
3	2.25	0.13	NSS-SO <sub>4</sub> <sup>2-</sup>	0.13	NH <sub>4</sub> <sup>+</sup>	0.10	Na					0.50
	2.24	0.19	Tot-SO <sub>4</sub> <sup>2-</sup>	0.10	Ox	0.07	NH <sub>4</sub> <sup>+</sup>					0.49
	2.25	0.17	Tot-SO <sub>4</sub> <sup>2-</sup>	0.11	NH <sub>4</sub> <sup>+</sup>	0.08	MSA					0.49
4	2.32	0.21	Tot-SO <sub>4</sub> <sup>2-</sup>	0.20	Ox	0.09	NH <sub>4</sub> <sup>+</sup>	-0.15	NO <sub>3</sub> <sup>-</sup>			0.52
	2.29	0.11	NSS-SO <sub>4</sub> <sup>2-</sup>	0.10	Ox	0.09	Na	0.08	NH <sub>4</sub> <sup>+</sup>			0.51
	2.31	0.11	NH <sub>4</sub> <sup>+</sup>	0.10	NSS	0.10	MSA	0.08	Na			0.51
5	2.10	0.13	Na	0.12	Ox	0.11	NSS-SO <sub>4</sub> <sup>2-</sup>	0.08	NH <sub>4</sub> <sup>+</sup>	-0.05	V	0.56
	2.40	0.23	Ox	0.13	NSS-SO <sub>4</sub> <sup>2-</sup>	0.10	NH <sub>4</sub> <sup>+</sup>	0.09	Na	-0.17	NO <sub>3</sub> <sup>-</sup>	0.55
	2.36	0.14	NH <sub>4</sub> <sup>+</sup>	0.14	MSA	0.12	NSS-SO <sub>4</sub> <sup>2-</sup>	0.07	Na	-0.08	NO <sub>3</sub> <sup>-</sup>	0.52

**Table 6.** Comparison of regressions containing NSS-SO<sub>4</sub><sup>2-</sup>, Na, and Ox.

No. of Predictors	Predictors ( $P_i$ ) and their respective coefficients ( $a_i$ )								$R_{\text{adj}}^2$
	$a_0$	$a_1$	$P_1$	$a_2$	$P_2$	$a_3$	$P_3$		
1	2.13	0.28	NSS-SO <sub>4</sub> <sup>2-</sup>					0.29	
2	2.12	0.23	NSS-SO <sub>4</sub> <sup>2-</sup>	0.12	Na			0.40	
	2.26	0.24	NSS-SO <sub>4</sub> <sup>2-</sup>	0.12	Ox			0.34	
3	2.22	0.22	NSS-SO <sub>4</sub> <sup>2-</sup>	0.10	Na	0.08	Ox	0.42	

droplets, the effect of giant CCN (GCCN), such as sea salt, is not as clear. Cloud microphysics studies suggest two mechanisms by which more sea salt leads to less  $N_d$ . (1) The large size and highly hygroscopic nature of sea salt causes these particles to activate into droplets before other smaller particles. This reduces the amount of available water vapor and creates unfavorable conditions for smaller particles to nucleate into droplets (e.g., Andreae and Rosenfeld, 2008). (2) GCCN nucleate into larger droplets as compared to CCN, which in turn are more likely to collide and coalesce with surrounding droplets. This combination of droplets creates larger but fewer droplets and ultimately leads to the formation of rain drops and precipitation (e.g., Feingold et al., 1999; Jung et al., 2015). Therefore, it is expected that the

negative correlation between GCCN and  $N_d$  should translate into a negative coefficient for Na (the sea salt tracer) in a multi-predictor regression equation. However, this behavior was not observed in this study. A plausible explanation for this discrepancy is that the effect of GCCN on  $N_d$  is highly dependent on conditions like LWC and  $N_d$  itself (e.g., Feingold et al., 1999) and that this study did not capture the appropriate conditions to observe this effect. However, McCoy et al. (2017) did observe a negative coefficient for sea salt and ascribed it to a simulation artifact caused by the intimate link between sea salt generation and wind speed (i.e., turbulence). An attempt to isolate the effects of sea salt and turbulence on  $N_d$  is provided in Sect. 3.1.1.

**Table 7.** Results of multivariable regressions from previous studies that have correlated  $N_d$  to mass concentrations. The regression corresponds to a model like Eq. (3). OM represents organic matter, SS represents sea salt, BC represents black carbon, and DU represents dust.

Reference	Predictors ( $P_i$ ) and their respective coefficients ( $a_i$ )										$R^2$	Cloud type
	$a_0$	$a_1$	$P_1$	$a_2$	$P_2$	$a_3$	$P_3$	$a_4$	$P_4$			
Menon et al. (2002)*	2.41	0.50	NSS-SO <sub>4</sub> <sup>2-</sup>	0.13	OM							Continental
	2.41	0.50	NSS-SO <sub>4</sub> <sup>2-</sup>	0.13	OM	0.05	SS					Marine
McCoy et al. (2017)	1.78	0.31	NSS-SO <sub>4</sub> <sup>2-</sup>	-0.19	SS	0.057	BC	0.031	DU	0.44		Marine stratocumulus (global average)
McCoy et al. (2018)	2.03	0.2	NSS-SO <sub>4</sub> <sup>2-</sup>	-0.04	SS	-0.03	BC	0	DU	0.08		Marine stratocumulus (just Californian coast)

\* This study obtains data from other studies and calculates organic matter.

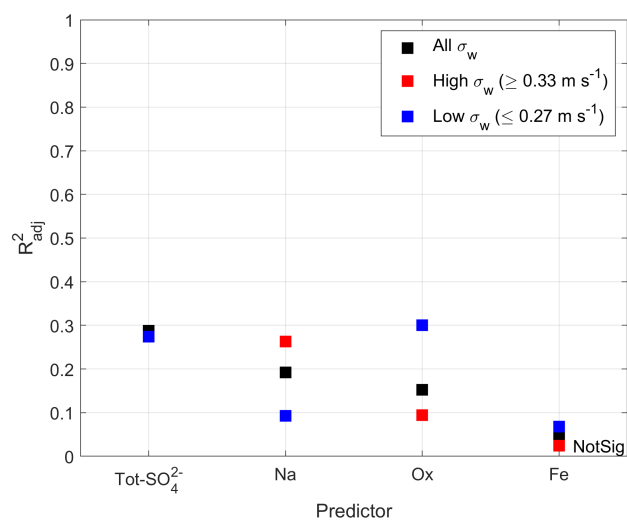
Menon et al. (2002) and McCoy et al. (2017, 2018) are among the few studies that have used multiple species to predict  $N_d$  (Table 7). Menon et al. (2002) used three species (sulfate, organic matter, and sea salt). McCoy et al. (2017, 2018) used five species (sulfate, sea salt, black carbon, organic carbon, and dust), but the 2017 study found the contribution of organic matter to be negligible. In order to intercompare results with previous studies, we selected species homologous to those of McCoy et al. (2017, 2018). We select NSS-SO<sub>4</sub><sup>2-</sup> for sulfate, Na for sea salt, oxalate for organic carbon, and Fe for dust. We did not measure a species analogous to black carbon. The subsequent analysis examines only these four species using single-predictor regressions.

### 3.3 Analysis of meteorological factors through binning

Historically, the effect that meteorological factors have on the composition- $N_d$  (or composition- $N_{CCN}$ ) empirical relationship has been examined by analyzing regressions after binning by turbulence (Leitch et al., 1996), cloud type (Leitch et al., 1992; Novakov and Penner, 1993), and region (McCoy et al., 2018). The following sections address the effects of turbulence, smoke influence, and location along cloud depth.

#### 3.3.1 Effect of turbulence

Building upon the work of Leitch et al. (1996), who studied how turbulence affects the correlation between Tot-SO<sub>4</sub><sup>2-</sup> and  $N_d$ , this study extends that analysis to examine four additional species. Similar to Leitch et al. (1996), this study quantified turbulence by the standard deviation of vertical wind speed ( $\sigma_w$ ). Our range of  $\sigma_w$  was 0.10–0.51 m s<sup>-1</sup>. Low turbulence was considered to be in the bottom 33rd percentile ( $\leq 0.27$  m s<sup>-1</sup>), whereas high turbulence was taken to be values in the top 33rd percentile ( $\geq 0.33$  m s<sup>-1</sup>). Leitch et al. (1996) considered low and high turbulence to be  $\sigma_w < 0.17$  m s<sup>-1</sup> and  $\sigma_w > 0.23$  m s<sup>-1</sup>, respectively, and it is worth noting that only 5 of our 385 samples are considered low tur-



**Figure 5.** Effect of turbulence (quantified using  $\sigma_w$ ) on the ability of a single species to predict  $N_d$ . For NSS-SO<sub>4</sub><sup>2-</sup>; the high (red) and low (blue)  $\sigma_w$  data points overlap. NotSig represents not statistically significant according to the definition in Sect. 2.5.

bulence according to the criterion of Leitch et al. (1996). Figure 5 and Table 8 show how  $R^2_{adj}$  depends on the predicting species and the turbulence regime; the scatterplots from which the  $R^2_{adj}$  values are taken are shown in Fig. S3.

For NSS-SO<sub>4</sub><sup>2-</sup>, there is no significant difference in  $R^2_{adj}$  when comparing all the points or by binning by  $\sigma_w$ . However, this is not the case for Tot-SO<sub>4</sub><sup>2-</sup>, in which there is a large difference in the degree of correlation ( $R^2_{adj} = 0.27$  and  $R^2_{adj} = 0.55$  for low  $\sigma_w$  and high  $\sigma_w$ , respectively). This is in agreement with Leitch et al. (1996), in which the correlation (albeit, not log–log) between Tot-SO<sub>4</sub><sup>2-</sup> and  $N_d$  yielded an  $R^2 = 0.53$  and  $R^2 = 0.91$  for low and high  $\sigma_w$ , respectively. The difference in the behavior between Tot-SO<sub>4</sub><sup>2-</sup> and NSS-SO<sub>4</sub><sup>2-</sup> hints that the sea salt contributions to SO<sub>4</sub><sup>2-</sup> (i.e.,

**Table 8.** Summary of the  $R_{\text{adj}}^2$  obtained when correlating mass concentration of a species to  $N_{\text{d}}$  under different atmospheric conditions.

Binning criterion	Data points considered	$R_{\text{adj}}^2$			
		NSS-SO <sub>4</sub> <sup>2-</sup>	Na	Ox	Fe
None	All	0.29	0.19	0.15	0.05
Turbulence	High $\sigma_w$	0.27	0.26	0.09	0.02 <sup>a</sup>
	Low $\sigma_w$	0.27	0.09	0.30	0.07
Smoke influence	No smoke	0.36	0.24	0.07	0.04
	Smoke	0.22	0.17	0.42	0.15
	NiCE <sup>b</sup>	0.36	0.46	0.60	0.28
	FASE <sup>b</sup>	0.18	0.13	0.41	0.12
Normalized cloud height	Top third	0.33	0.33	0.08	0.03
	Middle third	0.29	0.16	0.16	0.03
	Bottom third	0.17	0.10	0.29	0.20

<sup>a</sup> This  $R_{\text{adj}}^2$  has a  $p$  value > 0.05. <sup>b</sup> Only smoke-influenced samples in this campaign were considered.

ocean-derived species) are the ones affected by turbulence and hence explains the insensitivity that NSS-SO<sub>4</sub><sup>2-</sup> has to turbulence.

For Ox, the correlation improves at low turbulence ( $R_{\text{adj}}^2 = 0.30$ ) but not at high turbulence ( $R_{\text{adj}}^2 = 0.09$ ). We believe Ox behaves differently than Na, because it does not necessarily just enter the cloud from below via updrafts but rather it enters the cloud from above via entrainment of air from the free troposphere that can at times be enriched with organic species in the study region (Coggon et al., 2014; Crosbie et al., 2016; Hersey et al., 2009; Sorooshian et al., 2007). For Fe, all turbulence scenarios yield a low correlation between Fe and  $N_{\text{d}}$ , indicating that, overall, Fe is not a good predictor for  $N_{\text{d}}$ .

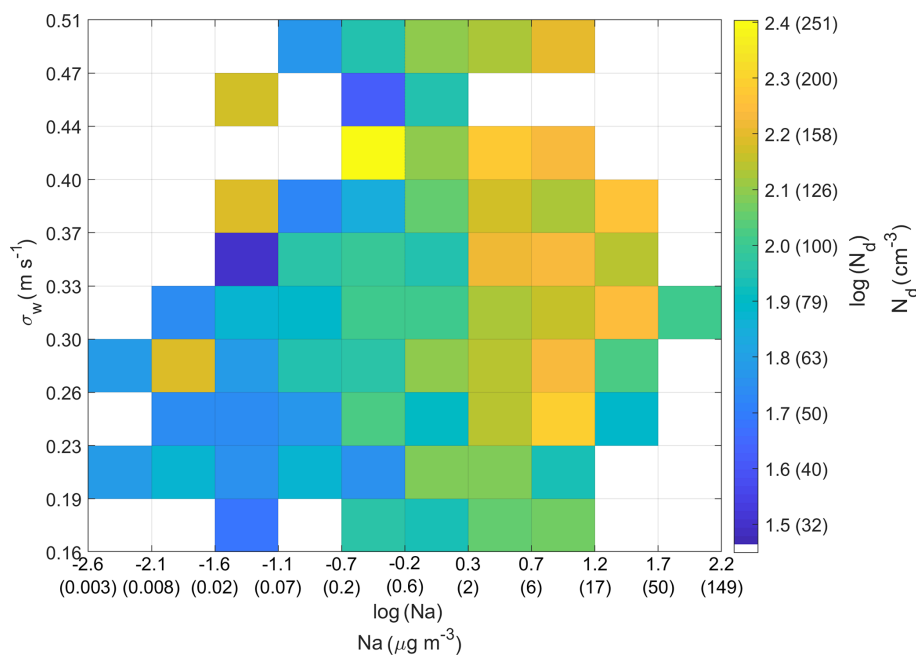
For Na, there is a better correlation at high turbulent conditions than at smooth conditions ( $R_{\text{adj}}^2 = 0.26$  and  $R_{\text{adj}}^2 = 0.09$  for high and low  $\sigma_w$ , respectively). This further strengthens the argument that turbulence plays an important role in the vertical transport of sea salt (and other ocean emissions) from the ocean surface to the cloud base. The present data set allows for deeper analysis into the entangled effects of sea salt and turbulence on  $N_{\text{d}}$ . More specifically, aerosol reanalysis products like those from MERRA-2 calculate the mass concentration of sea salt via parameterizations that link wind speed to sea salt emissions (Gong et al., 2003; Randles et al., 2017). Since wind speed affects turbulence, it follows that sea salt concentrations are not independent from turbulence, as turbulence is used to calculate sea salt concentrations. Subsequently, these sea salt concentrations are used to predict  $N_{\text{d}}$  (e.g., McCoy et al., 2017, 2018). The present study measured both sea salt (quantified by Na) and turbu-

lence (quantified by  $\sigma_w$ ) and thus offers an opportunity to try to isolate the effects of both factors on  $N_{\text{d}}$  (Fig. 6). Two results emerge. First, more turbulence is correlated to more sea salt, which is consistent with what the models predict (Randles et al., 2017). Second, at a fixed concentration of Na,  $N_{\text{d}}$  does not vary significantly with  $\sigma_w$ , as evidenced by a weak change in color. However, at a fixed value of  $\sigma_w$ ,  $N_{\text{d}}$  does vary significantly with Na, as evidenced by the noticeable change in color. Therefore, the independent measurement of both variables reveals that  $N_{\text{d}}$  is more sensitive to changes in Na than to changes in  $\sigma_w$ . We caution that  $\sigma_w$  is not obtained from below the cloud but from within the cloud during sampling time (Fig. S1).

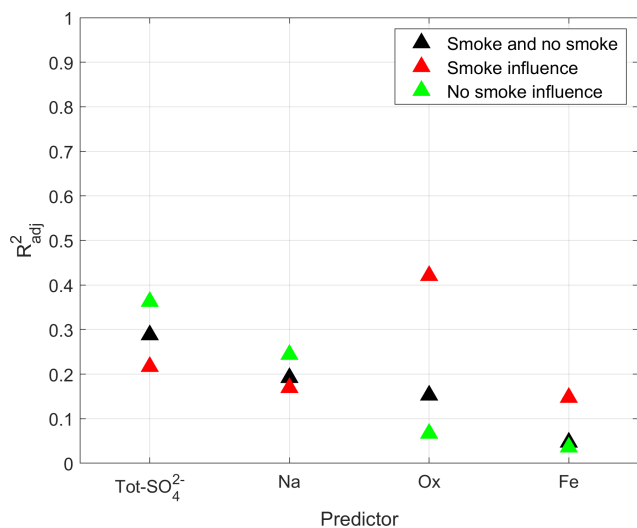
### 3.3.2 Effect of smoke influence

The clouds in the study region are affected by the smoke from wildfires (e.g., Dadashazar et al., 2019; Maudlin et al., 2015; Schlosser et al., 2017). As mentioned in Sect. 2.7, Mardi et al. (2018) used the same data set as this study and identified research flights (RFs) that contained smoke-influenced cloud soundings, namely, NiCE RFs 16–23 and FASE RFs 3–11 and 13–15. In this study, we considered that all cloud water samples collected during the aforementioned RFs were influenced by smoke. Furthermore, we did not distinguish if the smoke was above or below the cloud; this is an important caveat, as cloud microphysical properties seem to depend on the surrounding smoke vertical profile (e.g., Diamond et al., 2018; Koch and Del Genio, 2010). The correlation between  $N_{\text{d}}$  and composition as a function of smoke influence is shown in Fig. 7 and Table 8, and the scatterplots from which the  $R_{\text{adj}}^2$  values are taken are shown in Fig. S4. Species that are produced during wildfires exhibited an improvement in  $R_{\text{adj}}^2$  when considering only the smoke-influenced cases. The opposite is true for species not produced during wildfires. More specifically, Ox and Fe showed an increase in correlation for smoke-influenced conditions ( $R_{\text{adj}}^2 = 0.42$  and  $R_{\text{adj}}^2 = 0.15$  for Ox and Fe, respectively) and a small decrease for smoke-free conditions ( $R_{\text{adj}}^2 = 0.07$  and  $R_{\text{adj}}^2 = 0.04$  for Ox and Fe, respectively). This is most likely because Ox and Fe concentrations increase during wildfires (e.g., Maudlin et al., 2015) and thus contribute appreciably to the regional CCN during the summertime when wildfires are prevalent.

NSS-SO<sub>4</sub><sup>2-</sup> and Na showed a decrease in correlation for smoke-influenced conditions ( $R_{\text{adj}}^2 = 0.22$  and  $R_{\text{adj}}^2 = 0.17$  for NSS-SO<sub>4</sub><sup>2-</sup> and Na, respectively) and an increase for smoke-free conditions ( $R_{\text{adj}}^2 = 0.36$  and  $R_{\text{adj}}^2 = 0.24$  for NSS-SO<sub>4</sub><sup>2-</sup> and Na, respectively). We suspect this is because even though wildfires can produce NSS-SO<sub>4</sub><sup>2-</sup> (e.g., Reid et al., 1998) and Na (e.g., Hudson et al., 2004; Silva et al., 1999), these species are not produced as effectively as Ox or Fe. For example, Maudlin et al. (2015) measured aerosol mass concentration in the study region during both smoke-influenced and non-smoke-influenced conditions. They re-



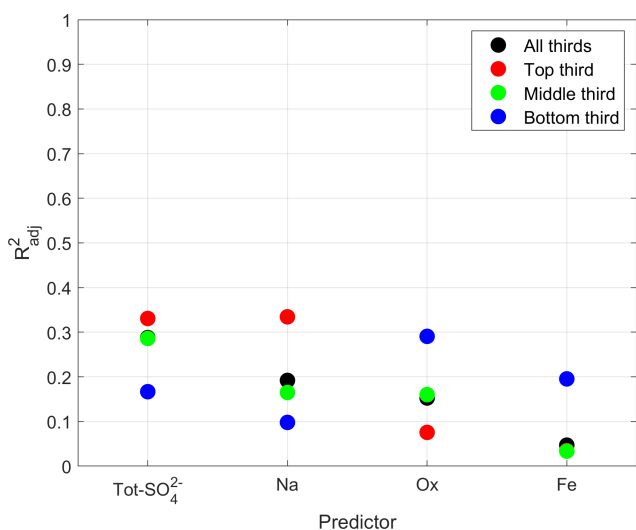
**Figure 6.** Heatmap showing the dependence of  $N_d$  on both  $\sigma_w$  and Na. The lower and upper bounds for the x axis, y axis, and color bar cover the entire range of  $\sigma_w$ , Na, and  $N_d$ , respectively. To assist in physical interpretation, the tick markings on the x axis and color bar show two numbers: those without parenthesis correspond to  $\log(\text{Na})$  or  $\log(N_d)$  and those within parenthesis correspond to Na or  $N_d$ , in their respective units.



**Figure 7.** Effect of the influence of smoke on the ability of a single species to predict  $N_d$ .

ported an increase in mass concentration for  $\text{NSS-SO}_4^{2-}$ , Na, Ox, and Fe to be 30 %, 120 %, 220 %, and 408 %, respectively, for submicron particles, and  $-2$  %,  $-28$  %, 164 %, and 97 %, respectively, for supermicrometer particles. Consequently, Ox and Fe are produced more in wildfires in the study region than  $\text{NSS-SO}_4^{2-}$  and Na.

The NiCE (2015) and FASE (2016) campaigns were influenced by smoke originating from different sources. NiCE was influenced by the Big Windy, Whiskey Complex, and Douglas Complex forest fires near the California–Oregon border, with a transport time of approximately 2 d to reach the base of aircraft operations in Marina and adjacent areas where most samples were collected (Maudlin et al., 2015). In contrast, FASE was influenced by the Soberanes Fire approximately 30 km southwest of the aircraft hangar (Braun et al., 2017). Hence, analyzing each campaign separately may provide some insights into the sensitivity of  $N_d$  to smoke from both different fuel types and with varying transport trajectories. NiCE fire data were linked to timber, grass, and shrub models, whereas those from FASE were associated with chaparral, tall grass, and timber (Braun et al., 2017; Mardi et al., 2018). The results are shown in Table 8 and Fig. S4. When comparing FASE to both campaigns combined, the prediction of  $N_d$  using  $\text{NSS-SO}_4^{2-}$ , Na, Ox, and Fe is not improved, resulting in  $\Delta R_{\text{adj}}^2$  values of  $-0.04$ ,  $-0.04$ ,  $0.01$ , and  $-0.03$ , respectively. However, when comparing NiCE to both campaigns combined, the prediction of  $N_d$  using  $\text{NSS-SO}_4^{2-}$ , Na, Ox, and Fe is significantly improved, resulting in  $\Delta R_{\text{adj}}^2$  values of 0.14, 0.29, 0.18, and 0.13, respectively. The difference between NiCE and FASE could be because different forest fires produce aerosols with varying aerosol chemical signatures and size distributions, as studies in the region have shown (Ma et al., 2019; Mardi et al., 2019). Alternatively, the difference could be due to the small sample size of NiCE (31



**Figure 8.** Effect of the influence of normalized cloud height on the ability of a single species to predict  $N_d$ . For Fe, the top third of cloud (red) data point overlaps with the middle and bottom thirds of cloud (green and blue) data points.

samples) as compared to FASE (136 samples) (Table 1). Certainly more research, including larger data sets, is warranted to investigate how different fuel types and plume aging times impact aerosol–cloud interactions.

### 3.3.3 Effect of in-cloud height

MacDonald et al. (2018) used the same data set as this study to show that the chemical composition of cloud water varies with height within a cloud. It is therefore reasonable that the  $N_d$ –chemical-composition relationship also varies with in-cloud height. The correlation between  $N_d$  and composition as a dependence of in-cloud height is shown in Fig. 8 and Table 8, and the scatterplots from which the  $R^2_{\text{adj}}$  are taken are shown in Fig. S5.

Ox and Fe exhibit a better correlation when focusing on the bottom third of the cloud ( $R^2_{\text{adj}} = 0.29$  and  $R^2_{\text{adj}} = 0.20$  for Ox and Fe, respectively). When focusing on the top third of the cloud, the correlation decreased for Ox ( $R^2_{\text{adj}} = 0.08$ ) and remained unchanged for Fe ( $R^2_{\text{adj}} = 0.03$ ). One possible hypothesis to explain why Ox and Fe are better predictors of  $N_d$  at cloud base is that smoke affects cloud microphysics ( $N_d$  and effective radius) more at cloud base than at cloud top, regardless of whether the smoke was above or below the cloud (Diamond et al., 2018; Mardi et al., 2019).

NSS-SO<sub>4</sub><sup>2-</sup> and Na exhibit a better correlation with  $N_d$  when focusing on the top third of the cloud ( $R^2_{\text{adj}} = 0.33$  and  $R^2_{\text{adj}} = 0.33$  for NSS-SO<sub>4</sub><sup>2-</sup> and Na, respectively). The correlation decreases when focusing on the bottom third of the cloud ( $R^2_{\text{adj}} = 0.17$  and  $R^2_{\text{adj}} = 0.10$  for NSS-SO<sub>4</sub><sup>2-</sup> and

Na, respectively). Tot-SO<sub>4</sub><sup>2-</sup> also follows this pattern ( $R^2_{\text{adj}} = 0.56$  and  $R^2_{\text{adj}} = 0.22$  for top and bottom, respectively).

It is not entirely clear why NSS-SO<sub>4</sub><sup>2-</sup> and Na would be better correlated with  $N_d$  in the top third of clouds. MacDonald et al. (2018) noted that the concentration of chemical species varies as a function of in-cloud height and is not the same for all species; the concentration of Na is greatest at cloud base, whereas that of NSS-SO<sub>4</sub><sup>2-</sup> and Ox are greatest mid-cloud. It would be expected that the vertical profile of concentration is related to the ability to predict  $N_d$  (i.e., that a larger concentration of a species leads to a better correlation with  $N_d$ ), but that expectation is not observed in these results. It is also interesting to point out that there is not much difference in  $R^2_{\text{adj}}$  when considering all cloud thirds versus only the middle third; this makes sense, as almost half of the cloud water samples (46 %) were collected in the middle third of the cloud.

The dependence of the correlation between chemical composition and  $N_d$  on in-cloud height is of relevance to remote sensing, which relies on satellite measurement of cloud-top properties such as cloud-top temperature to then calculate a constant  $N_d$  throughout the cloud depth (e.g., Grosvenor et al., 2018).

## 4 Conclusions

This study used a 4-year data set of airborne measurements collected in warm marine stratocumulus clouds off the California coast and analyzed the extent to which the chemical composition of cloud water can be used to predict  $N_d$ . A total of 80 species were filtered to 9 to examine the prediction of  $N_d$  using a single-species model, and then using a multi-species model. The nine species were subsequently filtered to four to examine how the four single-species models were affected by environmental factors, namely, turbulence, smoke influence, and vertical location within a cloud. The most important findings of this paper are the following.

The species that best predicted  $N_d$  is Tot-SO<sub>4</sub><sup>2-</sup> with  $R^2_{\text{adj}} = 0.40$ , followed by NH<sub>4</sub><sup>+</sup> ( $R^2_{\text{adj}} = 0.34$ ), NSS-SO<sub>4</sub><sup>2-</sup> ( $R^2_{\text{adj}} = 0.29$ ), MSA ( $R^2_{\text{adj}} = 0.26$ ), and NO<sub>3</sub><sup>-</sup> ( $R^2_{\text{adj}} = 0.24$ ).

The prediction of  $N_d$  can be improved by using a multi-species model. However, increasing the number of species caused the  $R^2_{\text{adj}}$  to asymptotically approach  $\sim 0.6$ . Furthermore, the regressions with six or more species became statistically insignificant.

Analyzing the three best correlations for each of the  $n$  species models (where  $n = 1$ –5) shows that the factors that constitute a good regression are a form of SO<sub>4</sub><sup>2-</sup> (total or non-sea-salt), an ocean emissions tracer, and an organic tracer.

Greater turbulence (approximated as the standard deviation of vertical wind speed) improves the ability of ocean-derived species to predict  $N_d$ , as observed when comparing regressions using turbulent data points versus all data points



for Tot-SO<sub>4</sub><sup>2-</sup> ( $\Delta R_{\text{adj}}^2 = 0.15$ ) and Na ( $\Delta R_{\text{adj}}^2 = 0.07$ ) but not for NSS-SO<sub>4</sub><sup>2-</sup> ( $\Delta R_{\text{adj}}^2 = -0.01$ ) or Ox ( $\Delta R_{\text{adj}}^2 = -0.06$ ).

The influence of smoke significantly affects those species that best predict  $N_{\text{d}}$ . Ox (a species known to be produced during biomass burning) was best correlated with  $N_{\text{d}}$  ( $R_{\text{adj}}^2 = 0.42$ ) under smoke-influenced conditions.

Vertical location within the cloud affects the ability to predict  $N_{\text{d}}$ . The species that are best correlated with  $N_{\text{d}}$  at cloud top are Tot-SO<sub>4</sub><sup>2-</sup> ( $R_{\text{adj}}^2 = 0.56$ ) and NSS-SO<sub>4</sub><sup>2-</sup> ( $R_{\text{adj}}^2 = 0.33$ ); those best correlated with  $N_{\text{d}}$  at cloud base are fire tracers such as Ox ( $R_{\text{adj}}^2 = 0.29$ ) and Fe ( $R_{\text{adj}}^2 = 0.20$ ), as it has been reported that the base of a cloud is more sensitive to the influence of smoke.

*Data availability.* All data used in this work can be found on the Figshare database (Sorooshian et al., 2017; [https://figshare.com/articles/A\\_Multi-Year\\_Data\\_Set\\_on\\_Aerosol-Cloud-Precipitation-Meteorology\\_Interactions\\_for\\_Marine\\_Stratocumulus\\_Clouds/5099983](https://figshare.com/articles/A_Multi-Year_Data_Set_on_Aerosol-Cloud-Precipitation-Meteorology_Interactions_for_Marine_Stratocumulus_Clouds/5099983), last access: 27 June 2020).

*Supplement.* The supplement related to this article is available online at: <https://doi.org/10.5194/acp-20-7645-2020-supplement>.

*Author contributions.* All coauthors contributed to some aspect of the data collection. ABM and AS conducted the data analysis and interpretation. ABM and AS prepared the article with contributions from all coauthors.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* Alexander B. MacDonald acknowledges support from the Mexican National Council for Science and Technology (CONACYT). We acknowledge Agilent Technologies for their support and Shane Snyder's laboratories for ICP-QQQ data.

*Financial support.* This research has been supported by the Office of Naval Research (grant nos. N00014-10-1-0811, N00014-11-1-0783, N00014-10-1-0200, N00014-04-1-0118, and N00014-16-1-35 2567), and the National Aeronautics and Space Administration (grant no. 80NSSC19K0442), in support of the ACTIVATE Earth Venture Suborbital-3 (EVS-3) investigation, which is funded by NASA's Earth Science Division and managed through the Earth System Science Pathfinder Program Office.

*Review statement.* This paper was edited by Lynn M. Russell and reviewed by two anonymous referees.

## References

- Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation 2. Multiple aerosol types, *J. Geophys. Res.*, 105, 6837–6844, <https://doi.org/10.1029/1999JD901161>, 2000.
- Ackerman, A. S., Kirkpatrick, M. P., Stevens, D. E., and Toon, O. B.: The impact of humidity above stratiform clouds on indirect aerosol climate forcing, *Nature*, 432, 1014–1017, <https://doi.org/10.1038/nature03174>, 2004.
- Adams, P. J. and Seinfeld, J. H.: Disproportionate impact of particulate emissions on global cloud condensation nuclei concentrations, *Geophys. Res. Lett.*, 30, 1239, <https://doi.org/10.1029/2002gl016303>, 2003.
- Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, *Science*, 245, 1227–1230, <https://doi.org/10.1126/science.245.4923.1227>, 1989.
- Andreae, M. O. and Rosenfeld, D.: Aerosol-cloud-precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, *Earth-Sci. Rev.*, 89, 13–41, <https://doi.org/10.1016/j.earscirev.2008.03.001>, 2008.
- ApSimon, H. M., Kruse, M., and Bell, J. N. B.: Ammonia emissions and their role in acid deposition, *Atmos. Environ.*, 21, 1939–1946, [https://doi.org/10.1016/0004-6981\(87\)90154-5](https://doi.org/10.1016/0004-6981(87)90154-5), 1987.
- Bauer, S. E., Tsigaridis, K., and Miller, R.: Significant atmospheric aerosol pollution caused by world food cultivation, *Geophys. Res. Lett.*, 43, 5394–5400, <https://doi.org/10.1002/2016GL068354>, 2016.
- Baumgardner, D., Jonsson, H., Dawson, W., O'Connor, D., and Newton, R.: The cloud, aerosol and precipitation spectrometer: A new instrument for cloud investigations, *Atmos. Res.*, 59, 251–264, [https://doi.org/10.1016/S0169-8095\(01\)00119-3](https://doi.org/10.1016/S0169-8095(01)00119-3), 2001.
- Behrenfeld, M. J., Moore, R. H., Hostetler, C. A., Graff, J., Gaube, P., Russell, L. M., Chen, G., Doney, S. C., Giovannoni, S., Liu, H., Proctor, C., Bolaños, L. M., Baetge, N., Davie-Martin, C., Westberry, T. K., Bates, T. S., Bell, T. G., Bidle, K. D., Boss, E. S., Brooks, S. D., Cairns, B., Carlson, C., Halsey, K., Harvey, E. L., Hu, C., Karp-Boss, L., Kleb, M., Menden-Deuer, S., Morison, F., Quinn, P. K., Scarino, A. J., Anderson, B., Chowdhary, J., Crosbie, E., Ferrare, R., Hair, J. W., Hu, Y., Janz, S., Redemann, J., Saltzman, E., Shook, M., Siegel, D. A., Wisthaler, A., Martin, M. Y., and Ziemba, L.: The North Atlantic Aerosol and Marine Ecosystem Study (NAAMES): Science motive and mission overview, *Front. Mar. Sci.*, 6, 1–25, <https://doi.org/10.3389/fmars.2019.00122>, 2019.
- Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Danaïu, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, *Rev. Geophys.*, 58, e2019RG000660, <https://doi.org/10.1029/2019rg000660>, 2020.
- Borys, R. D., Lowenthal, D. H., Wetzol, M. A., Herrera, F., Gonzalez, A., and Harris, J.: Chemical and microphysical properties of marine stratiform cloud in the North Atlantic, *J. Geophys. Res.-Atmos.*, 103, 22073–22085, <https://doi.org/10.1029/98JD02087>, 1998.

- Boucher, O. and Lohmann, U.: The sulfate-CCN-cloud albedo effect, *Tellus B*, 47, 281–300, <https://doi.org/10.3402/tellusb.v47i3.16048>, 1995.
- Bouwman, A. F., Lee, D. S., Asman, W. A. H., Dentener, F. J., Van Der Hoek, K. W., and Olivier, J. G. J.: A global high-resolution emission inventory for ammonia, *Global Biogeochem. Cy.*, 11, 561–587, <https://doi.org/10.1029/97GB02266>, 1997.
- Charlson, R. J., Lovelock, J. E., Andreae, M. O., and Warren, S. G.: Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate, *Nature*, 326, 655–661, <https://doi.org/10.1038/326655a0>, 1987.
- Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. J., Hansen, J. E., and Hofmann, D. J.: Climate forcing by anthropogenic aerosols, *Science*, 117, 423–430, <https://doi.org/10.1126/science.255.5043.423>, 1992.
- Chen, Y. and Penner, J. E.: Uncertainty analysis for estimates of the first indirect aerosol effect, *Atmos. Chem. Phys.*, 5, 2935–2948, <https://doi.org/10.5194/acp-5-2935-2005>, 2005.
- Chen, Y.-C., Christensen, M. W., Xue, L., Sorooshian, A., Stephens, G. L., Rasmussen, R. M., and Seinfeld, J. H.: Occurrence of lower cloud albedo in ship tracks, *Atmos. Chem. Phys.*, 12, 8223–8235, <https://doi.org/10.5194/acp-12-8223-2012>, 2012.
- Chuang, C. C. and Penner, J. E.: Effects of anthropogenic sulfate on cloud drop nucleation and optical properties, *Tellus B*, 47, 566–577, <https://doi.org/10.1034/j.1600-0889.47.issue5.4.x>, 1995.
- Chuang, C. C., Penner, J. E., and Edwards, L. L.: Nucleation Scavenging of Smoke Particles and Simulated Drop Size Distributions over Large Biomass Fires, *J. Atmos. Sci.*, 49, 1264–1276, [https://doi.org/10.1175/1520-0469\(1992\)049<1264:NSOSPA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1992)049<1264:NSOSPA>2.0.CO;2), 1992.
- Coggon, M. M., Sorooshian, A., Wang, Z., Metcalf, A. R., Frossard, A. A., Lin, J. J., Craven, J. S., Nenes, A., Jonsson, H. H., Russell, L. M., Flagan, R. C., and Seinfeld, J. H.: Ship impacts on the marine atmosphere: insights into the contribution of shipping emissions to the properties of marine aerosol and clouds, *Atmos. Chem. Phys.*, 12, 8439–8458, <https://doi.org/10.5194/acp-12-8439-2012>, 2012.
- Coggon, M. M., Sorooshian, A., Wang, Z., Craven, J. S., Metcalf, A. R., Lin, J. J., Nenes, A., Jonsson, H. H., Flagan, R. C., and Seinfeld, J. H.: Observations of continental biogenic impacts on marine aerosol and clouds off the coast of California, *J. Geophys. Res.-Atmos.*, 119, 6724–6748, <https://doi.org/10.1002/2013JD021228>, 2014.
- Crosbie, E., Wang, Z., Sorooshian, A., Chuang, P. Y., Craven, J. S., Coggon, M. M., Brunke, M., Zeng, X., Jonsson, H., Woods, R. K., Flagan, R. C., and Seinfeld, J. H.: Stratocumulus Cloud Clearings and Notable Thermodynamic and Aerosol Contrasts across the Clear–Cloudy Interface, *J. Atmos. Sci.*, 73, 1083–1099, <https://doi.org/10.1175/JAS-D-15-0137.1>, 2016.
- Dadashazar, H., Wang, Z., Crosbie, E., Brunke, M., Zeng, X., Jonsson, H., Woods, R. K., Flagan, R. C., Seinfeld, J. H., and Sorooshian, A.: Relationships between giant sea salt particles and clouds inferred from aircraft physicochemical data, *J. Geophys. Res.*, 122, 3421–3434, <https://doi.org/10.1002/2016JD026019>, 2017.
- Dadashazar, H., Ma, L., and Sorooshian, A.: Sources of pollution and interrelationships between aerosol and precipitation chemistry at a central California site, *Sci. Total Environ.*, 651, 1776–1787, <https://doi.org/10.1016/j.scitotenv.2018.10.086>, 2019.
- Diamond, M. S., Dobracki, A., Freitag, S., Small Griswold, J. D., Heikkila, A., Howell, S. G., Kacarab, M. E., Podolske, J. R., Saide, P. E., and Wood, R.: Time-dependent entrainment of smoke presents an observational challenge for assessing aerosol–cloud interactions over the southeast Atlantic Ocean, *Atmos. Chem. Phys.*, 18, 14623–14636, <https://doi.org/10.5194/acp-18-14623-2018>, 2018.
- Facchini, M. C., Mircea, M., Fuzzi, S., and Charlson, R. J.: Cloud albedo enhancement by surface-active organic solutes in growing droplets, *Nature*, 401, 257–259, <https://doi.org/10.1038/45758>, 1999.
- Feingold, G., Frisch, A. S., Stevens, B., and Cotton, W. R.: On the relationship among cloud turbulence, droplet formation and drizzle as viewed by Doppler radar, microwave radiometer and lidar, *J. Geophys. Res.-Atmos.*, 104, 22195–22203, <https://doi.org/10.1029/1999JD900482>, 1999.
- 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. Atmos. Sci.*, 56, 4100–4117, [https://doi.org/10.1175/1520-0469\(1999\)056<4100:TIOGCC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<4100:TIOGCC>2.0.CO;2), 1999.
- Fountoukis, C. and Nenes, A.: Continued development of a cloud droplet formation parameterization for global climate models, *J. Geophys. Res.-Atmos.*, 110, D11212, <https://doi.org/10.1029/2004JD005591>, 2005.
- Freund, R. J., Wilson, W. J., and Mohr, D. L.: *Statistical Methods*, 3 Edn., Academic Press, Burlington, MA, 824 pp., 2010.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Parityka, G., Pawson, S., Putman, W., Rienercker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *J. Climate*, 30, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>, 2017.
- Gerber, H., Arends, B. G., and Ackerman, A. S.: New microphysics sensor for aircraft use, *Atmos. Res.*, 31, 235–252, [https://doi.org/10.1016/0169-8095\(94\)90001-9](https://doi.org/10.1016/0169-8095(94)90001-9), 1994.
- Gerber, H., Frick, G., and Rodi, A. R.: Ground-based FSSP and PVM measurements of liquid water content, *J. Atmos. Ocean. Tech.*, 16, 1143–1149, [https://doi.org/10.1175/1520-0426\(1999\)016<1143:GBFAPM>2.0.CO;2](https://doi.org/10.1175/1520-0426(1999)016<1143:GBFAPM>2.0.CO;2), 1999.
- Ghan, S. J., Leung, L. R., and Easter, R. C.: Prediction of cloud droplet number in a general circulation model, *J. Geophys. Res.*, 102, 21777–21794, <https://doi.org/10.1029/97JD01810>, 1997.
- Ghan, S. J., Abdul-Razzak, H., Nenes, A., Ming, Y., Liu, X., Ovchinnikov, M., Shipway, B., Meskhidze, N., Xu, J., and Shi, X.: Droplet nucleation: Physically-based parameterizations and comparative evaluation, *J. Adv. Model. Earth Syst.*, 3, 1–34, <https://doi.org/10.1029/2011ms000074>, 2011.
- Gong, S. L.: A parameterization of sea-salt aerosol source function for sub- and super-micron particles, *Global Biogeochem. Cy.*, 17, 1097, <https://doi.org/2003gb002079>, 2003.
- Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold,

- G., Fridlind, A., Hünerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote Sensing of Droplet Number Concentration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, *Rev. Geophys.*, 56, 409–453, <https://doi.org/10.1029/2017RG000593>, 2018.
- Hegg, D. A.: Impact of gas-phase  $\text{HNO}_3$  and  $\text{NH}_3$  on microphysical processes in atmospheric clouds, *Geophys. Res. Lett.*, 27, 2201–2204, <https://doi.org/10.1029/1999GL011252>, 2000.
- Hegg, D. A. and Hobbs, P. V.: Cloud Water Chemistry and the production of sulfates in clouds, *Atmos. Environ.*, 15, 1597–1604, [https://doi.org/10.1016/0004-6981\(81\)90144-X](https://doi.org/10.1016/0004-6981(81)90144-X), 1981.
- Hegg, D. A. and Hobbs, P. V.: Sulfate and nitrate chemistry in cumuliform clouds, *Atmos. Environ.*, 20, 901–909, [https://doi.org/10.1016/0004-6981\(86\)90274-X](https://doi.org/10.1016/0004-6981(86)90274-X), 1986.
- Hegg, D. A., Ferek, R. J., and Hobbs, P. V.: Light scattering and cloud condensation nucleus activity of sulfate aerosol measured over the northeast Atlantic Ocean, *J. Geophys. Res.-Atmos.*, 98, 14887–14894, <https://doi.org/10.1029/93JD01615>, 1993.
- Hersey, S. P., Sorooshian, A., Murphy, S. M., Flagan, R. C., and Seinfeld, J. H.: Aerosol hygroscopicity in the marine atmosphere: a closure study using high-time-resolution, multiple-RH DASH-SP and size-resolved C-ToF-AMS data, *Atmos. Chem. Phys.*, 9, 2543–2554, <https://doi.org/10.5194/acp-9-2543-2009>, 2009.
- Hudson, P. K., Murphy, D. M., Cziczo, D. J., Thomson, D. S., de Gouw, J. A., Warneke, C., Holloway, J., Jost, H. J., and Hübner, G.: Biomass-burning particle measurements: Characteristics composition and chemical processing, *J. Geophys. Res.-Atmos.*, 109, D23S27, <https://doi.org/10.1029/2003JD004398>, 2004.
- Intergovernmental Panel on Climate Change: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp., 2007.
- Intergovernmental Panel on Climate Change: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., 2013.
- Jung, E., Albrecht, B. A., Jonsson, H. H., Chen, Y.-C., Seinfeld, J. H., Sorooshian, A., Metcalf, A. R., Song, S., Fang, M., and Russell, L. M.: Precipitation effects of giant cloud condensation nuclei artificially introduced into stratocumulus clouds, *Atmos. Chem. Phys.*, 15, 5645–5658, <https://doi.org/10.5194/acp-15-5645-2015>, 2015.
- Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., Brooks, N., Cao, J. J., Boyd, P. W., Duce, R. A., Hunter, K. A., Kawahata, H., Kubilay, N., and Liss, P. S.: Global Iron Connections Between Desert Dust, Ocean Biogeochemistry, and Climate, *Science*, 308, 67–71, <https://doi.org/10.1126/science.1105959>, 2005.
- Kahane, L. H.: Regression Basics, 2 Edn., Sage Publications, Thousand Oaks, CA, 234 pp., 2008.
- Koch, D. and Del Genio, A. D.: Black carbon semi-direct effects on cloud cover: review and synthesis, *Atmos. Chem. Phys.*, 10, 7685–7696, <https://doi.org/10.5194/acp-10-7685-2010>, 2010.
- Kulmala, M., Laaksonen, A., Korhonen, P., Vesala, T., Ahonen, T., and Barrett, J. C.: The effect of atmospheric nitric acid vapor on cloud condensation nucleus activation, *J. Geophys. Res.*, 98, 22949–22958, <https://doi.org/10.1029/93JD02070>, 1993.
- Lance, S., Nenes, A., Mazzoleni, C., Dubey, M. K., Gates, H., Varutbangkul, V., Rissman, T. A., Murphy, S. M., Sorooshian, A., Flagan, R. C., Seinfeld, J. H., Feingold, G., and Jonsson, H. H.: Cloud condensation nuclei activity, closure, and droplet growth kinetics of Houston aerosol during the Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS), *J. Geophys. Res.*, 114, D00F15, <https://doi.org/10.1029/2008jd011699>, 2009.
- Leaitch, W. R., Strapp, J. W., Wiebe, H. A., Anlauf, K. G. and Isaac, G. A.: Chemical and microphysical studies of nonprecipitating summer cloud in Ontario, Canada, *J. Geophys. Res.-Atmos.*, 91, 11821–11831, <https://doi.org/10.1029/JD091iD11p11821>, 1986.
- Leaitch, W. R., Isaac, G. A., Strapp, J. W., Banic, C. M., and Wiebe, H. A.: The relationship between cloud droplet number concentrations and anthropogenic pollution: observations and climatic implications, *J. Geophys. Res.-Atmos.*, 97, 2463–2474, <https://doi.org/10.1029/91JD02739>, 1992.
- Leaitch, W. R., Banic, C. M., Isaac, G. A., Couture, M. D., Liu, P. S. K., Gultepe, I., Li, S. M., Kleinman, L., Daum, P. H., and MacPherson, J. I.: Physical and chemical observations in marine stratus during the 1993 North Atlantic Regional Experiment: Factors controlling cloud droplet number concentrations, *J. Geophys. Res.-Atmos.*, 101, 29123–29135, <https://doi.org/10.1029/96JD01228>, 1996.
- Lowenthal, D. H. and Borys, R. D.: Sources of microphysical variation in marine stratiform clouds in the North Atlantic, *Geophys. Res. Lett.*, 27, 1491–1494, <https://doi.org/10.1029/1999GL010992>, 2000.
- Lowenthal, D. H., Borys, R. D., Choularton, T. W., Bower, K. N., Flynn, M. J., and Gallagher, M. W.: Parameterization of the cloud droplet–sulfate relationship, *Atmos. Environ.*, 38, 287–292, <https://doi.org/10.1016/j.atmosenv.2003.09.046>, 2004.
- Lu, M. L., Sorooshian, A., Jonsson, H. H., Feingold, G., Flagan, R. C., and Seinfeld, J. H.: Marine stratocumulus aerosol-cloud relationships in the MASE-II experiment: Precipitation susceptibility in eastern Pacific marine stratocumulus, *J. Geophys. Res.-Atmos.*, 114, 1–11, <https://doi.org/10.1029/2009JD012774>, 2009.
- Ma, L., Dadashazar, H., Braun, R. A., MacDonald, A. B., Aghdam, M. A., Maudlin, L. C., and Sorooshian, A.: Size-resolved characteristics of water-soluble particulate elements in a coastal area: Source identification, influence of wildfires, and diurnal variability, *Atmos. Environ.*, 206, 72–84, <https://doi.org/10.1016/j.atmosenv.2019.02.045>, 2019.
- MacDonald, A. B., Dadashazar, H., Chuang, P. Y., Crosbie, E., Wang, H., Wang, Z., Jonsson, H. H., Flagan, R. C., Seinfeld, J. H., and Sorooshian, A.: Characteristic Vertical Profiles of Cloud Water Composition in Marine Stratocumulus Clouds and Relationships with Precipitation, *J. Geophys. Res.-Atmos.*, 123, 3704–3723, <https://doi.org/10.1002/2017JD027900>, 2018.
- Mardi, A. H., Dadashazar, H., MacDonald, A. B., Braun, R. A., Crosbie, E., Xian, P., Thorsen, T. J., Coggon, M. M., Fenn, M.

- A., Ferrare, R. A., Hair, J. W., Woods, R. K., Jonsson, H. H., Flagan, R. C., Seinfeld, J. H., and Sorooshian, A.: Biomass Burning Plumes in the Vicinity of the California Coast: Airborne Characterization of Physicochemical Properties, Heating Rates, and Spatiotemporal Features, *J. Geophys. Res.-Atmos.*, 123, 13560–13582, <https://doi.org/10.1029/2018JD029134>, 2018.
- Mardi, A. H., Dadashazar, H., MacDonald, A. B., Crosbie, E., Coggon, M. M., Aghdam, M. A., Woods, R. K., Jonsson, H. H., Flagan, R. C., Seinfeld, J. H., and Sorooshian, A.: Effects of Biomass Burning on Stratocumulus Droplet Characteristics, Drizzle Rate, and Composition, *J. Geophys. Res.-Atmos.*, 124, 12301–12318, <https://doi.org/10.1029/2019JD031159>, 2019.
- Matsumoto, K., Tanaka, H., Nagao, I., and Ishizaka, Y.: Contribution of particulate sulfate and organic carbon to cloud condensation nuclei in the marine atmosphere, *Geophys. Res. Lett.*, 24, 655–658, <https://doi.org/10.1029/97GL00541>, 1997.
- Maudlin, L. C., Wang, Z., Jonsson, H. H., and Sorooshian, A.: Impact of wildfires on size-resolved aerosol composition at a coastal California site, *Atmos. Environ.*, 119, 59–68, <https://doi.org/10.1016/j.atmosenv.2015.08.039>, 2015.
- McCoy, D. T., Bender, F. A.-M., Mohrmann, J. K. C., Hartmann, D. L., Wood, R., and Grosvenor, D. P.: The global aerosol-cloud first indirect effect estimated using MODIS, MERRA, and AeroCom, *J. Geophys. Res.-Atmos.*, 122, 1779–1796, <https://doi.org/10.1002/2016JD026141>, 2017.
- McCoy, D. T., Bender, F. A.-M., Grosvenor, D. P., Mohrmann, J. K., Hartmann, D. L., Wood, R., and Field, P. R.: Predicting decadal trends in cloud droplet number concentration using reanalysis and satellite data, *Atmos. Chem. Phys.*, 18, 2035–2047, <https://doi.org/10.5194/acp-18-2035-2018>, 2018.
- Medina, J., Nenes, A., Sotiropoulou, R. E. P., Cottrell, L. D., Ziemba, L. D., Beckman, P. J., and Griffin, R. J.: Cloud condensation nuclei closure during the International Consortium for Atmospheric Research on Transport and Transformation 2004 campaign: Effects of size-resolved composition, *J. Geophys. Res.-Atmos.*, 112, D10S31, <https://doi.org/10.1029/2006JD007588>, 2007.
- Menon, S. and Saxena, V. K.: Role of sulfates in regional cloud–climate interactions, *Atmos. Res.*, 47, 299–315, [https://doi.org/10.1016/S0169-8095\(98\)00057-X](https://doi.org/10.1016/S0169-8095(98)00057-X), 1998.
- Menon, S., Genio, A. D. Del, Koch, D., and Tselioudis, G.: GCM Simulations of the Aerosol Indirect Effect: Sensitivity to Cloud Parameterization and Aerosol Burden, *J. Atmos. Sci.*, 59, 692–713, [https://doi.org/10.1175/1520-0469\(2002\)059<0692:gsotai>2.0.co;2](https://doi.org/10.1175/1520-0469(2002)059<0692:gsotai>2.0.co;2), 2002.
- Nenes, A., Charlson, R. J., Facchini, M. C., Kulmala, M., Laaksonen, A., and Seinfeld, J. H.: Can chemical effects on cloud droplet number rival the first indirect effect?, *Geophys. Res. Lett.*, 29, 1848, <https://doi.org/10.1029/2002gl015295>, 2002.
- Nenes, A. and Seinfeld, J. H.: Parameterization of cloud droplet formation in global climate models, *J. Geophys. Res.-Atmos.*, 108, 1–14, <https://doi.org/10.1029/2002jd002911>, 2003.
- Novakov, T. and Penner, J. E.: Large contribution of organic aerosols to cloud-condensation-nuclei concentrations, *Nature*, 365, 823–826, <https://doi.org/10.1038/365823a0>, 1993.
- Novakov, T., Rivera-Carpio, C., Penner, J. E., and Rogers, C. F.: The effect of anthropogenic sulfate aerosols on marine cloud droplet concentrations, *Tellus B*, 46, 132–141, <https://doi.org/10.3402/tellusb.v46i2.15758>, 1994.
- Partridge, D. G., Vrugt, J. A., Tunved, P., Ekman, A. M. L., Struthers, H., and Sorooshian, A.: Inverse modelling of cloud-aerosol interactions – Part 2: Sensitivity tests on liquid phase clouds using a Markov chain Monte Carlo based simulation approach, *Atmos. Chem. Phys.*, 12, 2823–2847, <https://doi.org/10.5194/acp-12-2823-2012>, 2012.
- Paulot, F., Jacob, D. J., Johnson, M. T., Bell, T. G., Baker, A. R., Keene, W. C., Lima, I. D., Doney, S. C., and Stock, C. A.: Global oceanic emission of ammonia: Constraints from seawater and atmospheric observations, *Global Biogeochem. Cy.*, 28, 1165–1178, <https://doi.org/10.1002/2015GB005106>, 2015.
- Prabhakar, G., Ervens, B., Wang, Z., Maudlin, L. C., Coggon, M. M., Jonsson, H. H., Seinfeld, J. H., and Sorooshian, A.: Sources of nitrate in stratocumulus cloud water: Airborne measurements during the 2011 E-PEACE and 2013 NiCE studies, *Atmos. Environ.*, 97, 166–173, <https://doi.org/10.1016/j.atmosenv.2014.08.019>, 2014.
- Pringle, K. J., Carslaw, K. S., Spracklen, D. V., Mann, G. M., and Chipperfield, M. P.: The relationship between aerosol and cloud drop number concentrations in a global aerosol microphysics model, *Atmos. Chem. Phys.*, 9, 4131–4144, <https://doi.org/10.5194/acp-9-4131-2009>, 2009.
- Pueschel, R. F., Valin, C. C., Castillo, R. C., Kadlecik, J. A., and Ganor, E.: Aerosols in polluted versus nonpolluted air masses: long-range transport and effects on clouds, *J. Appl. Meteorol. Clim.*, 25, 1908–1917, [https://doi.org/10.1175/1520-0450\(1986\)025<1908:AIPVNA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025<1908:AIPVNA>2.0.CO;2), 1986.
- Pye, H. O. T., Nenes, A., Alexander, B., Ault, A. P., Barth, M. C., Clegg, S. L., Collett Jr., J. L., Fahey, K. M., Hennigan, C. J., Herrmann, H., Kanakidou, M., Kelly, J. T., Ku, I.-T., McNeill, V. F., Riemer, N., Schaefer, T., Shi, G., Tilgner, A., Walker, J. T., Wang, T., Weber, R., Xing, J., Zaveri, R. A., and Zuend, A.: The acidity of atmospheric particles and clouds, *Atmos. Chem. Phys.*, 20, 4809–4888, <https://doi.org/10.5194/acp-20-4809-2020>, 2020.
- Quinn, P. K., Coffman, D. J., Johnson, J. E., Upchurch, L. M., and Bates, T. S.: Small fraction of marine cloud condensation nuclei made up of sea spray aerosol, *Nat. Geosci.*, 10, 674–679, <https://doi.org/10.1038/ngeo3003>, 2017.
- Randles, C. A., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., and Flynn, C. J.: The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation, *J. Climate*, 30, 6823–6850, <https://doi.org/10.1175/JCLI-D-16-0609.1>, 2017.
- Reid, J. S., Hobbs, P. V., Ferek, R. J., Blake, D. R., Martins, J. V., Dunlap, M. R., and Lioussé, C.: Physical, chemical, and optical properties of regional hazes dominated by smoke in Brazil, *J. Geophys. Res.-Atmos.*, 103, 32059–32080, <https://doi.org/10.1029/98JD00458>, 1998.
- Russell, L. M., Sorooshian, A., Seinfeld, J. H., Albrecht, B. A., Nenes, A., Ahlm, L., Chen, Y.-C., Coggon, M., Craven, J. S., Flagan, R. C., Frossard, A. A., Jonsson, H., Jung, E., Lin, J. J., Metcalf, A. R., Modini, R., Mülmenstädt, J., Roberts, G. C., Shingler, T., Song, S., Wang, Z., and WonaSchütz, A.: Eastern pacific emitted aerosol cloud experiment, *B. Am. Meteorol. Soc.*, 94, 709–729, <https://doi.org/10.1175/BAMS-D-12-00015.1>, 2013.
- Saxena, V. K. and Menon, S.: Sulfate-induced cooling in the southeastern US: An observational assessment, *Geophys. Res. Lett.*, 26, 2489–2492, <https://doi.org/10.1029/1999GL900555>, 1999.

- Schlosser, J. S., Braun, R. A., Bradley, T., Dadashazar, H., MacDonald, A. B., Aldhaif, A. A., Azadi Aghdam, M., Hossein Mardi, A., Xian, P., and Sorooshian, A.: Analysis of Aerosol Composition Data for Western United States Wildfires Between 2005–2015: Dust Emissions, Chloride Depletion, and Most Enhanced Aerosol Constituents, *J. Geophys. Res.-Atmos.*, 122, 8951–8966, <https://doi.org/10.1002/2017JD026547>, 2017.
- Schlosser, J. S., Dadashazar, H., Edwards, E.-L., Mardi, A. H., Prabhakar, G., Stahl, C., Jonsson, H. H. and Sorooshian, A.: Relationships between supermicrometer sea salt aerosol and marine boundary layer conditions: insights from repeated identical flight patterns, *J. Geophys. Res.-Atmos.*, 125, <https://doi.org/10.1029/2019JD032346>, 2020.
- Seinfeld, J. H. and Pandis, S. N.: *Atmospheric Chemistry and Physics*, 3 Edn., New York, NY, 1152 pp., 2016.
- Silva, P. J., Liu, D. Y., Noble, C. A., and Prather, K. A.: Size and chemical characterization of individual particles resulting from biomass burning of local Southern California species, *Environ. Sci. Technol.*, 33, 3068–3076, <https://doi.org/10.1021/es980544p>, 1999.
- Sorooshian, A., Lu, M. L., Brechtel, F. J., Jonsson, H., Feingold, G., Flagan, R. C., and Seinfeld, J. H.: On the source of organic acid aerosol layers above clouds, *Environ. Sci. Technol.*, 41, 4647–4654, <https://doi.org/10.1021/es0630442>, 2007.
- Sorooshian, A., Padró, L. T., Nenes, A., Feingold, G., McComiskey, A., Hersey, S. P., Gates, H., Jonsson, H. H., Miller, S. D., Stephens, G. L., Flagan, R. C., and Seinfeld, J. H.: On the link between ocean biota emissions, aerosol, and maritime clouds: Airborne, ground, and satellite measurements off the coast of California, *Global Biogeochem. Cy.*, 23, 1–15, <https://doi.org/10.1029/2009GB003464>, 2009.
- Sorooshian, A., MacDonald, A. B., Dadashazar, H., Bates, K. H., Coggon, M. M., Craven, J. S., Crosbie, E., Edwards, E.-L., Hersey, S. P., Hodas, N., Lin, J. J., Hossein Mardi, A., Marty, A. N., Maudlin, L. C., Metcalf, A. R., Murphy, S. M., Padro, L. T., Prabhakar, G., Rissman, T. A., Schlosser, J., Shingler, T., Varutbangkul, V., Wang, Z., Woods, R. K., Chuang, P. Y., Nenes, A., Jonsson, H. H., Flagan, R. C., and Seinfeld, J. H.: A multi-year data set on aerosol-cloud-precipitation-meteorology interactions for marine stratocumulus clouds, *figshare*, <https://doi.org/10.6084/m9.figshare.5099983.v10>, 2017.
- Sorooshian, A., MacDonald, A. B., Dadashazar, H., Bates, K. H., Coggon, M. M., Craven, J. S., Crosbie, E., Hersey, S. P., Hodas, N., Lin, J. J., Negrón Marty, A., Maudlin, L. C., Metcalf, A. R., Murphy, S. M., Padró, L. T., Prabhakar, G., Rissman, T. A., Shingler, T., Varutbangkul, V., Wang, Z., Woods, R. K., Chuang, P. Y., Nenes, A., Jonsson, H. H., Flagan, R. C., and Seinfeld, J. H.: A multi-year data set on aerosol-cloud-precipitation-meteorology interactions for marine stratocumulus clouds, *Sci. Data*, 5, 1–13, <https://doi.org/10.1038/sdata.2018.26>, 2018.
- Sorooshian, A., Anderson, B., Bauer, S. E., Braun, R. A., Cairns, B., Crosbie, E., Dadashazar, H., Diskin, G., Ferrare, R., Flagan, R. C., Hair, J., Hostetler, C., Jonsson, H. H., Kleb, M. M., Liu, H., Macdonald, A. B., McComiskey, A., Moore, R., Painemal, D., Russell, L. M., Seinfeld, J. H., Shook, M., Smith, W. L., Thornhill, K., Tselioudis, G., Wang, H., Zeng, X., Zhang, B., Ziemba, L., and Zuidema, P.: Aerosol–cloud–meteorology interaction airborne field investigations: Using lessons learned from the U.S. West coast in the design of activate off the U.S. East Coast, *B. Am. Meteorol. Soc.*, 100, 1511–1528, <https://doi.org/10.1175/BAMS-D-18-0100.1>, 2019.
- Sorooshian, A., Corral, A. F., Braun, R. A., Cairns, B., Crosbie, E., Ferrare, R., Hair, J., Kleb, M. M., Hossein Mardi, A., Maring, H., McComiskey, A., Moore, R., Painemal, D., Scarino, A. J., Schlosser, J., Shingler, T., Shook, M., Wang, H., Zeng, X., Ziemba, L. and Zuidema, P.: Atmospheric Research Over the Western North Atlantic Ocean Region and North American East Coast: A Review of Past Work and Challenges Ahead, *J. Geophys. Res. Atmos.*, 125, e2019JD031626, [doi:10.1029/2019JD031626](https://doi.org/10.1029/2019JD031626), 2020.
- Strapp, J. W., Leitch, W. R., and Liu, P. S. K.: Hydrated and dried aerosol-size-distribution measurements from the Particle Measuring Systems FSSP-300 probe and the deiced PCASP-100X probe, *J. Atmos. Ocean. Technol.*, 9, 548–555, [https://doi.org/10.1175/1520-0426\(1992\)009<0548:HADASD>2.0.CO;2](https://doi.org/10.1175/1520-0426(1992)009<0548:HADASD>2.0.CO;2), 1992.
- Twomey, S.: The nuclei of natural cloud formation part II: The supersaturation in natural clouds and the variation of cloud droplet concentration, *Geofis. Pura e Appl.*, 43, 243–249, <https://doi.org/10.1007/BF01993560>, 1959.
- Twomey, S.: The influence of pollution on the shortwave albedo of clouds, *J. Atmos. Sci.*, 34, 1149–1152, [https://doi.org/10.1175/1520-0469\(1977\)034<1149:TIOPOP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOP>2.0.CO;2), 1977.
- Van Dingenen, R., Raes, F., and Jensen, N. R.: Evidence for anthropogenic impact on number concentration and sulfate content of cloud-processed aerosol particles over the North Atlantic, *J. Geophys. Res.-Atmos.*, 100, 21057–21067, <https://doi.org/10.1029/95jd02141>, 1995.
- Wang, Z., Sorooshian, A., Prabhakar, G., Coggon, M. M., and Jonsson, H. H.: Impact of emissions from shipping, land, and the ocean on stratocumulus cloud water elemental composition during the 2011 E-PEACE field campaign, *Atmos. Environ.*, 89, 570–580, <https://doi.org/10.1016/j.atmosenv.2014.01.020>, 2014.
- Wang, Z., Mora Ramirez, M., Dadashazar, H., MacDonald, A. B., Crosbie, E., Bates, K. H., Coggon, M. M., Craven, J. S., Lynch, P., Campbell, J. R., Azadi Aghdam, M., Woods, R. K., Jonsson, H., Flagan, R. C., Seinfeld, J. H., and Sorooshian, A.: Contrasting cloud composition between coupled and decoupled marine boundary layer clouds, *J. Geophys. Res.-Atmos.*, 121, 11679–11691, <https://doi.org/10.1002/2016JD025695>, 2016.
- West, R. E. L., Stier, P., Jones, A., Johnson, C. E., Mann, G. W., Bellouin, N., Partridge, D. G., and Kipling, Z.: The importance of vertical velocity variability for estimates of the indirect aerosol effects, *Atmos. Chem. Phys.*, 14, 6369–6393, <https://doi.org/10.5194/acp-14-6369-2014>, 2014.
- Wonaschütz, A., Coggon, M., Sorooshian, A., Modini, R., Frossard, A. A., Ahlm, L., Mülmenstädt, J., Roberts, G. C., Russell, L. M., Dey, S., Brechtel, F. J., and Seinfeld, J. H.: Hygroscopic properties of smoke-generated organic aerosol particles emitted in the marine atmosphere, *Atmos. Chem. Phys.*, 13, 9819–9835, <https://doi.org/10.5194/acp-13-9819-2013>, 2013.
- Xue, H. and Feingold, G.: A modeling study of the effect of nitric acid on cloud properties, *J. Geophys. Res.*, 109, D18204, <https://doi.org/10.1029/2004JD004750>, 2004.