THE GLOBAL AEROSOL SYNTHESIS AND SCIENCE PROJECT (GASSP)
Measurements and Modeling to Reduce Uncertainty

C. L. Reddington, K. S. Carslaw, P. Stier, N. Schutgens, H. Coe, D. Liu, J. Allan, J. Browse, K. J. Pringle, L. A. Lee, M. Yoshioka, J. S. Johnson, L. A. Regayre, D. V. Spracklen, G. W. Mann, A. Clarke, M. Hermann, S. Henning, H. Wex, T. B. Kristensen, W. R. Leaitch, U. Pöschl, D. Rose, M. O. Andreae, J. Schmale, Y. Kondo, N. Oshima, J. P. Schwarz, A. Nenes, B. Anderson, G. C. Roberts, J. R. Snider, C. Leck, P. K. Quinn, X. Chi, A. Ding, J. L. Jimenez, and Q. Zhang

Novel methodologies for quantifying model uncertainty are combined with an extensive new database of in situ aerosol microphysical and chemical measurements to reduce uncertainty in aerosol effects on climate.

Earth's planetary radiative balance is strongly affected by atmospheric aerosol particles, which reflect and absorb solar radiation and influence the albedo and other properties of clouds. Air pollution has altered the properties of aerosols and caused a change in radiative balance, or radiative forcing, over the industrial period of between near 0 and -2 W m^{-2} (Boucher et al. 2013). This uncertainty has persisted through all Intergovernmental Panel on Climate Change (IPCC) assessment reports since 1996 and significantly limits our understanding of historical climate change and our confidence in climate change projections (Andreae et al. 2005; Seinfeld et al. 2016).

Changes in aerosols also have important effects on regional climate, atmospheric circulation, clouds, and precipitation (Shindell and Faluvegi 2009; Booth et al. 2012; Philipona et al. 2009; Wang et al. 2014; Kaufman et al. 2005; Stevens and Feingold 2009; Rosenfeld et al. 2008; Smith et al. 2016; Bollasina et al. 2011). Our ability to reliably quantify the long-term effects of aerosols on these climate processes ultimately rests on being able to reliably simulate aerosol properties and radiative effects on regional and global scales.

The uncertainty in aerosol radiative forcing has not fallen over the last 20 years despite substantial developments in model complexity, numerous model intercomparison projects, and enormous investments in global observing systems. While our knowledge about aerosols has improved, this has not been translated into more robust global models—that is, models that can make reliable predictions in spite of uncertainties in the underlying processes (see sidebar "Model uncertainty, constraint, and robustness" for definitions of terms related to model uncertainty). A recent assessment shows that global aerosol models have a very large spread of a factor of 2–30 (depending on region) in their simulations of climate-relevant aerosol properties (Mann et al. 2014).

The main challenge we face in reducing uncertainty is that models are now very complex, which can reduce their robustness (Knutti and Sedláček 2012). Early models simulated just the mass of various

aerosol chemical components, but many now simulate the full aerosol size distribution and all the associated microphysical processes in order to more realistically simulate how aerosols interact with solar radiation and clouds (Ghan and Schwartz 2007). This added complexity affects how we tackle the uncertainty. First, simulated aerosol properties are controlled by dozens of poorly quantified processes (Textor et al. 2006; Lee et al. 2013; Kipling et al. 2016; Kinne et al. 2006), which are difficult to observe in isolation. Second, a wider range of measurements is needed to evaluate the models. Third, the necessary aerosol microphysical and chemical properties cannot be measured using satellite remote sensing instruments (Stier 2016), so model evaluation must rely on sparse measurements from aircraft, ships, and ground stations.

This article describes the Global Aerosol Synthesis and Science Project (GASSP; http://gassp.org.uk/), which has four main objectives:

- understand and reduce the persistent uncertainty in aerosol models and the associated aerosol radiative forcing by constraining the spread of model simulations using a synthesis of aerosol measurements;
- attribute the reduction in uncertainty to particular measurements so that we can understand the "value of measurements" and identify where new measurements should be prioritized;
- understand model robustness by exploring whether models constrained by measurements remain reliable when used to predict under new conditions, which is an essential requirement for calculating radiative forcing; and
- exploit our enormous investments in aerosol measurements of microphysical and chemical properties to reduce model uncertainty as much as possible.

Figure 1 shows the four main activities in GASSP.

Objective 1 collects and harmonizes in situ aerosol measurements from aircraft, ships, and ground sites to produce an easily accessed dataset suitable for statistical constraint of global aerosol models. The first phase of GASSP has focused on constraining aerosol microphysical and chemical properties that are not well constrained by Earth-observation datasets (Stier 2015). Future extension to cloud microphysical properties would be highly valuable for evaluating modeled aerosol–cloud interaction and associated radiative forcing.

Objective 2 focuses on developing new methodologies for comparing models against sparse in situ measurements. Our research has shown that there are inherent uncertainties associated with using sparse measurement data, caused, for example, by spatial inhomogeneities that are not represented in the models (see section "Representativeness of sparse measurements and the importance for model constraint").

AFFILIATIONS: Reddington, Carslaw, Browse, Pringle, Lee, Yoshioka, Johnson, Regayre, and Spracklen—School of Earth and Environment, University of Leeds, Leeds, United Kingdom; Stier and Schutgens—Department of Physics, University of Oxford, Oxford, United Kingdom; Coe, Liu, and Allan—National Centre for Atmospheric Science, and School of Earth and Environmental Sciences, University of Manchester, Manchester, United Kingdom; Mann—National Centre for Atmospheric Science, and the School of Earth and Environment, University of Leeds, Leeds, United Kingdom; Clarke—University of Hawai'i at Mānoa, Honolulu, Hawaii; Hermann, Henning, Wex, and Kristensen—Leibniz-Institut für Troposphärenforschung, Leipzig, Germany; Leaitch—Environment and Climate Change Canada, Toronto, Ontario, Canada; Pöschl, Rose, and Andreae—Max Planck Institute for Chemistry, Mainz, Germany; Schmale—Paul Scherrer Institute, Villigen, Switzerland, and Centre for Ecology and Hydrology, Edinburgh, United Kingdom; Kondo—National Institute of Polar Research, Tachikawa, Tokyo, Japan; Oshima—Meteorological Research Institute, Tsukuba, Ibaraki, Japan; Schwarz—NOAA/Earth System Research Laboratory, Boulder, Colorado; Nenes—Georgia Institute of Technology, Atlanta, Georgia; Anderson—NASA Langley Research Center, Hampton, Virginia; Roberts—Energy, Environment and Water Research Center, The Cyprus Institute, Nicosia, Cyprus; Snider—Department of Atmospheric Science, University of Wyoming, Laramie, Wyoming; Leck—Department of Meteorology, Stockholm University, Stockholm, Sweden; Quinn—NOAA/Pacific Marine Environmental Laboratory, Seattle, Washington; Chi and Ding—School of Atmospheric Sciences, Nanjing University, Nanjing, China; Jimenez—University of Colorado, Boulder, Boulder, Colorado; Zhang—University of California, Davis, Davis, California

CORRESPONDING AUTHOR: Kenneth. S. Carslaw, k.s.carslaw@leeds.ac.uk

Objective 3 is about developing methodologies to understand and quantify model uncertainty. GASSP has focused on the spread of model predictions caused by uncertainties in the processes and inputs (like emissions) in a model—known as *parametric uncertainty*—because there are established statistical methodologies for quantifying and reducing this source of uncertainty (see sidebar "Model uncertainty, constraint, and robustness"). The spread of such simulations can be similar to the spread of several structurally different models (Lee et al. 2013; Mann

et al. 2014; Kipling et al. 2016), which is the measure of uncertainty used in many assessments (Kinne et al. 2006; Lamarque et al. 2013; Eyring et al. 2016).

Objective 4 brings everything together to assess how the observational constraint of aerosol microphysical properties affects the range of aerosol radiative forcing simulated by the model.

THE GASSP AEROSOL MEASUREMENT DATABASE. A synthesis of multiple measurements is a vital component of GASSP. As we describe in the
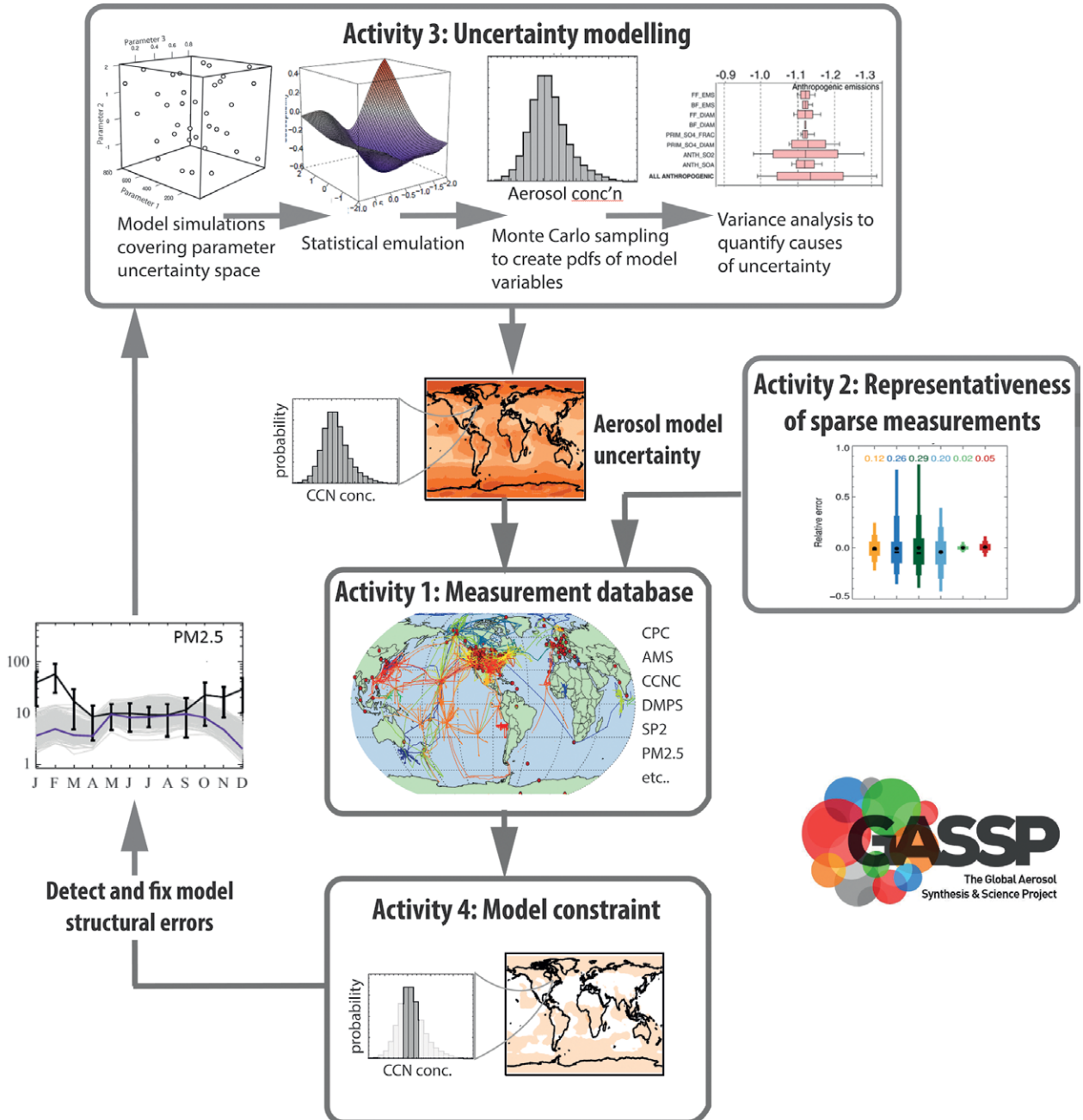

Fɪɢ. ɪ. Schematic of the overall design of GASSP.

TABLE 1. List of ground-based monitoring station networks with aerosol measurements: particle number concentration N, particle number size distribution (NSD), cloud condensation nucleus concentration (CCN), speciated mass concentration/composition (Comp), black carbon mass concentration (BC), and particle mass concentration less than 2.5 μm in diameter ($PM_{2.5}$). Short-term field campaign measurement data made at some of these stations are included in Table ES1.

Network acronym	Network name	Years	Location	Measurement
—	EBAS	1990–2014	Europe, North America, Arctic, Antarctica, Asia	N, NSD, Comp, BC, $PM_{2.5}$
ACTRIS CCN	Aerosols, Clouds, and Trace Gases Research Infrastructure CCN data	2010–14	Europe, Arctic, Brazil, South Korea	CCN
ARM	Atmospheric Radiation Measurement Program	1995–2015	Germany, United States, Brazil, Arctic, eastern Atlantic Ocean, China, India, Niger	N, NSD, CCN, Comp
AMS Global Database	Aerosol Mass Spectrometry Global Database	2000–11	Europe, North America, East Asia, South Africa, Atlantic Ocean, South America	Comp
IMPROVE	Interagency Monitoring of Protected Visual Environments	1990–2014	North America (United States)	$PM_{2.5}$
NAPS	National Air Pollution Surveillance Program	1990–2014	North America (Canada)	$PM_{2.5}$
A-PAD	Asia–Pacific Aerosol Database	2002–11	Asia	$PM_{2.5}$

TABLE 2. Definition of GASSP aerosol variables and the instruments from which they are derived.

Aerosol property	Variable	Example attributes	Meaning of attributes	Instruments (see Table 3)	Example measurements
Number concentration	N	Diameter D (nm)	Integral number concentration above a specified diameter where the upper size limit is essentially infinite	CPC	N_3, N_{10}
Number concentration in a specific size range	N	Diameters D_1 and D_2 (nm)	Integral number concentration of particles in a defined diameter range (NSD integral), where the upper size limit is small enough that it is not effectively infinite	SMPS, DMPS, APS, UHSAS, OPC, DMA	N_{50-750}
Number size distribution	NSD	Diameters of the bin edges	Particle number size distribution; particle concentration per size range	SMPS, DMPS, APS, UHSAS, OPC, DMA	—
Cloud condensation nucleus concentration	CCN	Supersaturation S (%)	Number concentration of CCN at a defined supersaturation	CCNC	$CCN_{0.2\%}$
Particle mass concentration	PM	Diameter D (μm)	Mass concentration of particulate matter with diameter smaller than a specified diameter	Gravimetric filter	PM_1, $PM_{2.5}$
Speciated mass concentration	Comp	—	Mass concentration of nonrefractory chemical components	AMS, ACSM, PILS	SO_4, Org, NO_3, Chl, NH_4
Black carbon	BC	Diameters of the bin edges	Mass and number concentration and size distribution of refractory BC	SP2	BC_{MASS}, BC_{NUM}

TABLE 3. Aerosol instruments used in GASSP.			
Acronym	**Instrument**	**Manufacturer**	**Reference(s)**
CPC	Condensation particle counter	Various (mostly TSI Inc.)	—
SMPS (DMPS)	Scanning (differential) mobility particle sizer	Various [e.g., TSI Inc. (Shoreview, MN), Grimm GmbH (Ainring, Germany)]	Wiedensohler et al. 2012
DMA	Differential mobility analyzer	Various	—
APS	Aerodynamic particle sizer	TSI Inc. (Shoreview, MN)	—
UHSAS	Ultra-high-sensitivity aerosol spectrometer	Droplet Measurement Technologies	—
PCASP	Passive Cavity Aerosol Spectrometer Probe (PCASP-100X)	Droplet Measurement Technologies	—
OPC	Optical particle counter	GRIMM Aerosol Technik GmbH and Co.KG	—
CCNC	Cloud condensation nuclei counter	Droplet Measurement Technologies, University of Wyoming	Roberts and Nenes 2005; Snider et al. 2006
—	Aethalometer	Magee Scientific	Hansen and Novakov 1989
PSAP	Particle soot absorption photometer	Radiance Research	—
MAAP	Multiangle absorption photometer	Thermo Scientific	Petzold and Schönlinner 2004
OCEC	Organic carbon elemental carbon analyser	Sunset Laboratories, Desert Research Institute	Birch and Cary 1996
AMS	Aerosol mass spectrometer	Aerodyne Research	Canagaratna et al. 2007
ACSM	Aerosol chemical speciation monitor	Aerodyne Research	Ng et al. 2011
PILS	Particle into liquid sampler	Brechtel Manufacturing Inc.	Weber et al. 2001
SP2	Single particle soot photometer	Droplet Measurement Technologies	Schwarz et al. 2006
—	Gravimetric filter analysis	GENT sampling unit, IMPROVE module A with cyclone inlet, Teflon filter	Hopke et al. 1997

"Model uncertainty and observational constraint" section, a diversity of measurements is a critical factor in constraining a complex model with many compensating sources of uncertainty.

The community's considerable investment in measurements of aerosol microphysical and chemical properties is significantly underexploited to evaluate and constrain climate models, mainly because many datasets are not available in a harmonized form (i.e., standardized format, time units, variable names, etc.) in common repositories. A key objective of GASSP was therefore to harmonize a large fraction of the world's in situ aerosol measurements and make them "user ready."

The number and sophistication of aerosol measurements has increased dramatically over the last 20 years, providing essential data to evaluate the latest model simulations of aerosol microphysical properties (see sidebar "Developments in aerosol microphysical measurements"). Long-term measurement programs have been established, such as the Global Atmosphere Watch (GAW) program; the Aerosols, Clouds, and Trace Gases Research Infrastructure (ACTRIS); and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network. These data are generally readily available (see Table 1) and have been used to evaluate multiple models (Mann et al. 2014). However, many measurements are made in short campaigns of typically a month duration. While some campaign data have been synthesized (Heintzenberg et al. 2000; Clarke and Kapustin 2002, 2010; Asmi et al. 2013) and used to evaluate multiple models (Mann et al. 2014), most data have been used to evaluate specific aspects of particular models under specific conditions, often called "golden days."

We collected measurements related to six broad aerosol properties important to the effects of aerosols on climate: particle number concentration, cloud condensation nucleus concentration, particle number size distribution, particle mass concentration less than 2.5 μm in diameter ($PM_{2.5}$), chemical composition, and black carbon (BC) mass concentration (Table 2). The

MODEL UNCERTAINTY, CONSTRAINT, AND ROBUSTNESS

Most simply, uncertainty can be defined as the spread of outputs in model simulations.

Structural uncertainty arises because there are different ways of representing the physical processes in a model because of insufficient knowledge, simplifications, different algorithms, or missing processes. Structural differences between models explain a large part of the range of predictions between different models, often evaluated through a multimodel ensemble (MME). The spread of outputs from a small number of models is often called the model diversity, and it is this diversity that has persisted through all IPCC assessments.

Process or parametric uncertainty is caused by uncertain values of model input parameters, like chemical-rate coefficients or emissions (Lee et al. 2013). Parametric uncertainty can be quantified using a set of simulations in which parameter values are systematically perturbed (a perturbed-parameter ensemble), an approach widely used in climate science (Allen et al. 2000; Pan et al. 1998; Lohmann and Ferrachat 2010; Haerter et al. 2009; Sexton et al. 2011; Shiogama et al. 2012; Yang et al. 2013). They are often combined with statistical emulation (see sidebar "Model emulation to enable dense sampling of uncertainty"), which enables model outputs to be generated for all parts of parameter space.

Constraint of a model is the process of reducing the spread of model simulations by comparing the simulations against measurements. There are various ways of defining plausible or acceptable model simulations, taking account of the measurement uncertainty and the model–measurement sampling uncertainty (see "Representativeness of sparse measurements and the importance for model constraint" section; Lee et al. 2016).

Model–measurement sampling uncertainty is the uncertainty associated with comparing a model with a measurement. Measurements sample only a small part of the atmosphere so uncertainty arises because the real world has greater spatial heterogeneity and temporal variability than is represented in a model.

A robust model is one that is reliable and can make useful predictions under different sets of conditions in spite of uncertainties in the model and its inputs (Knutti and Sedláček 2012). The robustness of a model can only be assessed if the uncertainties have been systematically explored.

---

measurements were made with 18 different aerosol instruments (Table 3), and within the six aerosol property classes there are about 70 different measured variables related to different particle size thresholds, cloud condensation nucleus supersaturations, etc. It is this diversity of measurements, compared to gas-phase chemistry (Sofen et al. 2016), that makes it so challenging to use a wide array of aerosol data in model evaluation.

The GASSP database currently contains measurements from 86 field campaigns (of a total of 119 collected) and long-term measurements from over 350 ground-based monitoring stations spanning 1990–2015 (see Table ES1 in the online supplement). It includes 20 ship campaigns, 16 ground station campaigns, 29 aircraft campaigns, and 21 campaigns involving multiple measurement platforms. We obtained data from 15 repositories (Table ES2) and, for 42 campaigns, through direct contact with investigators. The measurements include over 9,500 flight hours (over 13,000 instrument hours) and 22,000 ship hours (over 33,000 instrument hours). We estimate the total research investment in the aircraft measurements to be about $182 million (based on an average of $18,000 per flight hour plus $375,000 per detachment), not including personnel hours. The cost for ship measurements is about $57 million (based on $60,000 per day and $130,000 per cruise).

It took about 1.5 person years to collect and harmonize 52,000 data files from about 20 original file types and formats ranging from simple text files to standard formats like National Aeronautics and Space Administration (NASA) Ames (Stephens 2008) or International Consortium for Atmospheric Research on Transport and Transformation (ICARTT; Aknan et al. 2013). Considerable effort was put into merging files and correcting or standardizing units, variable names, error flags, missing value indicators, and time dimensions. We also collated metadata (as well as missing metadata from published papers) such as the location of the measurement, the sampling conditions (e.g., standard or ambient pressure), and information on the instrument type, detection limits, particle size definition, occurrence of clouds, and whether the particles were dried or not.

The data were converted to a standardized Network Common Data Form (netCDF) format (Pringle 2017). We included a large number of data fields or "tags" within the global attributes of the file, with many adopted from the National Center for Atmospheric Research (NCAR) Research Aviation Facility (RAF) file format (NCAR 2009), such as the bounding box and time coverage of the measurements. To make model evaluation more tractable, the information necessary to construct observed variables from model output variables is stored as GASSP file attributes.

GASSP uses data "levels" based on terminology widely used in remote sensing:
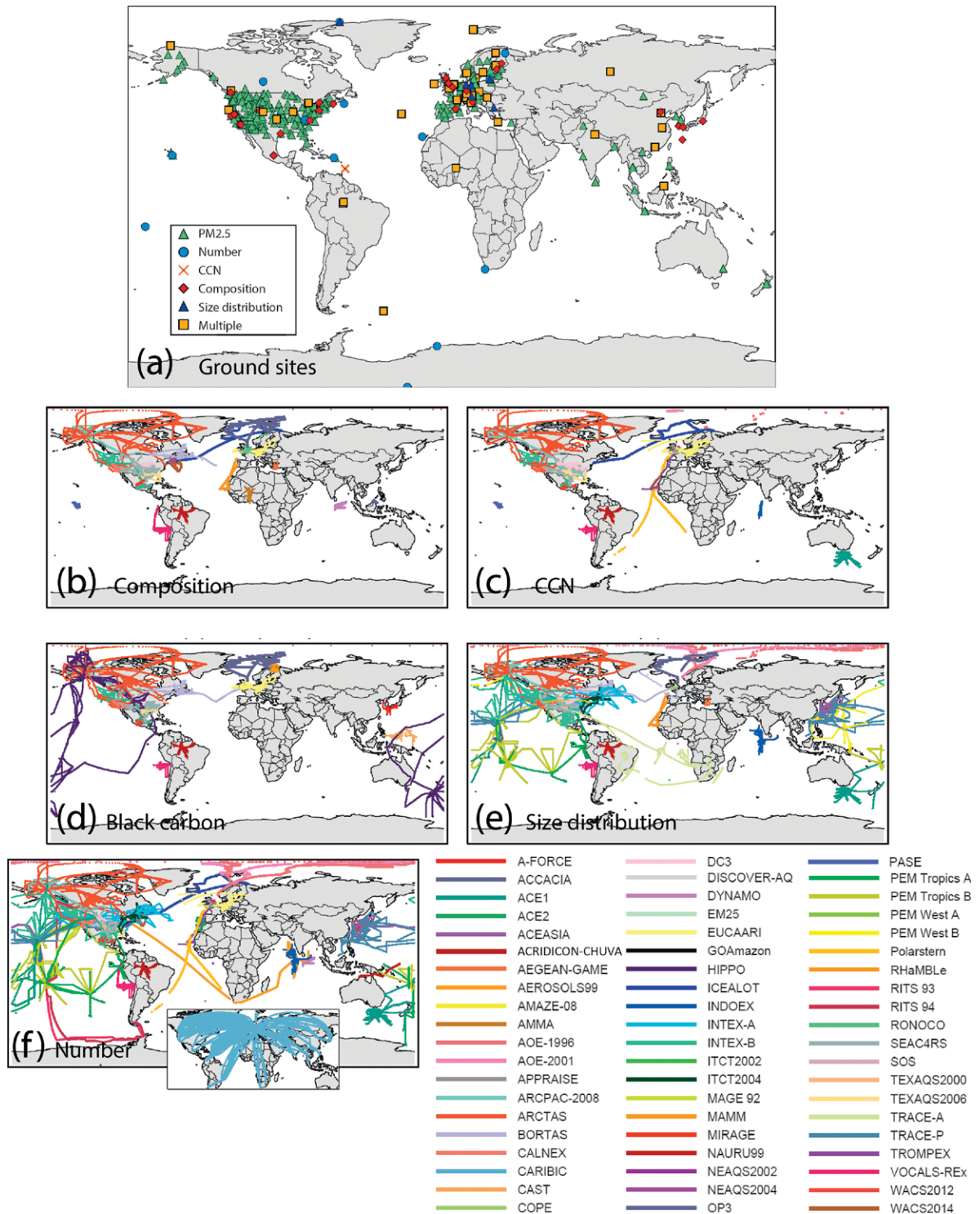
Fig. 2. The global distribution of key measurements in the GASSP database: number concentration from condensation particle counter (CPC) instruments; CCN; chemical composition from AMS and ACSM; size distributions from SMPS, DMPS, UHSAS, PCASP, DMA, and APS; and BC from SP2. (a) Ground-site measurements (sites within ±0.08° longitude and latitude are grouped under one symbol), (b) aircraft and ship measurements of chemical composition, (c) aircraft and ship measurements of CCN, (d) aircraft and ship measurements of BC, (e) aircraft and ship measurements of size distribution, and (f) aircraft and ship measurements of number concentration [Civil Aircraft for the Regular Investigation of the Atmosphere Based on an Instrument Container (CARIBIC) CPC data are shown in a separate map]. See sidebar "Developments in aerosol microphysical measurements" and Table 3 for a summary of instruments and abbreviations.

As the complexity of models increases, so does the demand for more advanced measurements. In the last 20 years many measurement techniques have been developed that are capable of measuring aerosol particle number, size, composition, and other properties with high time resolution and on a variety of platforms. A full list of instruments in GASSP is given in Table 3.

Mobility particle size spectrometers to measure size-resolved particle number concentrations have existed for a long time. International standards have been developed so that instruments can be operated consistently and those data used with confidence (Wiedensohler et al. 2012). The GASSP database includes measurements from scanning and differential mobility particle sizers (SMPS/DMPS), aerodynamic particle sizers (APS), differential mobility analyzers (DMA), ultra-high-sensitivity aerosol spectrometers (UHSAS), passive cavity aerosol spectrometer probes (PCASP), and optical particle counters (OPC).

Size distributions can be compared directly with models (Mann et al. 2014), but are often summarized in terms of the particle number concentration above a particular size (e.g., N_{50} for particles larger than 50-nm diameter). Variables like N_{50} and N_{100} are often used by modelers as representative of CCN concentrations (Mann et al. 2014).

Particle composition can now be measured in real time with AMS [and the related aerosol chemistry speciation monitor (ACSM)] (Canagaratna et al. 2007; Ng et al. 2011). The instruments provide quantitative data on the nonrefractory particle components, that is, organic matter, ammonium, sulfate, nitrate, and chloride.

Refractory BC mass can be measured in real time by SP2, which use laser-induced incandescence (Schwarz et al. 2006). High sensitivity allows airborne measurements even in pristine air. Previously, measurements had been made by combustion analysis or using optical absorption as a proxy, which involves serious artifact issues (Andreae and Gelencsér 2006).

CCN measurements were previously made using parallel plate systems (Snider et al. 2006, 2010). Now the continuous flow method (Roberts and Nenes 2005) commercialized by Droplet Measurement Technologies allows for faster measurements at supersaturations more representative of atmospheric conditions.

Level 0:
- The original raw datasets in original format.

Level 1:
- netCDF conversion of the original data file.
- Files contain the measurements of the aerosol variables (and names) along with the reported time stamp, position (latitude, longitude, altitude), and meteorology (air temperature, air pressure, relative humidity) variables, where available.
- Files are split by instrument.
- Measurement data are unchanged from original quality-assured data products supplied by originator.
- Data tags are added to metadata for additional information and to enable searching.

Level 2:
- Files are in GASSP-standardized netCDF format (including data tags).
- Time stamp is standardized to seconds since 1 January 1970 UTC (Unix time).
- Position and meteorology variables, where available, are given names compliant with Climate and Forecast (CF) metadata conventions (Eaton et al. 2011).
- Aerosol variable names are standardized.

Gridded "level 3" data formats (commonly provided with remote sensing datasets) and aerosol climatologies (Heintzenberg et al. 2000) allow for easy data handling. However, because requirements differ for specific applications, GASSP has instead supported the development of the Community Intercomparison Suite (CIS; www.cistools.net; Watson-Parris et al. 2016), an open-source command-line tool and Python library designed to read, aggregate, collocate, analyze, and visualize a wide range of ungridded and gridded datasets. CIS has a plug-in for NCAR-RAF and GASSP format data so the full GASSP level 2 database is accessible via CIS, allowing the in situ measurements to be collocated with model and many satellite products.

Figures 2 and 3 show the spatial and temporal distributions of the main measurement types. To give an impression of the density of measurements (albeit over several years), Fig. 4 shows particle concentrations from aircraft campaigns. North America is by far the most sampled region in terms of the number of airborne measurements, with more than 50% of the measurements conducted there. Most of the other airborne measurements have been made over Europe, the Pacific, and the Arctic (but less so the high Arctic), with less than 5% over the Southern Ocean, Africa, and most of Asia. Shipborne measurements of total particle concentrations are widespread, but measurements of cloud condensation nuclei (CCN), particle composition, and BC (see sidebar "Developments in aerosol microphysical measurements") are concentrated over
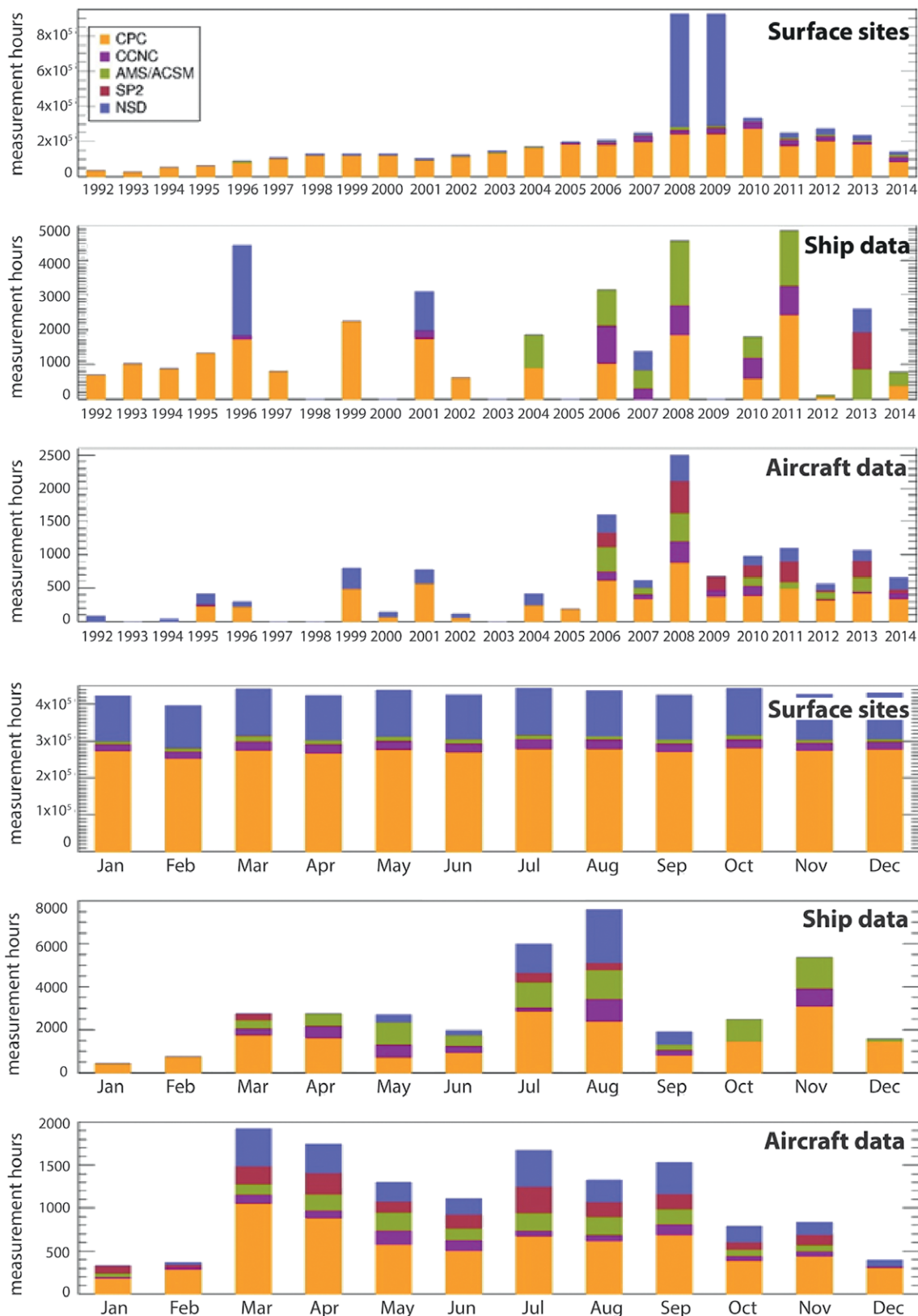
Fig. 3. Number of data hours by year and month for the main aerosol instrument types: Particle number concentration using CPC, CCN concentrations using CCN counters (CCNC), speciated mass concentrations using AMS or ACSM instruments, BC mass concentration using the SP2 instrument, and particle number size distribution (NSD) using SMPS, DMPS, UHSAS, PCASP, DMA, and APS. Instrument types are defined in Table 3.

► Fɪɢ. 4. The 3D distribution of measurements of particle number concentrations at standard temperature and pressure (STP) measured by CPC with varying cutoff sizes. Each data point represents a 5-min-average concentration. Data are shown for 20 aircraft field campaigns over several years: CARIBIC, SEAC4RS, TRACE-P, ACEASIA, ACE1, PEM tropics A, PEM tropics B, INTEX-NA, VOCALS-Rex, NEAQS-ITCT2004, ARCTAS, CALNEX, INDOEX, ARCPAC, TexAQS2006, ITCT2002, A-FORCE, DC3, PASE, and MIRAGE (see online supplement for campaign details).
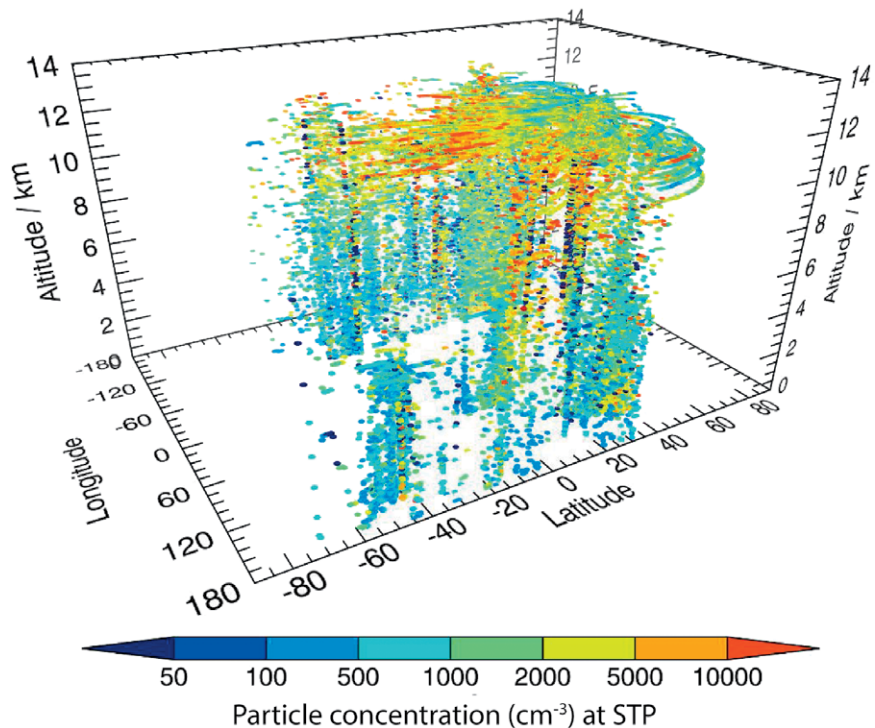
the Atlantic and Arctic. The availability of measurements from aircraft and ships varies substantially with the time of year, with only 8% of global measurements in winter (December–February).

Pristine (unpolluted) regions are considerably undersampled, especially on the continents. Such regions are informative about aerosol conditions in the preindustrial era (Hamilton et al. 2014), which are an important component of the aerosol–cloud radiative forcing uncertainty because cloud properties are very sensitive to aerosol uncertainties in clean conditions (Carslaw et al. 2013a).

Until 2004 almost all measurements were of particle concentration or size distribution, but since then
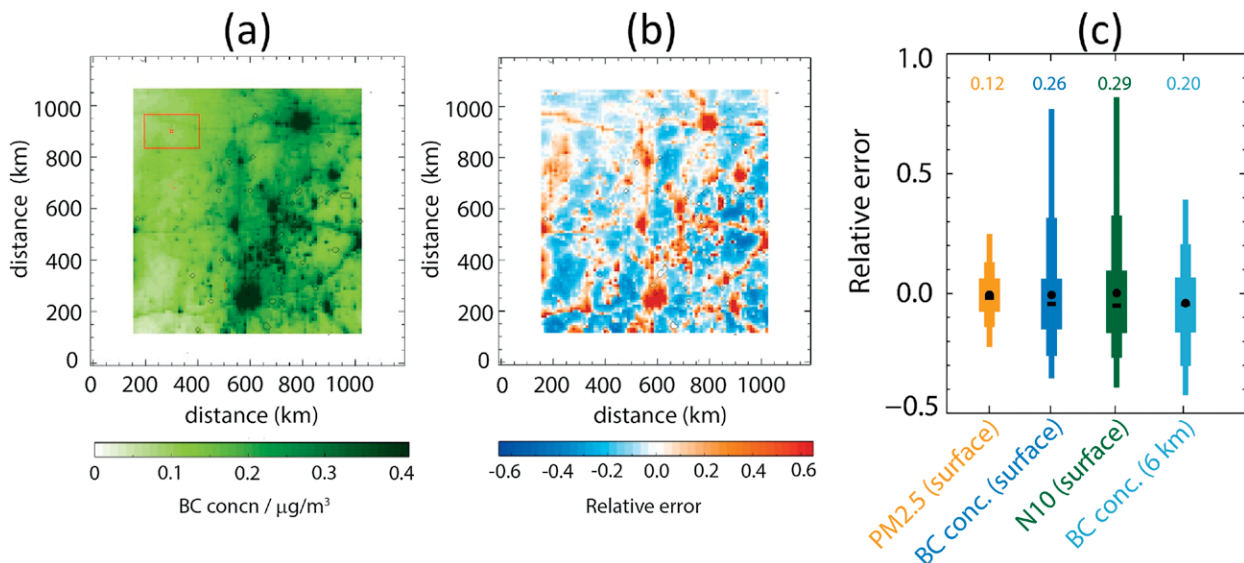
Fɪɢ. 5. Example of spatial sampling errors for a heterogeneous continental environment (Oklahoma). (a) Monthly average of a 10-km-resolution model field of surface BC concentrations representing measurement "reality." A typical climate model grid box is also shown by the red box. (b) The deduced relative model–measurement sampling error. (c) The distribution of sampling errors for four aerosol properties: $PM_{2.5}$, BC (at both the surface and 6 km), the number of particles larger than 10-nm diameter N_{10}. The different bars in (c) refer to the 50th, 82nd, and 96th interquantile ranges of the error distribution and the numbers refer to the standard deviation.

many measurements of speciated particle composition have been made by the Aerodyne aerosol mass spectrometer (AMS) and BC concentrations from the single particle soot photometer (SP2) instrument.

REPRESENTATIVENESS OF SPARSE MEASUREMENTS AND THE IMPORTANCE FOR MODEL CONSTRAINT.

Even the largest consistent collection of in situ aerosol measurements remains spatially and temporally sparse, raising important questions about how representative they are of what a low-resolution model simulates. Measurement sparseness can introduce measurement–model sampling uncertainties (see sidebar "Model uncertainty, constraint, and robustness").

In situ measurements at ground sites are particularly important data in GASSP. To quantify the potential spatial sampling errors in a typical continental environment, we used a high-resolution model to simulate "measurements" and compared them with spatially averaged fields representing a 100-km-scale global model grid cell (Schutgens et al. 2016a). Figure 5 shows that sampling errors can be as large as 80% when using monthly mean model output, but the errors are typically less than 30%. The error depends on the spatial heterogeneity of the aerosols, so particle number concentrations and BC mass concentrations have much larger sampling errors than PM$_{2.5}$ due to a greater heterogeneity in sources.

Our analysis suggests that spatial sampling errors significantly exceed measurement errors: measurement errors for PM$_{2.5}$ and BC mass concentrations are typically 15%, while typical instantaneous sampling errors (for data sampled every hour) are about 50% (Schutgens et al. 2016a). These errors can be reduced by about a factor of 3 by averaging over a month. However, there is a catch: long-term averaging increases the likelihood that model–measurement agreement is a result of compensating errors, such as underpredicting aerosol concentrations in
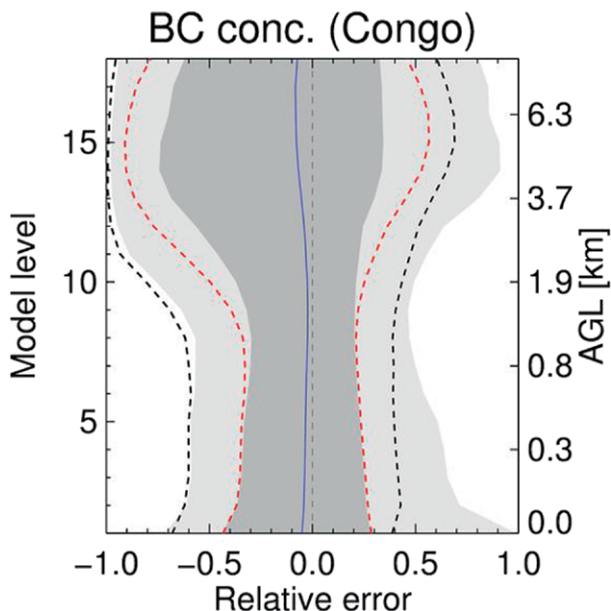

Fig. 6. Spatial sampling errors (1-h mean) for BC mass concentrations for typical flight campaigns over the Congo during the biomass-burning season. The gray shades indicate sampling errors for north–south tracks through global model grid boxes, with the shade indicating different quantile ranges (light: 2%–98%; medium: 9%–91%). The red dashed line shows the 9%–91% quantile range for east–west tracks, while the black dashed line represents point measurements.

polluted conditions but overpredicting them in clean conditions.

Another problem is that measurements are often discontinuous. Monthly mean sampling errors can be as high as 37% when measurements are discontinuous, with sometimes significant regionwide biases (Schutgens et al. 2016b). These increased errors can be largely mitigated by sampling the model to the measurement times before monthly averaging, for example, by using the CIS tools (Watson-Parris et al. 2016).

Flight campaign data present a particular challenge. In Fig. 6, we compare the spatial sampling
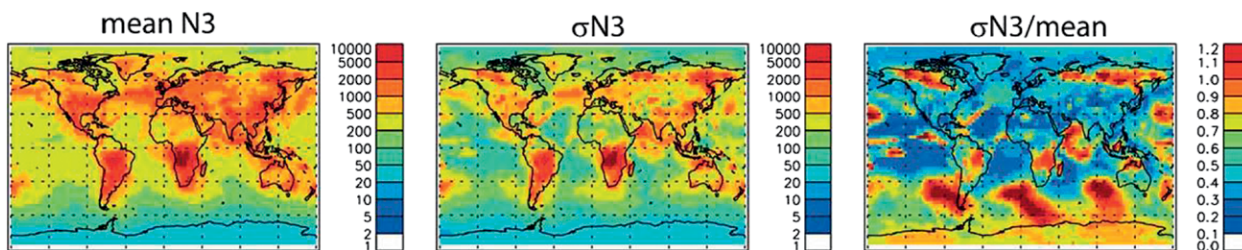
Fig. 7. (left) Calculated mean concentrations of particles larger than 3-nm diameter N_3 (cm^{-3}), (middle) the standard deviation in N_3 due to the parametric uncertainty in the model, and (right) the ratio of the standard deviation to the mean. The standard deviations were calculated by perturbing 28 aerosol model parameters (Carslaw et al. 2013b).
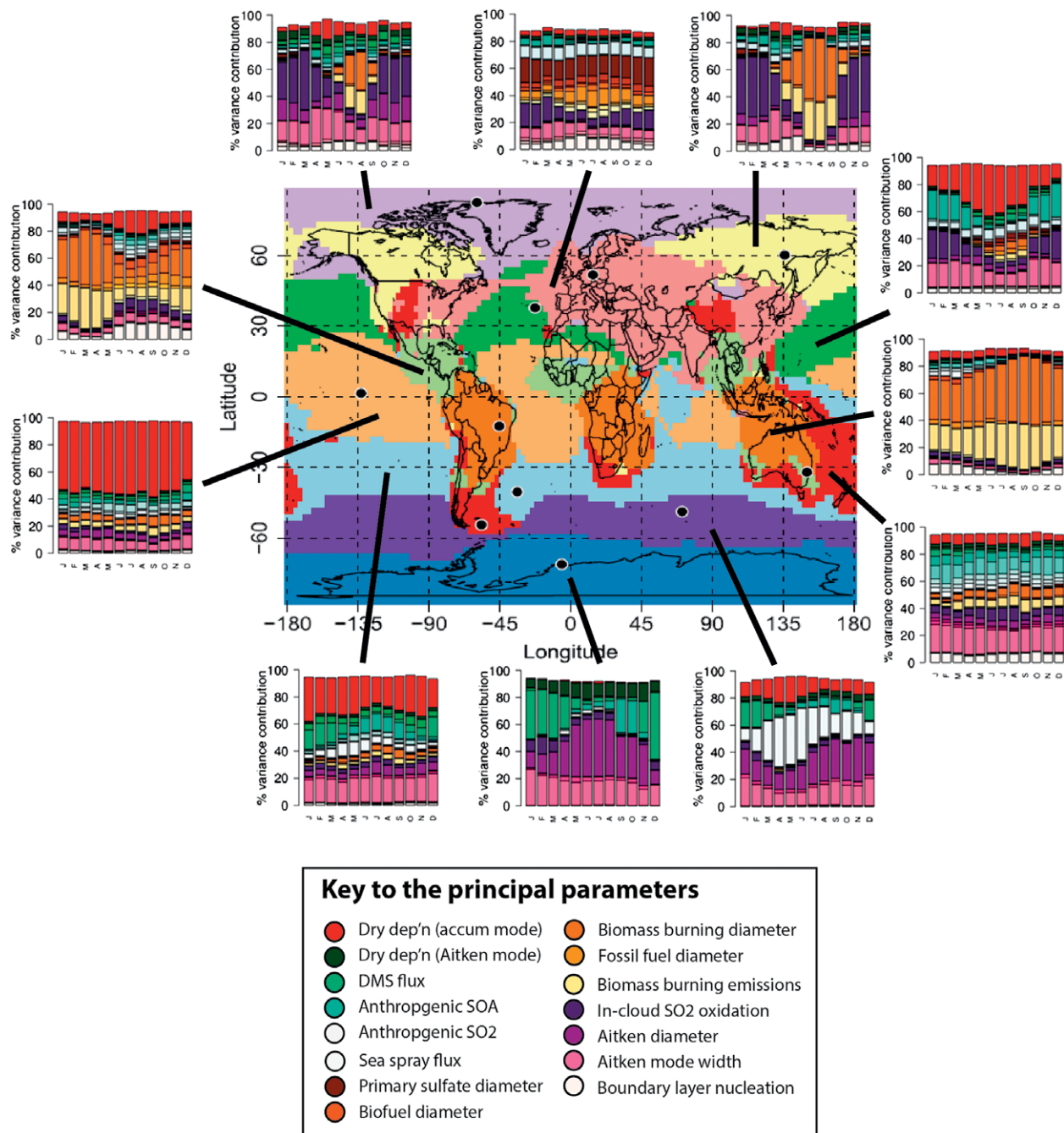
FIG. 8. Map of aerosol "uncertainty clusters" for the concentration of particles larger than 50-nm diameter N_{50}. The results were calculated using emulators, Monte Carlo sampling, and variance-based sensitivity analysis in each grid cell of the GLOMAP model (Lee et al. 2016). In each cluster the model parameters causing the uncertainty in N_{50} are similar and are shown by the bar chart thumbnails for each month. Each bar chart shows the fraction of N_{50} variance caused by each uncertain parameter, color coded according to the key [full figures are provided in Lee et al. (2016)]. Results are for annual-mean concentrations at the surface.

error distribution for instantaneous measurements from a flight campaign and a point measurement in a biomass-burning environment. While model sampling errors are smaller for flight measurements than for single point measurements (because multiple point measurements are averaged across a model grid box), the errors can still exceed 30%–50%. Campaigns in which aerosol plumes are deliberately followed will of course be particularly prone to model spatial sampling errors and biases, so they are not ideal for the statistical evaluation of global models.

Aperturbed-parameter ensemble (see sidebar "Model uncertainty, constraint, and robustness") of several hundred simulations of a climate model can only sample the multidimensional parameter uncertainty space extremely sparsely. Even with just two parameter settings for each dimension (a typical high–low perturbation approach), around 1 billion simulations are needed to cover 30 dimensions with all possible parameter combinations. In GASSP we therefore used emulators to enable a very dense sampling of the parameter space of the global model.

The purpose of emulation is to use the outputs from a perturbed parameter ensemble to generate continuous functions (multidimensional response surfaces) that describe how the model outputs vary across all the uncertain model parameters sampled by the simulations (the training data for the emulators). The emulators are validated by testing them against an additional set of model simulations.

Emulators can be built to describe the behavior of a model output at a specific location, the mean behavior over a region, or globally (Lee et al. 2012, 2013; Carslaw et al. 2013a; Regayre et al. 2014, 2015). An emulator is not a simplified model. Rather, it is a way to generate model output at untried parts of a particular model's parameter space.

In GASSP we used emulators to generate millions of "model variants" by sampling across the parameter space very densely in a Monte Carlo way. These model variants can be used to i) generate probability density functions to characterize uncertainty in a model output, ii) perform a sensitivity analysis to determine which processes affect the uncertainty, iii) compare against measurements so that plausible parts of parameter space can be found, and iv) test model robustness by using plausible parts of parameter space to quantify uncertainty in radiative forcing.

We have found that about 6–10 training simulations per parameter are needed to explore 30 dimensions of uncertainty in an aerosol–climate model, thus requiring about 180–300 simulations. When the meteorology is well defined (such as in a nudged climate model), then 1-yr simulations are sufficient (Lohmann and Ferrachat 2010).

Our overall conclusion is that the model–measurement sampling errors are likely to be nonnegligible compared to the model uncertainty that we are trying to constrain. Figure 7 shows the estimated standard deviation of monthly mean particle concentrations in the Global Model of Aerosol Processes (GLOMAP; Carslaw et al. 2013b; see "Model uncertainty and observational constraint" section). Over heterogeneous continental regions the standard deviation is about 50%–100%, so the spatial sampling errors for in situ measurements described above will limit the extent to which these model uncertainties can be reduced. Satellite measurements (e.g., of aerosol optical depth) may help alleviate the spatial sampling

problem associated with in situ measurements, although they are of course very limited in the aerosol microphysical properties they can constrain. A combination of in situ and satellite data will likely be most effective.

MODEL UNCERTAINTY AND OBSERVATIONAL CONSTRAINT.
The objective of this aspect of GASSP was to quantify global model uncertainty caused by uncertain processes and emissions

% reduction in N50 uncertainty

▶ FIG. 9. The estimated reduction in uncertainty in modeled concentrations of particles larger than 50-nm diameter N_{50} when an N_{50} measurement in central Europe is used to constrain the model. The reduction in uncertainty broadly mirrors the uncertainty cluster where anthropogenic aerosol parameters are the main source of uncertainty.
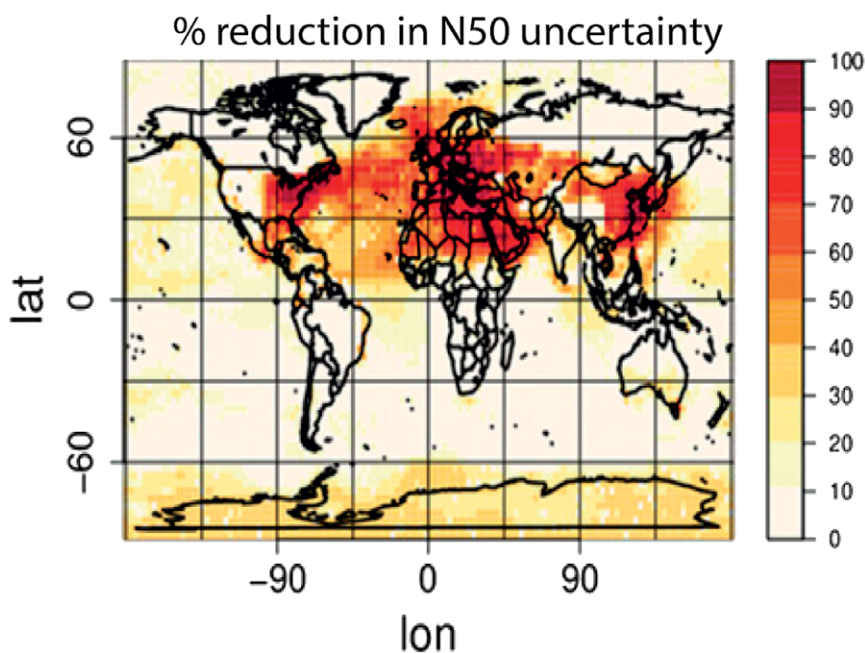
and then to explore how in situ aerosol measurements can help to reduce this uncertainty and determine what effect this has on aerosol radiative forcing.

Model simulations to quantify sensitivity and uncertainty. We used perturbed parameter ensembles of the aerosol model GLOMAP implemented in the chemical transport model Tropospheric Off-Line Model of Chemistry and Transport (TOMCAT) and the general circulation model Hadley Centre Global Environment Model (HadGEM) to quantify uncertainty. The GLOMAP–TOMCAT ensemble samples 28 uncertain aerosol parameters based on 168 simulations and has been used to explore uncertainty in present-day

cloud condensation nuclei concentrations (Lee et al. 2011, 2012, 2013), total particle concentrations (Fig. 7; Carslaw et al. 2013b), and aerosol–cloud radiative forcing since the preindustrial period (Carslaw et al. 2013a). An improved perturbed parameter ensemble explored radiative forcing regionally and over recent decades (Regayre et al. 2014, 2015). Two other ensembles of the climate model used to diagnose aerosol effective radiative forcing (Boucher et al. 2013) will be described elsewhere.

Even a few hundred simulations in a perturbed parameter ensemble are too few to adequately represent the model uncertainty because the simulations represent extremely sparse points in multidimensional
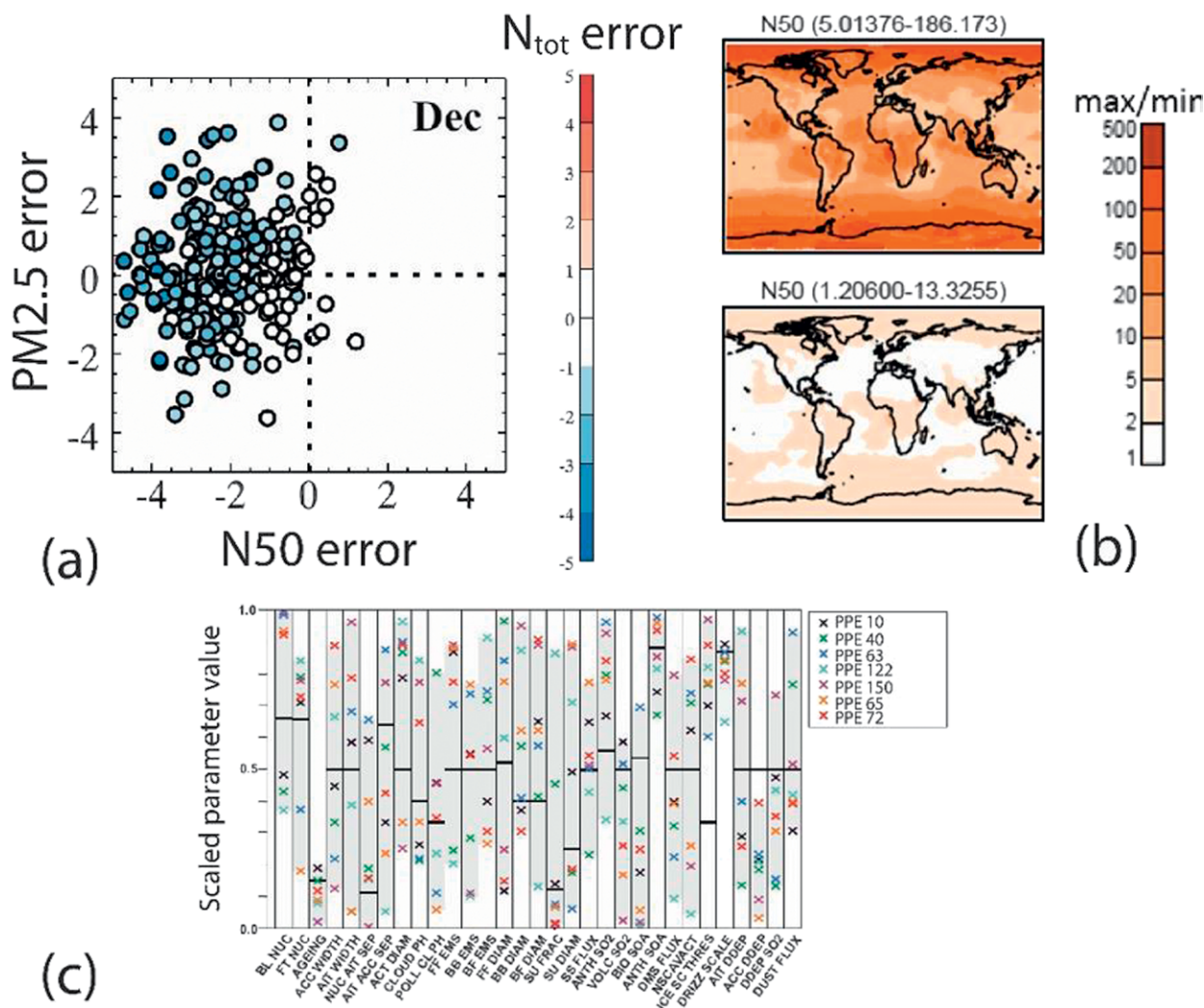

FIG. 10. Constraining the GLOMAP perturbed-parameter ensemble using multiple aerosol measurements. We used a selection of measurements of the total particle concentration N_{tot}, particle concentrations larger than 50-nm diameter N_{50}, BC concentrations, and particle mass concentration less than 2.5 μm in diameter (PM$_{2.5}$) in the boundary layer. (a) The relative error of each model simulation calculated across the measurements. (b) The N_{50} range (the maximum divided by the minimum in each grid cell) in the full ensemble and the best seven ensemble members. (c) The parameter values of the seven best ensemble members, showing that they are randomly distributed across parameter space.

parameter space. We therefore used the ensemble of simulations to build model emulators, which enable us to use Monte Carlo sampling to perform a full statistical analysis of the model (Lee et al. 2013). The sidebar "Model emulation to enable dense sampling of uncertainty" and Fig. 1 provide more details.

Relating aerosol measurements to causes of model uncertainty. One objective of GASSP is to understand which sources of uncertainty in a model can be constrained by particular measurements. As an example, Fig. 8 shows a map of the model parameters that control uncertainty in the number of particles larger than 50-nm dry diameter N_{50}, a variable that is commonly used to approximately represent cloud condensation nuclei concentrations in models (see sidebar "Developments in aerosol microphysical measurements"). The causes of uncertainty can be clustered into "uncertainty environments" within which the causes of uncertainty are similar (Lee et al. 2016). In Northern Hemisphere polluted regions (pink), measurements of N_{50} would help to constrain a range of aerosol microphysical parameters throughout the year, while in the boreal forest regions (pale yellow), N_{50} measurements would help to constrain aerosol processes in winter and fire emissions in summer.

The fact that model uncertainties can be clustered regionally suggests that large gaps in available measurements may not necessarily affect how well we can constrain modeled aerosols as long as we have measurements that are representative or characteristic of aerosols throughout these regions. To illustrate the constraint provided by even isolated measurements, Fig. 9 shows how much the N_{50} uncertainty could be reduced globally by constraining the model spread to match a measurement that is representative of aerosols over central Europe (Lee et al. 2016). Although this is only a single (monthly mean) measurement, Fig. 9 shows that it can have tremendous effect on model uncertainty provided its representativeness
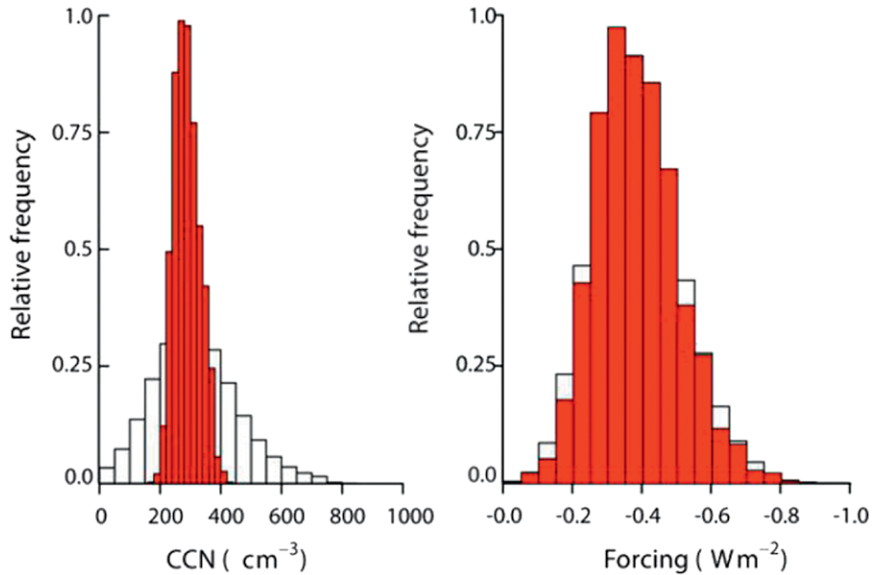

FIG. 11. Constraint of modeled **CCN** concentrations (represented by the number of particles larger than 50-nm dry diameter N_{50}) over the North Atlantic and the effect on the uncertainty in cloud albedo radiative forcing from 1750 to the present day. The white distribution shows the model spread before constraint and the red is after constraining **CCN** to lie within 300 cm^{-3} ± 30%, representing a typical measurement sampling uncertainty (see "Representativeness of sparse measurements and the importance for model constraint" section). From Lee et al. (2016).

of the region can be defined. Such model sensitivity data could in the future guide us to where new measurements should be made, and what effect the measurement will have on model uncertainty.

Observational constraint and model robustness. We have tried two approaches to constrain the spread of simulated aerosols. The first approach was to simply select a small number of simulations from the GLOMAP ensemble that best match the GASSP measurements (Fig. 10). A subset of the observationally most plausible simulations can be selected (those with a small average error compared to multiple measurements), which leads to a much narrower uncertainty range compared to the full ensemble (Fig. 10b).

Selecting the best runs from an ensemble produces a model that is close to aerosol measurements, but it is not sufficient to understand the reduction in aerosol radiative forcing uncertainty. This is because the best simulations lie very sparsely in multidimensional parameter space (Fig. 10c) so they are not statistically representative. Nevertheless, it is worth bearing in mind that in multimodel ensembles most modeling centers typically choose one model variant from the many that might be plausible (Kinne et al. 2006; Lamarque et al. 2013; Mann et al. 2014; Eyring et al. 2016).
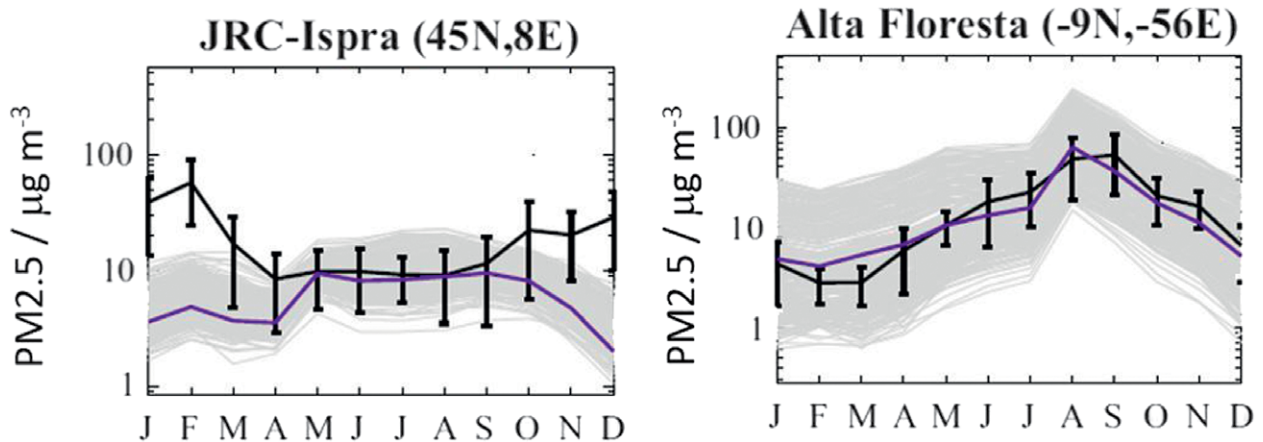
FIG. 12. An example of using a perturbed-parameter ensemble to detect a structural deficiency in the **GLOMAP** model. The black line and error bars show measured particle mass concentration less than 2.5 μm in diameter (**$PM_{2.5}$**). The gray lines show over 300 model simulations exploring combinations of 31 model parameters. The purple line shows the mean of the model ensemble.

Our second approach was to use emulators (see sidebar "Model emulation to enable dense sampling of uncertainty") to generate several million model variants that densely sample the 28-dimensional parameter space (Lee et al. 2016). The advantage of this approach is that we can find the observationally plausible regions of the model parameter space, rather than relying on just a few sparse simulations.

Figure 11 shows the result of selecting model variants from a set of 3,000,000 that match hypothetical measurements of N_{50} within a typical 30% sampling uncertainty over the North Atlantic (Lee et al. 2016). This relatively tight constraint of N_{50} concentrations has very little effect on the spread of the calculated cloud albedo radiative forcing, even though we have eliminated over 95% of the initial parameter space and the aerosol model is assumed to be the only source of uncertainty in the calculated forcing. Further work is now underway to relate all the measurement types in GASSP to their effectiveness at constraining the modeled radiative forcing.

This is an important discovery of the GASSP project—that even a tightly constrained global aerosol model can still generate a wide range of aerosol radiative forcings even though the change in aerosols directly causes the forcing, and we have assumed that the uncertainty in the simulated forcing comes only from the aerosol component of the model. The explanation is that the observationally plausible model variants lie in widely distributed parts of parameter space (just as in Fig. 10c). We describe these as "equifinal models" (Beven 2006). Only by sampling across the whole parameter space of the model has it been possible to demonstrate the existence of these equifinal models.

Identifying model structural errors. The comparison of a perturbed parameter ensemble with measurements provides a way to identify potential model structural errors. For example, Fig. 12 shows that the full GLOMAP ensemble, sampling 28 dimensions of parameter uncertainty, fails to reproduce the seasonal cycle of $PM_{2.5}$ at one site despite an order of magnitude model spread. This bias, which is apparent in other aerosol properties, can be attributed to poor representation of regional aerosol emissions. The extensive GASSP aerosol database combined with the perturbed parameter ensembles provides an optimum way to detect such structural errors.

SUMMARY AND FUTURE PROSPECTS.
GASSP set out to understand and reduce uncertainty in model simulations of aerosol radiative forcing caused by the uncertainties in the aerosol component of the model.

What have we learned about the nature and availability of aerosol in situ measurements in the context of model constraint? GASSP has created a harmonized dataset of nearly a quarter of a century of aerosol in situ measurements comprising over 46,000 measurement hours from aircraft and ships and from over 300 surface sites (Fig. 2). We intend to make these data readily accessible for wider use with an appropriate data protocol and with the permission of the data providers and other data centers. The database can, and should, be added to in the future.

Aerosol measurements are highly diverse and difficult to harmonize (the GASSP database includes 70 measured aerosol variables). However, this diversity is a strength because the aerosol system is complex

with many compensating uncertainties (Lee et al. 2016), so optimum model constraint is achieved with the greatest diversity of measurements. While seeking the most valuable measurements, we should also embrace measurement diversity and ensure that we have procedures that enable the measurement data to be ingested into models as easily as possible, such as GASSP has helped to facilitate.

The distribution of global aerosol in situ measurements is not optimum for constraining models. The overwhelming focus of aerosol measurement campaigns has been on "process-based studies" rather than to obtain a representative sample from different aerosol environments. Model uncertainty reduction would be accelerated if we put more effort into understanding the representativeness of measurements, how they characterize aerosol properties in different environments, and how they help to reduce the spread of plausible model simulations.

Pristine (unpolluted) environments are considerably undersampled. Analysis in GASSP shows that much of the uncertainty in radiative forcing stems from the properties of pristine aerosol environments (Carslaw et al. 2013a). We have some idea where near-pristine, preindustrial-like environments exist (Hamilton et al. 2014), such as the summertime Arctic, some remote boreal regions, and the western Pacific, which should provide some guide to where future measurements could be prioritized.

The use of sparse point measurements imposes limits on how much model uncertainty can be reduced because the real atmosphere has greater spatial and temporal variability than a low-resolution global model (Schutgens et al. 2016a,b). Typical spatial sampling errors of 30%–50% (Figs. 5, 6) are smaller than the current parametric uncertainty in single models (Fig. 7; Lee et al. 2013; Carslaw et al. 2013b) and multiple models (Mann et al. 2014), but may still be large enough to limit the reduction in radiative forcing uncertainty (Lee et al. 2016).

What have we learned about our ability to constrain radiative forcing uncertainty caused by uncertainty in aerosol models? Aerosol–climate models are close to becoming an overdetermined system with many interacting sources of uncertainty but a limited range of observations to constrain them (Haerter et al. 2009; Lohmann and Ferrachat 2010; Lee et al. 2016). Although the spread of aerosol model simulations can be constrained by measurements, there are many "model variants" that can achieve equally plausible model–measurement agreement, and these variants may simulate a wide range of aerosol forcings (Lee et al. 2016). The implication is that agreement of a single tuned model with measurements does not imply a robust model; that is, there are likely to be other plausible model variants that will simulate different aerosol radiative forcings. Such variants need to be identified if we are to quantify model uncertainty.

Given the complexity of the aerosol–cloud–climate system, there is unlikely to be a shortcut to reducing uncertainty. Rather, reduction in uncertainty will be achieved by simultaneously applying extensive and diverse observational constraints on the whole system. This means constraining aerosol, cloud, and radiation state variables as well as the relationships between them, which often relate to specific processes and the behavior of the system (Quaas et al. 2009; Feingold et al. 2016; Ghan et al. 2016).

What are the priorities for future research on aerosol measurements and model uncertainty? While there will always be a need to improve specific processes in models in order to eliminate structural errors, we argue that model development needs to be pursued in conjunction with intensified efforts to quantify and constrain model uncertainty. There is comparatively little effort devoted to understanding the radiative forcing problem from a system uncertainty point of view.

Further effort is needed to enable future aerosol measurements to be harmonized so that modelers and experimentalists can collaborate more easily to reduce model uncertainty. Network infrastructures like ACTRIS make a substantial contribution in this direction, but extensive campaign-type measurements should also be harmonized. Measurement standards and quality control are vital (Wiedensohler et al. 2012), but equally important is data harmonization to reduce the very large number of data formats that have proliferated (see "The GASSP aerosol measurement database" section).

We should prioritize new measurements in a targeted way with the specific objective of reducing model uncertainty. GASSP has shown that we can define representative "uncertainty environments" (Fig. 8), which help to establish the effect that specific measurements will have on model uncertainty (Fig. 9). Such information from models could help guide measurement strategies and priorities.

A greater number of longer-term measurements with greater global coverage, particularly in undersampled environments, would help to constrain global aerosol models. Development and deployment of low-cost sensors might be one way to achieve the required coverage.

A similar effort to GASSP dedicated to collecting and harmonizing cloud microphysical properties, such as droplet and ice crystal number concentrations, would be very valuable. While cloud droplet effective radii can be retrieved from satellite measurements, in situ measurements at cloud base, ideally along with updraft speed, would provide a strong constraint on modeled aerosol–cloud interactions.

A logical next step is to extend the statistical approach used in GASSP to multiple models and thereby merge efforts on parametric and structural uncertainty (Shiogama et al. 2014). Such an approach would be particularly useful for identifying the causes of model–observation bias because we will learn from the similarities and differences in the structural deficiencies across models.

Finally, greater progress will be made through closer cooperation of modeling and observational scientists, as increasingly achieved in projects like AeroCom and ACTRIS. Such interaction is particularly necessary in aerosol science because of the huge diversity and complexity of aerosol measurements. Greater interaction will enable a larger fraction of aerosol measurements to be used routinely in model evaluation and constraint, leading to a long-sought reduction in aerosol model uncertainty and enhancing our ability to simulate historical and future climates.

REFERENCES

Aknan, A., G. Chen, J. Crawford, and E. Williams, 2013: ICARTT data format. NASA, accessed 15 December 2015. [Available online at https://www-air.larc.nasa.gov/missions/etc/IcarttDataFormat.htm.]

Allen, M. R., P. A. Stott, J. F. B. Mitchell, R. Schnur, and T. L. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617–620, doi:10.1038/35036559.

Andreae, M. O., and A. Gelencsér, 2006: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. *Atmos. Chem. Phys.*, **6**, 3131–3148, doi:10.5194/acp-6-3131-2006.

——, C. D. Jones, and P. M. Cox, 2005: Strong present-day aerosol cooling implies a hot future. *Nature*, **435**, 1187–1190, doi:10.1038/nature03671.

Asmi, A., and Coauthors, 2013: Aerosol decadal trends—Part 2: In-situ aerosol particle number concentrations at GAW and ACTRIS stations. *Atmos. Chem. Phys.*, **13**, 895–916, doi:10.5194/acp-13-895-2013.

Beven, K., 2006: A manifesto for the equifinality thesis. *J. Hydrol.*, **320**, 18–36, doi:10.1016/j.jhydrol.2005.07.007.

Birch, M. E., and R. A. Cary, 1996: Elemental carbon-based method for monitoring occupational exposures to particulate diesel exhaust. *Aerosol Sci. Technol.*, **25**, 221–241, doi:10.1080/02786829608965393.

Bollasina, M. A., Y. Ming, and V. Ramaswamy, 2011: Anthropogenic aerosols and the summer monsoon. *Science*, **334**, 502–505, doi:10.1126/science.1204994.

Booth, B. B. B., N. J. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232, doi:10.1038/nature10946.

Boucher, O., and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 571–657.

Canagaratna, M. R., and Coauthors, 2007: Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. *Mass Spectrom. Rev.*, **26**, 185–222, doi:10.1002/mas.20115.

Carslaw, K. S., and Coauthors, 2013a: Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, **503**, 67–71, doi:10.1038/nature12674.

——, L. A. Lee, C. L. Reddington, G. W. Mann, and K. J. Pringle, 2013b: The magnitude and sources of uncertainty in global aerosol. *Faraday Discuss.*, **165**, 495–512, doi:10.1039/c3fd00043e.

Clarke, A., and V. Kapustin, 2002: A Pacific aerosol survey. Part I: A decade of data on particle production, transport, evolution, and mixing in the troposphere. *J. Atmos. Sci.*, **59**, 363–382, doi:10.1175/1520-0469(2002)059<0363:APASPI>2.0.CO;2.

——, and V. Kapustin, 2010: Hemispheric aerosol vertical profiles: Anthropogenic impacts on optical depth and cloud nuclei. *Science*, **329**, 1488–1492, doi:10.1126/science.1188838.

Eaton, B., and Coauthors, 2011: netCDF Climate and Forecast (CF) metadata conventions. Lawrence Livermore National Laboratory. [Available online at http://cfconventions.org/cf-conventions/v1.6.0/cf-conventions.html.]

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization.

Geosci. Model Dev., **9**, 1937–1958, doi:10.5194/gmd-9-1937-2016.

Feingold, G., A. McComiskey, T. Yamaguchi, J. S. Johnson, K. S. Carslaw, and K. S. Schmidt, 2016: New approaches to quantifying aerosol influence on the cloud radiative effect. *Proc. Natl. Acad. Sci. USA*, **113**, 5812–5819, doi:10.1073/pnas.1514035112.

Ghan, S. J., and S. E. Schwartz, 2007: Aerosol properties and processes: A path from field and laboratory measurements to global climate models. *Bull. Amer. Meteor. Soc.*, **88**, 1059–1083, doi:10.1175/BAMS-88-7-1059.

——, and Coauthors, 2016: Challenges in constraining anthropogenic aerosol effects on cloud radiative forcing using present-day spatiotemporal variability. *Proc. Natl. Acad. Sci. USA*, **113**, 5804–5811, doi:10.1073/pnas.1514036113; Corrigendum, **113**, E3049, doi:10.1073/pnas.1604888113.

Haerter, J. O., E. Roeckner, L. Tomassini, and J.-S. von Storch, 2009: Parametric uncertainty effects on aerosol radiative forcing. *Geophys. Res. Lett.*, **36**, L15707, doi:10.1029/2009GL039050.

Hamilton, D. S., L. A. Lee, K. J. Pringle, C. L. Reddington, D. V Spracklen, and K. S. Carslaw, 2014: Occurrence of pristine aerosol environments on a polluted planet. *Proc. Natl. Acad. Sci. USA*, **111**, 18 466–18 471, doi:10.1073/pnas.1415440111.

Hansen, D. A., and T. Novakov, 1989: Real-time measurements of the size fractionation of ambient black carbon aerosols at elevated humidities. *Aerosol Sci. Technol.*, **10**, 106–110, doi:10.1080/02786828908959225.

Heintzenberg, B. J., D. C. Covert, and R. Van Dingenen, 2000: Size distribution and chemical composition of marine aerosols: A compilation and review. *Tellus*, **52B**, 1104–1122, doi:10.3402/tellusb.v52i4.17090.

Hopke, P. K., Y. Xie, T. Raunemaa, S. Biegalski, S. Landsberger, W. Maenhaut, P. Artaxo, and D. Cohen, 1997: Characterization of the Gent Stacked Filter Unit PM10 Sampler. *Aerosol Sci. Technol.*, **27**, 726–735, doi:10.1080/02786829708965507.

Kaufman, Y. J., I. Koren, L. A. Remer, D. Rosenfeld, and Y. Rudich, 2005: The effect of smoke, dust, and pollution aerosol on shallow cloud development over the Atlantic Ocean. *Proc. Natl. Acad. Sci. USA*, **102**, 11 207–11 212, doi:10.1073/pnas.0505191102.

Kinne, S., and Coauthors, 2006: An AeroCom initial assessment—Optical properties in aerosol component modules of global models. *Atmos. Chem. Phys.*, **6**, 1815–1834, doi:10.5194/acp-6-1815-2006.

Kipling, Z., and Coauthors, 2016: What controls the vertical distribution of aerosol? Relationships between process sensitivity in HadGEM3–UKCA

and inter-model variation from AeroCom Phase II. *Atmos. Chem. Phys.*, **16**, 2221–2241, doi:10.5194/acp-16-2221-2016.

Knutti, R., and J. Sedláček, 2012: Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Climate Change*, **3**, 369–373, doi:10.1038/nclimate1716.

Lamarque, J.-F., and Coauthors, 2013: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): Overview and description of models, simulations and climate diagnostics. *Geosci. Model Dev.*, **6**, 179–206, doi:10.5194/gmd-6-179-2013.

Lee, L. A., K. S. Carslaw, K. J. Pringle, G. W. Mann, and D. V. Spracklen, 2011: Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmos. Chem. Phys.*, **11**, 12253–12273, doi:10.5194/acp-11-12253-2011.

——, ——, ——, and ——, 2012: Mapping the uncertainty in global CCN using emulation. *Atmos. Chem. Phys.*, **12**, 9739–9751, doi:10.5194/acp-12-9739-2012.

——, K. J. Pringle, C. L. Reddington, G. W. Mann, P. Stier, D. V. Spracklen, J. R. Pierce, and K. S. Carslaw, 2013: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei. *Atmos. Chem. Phys.*, **13**, 8879–8914, doi:10.5194/acp-13-8879-2013.

——, C. L. Reddington, and K. S. Carslaw, 2016: On the relationship between aerosol model uncertainty and radiative forcing uncertainty. *Proc. Natl. Acad. Sci. USA*, **113**, 5820–5827, doi:10.1073/pnas.1507050113.

Lohmann, U., and S. Ferrachat, 2010: Impact of parametric uncertainties on the present-day climate and on the anthropogenic aerosol effect. *Atmos. Chem. Phys.*, **10**, 11373–11383, doi:10.5194/acp-10-11373-2010.

Mann, G. W., and Coauthors, 2014: Intercomparison and evaluation of global aerosol microphysical properties among AeroCom models of a range of complexity. *Atmos. Chem. Phys.*, **14**, 4679–4713, doi:10.5194/acp-14-4679-2014.

NCAR, 2009: NCAR-RAF netCDF conventions. Accessed 15 December 2015. [Available online at www.eol.ucar.edu/raf/software/netCDF.html.]

Ng, N. L., and Coauthors, 2011: An Aerosol Chemical Speciation Monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol. *Aerosol Sci. Technol.*, **45**, 780–794, doi:10.1080/02786826.2011.560211.

Pan, W., M. A. Tatang, G. J. McRae, and R. G. Prinn, 1998: Uncertainty analysis of indirect radiative by anthropogenic sulfate aerosols forcing. *J. Geophys. Res.*, **103**, 3815–3823, doi:10.1029/97JD02809.

Petzold, A., and M. Schönlinner, 2004: Multi-angle absorption photometry—A new method for the measurement of aerosol light absorption and atmospheric black carbon. *J. Aerosol Sci.*, **35**, 421–441, doi:10.1016/j.jaerosci.2003.09.005.

Philipona, R., K. Behrens, and C. Ruckstuhl, 2009: How declining aerosols and rising greenhouse gases forced rapid warming in Europe since the 1980s. *Geophys. Res. Lett.*, **36**, L02806, doi:10.1029/2008GL036350.

Pringle, K. J., 2017: GASSP processing scripts. Accessed 15 December 2015. [Available online at https://github.com/KirstyPringle/GASSP_Processing.]

Quaas, J., Y. Ming, S. Menon, T. Takemura, M. Wang, J. E. Penner, A. Gettelman, and U. Lohmann, 2009: Aerosol indirect effects—General circulation model intercomparison and evaluation with satellite data. *Atmos. Chem. Phys.*, **9**, 8697–8717, doi:10.5194/acp-9-8697-2009.

Regayre, L. A., and Coauthors, 2014: Uncertainty in the magnitude of aerosol-cloud radiative forcing over recent decades. *Geophys. Res. Lett.*, **41**, 9040–9049, doi:10.1002/2014GL062029.

——, and Coauthors, 2015: The climatic importance of uncertainties in regional aerosol–cloud radiative forcings over recent decades. *J. Climate*, **28**, 6589–6607, doi:10.1175/JCLI-D-15-0127.1.

Roberts, G. C., and A. Nenes, 2005: A continuous-flow streamwise thermal-gradient CCN chamber for atmospheric measurements. *Aerosol Sci. Technol.*, **39**, 206–221, doi:10.1080/027868290913988.

Rosenfeld, D., U. Lohmann, G. B. Raga, C. D. O. Dowd, M. Kulmala, S. Fuzzi, A. Reissell, and M. O. Andreae, 2008: Flood or drought: How do aerosols affect precipitation? *Science*, **321**, 1309–1314, doi:10.1126/science.1160606.

Schutgens, N. J., E. Gryspeerdt, N. Weigum, S. Tsyro, D. Goto, M. Schulz, and P. Stier, 2016a: Will a perfect model agree with perfect observations? The impact of spatial sampling. *Atmos. Chem. Phys.*, **16**, 6335–6353, doi:10.5194/acp-16-6335-2016.

——, D. G. Partridge, and P. Stier, 2016b: The importance of temporal collocation for the evaluation of aerosol models with observations. *Atmos. Chem. Phys.*, **16**, 1065–1079, doi:10.5194/acp-16-1065-2016.

Schwarz, J. P., and Coauthors, 2006: Single-particle measurements of midlatitude black carbon and light-scattering aerosols from the boundary layer to the lower stratosphere. *J. Geophys. Res.*, **111**, D16207, doi:10.1029/2006JD007076.

Seinfeld, J. H., and Coauthors, 2016: Improving our fundamental understanding of the role of aerosol-cloud interactions in the climate system. *Proc.*

Natl. Acad. Sci. USA, **113**, 5781–5790, doi:10.1073 /pnas.1514043113.

Sexton, D. M. H., J. M. Murphy, M. Collins, and M. J. Webb, 2011: Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Climate Dyn.*, **38**, 2513–2542, doi:10.1007 /s00382-011-1208-9.

Shindell, D., and G. Faluvegi, 2009: Climate response to regional radiative forcing during the twentieth century. *Nat. Geosci.*, **2**, 294–300, doi:10.1038/ngeo473.

Shiogama, H., and Coauthors, 2012: Perturbed physics ensemble using the MIROC5 coupled atmosphere–ocean GCM without flux corrections: Experimental design and results. *Climate Dyn.*, **39**, 3041–3056, doi:10.1007/s00382-012-1441-x.

——, M. Watanabe, T. Ogura, T. Yokohata, and M. Kimoto, 2014: Multi-parameter multi-physics ensemble (MPMPE): A new approach exploring the uncertainties of climate sensitivity. *Atmos. Sci. Lett.*, **15**, 97–102, doi:10.1002/asl2.472.

Smith, D. M., and Coauthors, 2016: Role of volcanic and anthropogenic aerosols in the recent global surface warming slowdown. *Nat. Climate Change*, **6**, 936–940, doi:10.1038/nclimate3058.

Snider, J., M. Petters, P. Wechsler, and P. Liu, 2006: Supersaturation in the Wyoming CCN instrument. *J. Atmos. Oceanic Technol.*, **23**, 1323–1339, doi:10.1175 /JTECH1916.1.

——, and Coauthors, 2010: Intercomparison of cloud condensation nuclei and hygroscopic fraction measurements: Coated soot particles investigated during the LACIS Experiment in November (LExNo). *J. Geophys. Res.*, **115**, D11205, doi:10.1029/2009JD012618.

Sofen, E. D., and Coauthors, 2016: Gridded global surface ozone metrics for atmospheric chemistry model evaluation. *Earth Syst. Sci. Data*, **8**, 41–59, doi:10.5194 /essd-8-41-2016.

Stephens, A., 2008: NASA Ames format for data exchange: Documentation. British Atmospheric Data Centre, accessed 15 December 2015. [Available online at http://cedadocs.badc.rl.ac.uk/73/.]

Stevens, B., and G. Feingold, 2009: Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, **461**, 607–613, doi:10.1038/nature08281.

Stier, P., 2015: Limitations of passive satellite remote sensing to constrain global cloud condensation nuclei. *Atmos. Chem. Phys.*, **16**, 6595–6607, doi:10.5194/acp -16-6595-2016.

——, 2016: Limitations of passive remote sensing to constrain global cloud condensation nuclei. *Atmos. Chem. Phys.*, **16**, 6595–6607, doi:10.5194/acp-16-6595-2016.

Textor, C., and Coauthors, 2006: Analysis and quantification of the diversities of aerosol life cycles within AeroCom. *Atmos. Chem. Phys.*, **6**, 1777–1813, doi:10.5194/acp-6-1777-2006.

Wang, Y., and Coauthors, 2014: Assessing the effects of anthropogenic aerosols on Pacific storm track using a multiscale global climate model. *Proc. Natl. Acad. Sci. USA*, **111**, 6894–6899, doi:10.1073/pnas.1403364111.

Watson-Parris, D., N. Schutgens, N. Cook, Z. Kipling, P. Kershaw, E. Gryspeerdt, B. Lawrence, and P. Stier, 2016: Community Intercomparison Suite (CIS) v1.3.2: A tool for intercomparing models and observations. *Geosci. Model Dev.*, **9**, 3093–3110, doi:10.5194/gmd -9-3093-2016.

Weber, R. J., D. Orsini, Y. Daun, Y. Lee, P. J. Klotz, and F. Brechtel, 2001: A particle-into-liquid collector for rapid measurement of aerosol bulk chemical composition. *Aerosol Sci. Technol.*, **35**, 718–727, doi:10.1080/02786820152546761.

Wiedensohler, A., and Coauthors, 2012: Mobility particle size spectrometers: Harmonization of technical standards and data structure to facilitate high quality long-term observations of atmospheric particle number size distributions. *Atmos. Meas. Tech.*, **5**, 657–685, doi:10.5194/amt-5-657-2012.

Yang, B., and Coauthors, 2013: Uncertainty quantification and parameter tuning in the CAM5 Zhang-McFarlane convection scheme and impact of improved convection on the global circulation and climate. *J. Geophys. Res. Atmos.*, **118**, 395–415, doi:10.1029/2012JD018213.

Find out from the authoritative source

for definitions of meteorological terms.

[What's a dust devil?]
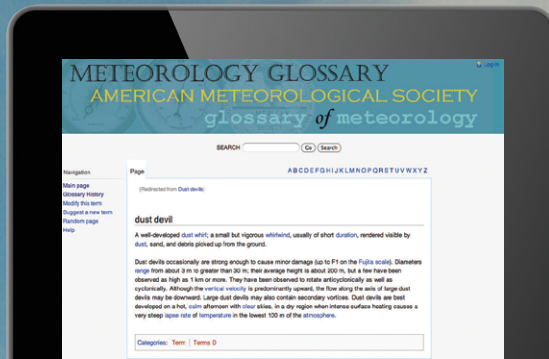
THE AMERICAN METEOROLOGICAL SOCIETY
Online Glossary of Meteorology

With over 12,000 meteorological terms, you'll be able to look up definitions online any time, any place, anywhere.

http://glossary.ametsoc.org/wiki

Also available in hardcover and CD formats at the AMS Bookstore, www.ametsoc.org/amsbookstore.