**Comparison of Aerosol Optical Property Climatology**

**Between In-situ Surface Measurements and Model Simulations**

**Abstract**

In the spring of 2015 AeroCom initiated the INSITU project designed to evaluate the performance of a suite of AeroCom aerosol models with surface, in-situ aerosol optical observations from the approximately 70 surface sites which have submitted their data to the World Data Centre for Aerosols (http://ebas.nilu.no/). Here, results from the first phase of analysis investigating how well models reproduce observed aerosol climatologies on a variety of time scales are presented. The results suggest that the combined model median over-estimates aerosol absorption and under-estimates aerosol scattering relative to the in-situ observations. This results in the simulated aerosol single scattering albedo being lower than the observed single scattering albedo. Additionally, the models suggest that the atmospheric aerosol has a greater contribution from coarse aerosol at most surface sites (as indicated by lower scattering Angstrom exponents) than is indicated by the in-situ measurements. The ability of the models to reproduce the observed systematic relationships among aerosol optical properties shows blah blah. Reasons for the observed discrepancies include blah blah. Something about biases…

**1. Introduction, Background and Literature Review**

Atmospheric aerosol particles are spatially and temporally inhomogeneous, due to variations in source strength, location and seasonality as well as changes due atmospheric transport and processing. Even if the variability in space and time is disregarded, aerosol particles are complex – they can vary in size, shape, composition and number. The physical and chemical characteristics of aerosol particles determine their optical properties (e.g., how they scatter and absorb light) and thus control their radiative effect in the atmosphere. The radiative effects of atmospheric aerosol particles are a key element to understanding climate forcing (IPCC, 2013).

Global climate models provide a method for evaluating aerosol effects on climate. By incorporating information on sources (e.g., emissions inventories), transport (e.g., meteorology) and parameterizations of aerosol properties and processes (e.g., hygroscopicity, removal rates, etc.) models are able to simulate location, amount and characteristics of aerosol particles across the globe and evaluate the resulting aerosol influence on the radiative energy balance. While model input (e.g., temporal and spatial resolution, complexity of parameterizations, etc.) continues to improve, it is important to recognize that there are many assumptions utilized in model simulations and that, therefore, model output needs to be evaluated against high-quality observational datasets.

The need for model evaluation has long been recognized and many studies have compared model output with observational data sets to identify the ability of models to simulate atmospheric constituents/characteristics as a function of space and time (e.g., references). AeroCom (Aerosol Comparisons between Observations and Models, http://aerocom.met.no/) was initiated in order to evaluate similarities and differences amongst multiple-models simulating global aerosol properties as well as to provide a framework for model/measurement comparisons of aerosol properties (Kinne et al., 2006). Many of the AeroCom multi-model comparisons between model simulations and observational data have been performed using remote sensing measurements of aerosol optical properties (i.e., from satellite retrievals and the AERONET surface network of sunphotometers) (e.g., Kinne et al., 2006; Koffi et al., 2012; 2016; Kim et al., 2014; Lacagnina et al., 2015; Sand et al., 2017) because of the global coverage provided by satellites and the relatively dense global network of AERONET sites.

Another approach is to compare model simulations with long-term in-situ measurements made at the surface. Some of the multi-model model/surface aerosol measurement comparisons performed by the AeroCom community include studies focused on organics (Tsigaridis et al., 2014), aerosol microphysics (Mann et al., 2014), dust (Huneeus et al., 2011), and black carbon (BC, Koch et al., 2009). Something about what these comparisons have shown? Something about how the models move chemical components around and then derive optical properties from them to link with next paragraph.

Less effort however has been put into evaluating simulated aerosol optical properties with in-situ surface measurements of these properties. This is due in part because in-situ measurements tend to have sparse global coverage relative to satellite measurements and the AERONET surface network. *Additional reasons include access to and understanding of the available in-situ measurements.* Some black carbon model/measurement comparisons have have converted in-situ surface measurements of aerosol light absorption (an optical property) to ‘equivalent black carbon’ (a chemical property) for comparison with modelled values of BC (e.g., Gilardoni et al., 2011; Skeie et al., 2011; Eckhardt et al., 2015). Vignati et al (2010) suggest that, when the focus is on aerosol optical properties (i.e., climate-related), it may be more appropriate to calculate the optical property from the model simulation for direct comparison with the measurements. Alvarado et al. (2016) compared aircraft measurements of aerosol absorption and scattering with model simulations in a closure study. Glantz et al (2014) used long-term surface scattering measurements to estimate AOD for comparison with model simulations of AOD but did not directly compare measured scattering with model simulations scattering at the surface.

Despite the sparseness of in-situ surface measurements, there are some significant advantages to utilizing in-situ measurements, particularly from long-term surface sites. First, these measurements can be directly referenced to physical standards. Blah blah. Second, in-situ measurements of some parameters (e.g., absorption and single scattering albedo) can be made reliably at lower aerosol loadings than is possible with remote sensing or satellite retrievals. For example, AERONET retrievals of column single scattering albedo require AOD at 440 nm to be greater than 0.4 (e.g., Holben et al., 2006). This level of AOD is not representative of the vast majority of global annual average AOD (e.g., Andrews et al., 2017). Similarly, Sheridan et al. (2012) showed that CALIPSO wasn’t sensitive to aerosol extinction below ~25 Mm-1. Third, surface in-situ measurements can be made continuously – they are not affected by the presence of clouds, orbit schedule or lack of incoming solar radiation (i.e., night!) Fourth, evaluation with surface measurements mean that the simulated location of the aerosol is more constrained – in contrast, column comparisons constrain in the horizontal dimension but not in the vertical. It should be noted that there are also studies investigating model simulations of the vertical profiles of aerosols (e.g., Koch et al., 2009; Skeie et al., 2011; Samset et al., 2014) which can further constrain model simulations, although many aircraft campaigns are of limited temporal span which can complicate comparisons with models.

The primary objective of AeroCom is to improve modeling of aerosols and thus improve predictions of the aerosol-related climate effects. Because, ultimately, the climate effects of aerosol particles are caused by their optical properties, this paper focuses on the use of in-situ surface aerosol optical property measurements to evaluate the performance and biases of a suite of AeroCom aerosol models. The scientific question to be answered is:

* How well do modelled global aerosol optical property climatologies compare (temporally and spatially) to climatologies derived from measurements of in-situ surface aerosol optical properties? Specifically what are the Biases in terms or parameter, region, aerosol type

This is the first (to our knowledge) global model evaluation using surface in-situ aerosol optical measurements. The goal, then, is to inform iterative improvements to model aerosol modules in order to better the predictive capability of global climate models. To support this goal, the in-situ data used in this analysis has been made available as a benchmark data set for future analysis as models evolve.

**2. Methods**

*2.1 Description of in-situ aerosol optical property data*

The data used in this study consists of surface in-situ aerosol light scattering coefficients and aerosol light absorption coefficients. The measurements are typically made following GAW protocols (WMO, 2016) meaning the measurements are made at controlled low humidity (usually RH<40%) and have appropriate instrumental corrections applied as described below. Figure 1 is a map of the surface in-situ sites used in this study. Table S1 in supplemental materials provides more information about the surface sites used in this study.

The Level 2, hourly-averaged QA/QC’d surface in-situ aerosol optical property data were downloaded from EBAS (Tørseth et al., 2012; <http://ebas.nilu.no/>) and underwent further review and interaction with the data providers as previous efforts involving multi-site analyses have shown external data review to be extremely helpful (Asmi et al., 2013; Collaud Coen et al 2013; Andrews et al., 2011). The hourly data are averaged to coarser time resolutions for comparison with less temporally resolved model output. A few additional data sets comprising at least 1 year of measurements were obtained from data mentors (see supplemental materials for details).

The light scattering measurements are made by integrating nephelometers (either TSI model#3563 or various models of the Ecotech nephelometer). The TSI nephelometer is a spectral instrument (450, 550 and 700 nm) allowing the calculation of wavelength dependence of scattering. Most sites with Ecotech nephelometers also submitted spectral scattering (450, 525 and 635 nm) although some sites operated single wavelength Ecotech instruments. TSI nephelometer data were corrected for known instrument non-idealities (truncation, light source) using the Anderson and Ogren (1998) method. Ecotech nephelometer data were corrected using Mueller et al. (2011).

The light absorption measurements are made using a variety of filter-based instruments including the Multi-Angle Absorption Photometer (MAAP, Thermo, Inc.); the Particle Soot Absorption Photometer (PSAP, Radiance Research) and the Continuous Light Absorption Photometer (CLAP, NOAA’s extended sampling time version of the PSAP, Ogren et al., 2017). The MAAP and original PSAP are single wavelength instruments providing light absorption at 635 nm and 550 nm, respectively. The CLAP and newer versions of the PSAP are multi-wavelength instruments (PSAP: 467, 530 and 660 nm; CLAP: 467, 528 and 652 nm). The PSAP and CLAP are corrected for scattering artifacts, etc. (e.g., Bond et al., 1999, Virkkula et al., 2005; 2010, Ogren et al., 2010). Aethalometer data were not used as, when the project was initiated, EBAS did not have an approved Level 2 data format for reporting corrected, QA/QC’d aerosol absorption from aethalometer measurements. (An aethalometer correction and is now available for reporting absorption from the AE31 aethalometer model, but it many cases the corrected data are not yet submitted.)

From the available measurements of aerosol light absorption and scattering additional parameters can be calculated. Here we will focus on two: the aerosol single scattering albedo (SSA) and the scattering Angstrom exponent (SAE).

 (1) SAE = log(1/2)/log(2/1)

and

 (2) SSA = sp,i/(sp,i + ap,i)

where  is the aerosol light scattering or absorption (subscript ‘sp’ subscript ‘ap’, respectively) at wavelength i (i).

SAE is a proxy for the aerosol size distribution. Larger SAE values (SAE ~2) indicate that there is a larger scattering contribution from sub-micrometer particles while smaller SAE values (SAE<1.5) indicate larger particles (>1 m) have a greater contribution to the observed light scattering (reference). Delene and Ogren (2002) showed that SAE strongly correlated with the fine mode fraction, which is useful as the fine mode fraction is often used as a proxy for anthropogenic aerosol.

The SSA provides information on the contribution of aerosol absorption to aerosol extinction. Values of SSA close to unity represent a primarily scattering aerosol (little to no absorption) such as sea salt. As SSA decreases from 1 this indicates that relatively more absorption is contributing to extinction. The single scattering albedo of black carbon is approximately 0.3 (Bond and Bergstrom, 2005).



**Figure 1.** Map of sites with in-situ aerosol scattering and absorption. Need to update with actual sites used. Inset for Europe?

*2.2 Description of AeroCom models and data*

Model output was requested from the AeroCom community of aerosol modelers for the INSITU experiment. Modelers provided aerosol extinction and absorption data at dry conditions (relative humidity = 0%) at one or more wavelengths (440, 550, 870 nm). All models provided daily averaged data, and some models provided 3h-averaged data as well. The specific output requested was: https://wiki.met.no/\_media/aerocom/INSITU\_AeroComPIII\_description.pdf. Table 1 provides a list of the models which provided output either for this Phase III AeroCom data call or had previously provided relevant data for an AeroCom Phase II data request.

**Table 1.** Description of models used in analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model name**Grid size (°lon x °lat)Output Year | **Citation(s)** | **MET** | **Mixing** | **Something****Else****Chemistry?****Emissions?** |
| **TM5**3.0° x 2.0°2010 | Van Noije et al., 2014 | Offline (ERA-Interim) | Internal mixing within modes |  |
| **GEOS-Chem**2.4° x 2.0°2010 | Bey et al. (2001) | offline chemistry-transport model(GEOS-5) | External mixing |  |
| **CAM5**2.4° x 1.9°2010 | Liu et al. (2012) Ghan et al. (2012) | horizontal winds nudged towards ERA-Interim reanalysis | Internal mixing by volume |  |
| **ECHAM6-SALSA**1.8° x 1.9°2010 | Bergman et al. 2012.Laakso et al., 2016 | Nudged towards ERA-Interim reanalysis data | Internal mixing by volume |  |
| **GEOS5-Globase**1.25° x 1°2010 | Chin et al., 2002, 2014Colarco et al. 2010 | runs in “re-play” mode; MERRA met analysis | External mixing |  |
| **GEOS5-MERRAero**0.6° x 0.5°2010 | Buchard et al 2015, 2016 | Driven by meteorology from the MERRA-1 reanalysis | External mixing |  |
| **OsloCTM2**2.8° x 2.8°2008 | Myhre et al. 2013Skeie et al 2011 | Offline meteorology from IFS ECMWF | Need to look at aerocom wiki |  |
| **GOCART**2.5° x 2.0°2006 | Chin et al., 2009 | Offline CTM. | External mixing |  |
| **MPIHAM\_v1**1.8° x 0.9°2006 | Waiting for Kai |  |  |  |
| **MPIHAM\_v2 (HCA)**1.8° x 0.9°2006 | Zhang et al. 2012 | Nudged to ECMWF analysis | Volume weighted mixing |  |
| **SPRINTARS**1.1° x 1.1°2006 | Takemura et al. 2005 | Nudged by the NCEP/NCAR reanalysis. | External mixing |  |
| **EMEP**0.5° x 0.5°2010 | Simpson et al. 2012 | off-line 3-hourly fields from ECMWF-IFS | External mixing |  |
| **OsloCAM5.3**1.25° x .9°2010 | Grini/Kirkevåg et al. (in prep) | Nudged to ERA interim meteorology | Internal mixing by volume |  |
| **ECMWF** |  |  |  |  |
| **Philip stier’s model?** |  |  |  |  |

* 1. *Matching of model and measurement*

The model simulation data provided simplifies comparisons with the in-situ measurements. Nonetheless, some additional data treatment and assumptions were required to get the most out of the comparisons. These are described below.

Simulated aerosol extinction and absorption from the closest model gridpoint to each site’s location was utilized for comparison to the measurement data (same as Eckhardt et al., 2015). Additionally, the values chosen were the lowest level from the model – the surface layer – as all measurement sites are surface sites. Using model surface data for sites located in topographically complex terrain may problematic as discussed in section XXX. To compare with the in-situ measurements of aerosol scattering, a model aerosol scattering value was calculated by subtracting simulated absorption from simulated extinction.

Each model provided data for a single year (2006, 2008 or 2010 – see Table 1), however, more than half of the available in-situ data sets began during or after 2008. The decision was made to utilize the full time series for each measurement location for the comparisons with the model data shown in Section 3. Doing this assumes that the inter-annual variability of the measurements is minimal. The validity of this approach is explored using long-term data sets encompassing all three years as described HERE.

Schutgens et al. (XXXX) demonstrate the importance of matching temporal sampling when doing these types of comparisons. The in-situ measurements are collecting data continuously, day and night, clear or cloudy, however, gaps in the data may occur due to instrument malfunctions, power outages and/or downtime for maintenance. For averaged comparisons with the model data 80% in-situ data coverage was required (e.g., 292 days of in-situ data to represent an annual average, 73 days of in-situ data to represent a seasonal average and 24 days of in-situ data to represent a monthly average).

In the EBAS archive, the clean, mentor-edited data sets are provided at standard temperature and pressure (Tstd=0 C, Pstd=1013 mb), while the models simulate data at ambient T and P. Because not all model or in-situ data sets provided the necessary T&P values, an ambient pressure value was estimated based on site elevation, and an annual average temperature was assumed and these values were used to adjust the in-situ measurements to ambient TP.

Many of the in-situ instruments utilized in this comparison are spectral instruments so an Angstrom adjustment was used to adjust both the scattering and absorption values to the wavelengths simulated by the models. For single wavelength absorption measurements a 1/wavelength dependence was assumed – this is equivalent to assuming an absorption Angstrom exponent of 1. For single wavelength scattering measurements, a scattering Angstrom exponent of XXXX was assumed.

* 1. *Model performance*

In addition to showing simple model/measurement comparisons (e.g., annual value scattering plots and seasonal variability) Taylor diagrams were also used to assess model performance (Taylor, 2001). A Taylor diagram is a means of plotting model/measurement comparisons on a single plot in terms of three statistical parameters: standard deviation, root mean square error(?) and correlation in order to see if there are patterns in the comparison. Data can be normalized so that disparate data sets (e.g., covering several orders of magnitude) can be shown on same plot. other methods for evaluating/quantifying model performance (quilt plots?)

**3. Results**

This section presents model/measurement comparisons on various time scales and Taylor plots to evaluate model skill. Annual, monthly and sub-daily time scales are evaluated. The comparisons are segregated by site category (coastal, continental, mountain and polar), in order to look for patterns as function of broad aerosol types and location. In Section 3.1, annual comparisons are evaluated in terms of point to point comparisons (scatter plots) of annual values and global context (maps showing model/measurement differences). In section 3.2, monthly comparisons enable assessment of how well models simulate effects of seasonality in sources and transport. Section 3.3 presents lag-autocorrelation plots to evaluate whether models can reproduced observed persistence of aerosols. The Taylor plots presented in Section 3.4 provide an overall view of model skill in predicting aerosol optical properties.

3.1 Annual comparisons

Figure 2 shows the relationship between annual median observations and model simulations of several aerosol optical properties (absorption, scattering, single scattering albedo and Angstrom exponent) for all models listed in Table 1. Figure 2a suggests that models tend to over-predict aerosol absorption - the median values of model simulated absorption tend to lie above the 1:1 line. This is the opposite of what is found when modelled aerosol absorption optical depth (AAOD) is compared with AAOD obtained from AERONET retrievals (e.g., Bond et al., 2013 and references there in). This will be further examined in the discussion section. In contrast to the absorption comparison, Figure 2b indicates that the median simulated scattering values tend to be lower than the in-situ observations, with the model median primarily falling below the 1:1 line. Note that Figures 2a and 2b have log axes so departures from the 1:1 line may indicate significant differences. Comparing the range in model simulations (the vertical bars), there appears to be more model diversity in aerosol absorption than there is in aerosol scattering.

There aren’t many obvious relationships between site type and model/observation differences, probably because the site type designations (polar, coastal, mountain, continental) cover such a wide range of sampling conditions (e.g., continental Europe and rural Illinois fall into the same category). One notable exception is that the model medians tend to underestimate the observed aerosol loading (both absorption and scattering) at polar sites (gray data points on Figure 2). It is also comforting that the models are fairly consistent in their over-predictions of aerosol loading at high altitude locations as high elevation sites typically would be expected to be cleaner (less aerosol) than the surrounding lower elevation terrain. This demonstrates the effect of topography on sub-grid variability when comparing the simulated ‘surface’ values when the grid boxes encompass complex terrain.

|  |  |
| --- | --- |
|  |   |
|  |   |

**Figure 2.** Comparison of annual in-situ and model aerosol optical properties colored by site type. (a) Aerosol absorption at 550 nm; (b) Aerosol scattering at 550 nm (c) Single scattering albedo at 550 nm (d) Scattering Angstrom exponent for 440/550 wavelength pair. Colors indicate site type. Horizontal bars indicate uncertainty in in-situ measurement; vertical bars indicate range in model output over the X models in Table 1. Solid gray line is 1:1 line while dashed gray lines in (a) and (b) indicates factor of XX range.

Overall estimate of model bias about combining by site type for each model (by site, by season ,latitude). Could also combine sites by latitude and see if there are latitudinal biases (might not work cuz lacking sites in some latitude bands. Get rid of mountain sites. Look for ‘orog\_ht’ to eliminate sharp mtns. Could see if biases in surface are similar to biases in AOD (e.g. koffi)

Marine, dusty, polluted, continental background, polar 🡨types to look for bias.

Look at michael’s presentation for ideas about heatmaps showing model bias

The over-prediction of absorption and under-prediction of scattering by models leads to the models simulating a lower single scattering albedo (darker aerosol) relative to the SSA calculated from in-situ surface measurements (Figure 2c).Both models and measurements find less absorbing aerosol (i.e., higher SSA) at coastal and polar sites than is found at most continental sites, This is consistent with a sea salt dominated aerosol at coastal sites and distance from combustion sources at polar sites.

Figure 2d suggests that, based on scattering Angstrom exponent values, models tend to simulate an aerosol that has a larger coarse aerosol contribution than indicated by the spectral in-situ observations at most sites. The lowest Angstrom exponents (largest contribution from coarse aerosol) are found at coastal sites for both models and measurements, again consistent with sea salt dominated aerosol. The model diversity in Angstrom exponent is also interesting – the range in simulated values for coastal and polar sites is approximated twice the range simulated for Angstrom exponent at continental sites

With the exception of the ECHAM6 simulation, most models predict higher absorption than is observed at the in-situ measurement sites (red-colored points). This is particularly noticeable at the cluster of sites in continental Europe and for most models and sites across the continental US. Only at some of the sites on either side of the Atlantic (primarily coastal sites) are the in-situ absorption measurement values consistently larger than the model values (blue-colored points). **Does the lack of polar pattern indicate an improvement since Koch? What about more recent Eckhardt paper? Make annual and seasonal heatmaps**

In contrast to the absorption comparison, the models tend to simulate lower scattering than the in-situ measurements at most US and Asian sites (Figure X). The exceptions are a few high elevation sites (WLG, LLN, PYR) in Asia where several of the models suggest higher annual scattering than is observed. Figure X suggests that simulated scattering values for continental Europe are more likely to be higher than the in-situ observations regardless of station altitude. As with the absorption simulations, ECHAM6 is an outlier – the model consistently underpredicts measured aerosol light scattering at almost all sites. Consistent with Figure 2b, model simulations of scattering in the Arctic tend to be low relative to measurements, although several models (EMEP, GEOSCHEM, GLOBASE, MERRAero, and OSLOCAM5 and TM5) simulate higher values in the European Arctic than are observed.

Figure X is consistent with Figure 2c, in that the models tend to predict a lower single scattering albedo than obtained from the measurements. This is particularly true in the 30-60o N latitude band. The exceptions tend to be at coastal and high elevation sites. Again, ECHAM6 is an outlier in that simulated values of SSA are higher than observations at most sites. The large SSA differences observed in continental Europe represent the absorption and scattering hotspots shown in Figures 3 and 4. Many of the differences in SSA between simulation and measurement fall in the -10-30% difference range. For those cases, the model SSA may be 0.1 to 0.3 lower than in-situ SSA. Differences of this magnitude can change the direction of forcing depending on the underlying surface (**e.g., Randles et al??**).

Figure X shows percent difference plots for the scattering Angstrom exponent. No plots were made for four models (MPIHAM\_v1, MPIHAM\_v2, SPRINTARS and GOCART) because the Phase II control project did not request spectral values of surface optical properties. There is a fairly consistent picture that the models tend to simulate larger aerosol in Europe and most sites in North America than obtained from in-situ measurements. The locations where the measurements suggest the presence of larger particles than simulated by the model) are coastal sites (e.g., Cape Point, South Africa) and/or sites that may have a strong dust influence (e.g., Niamey in sub-Saharan Africa).

*3.2 Seasonal climatologies*

Making annual comparisons is a good first step for evaluating how models and measurements compare, but higher frequency comparisons can provide additional information about what may be contributing to discrepancies in annual comparisons (e.g., whether seasonal sources like forest fires are having too strong of an impact). Figures xxx display seasonal values for both measurements and model simulations of various parameters organized by site type.

* 1. *Lag-autocorrelation*

Lag-autocorrelation (LA) plots are a means of capturing temporal variability on various time scales (e.g., Anderson et al., 2003). Here LA plots are used to compare the diurnal variability exhibited by the high time resolution (3h) data simulated by several models ('echam6', 'tm5\_insitu', 'geoschem', 'merraero') with the hourly averaged in-situ measurement data. The shape of the LA curves provides information on aerosol temporal variability and persistence (Heintzenberg et al., 2004) and, on the 5-day time scale studied here, indicates the presence/absence of diurnal cycles. Autocorrelation analysis of high time resolution data can provide information about atmospheric processes and transport (e.g., new particle formation, upslope/downslope flow, etc.).

Figure 11 provides representative examples of LA plots for scattering and absorption for 4 sites. It is fairly common for models to predict strong diurnal oscillations in both scattering and absorption when none are observed in the in-situ data. In the short term (lag = 0-12 h) models tend to simulate more persistence than observed, while further out (lag = 24-120 h) the models may indicate more, less or about the same level of persistence. MERRAero and ECHAM6 tend to better represent the observed high frequency variability than do GEOSCHEM and TM5. GEOSCHEM and TM5 tend to simulate too much diurnal variability relative to the other two models.

In general the polar and coastal sites exhibit little diurnal variability in scattering or absorption and the models are able to represent this. Most models simulate the presence of diurnal cycle at mountain and continental sites which is consistent with the observations. However, the models are more likely to simulate diurnal flow at mountain and continental sites than is observed and the magnitude of the diurnal cycle at those site types tends to be stronger than is observed. Both models and measurements suggest there is more likely to be a diurnal cycle in absorption than scattering. The LA plots show that several of the more anthropogenically-influenced sites (e.g., APP, DEM, FKB, UGR) exhibit two diurnal peaks in absorption (likely corresponding to morning and evening traffic patterns), but none of the models capture the effects of such local sources (not shown).

Comment on whether should/shouldn’t do lag autocorr for SSA and SAE. Calculated values from in-situ measurements will be affected by noise, especially at clean sites because are ratios of small numbers. This means autocorrelations will be lower than reality.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |
| (g) | (h) |

**Figure 11.** Examples of lag autocorrelation for absorption (left column) and scattering (right column) at (a) and (b) BND (Bondville, US); (c) and (d) ALT (Alert, Canada); (e) and (f), CPR (Cape San Juan, Puerto Rico); (g) and (h) JFJ (Jungfraujoch, Switzerland) . Lag (k) is the time between measurements being compared; ‘r(k)’ is the lag autocorrelation statistic. Colored dotted lines are model results, solid black line is in-situ measurements. Red=MERRAero, Green=TM5, Blue=echam6 Yellow=Geoschem.

* + 1. *Model performance scores based on comparison to observations*

Taylor diagrams were made for several groupings of the data. Figure 12a shows a Taylor diagram for annual aerosol scattering coefficient climatology when the sites were grouped by type (coastal, continental, mountain, polar) and compared against the model results for those groupings. The Taylor diagram suggests several patterns. Models tend to simulate less variability (as indicated by standard deviation) than the in-situ measurements, although interestingly that is not the case for the mountain sites. For mountain sites, many models simulate standard deviations similar to those observed, but interpreting this is complicated by the use of model surface simulations rather than model simulations at the measurement elevation level. The model simulations of scattering appear to be more highly correlated with measurements at marine and continental sites and less correlated with polar and mountain site measurements. Indeed, CAM5(?) simulated output is inversely correlated with the mountain measurements (since we haven’t selected the appropriate model height for comparisons that is not saying anything positive/negative about CAM5). Marine sites are also best represented in terms of RMS.

|  |  |
| --- | --- |
| (a) | (b) |
| **Figure 12.** Taylor diagram of for (a) scattering and (b) absorption. Colors indicated site type (all data = black, coastal = blue, continental = red, mountain = yellow, polar = purple). Shapes indicate sign of correlation (triangle=positive, square=negative). Numbers indicate model (1=CAM5, 2=Merraero, 3=echam6, 4=TM5, 5=GEOSCHEM, 6=globas, 7=osloCAM5, 8=emep, 9=osloCTM, 10=sprintars, 11=mpiham, 12=mpiham\_HCA, 13=gocart |

Figure 12b shows the Taylor diagram for aerosol absorption segregated by site type. The correlation values to those for scattering for each site type, but the standard deviation across models is much broader than observed for scattering. This suggests there is much more diversity in how models simulate absorption than in how they simulate scattering. Again, in terms of correlation, models are best able to simulate the coastal site values and least able to simulate polar and mountain sites.

Taylor diagrams can also be used to determine if there are seasons when models are better able to capture observed statistics (not shown). When considering all data (i.e., not segregated by site), the models tend to do better at matching observed scattering variability in summer and winter. The correlation of modelled and measured scattering is not obviously seasonally dependent. In contrast, for absorption, how well models simulate variability is very model dependent. Further, there is a strong seasonal dependence in how well model simulations are correlated with observations – almost all models are better correlated with observations in the spring and summer than in the fall or winter. Does this suggest the importance of local emissions (and domestic heating type emissions) in the winter that may need to be better resolved by model?

Based on Taylor plot assessments, all models underestimate the variability of SSA and SAE relative to observations and tend to be poorly correlated across all sites and split by site type. The highest correlations between model and measurements for SSA occur for some models at Arctic sites. The worst correlations tend to be for continental sites. There is no obvious seasonal dependence of model skill for SSA. Model skill in predicting SAE is highest for coastal sites – most models are positively correlated with observations and best represent observed variability (although all models underestimate SAE variability regardless of site type). Models tend to be negatively correlated with observations for SAE at continental, mountain and polar sites. When all sites are considered, but split by season, the correlations between models and measurements for SAE are positive (likely reflecting that the dichotomy in SAE of continental and coastal sites is captured by the models). There is however no seasonal dependence of model skill for SAE.

Make table(s) or figure(s) showing model R2, slope?, std dev? Other? for each site and each model. Figure could be like quilt plot colored by R2 value instead of mod-meas/mod diff.

**4. Discussion**

* Summarize variables that models simulate exceptionally well or poorly

An unexpected finding of this analysis is that models tend to over-estimate absorption relative to observations. That’s inconsistent with almost every recent model/measurement comparison related to column absorbing aerosol, i.e., AAOD (e.g., Stier et al., 2007; Bond et al., 2013; Lacagnina et al., 2015). Andrews et al. (2017) suggest that AAOD retrievals from remote sensing tend to be higher than AAOD calculated from in-situ profiles although they were unable to assess whether the in-situ measurements were biased low or the column retrievals were biased high. They noted that because of this, the assumption that modelled absorption needed to be upscaled to match AAOD retrievals should be revisited.

BC model/measurement comparisons tend to be less consistent. Samset et al. (2014) found that models tend to over-estimate BC throughout the vertical profile in remote regions based on airborne measurements of BC. They also noted that models were in good agreement with some measurements over Japan and attributed measured BC being greater than modelled BC during some flight campaigns in the Arctic to forest fire emissions not captured by the models. Koch et al. (2009) also found a similar result to ours - that models overestimated surface BC at most locations in Europe and many in the US as well but noted that the models underestimated BC in Asia which they attributed to out-of-date emissions in that region (not counting mountain sites in asia we see a mixed bag – model higher/lower/same as observations depending on site) Gilardoni et al 2011 compare BC at several surface sites (derived from absorption measurements (THD, BND,ALT,IPR, JFJ) and aren’t able to come to conclusions about how well model/meas agree but suggest a few tests to assess. Vignati also caveat their attempts with model measurement comparisons of BC.

So what – what are implications for rad forcing if messing up surface? Other implications…

* Something about temporal stuff
* Models tend to over-estimate diurnal variability but under-estimate annual variability
* bad winter absorption in Europe –local sources seasonality of local sources not capture or stagnant BL air not captured?
* Compare these model/in-situ surface measurement comparisons to other model/optical property comparisons in the literature (How do these results compare to column comparisons? Aerosol number comparisons? Aerosol species comparisons?)
	+ Eckhardt, Koch, Breider Browse meas/model comparisons in Arctic
	+ Heald et al 2011 models underestimate OA in profiles compared to flight compaign data – could contribute to scat underestimate. Models tend to do ok for sulfate (w/in observed range). Saw differences in model skill for pollution, biomass burning, remote measurements
	+ Skeie et al 2011
	+ Alvarado et al 2016 – model/meas optical closure depends on filter-based correction (they use lack 2008 which lack actually cautions against) and model parameterizations. Model parameterization points are that assumptions about density, mixing and size dist of BC in models should be evaluated. Also note Sinha paper which suggests based on co-located measurements that if anything psap absorption is enhanced due to particles associated with BC and should be lowered (Lack correction increases absorption). Alvarod shows that modelled absorption>measured absorption when using more standard ‘virkkula’ correction (consistent with us)
	+ Glantz et al 2014 used dry insitu scattering measurements at ZEP to estimate AOD (assuming hygroscopic growth) rather than direct comparison of meas and model.
	+ Pan et al 2015 – notes underestimate of AOD in asia due to models not including any or enough nitrate and having too low emissions for BB fires and ‘biofuel’
	+ Crippa et al 2016 WRF-chem over-estimates northAm AOD and underestimates SAE- attributes SAE to potential issues with modeled size dist
	+ Lacagnina et al., 2015. Compares SSA and AOD from remote sensing to aerocom. Aerocom sees higher SSA (lower AAOD) than observations, Aerocom sees AOD equal or slightly less than observations

*4.X Constraints on measurement model comparisons*

In general, the models simulated less scattering and larger particles (as indicated by scattering Angstrom exponent) than observed. Here we address two measurement conditions that could contribute to both of these observations (measurement size cut and sample RH). Further, we discuss issues with our choice of using surface (lowest model level) model data to compare with the surface insitu measurements and sub-grid variability.

*4.X.1 Issues with RH*

In this study, modelled optical properties at dry (RH=0%) were compared with measurements which are considered dry (RH<40%). This discrepancy in RH could be one contributor to the models’ tendency to simulate lower scattering than is observed. Particles measured at low, but not absolutely dry RH, may have water associated with them (e.g., Zieger et al., 2014) and the scattering due to the aerosol water would be included in the measured scattering values. The RH of the sample can also affect the diameter of the particles, with higher RH leading to larger particles. If this were the case, the measured SAE values should be lower than the modelled values but that is the opposite of what was observed.

* + AOD comparisons assume that models get both aerosol and RH right – Roelofs et al 2010 show issues with RH affect mod/meas comparison of AOD

*4.X.2 Measurement size cut*

In contrast, many sites had a maximum inlet size cut of 10um (corresponding to an aerodynamic diameter of ~7um) meaning particles larger than this size are not sampled or sampled with lower efficiency. If the measurements were consistently under-sampling coarse aerosol, it would lead to the measured scattering Angstrom exponent being higher than the modelled values which simulate all aerosol. This is consistent with the model/measurement SAE comparisons presented here. Need a tie together sentence

* + Emissions underestimate large particle? (Matthias 2008)

*4.X.3 Surface level simulations*

Some model/measurement discrepancies may also be related to using the model surface (lowest pressure level) value for comparison with each site. While this would have the greatest impact on high elevation mountain site comparisons, it could also influence model/measurement discrepancies at lower elevation sites depending on topography. However, this should only affect parameters related to aerosol amount (e.g., absorption and scattering) and not on parameters that are independent of aerosol loading (e.g., SAE and SSA). Andrews et al. (2004) and Sheridan et al. (2012) present vertical profiles of SAE and SSA derived from multi-year, low RH, airborne in-situ measurements, suggesting that, in the boundary layer, SAE and SSA should be relatively invariant with altitude.

*4.X.4 Subgrid variability*

* Sub-grid variability (cite Benkowitz and Schwartz, 1997) Wang et al 2018
* Do we see better agreement w/smaller grid models?

*4.XX Suggestions for model improvements*

* Discuss what model processes might need improvement based on discrepancies in simulating aerosol optical property climatologies.
	+ Conversion from chemical to optical properties (Alvarado)
	+ Sources/stagnation in winter?
	+ Browse et al scavenging for arctic, lifetimes in general (Kristiansen (SO4), Samset (BC))
	+ Heald - More organics to improve scattering?
	+ Size distribution related issues (Alvarado closure experiment)

*4.XXX Future work*

* + Effect of RH of comparisons 0 vs 40%
	+ How well do models represent observed co-variance of aerosol properties?
	+ Hygroscopicity
* What unique information will this paper contribute to the field: constrain aerosol climatologies of models at surface, suggest magnitude and sign of biases, establish benchmark dataset that can be utilized by individual modelers to constrain models
* RQ2: What specific model runs (e.g., additional model output variables, temporal resolution, parameterizations, etc.) are proposed to better compare to observations and, perhaps, improve simulations? (Ideally, proposed runs would be used in additional comparative analyses, part of a larger iterative project to understand model/measurement discrepancies and actively help aerosol models to better represent surface aerosol measurements)
* Suggestions/requests for specific model runs that could be used to implement a long-term project comparing model and in-situ surface optical property measurements
* Outline of long-term project with iterative steps to adjust models based on results of comparison, re-do comparative analyses, adjust models based on results, repeat, etc.

**5. Conclusions**

* + Results
	+ Future work/needs
	+ *Description of benchmark dataset and availability*

*Acknowledgements*

**7. References**

Alvarado et al., 2016

Anderson et al 2003

Anderson, T. L., Ogren, J.A., “Determining aerosol radiative properties using the TSI 3563 integrating nephelometer,” Aerosol Sci. Technol., 29, 57– 59, 1998.

Andrews, E., et al., 2004 IAP paper

Andrews, E., Ogren, J.A., Kinne, S., Samset, B., “Comparison of AOD, AAOD and column single scattering albedo from AERONET retrievals and in situ profiling measurements,” Atmos. Chem. Phys., 17, 6041-6072, <http://www.atmos-chem-phys.net/17/6041/2017/>, 2017.

Andrews, E., Ogren, J.A., Bonasoni, P., Marinoni, A., Cuevas, E., Rodriguez, S., Sun, J.Y., Jaffe, D.A., Fischer, E.V., Baltensperger, U., Weingartner, E., Collaud Coen, M., Sharma, S., Macdonald, A.M., Leaitch, W.R., Lin, N.-H., Laj, P., Arsov, T., Kalapov, I., Jefferson, A., Sheridan, P.J., “Climatology of aerosol radiative properties in the free troposphere,” *Atmospheric Research*, 102, 365-393, 2011.

Asmi, A., Collaud Coen, M., Ogren, J.A., Andrews, E., Sheridan, P., Jefferson, A., Weingartner, E., Baltensperger, U., Bukowiecki, N., Lihavainen, H., Kivekäs, N., Asmi, E., Aalto, P.P., Kulmala, M., Wiedensohler, A., Birmili, W., Hamed, A., O'Dowd, C., Jennings, S.G., Weller, R., Flentje, H., Mari Fjaeraa, A., Fiebig, M., Lund Myhre, C., Hallar, A.G., and Laj,P., “Aerosol decadal trends – Part 2: In-situ aerosol particle number concentrations at GAW and ACTRIS stations,” *Atmos. Chem. Phys.,* 13, 895-916, 2013.

Bond, T. C., Anderson, T. L., and Campbell, D.: Calibration and intercomparison of filter-based measurements of visible light absorption by aerosols, Aerosol Sci. Technol., 30, 582–600, doi:10.1080/027868299304435, 1999.

Bond, T.C. and Bergstrom, R.W., “Light Absorption by Carbonaceous Particles: An Investigative Review,” Aerosol Science and Technology, 39, 1–41, doi: 10.1080/02786820500421521 2005.

Collaud-Coen, M., Andrews, E., Asmi, A., Baltensperger, U., Bukowiecki, N., Day, D., Fiebig, M., Fjaeraa, A.M., Flentje, H., Hyvärinen, A., Jefferson, A., Jennings, S.G., Kouvarakis, G., Lihavainen, H., Lund Myhre, C., Malm, W.C., Mihapopoulos, N., Molenar, J.V., O'Dowd, C., Ogren, J.A., Schichtel, B.A., Sheridan, P., Virkkula, A., Weingartner, E., Weller, R., and Laj, P.,” Aerosol decadal trends – Part 1: In-situ optical measurements at GAW and IMPROVE stations,” *Atmos. Chem. Phys.,* 13, 869-894, 2013.

Breider, T. J., et al. "Annual distributions and sources of Arctic aerosol components, aerosol optical depth, and aerosol absorption." *Journal of Geophysical Research: Atmospheres* 119.7 (2014): 4107-4124. *{Model measurement comparisons of aerosols and ozone in the Arctic}*

Breider, TJ et al. trends in arctic

Delene, D. J. and Ogren, J. A., “Variability of aerosol optical properties at four North American surface monitoring sites,” J. Atmos. Sci., 59, 1135–1150, 2002.

Eckhardt, S., et al., “Current model capabilities for simulating black carbon and sulfate concentrations in the Arctic atmosphere: a multi-model evaluation using a comprehensive measurement data set,” Atmos. Chem. Phys., 15, 9413–9433, 2015.

Ginoux, P., Horowitz, L. W., Ramaswamy, V., Geogdzhayev, I. V., Holben, B. N., Stenchikov, G., Tie, X., “Evaluation of aerosol distribution and optical depth in the Geophysical Fluid Dynamics Laboratory coupled model CM2.1 for present climate,” J. Geophys. Res., 111, doi:10.1029/2005JD006707, 2006.

Glantz, Paul, et al. "Remote sensing of aerosols in the Arctic for an evaluation of global climate model simulations." *Journal of Geophysical Research: Atmospheres* 119.13 (2014): 8169-8188.*{comparison between NorESM model and measurements at Zeppelin}*

Heintzenberg, J., Birmili, W., Wiedensohler, A., Nowak, A., Tuch, T., “Structure, variability and persistence of the submicrometre marine aerosol,” Tellus, 56B, 357–367, 2004.

Holben, B. N., Eck, T. F., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., Giles, D., Dubovik, O.: AERONET’s Version 2.0 quality assurance criteria, http://aeronet.gsfc.nasa.gov/new\_web/ Documents/AERONETcriteria\_final1.pdf (last access: 10 March 2014), 2006.

Huneeus et al., 2011 (dust);

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, doi:10.1017/CBO9781107415324.

Kim et al., 2014

Kinne, S., et al., “An AeroCom initial assessment – optical properties in aerosol component modules of global models,” Atmos. Chem. Phys., 6, 1-20, 2006.

Koch, D., et al “Evaluation of black carbon estimations in global aerosol models,” Atmos. Chem. Phys., 9, 9001–9026, 2009. *{highly cited paper about difficulty modeling BC in arctic (among other places), covered in Wang 2014}*

Koffi et al., 2012

Lacagnina et al., 2015;

Mann et al., “Intercomparison and evaluation of global aerosol microphysical properties among AeroCom models of a range of complexity,” Atmos. Chem. Phys., 14, 4679–4713, 2014.

Myhre, G. et al., “Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations,” Atmos. Chem. Phys., 13, 1853–1877, 2013.

Mueller, T., Laborde, M., Kassell, G., Wiedensohler, A., “Design and performance of a three-wavelength LED-based total scatter and backscatter integrating nephelometer”, Atmos. Meas. Tech., 4, 1291–1303, 2011.

Ogren, J. A., “Comment on “Calibration and Intercomparison of Filter-Based Measurements of Visible Light Absorption by Aerosols”, Aerosol Sci. Technol., 44, 589–591, doi:10.1080/02786826.2010.482111, 2010.

Ogren, J.A., “A systematic approach to in-situ observation of aerosol properties,” In: *Aerosol Forcing of Climate*, eds. R. Charlson and J. Heintzenberg, John Wiley & Sons, Ltd., 215-226, 1995.

Ogren, J.A., Wendell, J., Andrews, E., Sheridan, P.J.,"Continuous Light Absorption Photometer for Long-Term Studies," Atmos Meas. Tech.,10, 4805-4818, doi:10.5194/amt-10-4805-2017, 2017.

Pandolfi, M., Cusack, M., Alastuey, A., Querol, X., “Variability of aerosol optical properties in the Western Mediterranean Basin,” Atmos. Chem. Phys., 11, 8189–8203, 2011.

Petzold maap paper?

Samset et al., 2014 (BC);

Schutgens et al., 2016 comparing spatial/temporal consistent data sets

Sheridan, P.J., Delene, D.J., and Ogren, J.A., “Four years of continuous surface aerosol measurements from the Department of Energy’s Atmospheric Radiation Measurement Program Southern Great Plains Cloud and Radiation Testbed site,” J. Geophys. Res., 106, 20735-20747, 2001.

Sheridan, P.J., et al., 2012 AAO paper

Sherman, J.P., Sheridan, P.J., Ogren, J.A., Andrews, E., Hageman, D.C., Schmeisser, L., Jefferson, A., Sharma, A., “A multi-year study of lower tropospheric aerosol variability and systematic relationships from four North American regions,” Atmos. Chem. Phys., 15, 12487-12517, 2015.

Shindell, D. T., et al. "A multi-model assessment of pollution transport to the Arctic." *Atmospheric Chemistry and Physics* 8.17 (2008): 5353-5372.*{Sensitivity of aerosol and gas concentrations in Arctic to changes in emissions using multiple models}*

Simpson, D., et al., “The EMEP MSC-W chemical transport model – technical description,” Atmos. Chem. Phys., 12, 7825-7865, 2012.

Taylor, K.E., “Summarizing multiple aspects of model performance in a single diagram,” J. Geophys. Res., 106, 7183-7192, 2001.

Tørseth and Coauthors, “Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972-2009,” Atmos. Chem. Phys., 12, 5447-5481, doi:10.5194/acp-12-5447-2012, 2012.

Tsigaridis, K. et al. “The AeroCom evaluation and intercomparison of organic aerosol in global models,” Atmos. Chem. Phys., 14, 10845–10895, 2014.

Virkkula PSAP papers (original and correction)

Wang 2014 paper site representativeness

WMO, WMO/GAW Aerosol Measurement Procedures, Guidelines, and Recommendations, (WMO TD No. 1177) - GAW Report No. 227, World Meteorological Organization, Geneva, <https://library.wmo.int/opac/doc_num.php?explnum_id=3073>, 2016.

**Model papers in table 1 – need to sort and complete references**

Bergman, T., Kerminen, V.-M., Korhonen, H., Lehtinen, K. J., Makkonen, R., Arola, A., Mielonen, T., Romakkaniemi, S., Kulmala, M., and Kokkola, H.: Evaluation of the sectional aerosol microphysics module SALSA implementation in ECHAM5-HAM aerosol-climate model, Geosci. Model Dev., 5, 845-868, doi:10.5194/gmd-5-845-2012, 2012.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. Field, A. M. Fiore, Q. Li, H. Liu, L. J. Mickley, and M. Schultz, Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106, 23,073-23,096, 2001.

Buchard, V., and Coauthors, 2015: Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis. *Atmospheric Chemistry and* Physics, 15 (10), 5743–5760, doi:10.5194/acp-15-5743-2015.

Buchard, V., and Coauthors, 2016a: Evaluation of the surface PM2.5 in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States. Atmospheric Environment, 125, 100–111, doi:10.1016/j.atmosenv.2015.11.004.

Chin, M., P. Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, et al., Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sun photometer measurements, J. Atmos. Sci., 59, 461–483, 2002.

Chin, M., T. Diehl, Q. Tian, J. M. Prospero, R. A. Kahn, A. Remer, H. Yu, A. M. Sayer, H. Bian, et al., Multi-decadal variations of atmospheric aerosols from 1980 to 2009: sources and regional trends, Atmos. Chem. Phys., 14, 3657-3690, doi:10.5194/acp-14-3657-2014, 2014.

Chin, M., T. Diehl, O. Dubovik, T. F. Eck, B. N. Holben, A. Sinyuk, and D. G. Streets, Light absorption by pollution, dust and biomass burning aerosols: A global model study and evaluation with AERONET data, Ann. Geophys., 27, 3439-3464, 2009.

Colarco, P. R., A. da Silva, M. Chin, T. Diehl: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, J. Geophys. Res., 115, D14207, doi:10.1029/2009JD012820, 2010.

S. J. Ghan, X. Liu, R. C. Easter, R. Zaveri, P. J. Rasch, J.-H. Yoon, and B. Eaton, 2012: Toward a Minimal Representation of Aerosols in Climate Models: Comparative Decomposition of Aerosol Direct, Semidirect, and Indirect Radiative Forcing. J. Climate, 25, 6461–6476, doi: 10.1175/JCLI-D-11-00650.1.

Laakso, A., Kokkola, H., Partanen, A.-I., Niemeier, U., Timmreck, C., Lehtinen, K. E. J., Hakkarainen, H., and Korhonen, H.: Radiative and climate impacts of a large volcanic eruption during stratospheric sulfur geoengineering, Atmos. Chem. Phys., 16, 305-323, doi:10.5194/acp-16-305-2016, 2016.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G., and Mitchell, D.: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, Geosci. Model Dev., 5, 709-739, doi:10.5194/gmd-5-709-2012, 2012.

Myhre, G., et al. (2013), Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, *Atmos. Chem. Phys.*, *13*(4), 1853-1877

Skeie, R. B., et al. (2011), Anthropogenic radiative forcing time series from pre-industrial times until 2010, *Atmos. Chem. Phys.*, *11*(22), 11827-11857.

Takemura, T., T. Nozawa, S. Emori, T. Y. Nakajima, and T. Nakajima, 2005: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model. J. Geophys. Res., 110, D02202, doi:10.1029/2004JD005029.

Van Noije et al., Simulation of tropospheric chemistry and aerosols with the climate model EC-Earth, Geosci. Model Dev., 2014

**Appendix - Supplementary Materials**

**Table S1.** List of sites used in this study

|  |  |  |
| --- | --- | --- |
| **Station ID****Station name****‘Type’** | **Country** **Lat Long Elev** | **Instruments (dates)** |
| ALTAlert‘Polar’ | Canada | TSI nephPSAP-3wCLAP-3w |
| AMYAnmyeon-do‘coastal’ | South Korea | TSI nephPSAP-3wCLAP-3w |
| ANBAnnaberg-Buchholz‘continental’ | Germany | MAAP |
| APPAppalachian State‘continental’ | USA | TSI nephPSAP-3wCLAP-3w |
| APTAspvreten‘continental’ | Sweden | PSAP‘ITM custom’ |
| ARNEl Arenosillo‘coastal’ | Spain | TSI nephCLAP-3w |
| BEOBEO-Moussala‘mountain’ | Bulgaria | TSI nephCLAP-3w |
| BIRBirkenes‘continental’ | Norway | TSI nephPSAP-1w🡨 |
| BKTBukit Kototabang‘continental’ | Indonesia | Ecotech neph |
| BNDBondville‘continental’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| BRWBarrow‘polar’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| BSLBosel‘continental’ | Germany | MAAP |
| CESCabauw‘coastal’ | The Netherlands | TSI nephMAAP |
| CGOCape Grim‘coastal’ | Australia | Ecotech nephMAAP🡨 |
| CHCChacaltaya‘mountain’ | Bolivia | Ecotech nephMAAP |
| CMNMonte Cimone‘mountain’ | Italy | Ecotech neph (says TSI)🡨 MAAP |
| CPRCape San Juan‘coastal’ | USA (Puerto Rico) | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| CPTCape Point‘coastal’ | South Africa | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| DEMDemokritos‘continental’ | Greece | Ecotech nephAE31🡨 |
| DMVDanum Valley‘continental’ | Malaysia | Ecotech nephMAAP |
| EGBEgbert‘continental’ | Canada | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| ETLEast Trout Lake‘continental’ | Canada | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| FIKFinokalia‘coastal’ | Greece | RR neph🡨Ecotech (green)🡨AE22, AE31🡨 |
| FKBHesselbach‘continental’ | Germany | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| GSNGosan‘coastal’ | South Korea | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| GRWGraciosa‘coastal’ | Portugal (Azores) | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| HFEShouxian‘continental’ | China | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| HPBHohenpeissenberg‘continental’ | Germany | TSI nephMAAP |
| HYYHyytiala‘continental’ | Finland | TSI nephbap data is missing - weird |
| IPRIspra‘continental’ | Italy | TSI nephMAAP |
| IZAIzana‘mountain’ | Spain (Tenerife) | TSI nephMAAP |
| JFJJungfraujoch‘mountain’ | Switzerland | TSI nephMAAP |
| KPSK-puszta‘continental’ | Hungary | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| LEILeipzig‘continental’ | Germany | MAAP |
| LEWLeipzig-West‘continental’ | Germany | MAAP |
| LLNMt. Lulin‘mountain’ | Taiwan | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| MANManacapuro‘continental’ | Brazil | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| MAOManaus‘continental’ | Brazil | Don’t have data  |
| MHDMace Head‘coastal’ | Ireland | TSI nephMAAP |
| MLOMauna Loa‘mountain’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| MPZMelpitz‘continental’ | Germany | TSI nephMAAP |
| MSAMontsec‘mountain’ | Spain | Ecotech nephMAAP |
| MSYMontseny‘continental’ | Spain | Ecotech nephMAAP |
| NIMNiamey‘continental’ | Niger | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| NMYNeumayer‘polar’ | Antarctica (Germany) | TSI nephMAAP |
| OPEObs. Perenne de L’Environ.‘continental’ | France | EcotechAeth31🡨 |
| PALPallas‘polar’ | Finland | TSI nephMAAP |
| PGHNainital‘continental’ | India | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| PLAPreila‘continental’ | Lituania | TSI neph |
| PUYPuy de Dome‘mountain’ | France | TSI nephMAAP |
| PVCCape Cod‘coastal’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| PYEPoint Reyes‘coastal’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| RSLResolute‘polar’ | Canada | TSI nephCLAP-3w |
| SGPSouthern Great Plains‘continental’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| SIRSIRTA‘continental’ | France | Ecotech neph (blue only) |
| SPLStorm Peak‘mountain’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| SPOSouth Pole‘polar’ | USA (Antarctica) | TSI neph |
| SSLSchauinsland‘continental’ | Germany | MAAP |
| SUMSummit‘polar’ | Greenland | TSI nephCLAP-3w |
| THDTrinidad Head‘coastal’ | USA | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| TIKTiksi‘polar’ | Russia | TSI nephMAAP |
| TLLEl Tololo‘mountain’ | Chile | Ecotech neph |
| TRLTrollhaugen‘polar’ | Norway (Antarctica) | TSI nephPSAP |
| TRSTroll‘polar’ | Norway (Antarctica | TSI nephPSAP |
| VAVVavihill‘continental’ | Sweden | Ecotech nephPSAP-1w |
| WALWaldhof‘continental’ | Germany | MAAP |
| WHIMt Whistler‘mountain’ | Canada | TSI nephPSAP-1wPSAP-3wCLAP-3w |
| WLGMt Waliguan‘mountain’ | China | TSI nephPSAP-1wPSAP-3w |
| WSASable Island‘coastal’ | Canada | TSI nephPSAP-1w |
| ZEPZeppelin‘polar’ | Norway | TSI nephPSAP |
| ZSFZugspitze‘mountain’ | Germany | TSI nephMAAP |

|  |
| --- |
|  |

**Figure S1** Modelled annual median **(mean?)** aerosol scattering coefficient at 550 nm (No MerrAero, as only data at site locations was provided). Need to rearrange so models in same order as Figure 3b.

**Figure S2** Modelled annual median **(mean?)** aerosol absorption coefficient at 550 nm (No MerrAero, as only data at site locations was provided). Need to rearrange so models in same order as Figure 4b

**Figure S3** Modelled annual median **(mean?)** aerosol single scattering albedo at 550 nm (No MerrAero, as only data at site locations was provided). Need to rearrange so models in same order as Figure 3b.

**Figure S4** Modelled annual median **(mean?)** aerosol scattering Ångström exponent for 440/550(?) nm wavelength pair. (No MerrAero, as only data at site locations was provided). Need to rearrange so models in same order as Figure 3b.