The importance of temporal collocation for the evaluation of aerosol models with observations

N. A. J. Schutgens[1], D. G. Partridge[2,3], and P. Stier[1]

[1]Department of Physics, University of Oxford, Parks road, Oxford OX1 3PU, England
[2]Department of Environmental Science and Analytical Chemistry, Stockholm University, Stockholm, Sweden
[3]Bert Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

Correspondence to: N. A. J. Schutgens (schutgens@physics.ox.ac.uk)

Abstract. It is often implicitly assumed that over suitably long periods the mean of observations and models should be comparable, even if they have different temporal sampling. We assess the errors incurred due to ignoring temporal sampling and show that they are of similar magnitude as (but smaller than) actual model errors (20–60 %).

Using temporal sampling from remote-sensing data sets, the satellite imager MODIS (MODerate resolution Imaging Spectroradiometer) and the ground-based sun photometer network AERONET (AErosol Robotic NETwork), and three different global aerosol models, we compare annual and monthly averages of full model data to sampled model data. Our results show that sampling errors as large as 100 % in AOT (aerosol optical thickness), 0.4 in AE (Ångström Exponent) and 0.05 in SSA (single scattering albedo) are possible. Even in daily averages, sampling errors can be significant. Moreover these sampling errors are often correlated over long distances giving rise to artificial contrasts between pristine and polluted events and regions. Additionally, we provide evidence that suggests that models will underestimate these errors. To prevent sampling errors, model data should be temporally collocated to the observations before any analysis is made.

We also discuss how this work has consequences for in situ measurements (e.g. aircraft campaigns or surface measurements) in model evaluation.

Although this study is framed in the context of model evaluation, it has a clear and direct relevance to climatologies derived from observational data sets.
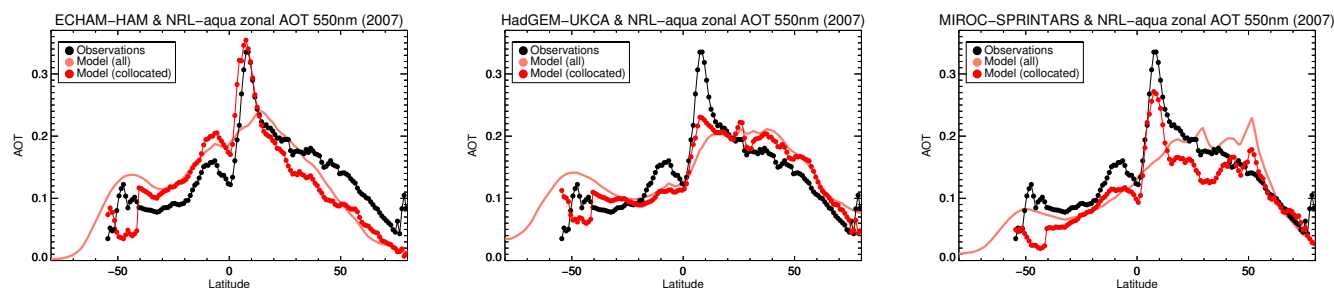
1 Introduction

In the past decades, the role of atmospheric aerosol in the Earth's climate and biosphere has become clearer. Aerosols can change the global radiation budget directly (Angstrom, 1962) and indirectly (Twomey, 1974; Albrecht, 1989). They can affect the temperature structure of the atmosphere (Hansen et al., 1997; Lohmann and Feichter, 2005) and may have consequences for the hydrological cycle (Lohmann and Feichter, 1997). Dust aerosols may transport nutrients for the biosphere over long distances (Vink and Measures, 2001; McTainsh and Strong, 2007; Maher et al., 2010; Lequy et al., 2012) and anthropogenic aerosols can pose health hazards for humans (Dockery et al., 1993; Brunekreef and Holgate, 2002; Ezzati et al., 2002; Smith et al., 2009; Beelen et al., 2013).

One approach to explore the role of aerosol is through (global) models. A lot of effort has gone into evaluating global aerosol models against observations (Kinne et al., 2006; Schulz et al., 2006; Textor et al., 2006, 2007; Huneeus et al., 2011; Koch et al., 2009; Quaas et al., 2009; Koffi et al., 2012) in the framework of the AEROCOM (AEROsol Comparisons between Observations and Models) community (http://aerocom.met.no). Although this has allowed the community to identify weaknesses in models (and has driven further model development), less effort has gone into developing best practices when performing such an evaluation.

When evaluating models with observations, one problem is the different temporal sampling of the data sets. While models can in principle provide data at any time, observations may not be available for much of a day or even extended periods (up to several weeks) due to, e.g. satellite overpass

Table 1. Models used in this study.

	ECHAM-HAM	HadGEM-UKCA	MIROC-SPRINTARS
Version	ECHAM6.1-HAM2.2	HadGEM3-A-GLOMAP UKCA 8.4	MIROC3.2-SPRINTARS3.84
Resolution	$1.9° \times 1.9°$, 31 levels (T63L31)	$1.9° \times 1.25°$, 85 levels (N96L85)	$2.8° \times 2.8°$, 20 levels (T42L20)
Reanalysis	ERA-interim	ERA-interim	NCEP
Natural emissions	model time-step	model time-step	model time-step
Anthropogenic emissions	monthly	monthly	monthly
Wildfire emissions	daily	monthly	monthly

Figure 1. Yearly and zonal averages of AOT from the observations (MODIS-NRL Aqua) and the models. The latter are averaged in two different ways: using either all model data or only those sampled at the observation times.

times, cloudiness, low-light conditions or other unfavourable circumstances as well as instrument malfunction or maintenance. Although limited temporal sampling is clearly an issue for remote-sensing data sets, in situ data may likewise be affected. This temporal sampling issue is however ignored in many studies where it is implicitly assumed that suitably long averaging periods (months or years) will put model and observational data sets on equal footing.

In this paper, we will study errors due to temporal sampling errors by applying temporal samplings from actual observational data sets to global model data. The temporal evolution of aerosol has received some interest before: Anderson et al. (2003) used hourly nephelometer measurements of dry extinction to study timescales of aerosol evolution over one polluted and one remote site. They found timescales from 2 to 48 h, with aerosol often becoming uncorrelated over a day and a half. Kaufman et al. (2000) used AERONET (AErosol Robotic NETwork) AOT (Aerosol Optical Thickness) data to assess whether a single daily overpass from either Terra or Aqua satellite would allow an accurate estimate of daily-averaged AOT. They concluded this was indeed possible, suggesting that aerosol diurnal cycles are small. Smirnov (2002), also using AERONET AOT, showed that diurnal cycles of 10 to 40 % existed near polluted and biomass burning sources while diurnal cycles were insignificant over ocean. Remer et al. (2006) put an upper limit of 2.5 % on regional AOT changes over 3 hours by comparing MODIS (MODerate resolution Imaging Spectroradiometer) Aqua and Terra aerosol observations over the remote oceans. Geogdzhayev et al. (2014) studied the ability of MODIS Aqua and Terra to assess regional and monthly AOT using model data and MODIS L2 temporal samplings. They found differences in monthly, regional AOT of 0.01–0.02 over ocean and of 0.03–0.09 over land. The model data consisted of daily averages so no investigation of diurnal cycles was possible.

The previous studies highlight aspects of the temporal evolution of aerosol but always with the purpose of deriving climatologies from observational data sets. Temporal sampling of aerosol observations in the context of model evaluation has received little interest, although Sayer et al. (2010) pointed out that differences in AOT of up to 0.1 in monthly and annual regional averages were possible when comparing model data to AATSR (Advanced Along-Track Scanning Radiometer) satellite observations.

Global models are routinely compared to a host of remote-sensing data, from satellites and ground sites. Within AEROCOM, standard practice is to use daily averaged model data in these comparisons but in the literature monthly or yearly averages are used as well. With regard to temporal sampling issues, it seems that many questions are still left unanswered: what magnitude of temporal sampling error is possible on daily, monthly and yearly timescales, for individual grid points? Do these errors exhibit distinct spatial patterns? Do aerosol diurnal cycles impact temporal sampling errors? Can models represent aerosol temporal variability well enough to be used to assess sampling errors? Are intensive aerosol properties like AE (Ångström Exponent) and SSA (single scattering albedo) less affected than extensive aerosol properties like AOT? What controls the size of these sampling errors? And finally, how do temporal sampling errors compare to model errors and observational errors? In this paper, we make a first attempt at answering these questions.
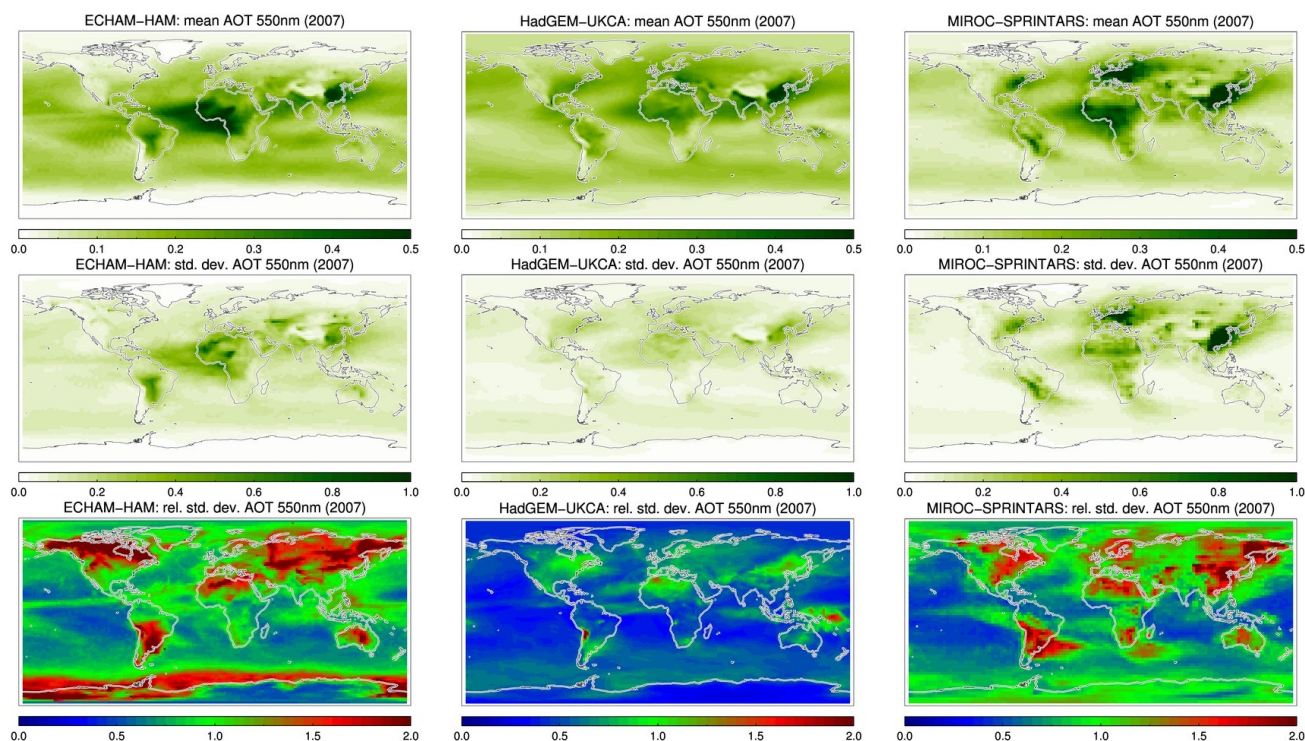
Figure 2. Global maps of yearly mean, standard deviation, and relative standard deviation in AOT.

In Sect. 2, we give a brief overview of the models that will be used, followed by a similarly brief overview of the observational data sets in Sect. 3. Our methodology is explained in more detail in Sect. 4, but the bulk of our paper is dedicated to the results in Sect. 5. We will start by describing simulated observations, their temporal variability and how they compare to actual observations in Sect. 5.2. Next, we will discuss the spatial patterns in temporal sampling errors. A fuller explanation of what controls temporal sampling errors is given in Sect. 5.4. These errors are compared to observational errors and models errors in Sect. 5.5. The paper is wrapped up with a summary (Sect. 6) of our results.

2 Models

Three global aerosol models, ECHAM-HAM, HadGEM-UKCA and MIROC-SPRINTARS, see also Table 1, are used in this paper to provide time-varying aerosol fields.

The global aerosol model ECHAM-HAM consists of the aerosol module HAM (Stier et al., 2005, 2007; Zhang et al., 2012) coupled to the atmospheric general circulation model ECHAM (Roeckner et al., 2003, 2006). It solves the prognostic equations for vorticity, divergence, surface pressure and temperature using spherical harmonics with triangular truncation. Aerosols are advected with a flux-form semi-Lagrangian transport scheme (Lin and Rood, 1996) on a Gaussian grid. The aerosol module HAM calculates the
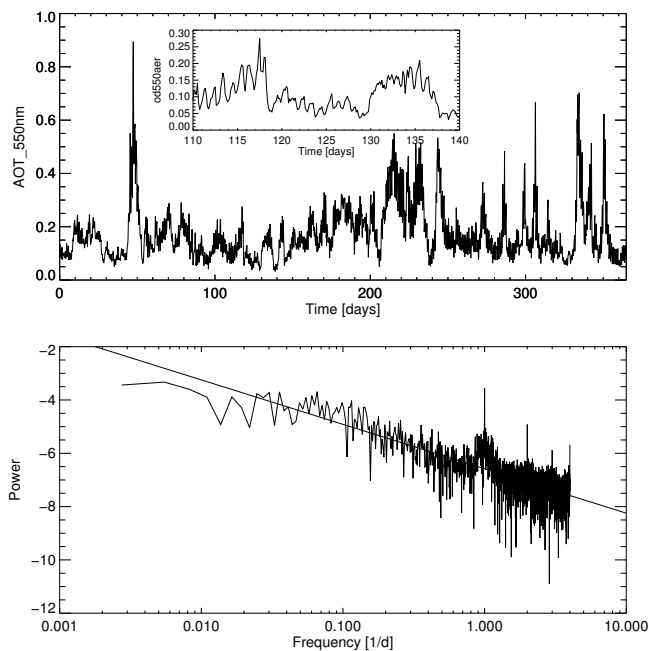

Figure 3. AOT time series and its associated power spectrum at $75°$ W, $0°$ N in HadGEM-UKCA. For this power spectrum, the ratio of power at 1 day over total power is ~ 0.01. The straight line in the lower plot has a slope of $-5/3$.

Table 2. Median ratio of daily AOT variation δ_o/δ_m across all AERONET sites. Only sites with more than two observations for at least 24 days during 3 months were used. Since the estimated error in AERONET AOT measurements is 0.01, the error in measured δ_o can be no more than $\sqrt{2 \times 0.01^2}$.

δ_o	ECHAM-HAM	HadGEM-UKCA	MIROC-SPRINTARS
≥ 0	2.48	2.36	2.08
$\geq \sqrt{2 \times 0.01^2}$	3.66	3.75	2.95
$\geq 2\sqrt{2 \times 0.01^2}$	4.19	4.95	3.35

global evolution of five aerosol species: sulphate, particulate organic matter, black carbon, sea salt and dust. These species are the constituents of both internally and externally mixed aerosol particles whose size distribution is represented by seven uni-modal log-normal distributions called modes. These seven modes describe four size classes (nucleation, Aitken, accumulation and coarse) and two hygroscopic classes (hydrophobic and hydrophilic). HAM uses the two-moment M7 aerosol microphysics scheme (Vignati et al., 2004).

The global aerosol model HadGEM-UKCA (Bellouin et al., 2013) uses the UKCA aerosol and chemistry schemes with the third generation of the Hadley Centre Global Environmental Model (Hewitt et al., 2011) developed at the UK Met Office. This circulation model is non-hydrostatic and uses a semi-Lagrangian transport scheme. UKCA calculates the evolution of five aerosol species, sulphate, particulate organic matter, black carbon, sea salt and dust, in both internally and externally mixed particles. The aerosol scheme in UKCA is based on the Global Model of Aerosol Processes (GLOMAP-mode, Mann et al., 2010) which is similar to the M7 framework. The main exception is that dust is currently calculated separately using six size bins. UKCA hence only considers five modes.

The global aerosol model MIROC-SPRINTARS uses the aerosol module SPRINTARS (Takemura et al., 2000, 2002; Takemura, 2005) in conjunction with the atmospheric general circulation model MIROC. SPRINTARS calculates the global evolution of three externally mixed aerosol species: sulphate, sea salt and dust; and two species that can be either internally or externally mixed: organic matter and black carbon. In contrast to HAM and UKCA, SPRINTARS uses a single moment scheme and carries only mass as a prognostic variable. The size distributions of sulphate, organic matter and black carbon are fixed but for sea salt and dust a bin scheme is used.

All three models were run for the year 2007, with nudged meteorology, prescribed sea surface temperatures and appropriate emission inventories (see Table 1).

3 Observations

We used the temporal sampling from several remote-sensing data sets, both satellite and ground sites. The satellite sensor is the MODIS wide-swath (~ 2330 km) multi-channel imager onboard the Aqua satellite (local equator crossing time 1:30 p.m., repeated view period: 1 to 2 days depending on latitude). The ground sites constitute the AERONET network of sun-photometers that measure solar transmittance with nominally a 15 min time resolution.

We chose not to use the original observations but data sets that have undergone further quality checks and spatiotemporal aggregation. It seems reasonable to study temporal sampling issues for data sets that have been optimised for actual model evaluations.

The satellite remote-sensing data are the MODIS Aqua L3 NRL (Naval Research Laboratory) data (Zhang and Reid, 2006; Shi et al., 2011; Hyer et al., 2011) for 2007. These are based on official MODIS Aqua L2 Coll. 5 data (see Remer et al. (2005) for a discussion of the very similar over-ocean retrieval for Coll. 4 and Levy et al. (2007) for Coll. 5 over-land retrievals) that have been subjected to further quality checks, empirical corrections and spatial aggregation. The data comprise of AOT and its error estimate for 1° by 1° grid boxes at a 6-hourly resolution. We have analysed MODIS Terra as well but since it does not alter our conclusions, MODIS Terra observations will not be used in this paper.

We also use the MODIS Aqua L3 AORI (Atmosphere and Ocean Research Institute) data (Schutgens et al., 2013) for 2007. These are based on MODIS Aqua L2 data that have been subjected to further quality checks, empirical corrections and spatial aggregation. The AORI data differ from the NRL data as they use different empirical corrections and include both AOT and AE. Unlike the NRL data, AORI data are available only over ocean. In a preliminary analysis against Maritime Aerosol Network observations, we found that the AORI data suffered less from the high bias in AOT at low AOT values than the NRL data. The AORI AOT and AE data and its error estimates are available for 1° by 1° grid boxes at a 6-hourly resolution. MODIS Terra data will be ignored, again because it does not alter our conclusions.

Not applying the extra quality checks in the AORI product almost doubles the number of observations. While this does reduce the temporal sampling error somewhat (primarily in yearly averages), it does not alter our conclusions fundamentally. This also suggests that using Coll. 6 instead of Coll. 5 will have a minimal impact.

The other two remote-sensing data sets are ground-based: AERONET Direct Sun lev 2.0 (Holben et al., 1998; Eck et al., 1999; Schmid et al., 1999) and AERONET Version 2 Inversions lev 2.0 (Dubovik and King, 2000; Dubovik et al., 2000) for 2007. Observations at each site were averaged over 6 h, each 6 h. AERONET data used in this study are AOT, AE and SSA.
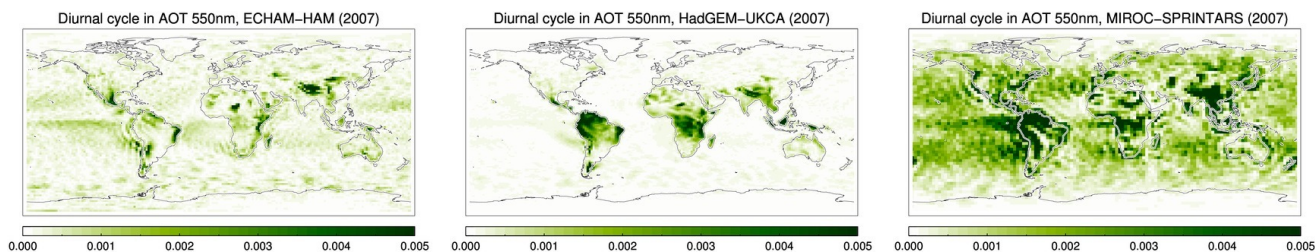
Figure 4. Global maps of the relative importance of the diurnal cycle: the ratio of power at a period of 1 day to total power for AOT time series.
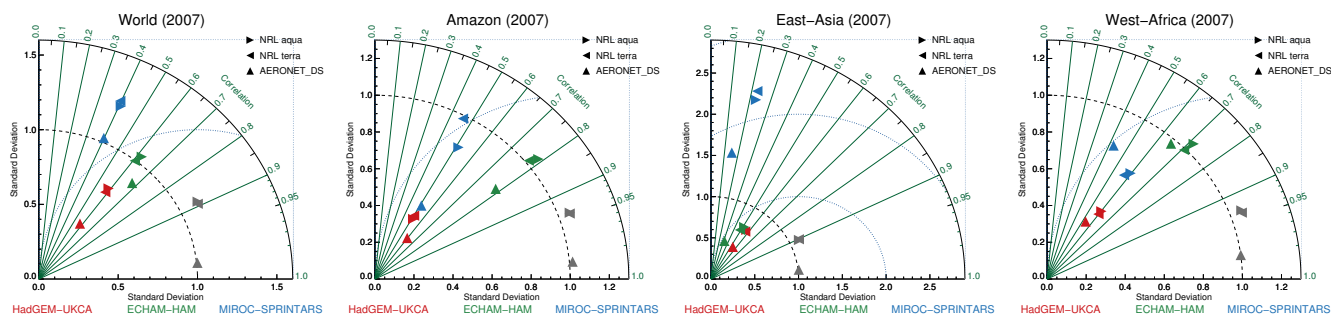

Figure 5. Taylor plots showing the standard deviation and correlation of modelled AOT against MODIS-NRL and AERONET observations. The dark grey symbols refer to observations randomly perturbed with estimated observational errors.

As the MODIS-NRL product is provided on a regular 6-hourly grid, all other observations were similarly put on a regular 6-hourly grid. Using a 3-hourly grid for AERONET hardly affected our results. Moreover it was just as likely to (slightly) increase sampling errors as (slightly) decrease them.

4 Method

The 3-hourly model data were linearly interpolated to the locations of actual observations. These model values v were then sampled at the times of the actual observations to generate simulated observations.

The straight average

$$\overline{v} = N^{-1} \sum_{i=1}^{i=N} v_i, \tag{1}$$

(where i represents time and N the number of values over a day, a month or a year) is the average normally produced by models and often used in model evaluations.

The sampled average

$$\widetilde{v} = \left(\sum_{i=1}^{i=N} f_i\right)^{-1} \sum_{i=1}^{i=N} f_i v_i, \tag{2}$$

is a model-simulated observation (daily, monthly or yearly mean) where f_i represent the observational sampling (taken

from actual observational data sets):

$$f_i = \begin{cases} 0 & \text{if no observation present at time } i \\ 1 & \text{if observation present at time } i \end{cases} \tag{3}$$

The impact of not collocating model data with observations is now given by the error

$$\epsilon = \overline{v} - \widetilde{v}, \tag{4}$$

where we consider $\widetilde{v}$ the reference value. After all, actual observations are (hopefully) a proxy for the truth and used to assess models. Note that this error can be rewritten as

$$\epsilon = N^{-1} \sum_{i=1}^{i=N} \left(1 - \left(\frac{\sum_{j=1}^{j=N} f_j}{N}\right)^{-1} f_i\right) v_i. \tag{5}$$

The temporal sampling error depends on the statistics of the time series v_i as well as the statistics of observations times f_i and in particular the amount of observational temporal coverage

$$C = \frac{\sum_{j=1}^{j=N} f_j}{N}. \tag{6}$$

5 Results

5.1 Does temporal sampling matter?

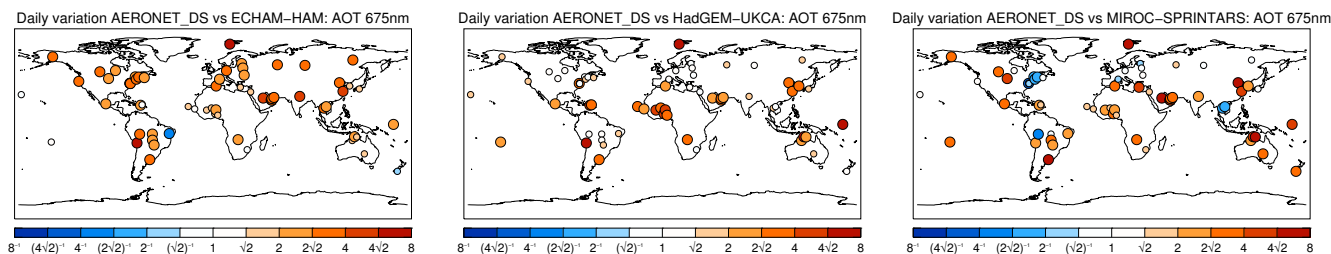Figure 1 shows yearly and zonally averaged AOT from both observations (MODIS-NRL Aqua) and models (either collo-

Figure 6. Median ratio of daily difference in maximum and minimum AOT δ_o/δ_m, either observed or modelled, for all AERONET sites. Smaller circles indicate sites where $0.5 \leq$ median $\delta_o/\delta_m \leq 2$. Only sites with more than two observations per day for at least 24 days during 3 or more separate months were used.
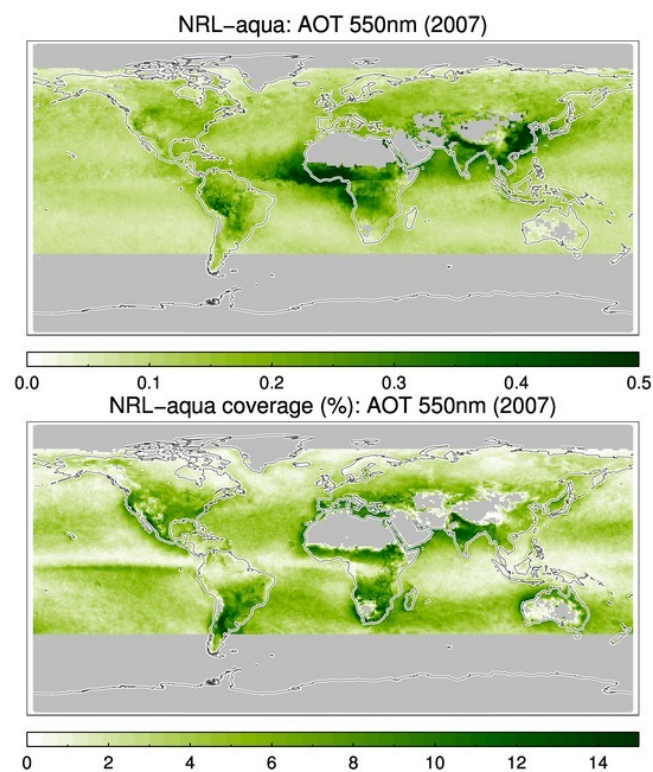

Figure 7. Top panel shows the 2007 yearly mean AOT of MODIS-NRL Aqua. Bottom panel shows the temporal coverage (%) of observations during that year. Grey colours indicate no observations available at all.

cated with the observations or not). The difference between the straight model average (solid red line) and the sampled model average (dotted red line) shows the impact sampling has on model averages. The difference between the sampled model average (dotted red line) and observations (dotted black line) is due to model errors (and observational errors). Clearly, sampling errors can be as big as model errors. We also see that resampling the model improves agreement with some prominent features in the observations.

5.2 Temporal evolution of modelled AOT

To better understand the temporal sampling errors, we first need to consider the models' temporal evolution. Figure 2 shows the mean, standard deviation and their ratio (relative standard deviation) of the modelled AOT time series. All three models agree in global mean AOT patterns (elevated AOT near natural sources like the Sahara and Southern Ocean and anthropogenic sources like Europe and East Asia), but there are substantial differences: HadGEM-UKCA shows lower mean AOT over the Sahara, ECHAM-HAM shows lower AOT over Europe and MIROC-SPRINTARS shows lower AOT over the Southern Ocean. Similarly, there are differences in standard deviation: HadGEM-UKCA has its largest standard deviation in AOT over East-Asia, ECHAM-HAM over the Sahara and East-Asia and MIROC-SPRINTARS over East-Asia and Eastern Europe. The relative standard deviation (standard deviation divided by the mean) also differs among these models, even though all three had their meteorology nudged to reanalysis data, suggesting there are fundamental differences in their time evolution due to the aerosol modelling itself. In particular, HadGEM-UKCA shows globally lower relative standard deviation than the other two models, while ECHAM-HAM and MIROC-SPRINTARS differ in the spatial patterns of the relative standard deviation. Regionally, relative standard deviation can easily differ by more than a factor of 2 among these three models.

AOT time series in our models are characterised by power-law power spectra with an exponent close to Kolmogorov's $-5/3$ constant, see Fig. 3 as an example. Sometimes this power spectrum flattens out into white noise at large periods due to synoptic-scale weather systems. In any case, AOT variability is largest for the longest periods. Diurnal cycles are not very strong for these models and their global patterns and magnitudes are very different (Fig. 4). As the modelled diurnal cycles are quite uncertain, so will our (later) conclusions regarding diurnal sampling errors.

We present a comparison of the (collocated) models against actual AOT observations as Taylor plots in Fig. 5. A Taylor plot (Taylor, 2001) graphically represents the variabil-
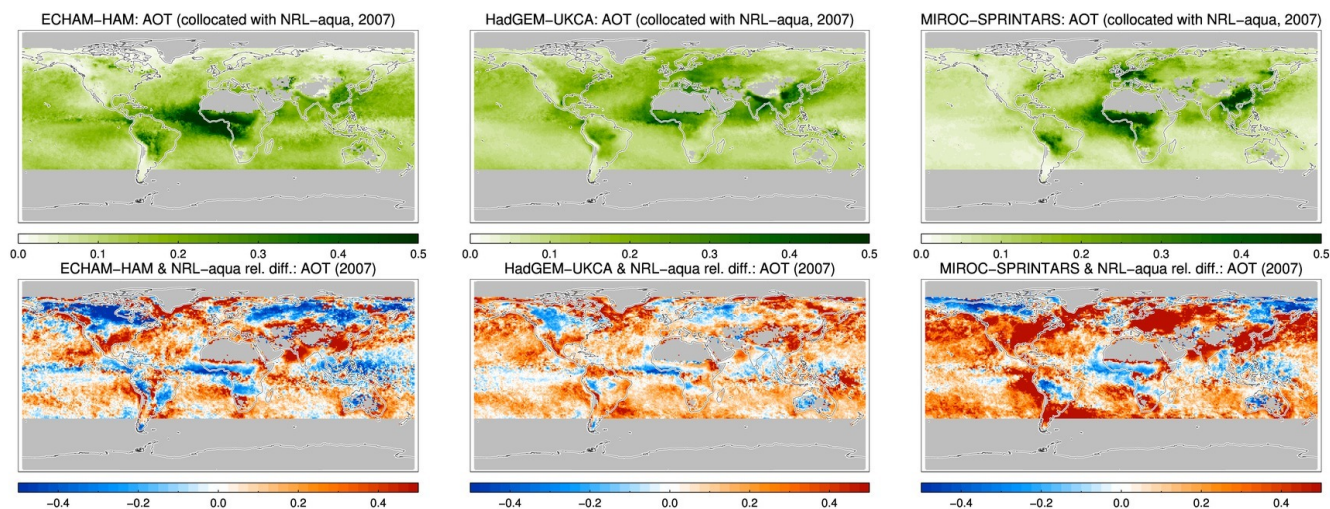
Figure 8. Top row shows global map of yearly mean of model AOT collocated with MODIS-NRL Aqua AOT data. Bottom row shows the relative temporal sampling error using MODIS-NRL Aqua AOT sampling. Grey colors indicate no observations available.
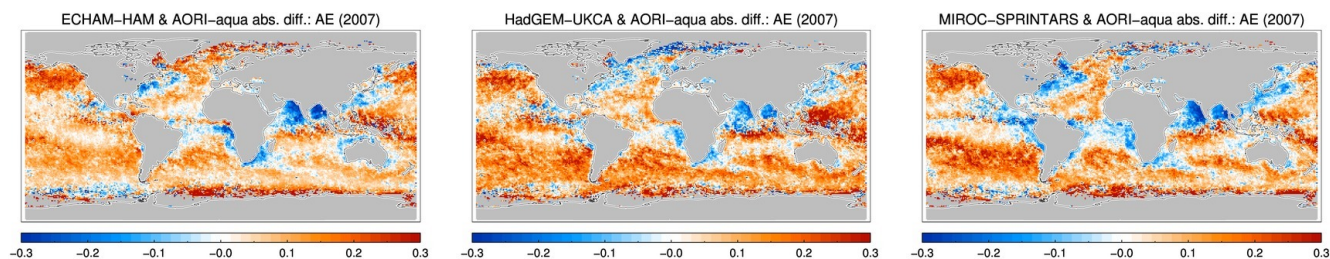

Figure 9. Global map of absolute temporal sampling error using MODIS-AORI Aqua AE sampling. Grey colors indicate no observations available.

ity of the model data and its correlation with observations and is especially suited to analyse the models' temporal evolution. Here we show them for four different regions: the entire Earth, the Amazon, East Asia and Western Africa. Mostly the models underestimate observed variability (the exception is MIROC-SPRINTARS which often overestimates it). Note that observational error has only a small impact on these statistics.

Assessing the model's diurnal cycle from remote-sensing observations that require daylight is difficult, but a simple analysis that is appropriate to this paper's focus is to consider daily variation. This daily variation is not necessarily the result of a diurnal cycle but may also be caused by, e.g. synoptic weather patterns. We compare the difference δ between maximum and minimum AOT over a day for all AERONET sites for both observations δ_o and models δ_m. Figure 6 shows the median δ_o / δ_m for each AERONET site throughout the year, while Table 2 presents the median across all sites (in this case, medians provide a more conservative and lower estimate than means). Over a period of a day, observations tend to show more variability than the models. Various sensitivity studies (e.g. δ_o / δ_m as a function of δ_o as shown in Table 2) suggest that this is not due to observational error. It is likely

due to at least two factors: first an underestimation of modelled AOT over the continents that limits absolute daily variation; secondly, modelled AOT time series show very strong auto-correlations (> 0.8) over 6 h, further reducing potential daily variation. Consequently, we suggest that these models underestimate daily variation in AOT.

Similarly, there are substantial differences in the time series of AE and SSA among the three models (not shown). In particular, ECHAM-HAM shows a larger standard deviation in AE and SSA over the continents than the other two models, while HadGEM-UKCA shows a very small standard deviation in SSA over the oceans.

5.3 Spatial patterns in temporal sampling errors

Yearly averaged MODIS-NRL Aqua AOT and its temporal coverage are shown in Fig. 7. Coverage is especially low over regions with high cloudiness and land areas with strong variations in surface albedo. In Fig. 8, we show yearly averages of the modelled AOT collocated with MODIS-NRL Aqua. The relative sampling error (Eq. 4) when using straight model averages is also shown. There are several things to note. First, temporal sampling errors can easily cause AOT
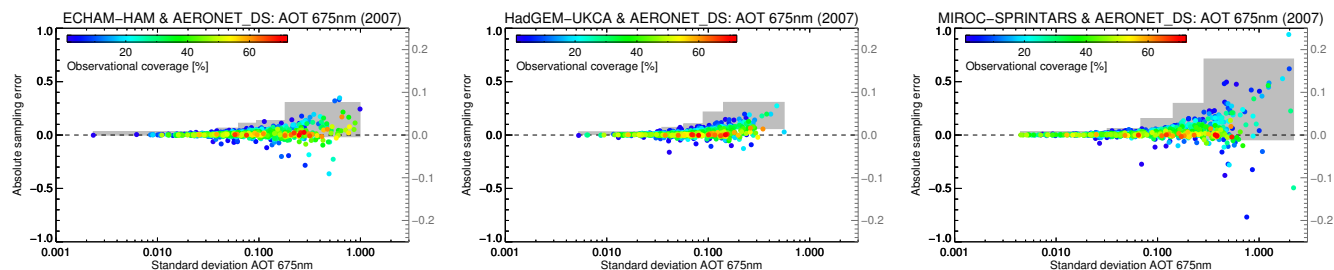
Figure 10. Absolute temporal sampling error for monthly mean AOT using AERONET sampling, as a function of time series standard deviation and observational coverage. The grey boxes in the background (axis on the right), show the 25 and 75 % quantiles of those errors for equally sized sub-samples.
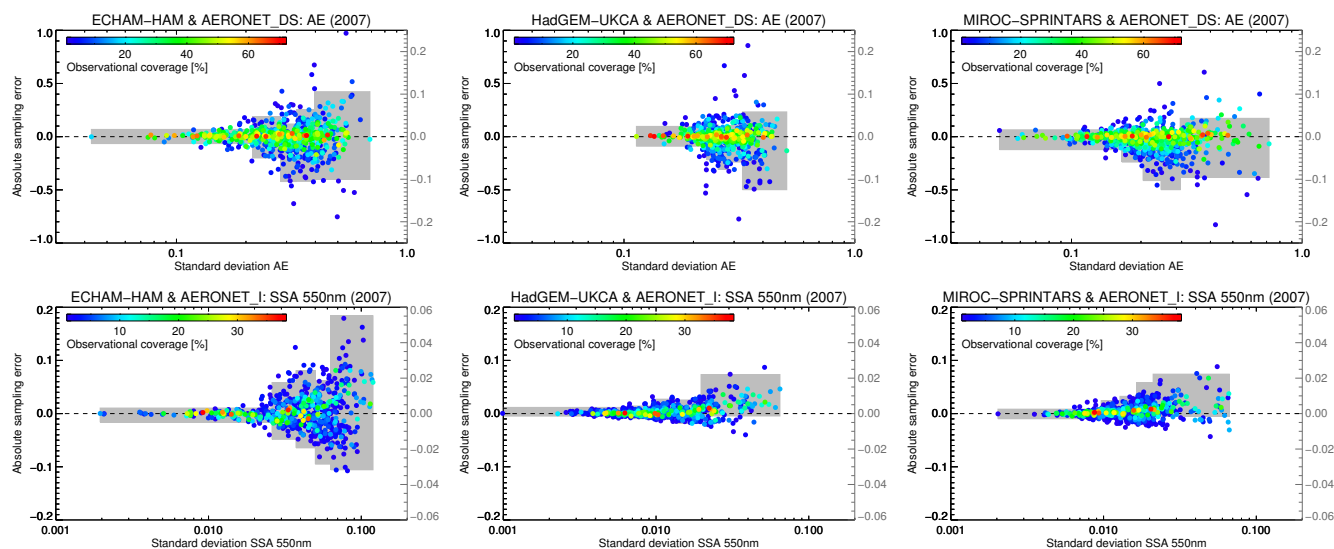

Figure 11. Absolute temporal sampling error for monthly mean AE and SSA using AERONET sampling, as a function of time series standard deviation and observational coverage. The grey boxes in the background (axis on the right), show the 25 and 75 % quantiles of those errors for equally sized sub-samples.

to be over or under-estimated by as much as 50 %. Secondly, these temporal sampling errors show substantial spatial correlation: entire regions show similar errors greatly reducing the possibility of reducing temporal sampling errors through spatial averaging (see, e.g. Fig. 1). Thirdly, different models predict different error patterns. Finally, we point out that observational coverage only partly explains sampling errors. For instance, Fig. 7 suggests that India is far better sampled (temporally) than the Bay of Bengal, yet relative errors are larger in India, at least in the ECHAM-HAM and MIROC-SPRINTARS experiments.

Obviously, the temporal sampling error in yearly averages may be strongly impacted by several months of no observations at all. Over land, snow cover (combined with cloud cover) may preclude meaningful observations and, e.g. Siberia is not observed at all during the Northern Hemisphere winter. Even over ocean, high latitudes may not be observed during part of the year due to low SZA (Solar Zenith Angle) or ice cover. We have inspected global maps of this

temporal sampling error for monthly means, but found that errors can be just as large as for yearly means, due to the 1 to 2-day period between repeated views by MODIS and the possibility of extended cloud cover. In particular, for almost every location on Earth observed by MODIS, we found at least 1 month when temporal sampling errors exceeded 100 %. A more in-depth analysis of the impact of averaging periods will follow later (Sect. 5.5).

Lastly, the absolute temporal sampling error in yearly averaged AE from the MODIS-AORI data set is shown in Fig. 9. The observational coverage of AE is less than that of AOT, due to more stringent quality checks (not shown). Sampling errors rarely exceed ±0.3. However, a very distinct spatial pattern can be seen for all three models: near coastlines, especially in polluted outflow regions, the error tends to be strongly negative while over the remote oceans it tends to be strongly positive. If one does not properly collocate AE model data with observations, this would create an artificial contrast up to ~ 0.6 in AE between the continents and
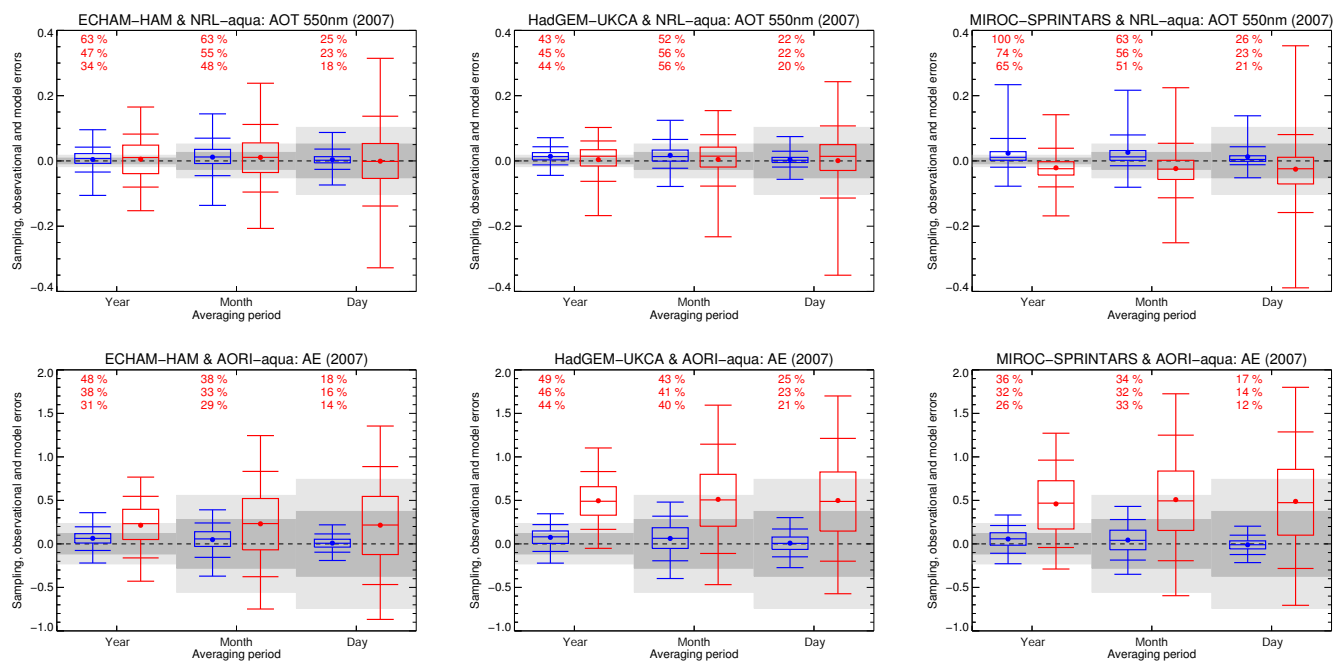
Figure 12. Box-whisker plots of temporal sampling errors, observational and model errors. For MODIS Aqua, sampling errors (blue), observational errors (grey) and model errors (red) are shown side by side as a function of averaging period (yearly, monthly, daily). The box-whisker plot shows the 2, 9, 25, 50, 75, 91 and 98 % quantiles of these errors, the dot shows the mean error. The grey shading shows the 9, 25, 75 and 91 % quantiles of the observational errors. The red numbers above each blue box-whisker show how the different quantile ranges (from top to bottom: 2–98, 9–91 and 25–75 %) compare between the sampling error and the model errors.
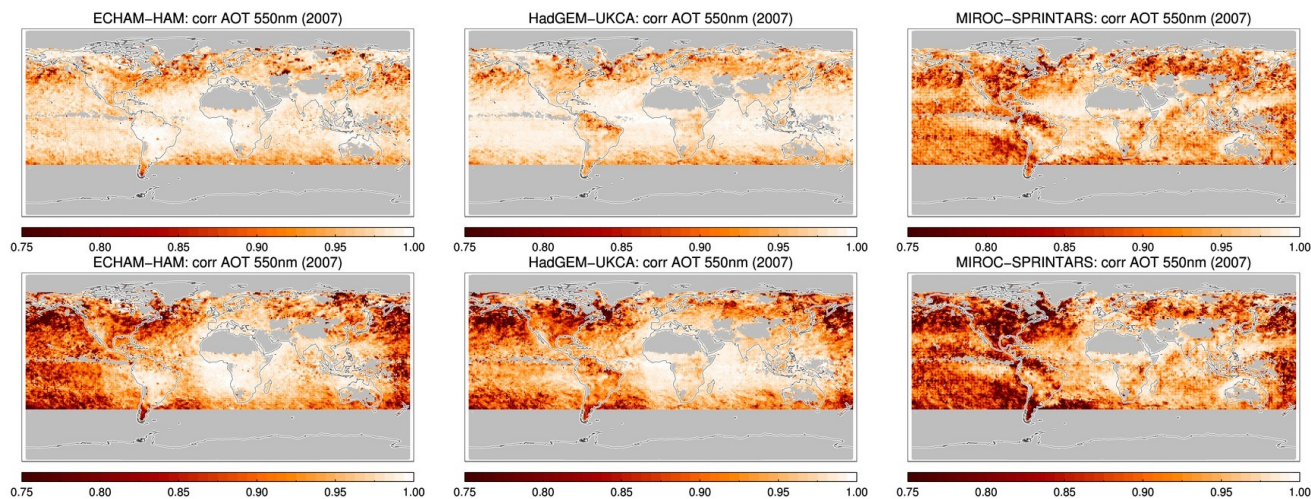

Figure 13. Correlation between model AOT at the time of MODIS-NRL Aqua observation and a 24 h model AOT average. Top: 24 h model average uses overpass time as the centre for the 24 h average; bottom: 24 h average defined according to UTC. Note how in the bottom plot, correlations drop off towards the date-line.

the remote oceans. The reason is that AE observations require a minimum AOT to be successful (see Schutgens et al., 2013). In areas with polluted outflows, low AOT correspond to the sea-salt background with low AE and observations will mostly sample the outflow with high AE, leading to negative sampling errors. Conversely, over the remote oceans high

wind-speeds lead to high AOT and low AE (larger particles) that will be more often observed than the lower AOT and higher AE (smaller particles) that occur at low wind-speeds.

The last paragraph suggests that minimum AOT requirements in retrievals can have large impacts on temporal sampling errors. While this was shown for the AORI data set,
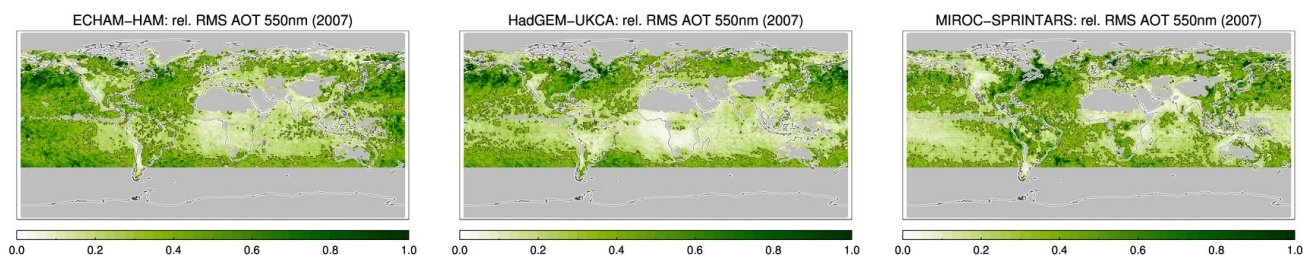
Figure 14. Ratio of the RMS (root mean squared) temporal sampling daily error to RMS model daily error, i.c. MODIS-NRL Aqua. The daily errors were defined according to UTC. The thin black contour represents a ratio of 0.3.
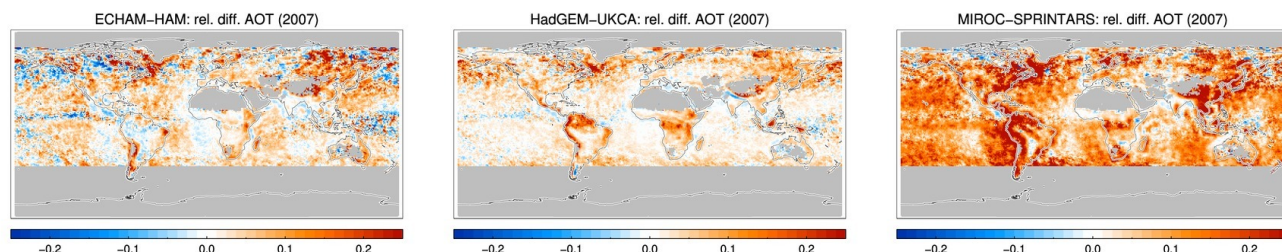

Figure 15. Relative temporal sampling error using MODIS-NRL Aqua AOT sampling. The yearly averages are constructed from daily data, excluding days for which no observations were present. This figure can be compared to the bottom row of Fig. 8 (but note that the colour bar has half its range). Grey colors indicate no observations available.

already the official MODIS Coll. 5 product shows a strong dependence of the quality of AE retrievals on AOT values, as might be expected. See also Fig. 11 in Schutgens et al. (2013).

5.4 Drivers of temporal sampling errors

In Fig. 10, we show temporal sampling errors for monthly averaged AERONET AOT (similar plots can be shown for MODIS). Sampling errors increase with the standard deviation of the time series, but decrease with increasing observational temporal coverage. Still, even for relatively high coverage large sampling errors are possible. We surmise that the clear positive bias is due to cases of high relative humidity that will tend to (1) prevent observations due to cloudiness; (2) increase model AOT due to wet growth. Randomising the observation times (instead of using actual AERONET times) in this analysis removes this bias and also reduces the spread in the errors by 15 to 45 % (depending on the model). Thus it appears that correlations between the time series and the observation times (see Eq. 5) also affect these sampling errors.

Similar figures for monthly-averaged AERONET AE and SSA are shown in Fig. 11. They confirm previous results and mostly serve to show the magnitude of the sampling errors that can possibly occur. The positive bias for SSA is likely due to a minimum AOT in the AERONET Inversion data files (all inverted SSA values have AOT $\geq$ 0.2). As a result, the sampled model data will mostly sample pollution or dust events with low SSA while the straight model data also include the aerosol background state with higher SSA.

5.5 Temporal sampling errors in context

To put these sampling errors in context, we will compare them to model errors and observational errors. The model errors are defined as the difference of collocated model values from actual observations. Strictly speaking this "model error" includes observational errors as well but unfortunately they can not be separated. However, estimates of this observational error's standard deviation are part of our observational data sets (we will assume this error distribution to be an unbiased Gaussian). The model error (minus the observational error) is the signal we are interested in when evaluating models.

Figure 12 shows a comparison of the above errors for yearly, monthly and daily averages of MODIS data (AOT and AE). For both AOT and AE, we see that observational errors and model errors decrease as averaging periods increase. In contrast, sampling errors behave differently, with the largest errors typically occurring for monthly averages and the smallest errors for daily averages. The models differ both in the sampling errors and how they compare to the model errors. But a few conclusions seem obvious: (1) for yearly and monthly averages, sampling errors will contribute considerably (typically 30–60 %) to model errors (defined as the difference between a model and observations) if no collocation is performed; (2) sampling errors in AOT are comparable to or larger than observational errors for monthly or yearly averages; (3) for daily averages, these sampling errors appear to be smaller.
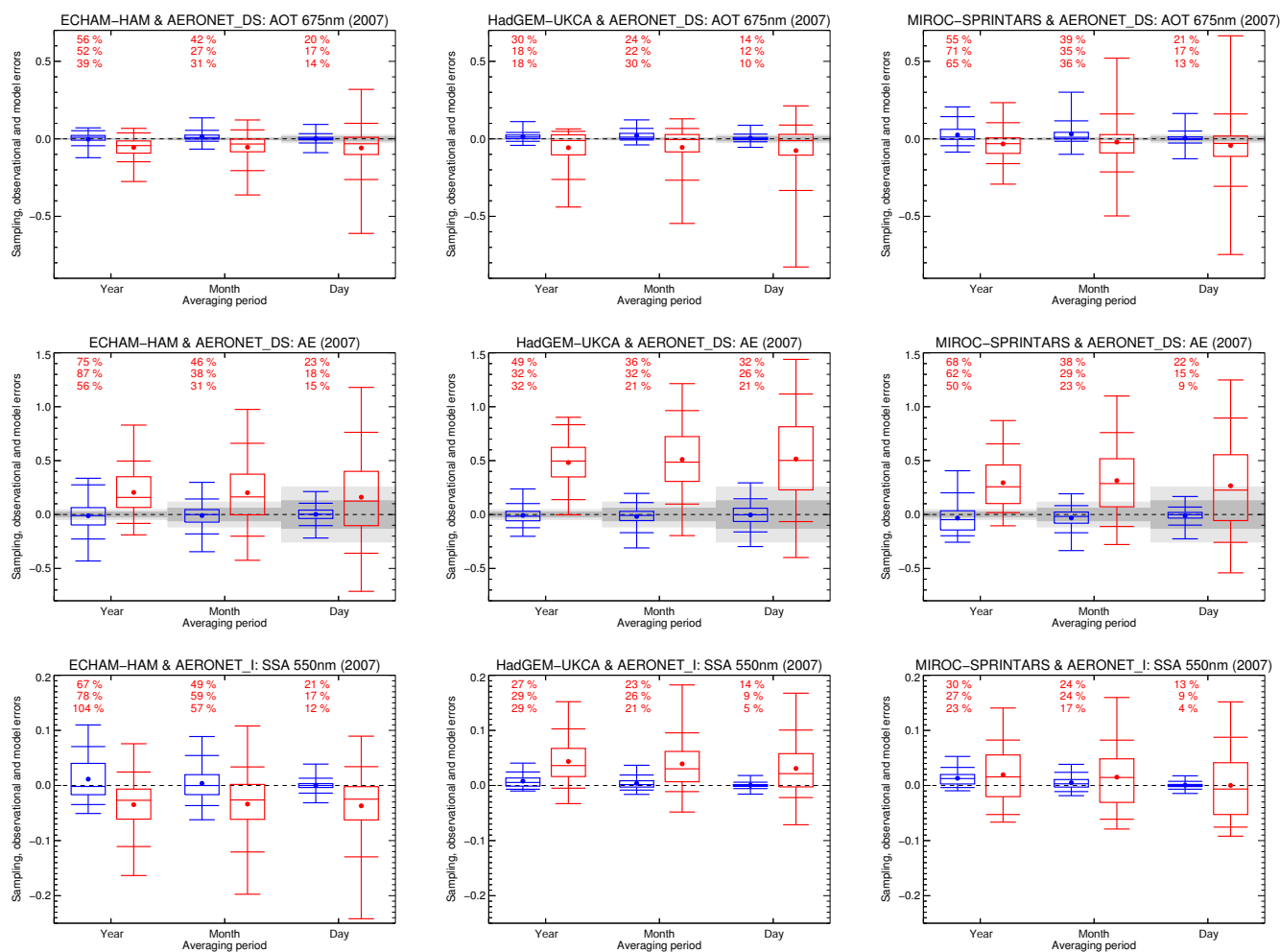
Figure 16. Box-whisker plots of temporal sampling errors, observational and model errors. For AERONET, sampling errors (blue), observational errors (grey) and model errors (red) are shown side by side as a function of averaging period (yearly, monthly, daily). The box-whisker plot shows the 2, 9, 25, 50, 75, 91 and 98 % quantiles of these errors, the dot shows the mean error. The grey shading shows the 9, 25, 75 and 91 % quantiles of the observational errors. The red numbers above each blue box-whisker show how the different quantile ranges (from top to bottom: 2–98, 9–91 and 25–75 %) compare between the sampling error and the model errors. For SSA, no observational error estimates were available.

This last point is important as standard practice for AEROCOM models is to save data as daily averages. Doing so would incur sampling errors of at most 25 % of the model error. There are a few caveats however. We have already pointed out that models seem to under-estimate AOT variability, especially on daily timescales (see Sect. 5.2). Thus we are likely to underestimate daily sampling errors. Furthermore, the AEROCOM daily averages are defined according to UTC. This means that, e.g. satellite observations (made close to local noon time) near the date-line are not as representative of the model day as observations near the Greenwich meridian. The consequence of this is shown in Fig. 13, where correlations between daily model averages and simulated observations clearly drop off towards the date-line. A box-whisker plot as in Fig. 12 also hides the fact that in large

parts of the world daily sampling errors are in magnitude at least 30 % of the daily model errors (see Fig. 14).

It is interesting to consider what temporal sampling errors remain if an annual average is constructed from daily model averages and observations, excluding those days for which no observations exist. Figure 15 shows that this does indeed reduce sampling errors compared to the case of straight annual averages (as in Fig. 8) but we still find typical root mean square errors of 12, 7 and 17 % for ECHAM-HAM, HadGEM-UKCA and MIROC-SPRINTARS. Interestingly, such a procedure results in mostly positive errors because daily minimum AOT (as modelled) occurs near the time of observation. Note also that the error patterns correlate somewhat with the diurnal cycles in these models, see Fig. 4.

Any analysis of model data attempting to look at shorter timescales (e.g. scatter plots, time series) will be further neg-

atively affected when using daily data. To prevent this, models need to be resampled to the observation times.

Figure 16 shows a comparison of the above errors for yearly, monthly and daily averages of AERONET data (AOT, AE and SSA). The overall picture is much the same as Fig. 12: model errors and observational errors decrease as averaging increases, while sampling errors are the largest for either yearly or monthly averages. Note that in the case of AERONET AOT, already for daily averages, sampling errors are larger than observational errors.

6   Summary

Although model data and observations usually have very different temporal sampling, researchers often assume this does not affect their analyses. Consequently, monthly or yearly means are compared without first temporally collocating the model data with the observations. We have assessed potential errors resulting from this practice and shown them to be significant. We would like to point out that the practice of temporal collocation of data sets is very normal in the remote-sensing and data assimilation communities, but less so in the modelling community.

Our analysis is based on the temporal sampling of several oft used remote-sensing data sets (MODIS, AERONET) and data from three different models: ECHAM-HAM, HadGEM-UKCA and MIROC-SPRINTARS. We define the temporal sampling error as the difference in yearly, monthly or daily means of the full model data set and the sampled model data sets. Three different models were used as we found that current models differ in their temporal evolution even though yearly mean global patterns are fairly similar and meteorological fields were nudged to reanalysis data. Although this study is framed in the context of model evaluation, it has a clear and direct relevance to climatologies derived from observational data sets.

We find that temporal sampling errors in yearly and monthly means can amount to 100 % in AOT, 0.4 in AE and 0.05 in SSA. In addition, these sampling errors are comparable to model errors (30–60 % for MODIS, 20–90 % for AERONET). While daily averages incur smaller sampling errors, we argue (based on observations) that our models very likely underestimate daily variation and hence sampling errors. In addition, model daily averages are defined through UTC while remote-sensing observations are only possible during local daylight. This introduces a marked longitudinal dependence in temporal sampling errors. Annual averages constructed from daily averages still exhibit temporal sampling errors although at reduced levels (we estimate typical errors of 7–17 %).

Temporal sampling errors will affect model evaluation of not only AOT, AE and SSA but also derived properties like direct aerosol radiative forcing, if data are not properly collocated with the observations. Our analysis should also provide caution to researchers using in situ data from, e.g. ground sites. While such data often have nominally high measurement frequencies (hours), significant gaps (days to weeks) in temporal coverage may be present. Much will depend on the observational sampling and the nature of the observed time series. For instance, $PM_{2.5}$ filter measurements by EMEP (European Monitoring and Evaluation Network, http://www.emep.int) are time-integrated measurements that can probably be compared to models without temporal sampling issues. But several EUSAAR (European Supersites for Atmospheric Aerosol Research, http://www.eusaar.net) stations have significant intermittency in their observations. Flight campaign data obviously require temporal collocation to make any sense in a model evaluation context. Also, unlike remotely sensed column properties like AOT, in situ measurements will be affected by vertical transport. This may possibly exacerbate any temporal sampling issues.

As a corollary to our study, we have shown that the temporal evolution of models can be very different (although yearly means are quite similar). We suggest that more effort should be made to understand modelled and observed time series. It is likely that such time series contain finger prints of individual aerosol processes and provide new ways to evaluate models. In particular we point out the large differences we found in diurnal cycles (see Fig. 4) and the research opportunities that may lie in observational data sets (e.g. from geostationary satellites) that resolve this diurnal cycle.

Finally, if there is one error in the field of aerosol modelling over which we have direct control and that can be confidently eliminated, it is the error due to incongruous temporal sampling of model data and observations. Not doing so barely saves time or resources but risks compromising any model evaluation. For instance, we have shown that the latitudinal pattern of yearly, zonal averages of collocated model AOT agrees better with MODIS observations than the normal yearly, zonal model average. The collocation of model data with observations can either be done on-line through observation simulators or off-line through sampling of high frequency model output. A generic tool to support the latter operation is freely available as the Community Intercomparison Suite (http://www.cistools.net).

References

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, 1989.

Anderson, T. E., Charlson, R. J., Winker, D. M., Ogren, J. A., and Holmen, K.: Mesoscale Variations of Tropospheric Aerosols, J. Atmos. Sci., 60, 119–136, 2003.

Angstrom, B. A.: Atmospheric turbidity , global illumination and planetary albedo of the earth, Tellus, XIV, 435–450, 1962.

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A. S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N. L., Östenson, C.-G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K. T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P. H., Meliefste, K., Wang, M., Bueno-de Mesquita, B., Sugiri, D., Krämer, U., Heinrich, J., de Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Künzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M.-Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B., and Hoek, G.: Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project, The Lancet, 6736, 1–11, doi:10.1016/S0140-6736(13)62158-3, 2013.

Bellouin, N., Mann, G. W., Woodhouse, M. T., Johnson, C., Carslaw, K. S., and Dalvi, M.: Impact of the modal aerosol scheme GLOMAP-mode on aerosol forcing in the Hadley Centre Global Environmental Model, Atmos. Chem. Phys., 13, 3027–3044, doi:10.5194/acp-13-3027-2013, 2013.

Brunekreef, B. and Holgate, S. T.: Air pollution and health., Lancet, 360, 1233–42, doi:10.1016/S0140-6736(02)11274-8, 2002.

Dockery, D., Pope, A., Xu, X., Spengler, J., Ware, J., Fay, M., Ferris, B., and Speizer, F.: An association between air pollution and mortality in six U.S. cities, New Engl. J. Med., 329, 1753–1759, 1993.

Dubovik, O. and King, M. D.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, J. Geophys. Res., 105, 20673, doi:10.1029/2000JD900282, 2000.

Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., and Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, J. Geophys. Res., 105, 9791–9806, doi:10.1029/2000JD900040, 2000.

Eck, T. F., Holben, B. N., Reid, J. S., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, J. Geophys. Res., 104, 31333–31349, 1999.

Ezzati, M., Lopez, A. D., Rodgers, A., Vander Hoorn, S., and Murray, C. J. L.: Selected major risk factors and global and regional burden of disease., Lancet, 360, 1347–1360, doi:10.1016/S0140-6736(02)11403-6, 2002.

Geogdzhayev, I., Cairns, B., Mishchenko, M. I., Tsigaridis, K., and van Noije, T.: Model-based estimation of sampling-caused uncertainty in aerosol remote sensing for climate research applications, Q. J. Roy. Meteor. Soc., 140, 2353–2363, doi:10.1002/qj.2305, 2014.

Hansen, J., Sato, M., and Ruedy, R.: Radiative forcing and climate response, J. Geophys. Res., 102, 6831–6864, 1997.

Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geosci. Model Dev., 4, 223–253, doi:10.5194/gmd-4-223-2011, 2011.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET-A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sens. Environ., 66, 1–16, 1998.

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, A., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morcrette, J.-J., Myhre, G., Penner, J., Perlwitz, J., Stier, P., Takemura, T., and Zender, C. S.: Global dust model intercomparison in AeroCom phase I, Atmos. Chem. Phys., 11, 7781–7816, doi:10.5194/acp-11-7781-2011, 2011.

Hyer, E. J., Reid, J. S., and Zhang, J.: An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals, Atmos. Meas. Tech., 4, 379–408, doi:10.5194/amt-4-379-2011, 2011.

Kaufman, Y. J., Holben, B. N., Tanre, D., Slutsker, I., Smimov, A., and Eck, T. F.: Will aerosol measurements from Terra and Aqua polar orbiting satellites represent the daily aerosol abundance and properties?, Geophys. Res. Lett., 27, 3861–3864, 2000.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T. F., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kirkevåg, A., Kloster, S., Koch, D., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Lesins, G., Liu, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., and Tie, X.: An AeroCom initial assessment – optical properties in aerosol component modules of global models, At-

mos. Chem. Phys., 6, 1815-1834, doi:10.5194/acp-6-1815-2006, 2006.

Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J. R., Balkanski, Y., Bauer, S., Berntsen, T., Bond, T. C., Boucher, O., Chin, M., Clarke, A., De Luca, N., Dentener, F., Diehl, T., Dubovik, O., Easter, R., Fahey, D. W., Feichter, J., Fillmore, D., Freitag, S., Ghan, S., Ginoux, P., Gong, S., Horowitz, L., Iversen, T., Kirkevåg, A., Klimont, Z., Kondo, Y., Krol, M., Liu, X., Miller, R., Montanaro, V., Moteki, N., Myhre, G., Penner, J. E., Perlwitz, J., Pitari, G., Reddy, S., Sahu, L., Sakamoto, H., Schuster, G., Schwarz, J. P., Seland, Ø., Stier, P., Takegawa, N., Takemura, T., Textor, C., van Aardenne, J. A., and Zhao, Y.: Evaluation of black carbon estimations in global aerosol models, Atmos. Chem. Phys., 9, 9001–9026, doi:10.5194/acp-9-9001-2009, 2009.

Koffi, B., Schulz, M., Bréon, F.-M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M., Collins, W. D., Dentener, F., Diehl, T., Easter, R., Ghan, S., Ginoux, P., Gong, S., Horowitz, L. W., Iversen, T., Kirkevåg, A., Koch, D., Krol, M., Myhre, G., Stier, P., and Takemura, T.: Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results, J. Geophys. Res., 117, D10201, doi:10.1029/2011JD016858, 2012.

Lequy, É., Conil, S., and Turpault, M.-P.: Impacts of Aeolian dust deposition on European forest sustainability: A review, Forest Ecol. Manage., 267, 240–252, doi:10.1016/j.foreco.2011.12.005, 2012.

Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J.: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, J. Geophys. Res., 112, 1–21, doi:10.1029/2006JD007811, 2007.

Lin, S.-J. and Rood, R. B.: Multidimensional flux-form semi-Lagrangian transport schemes, Mon. Weather Rev., 124, 2046–2070, 1996.

Lohmann, U. and Feichter, J.: Impact of sulfate aerosols on albedo and lifetime of clouds: A sensitivity study with the ECHAM4 GCM, J. Geophys. Res., 102, 13685–13700, 1997.

Lohmann, U. and Feichter, J.: Global indirect aerosol effects: a review, Atmos. Chem. Phys., 5, 715–737, doi:10.5194/acp-5-715-2005, 2005.

Maher, B., Prospero, J., Mackie, D., Gaiero, D., Hesse, P., and Balkanski, Y.: Global connections between aeolian dust, climate and ocean biogeochemistry at the present day and at the last glacial maximum, Earth-Sci. Rev., 99, 61–97, doi:10.1016/j.earscirev.2009.12.001, 2010.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, Geosci. Model Dev., 3, 519–551, doi:10.5194/gmd-3-519-2010, 2010.

McTainsh, G. and Strong, C.: The role of aeolian dust in ecosystems, Geomorphology, 89, 39–54, doi:10.1016/j.geomorph.2006.07.028, 2007.

Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., Gettelman, A., Lohmann, U., Bellouin, N., Boucher, O., Sayer, A. M., Thomas, G. E., McComiskey, A., Feingold, G., Hoose, C., Kristjánsson, J. E., Liu, X., Balkanski, Y., Donner, L.

J., Ginoux, P. A., Stier, P., Grandey, B., Feichter, J., Sednev, I., Bauer, S. E., Koch, D., Grainger, R. G., Kirkevåg, A., Iversen, T., Seland, Ø., Easter, R., Ghan, S. J., Rasch, P. J., Morrison, H., Lamarque, J.-F., Iacono, M. J., Kinne, S., and Schulz, M.: 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, 2009.

Remer, L., Kaufman, Y., Tanre, D., Mattoo, S., Chu, D., Martins, J., Li, R.-R., Ichoku, C., Levy, R., Kleidman, R., Eck, T., Vermote, E., and Holben, B.: The MODIS Aerosol Algorithm, Products, and Validation, J. Atmos. Sci., 62, 947–973, 2005.

Remer, L., Kaufman, Y., and Kleidman, R.: Comparison of Three Years of Terra and Aqua MODIS Aerosol Optical Thickness Over the Global Oceans, IEEE Geosci. Remote Sens. Lett., 3, 537–540, doi:10.1109/LGRS.2006.879562, 2006.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5, part I: model description, Tech. Rep. 349, Max Planck Institute for Meteorology, Hamburg, 2003.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model, J. Climate, 19, 3771–3791, 2006.

Sayer, A. M., Thomas, G. E., Palmer, P. I., and Grainger, R. G.: Some implications of sampling choices on comparisons between satellite and model aerosol optical depth fields, Atmos. Chem. Phys., 10, 10705–10716, doi:10.5194/acp-10-10705-2010, 2010.

Schmid, B., Michalsky, J., Halthore, R., Beauharnois, M., Harnson, L., Livingston, J., Russell, P., Holben, B., Eck, T., and Smirnov, A.: Comparison of Aerosol Optical Depth from Four Solar Radiometers During the Fall 1997 ARM Intensive Observation Period, Geophys. Res. Lett., 26, 2725–2728, 1999.

Schulz, M., Textor, C., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Dentener, F., Guibert, S., Isaksen, I. S. A., Iversen, T., Koch, D., Kirkevåg, A., Liu, X., Montanaro, V., Myhre, G., Penner, J. E., Pitari, G., Reddy, S., Seland, Ø., Stier, P., and Takemura, T.: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations, Atmos. Chem. Phys., 6, 5225–5246, doi:10.5194/acp-6-5225-2006, 2006.

Schutgens, N. A. J., Nakata, M., and Nakajima, T.: Validation and empirical correction of MODIS AOT and AE over ocean, Atmos. Meas. Tech., 6, 2455–2475, doi:10.5194/amt-6-2455-2013, 2013.

Shi, Y., Zhang, J., Reid, J. S., Holben, B., Hyer, E. J., and Curtis, C.: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation, Atmos. Chem. Phys., 11, 557–565, doi:10.5194/acp-11-557-2011, 2011.

Smirnov, A.: Diurnal variability of aerosol optical depth observed at AERONET (Aerosol Robotic Network) sites, Geophys. Res. Lett., 29, 2115, doi:10.1029/2002GL016305, 2002.

Smith, K. R., Jerrett, M., Anderson, H. R., Burnett, R. T., Stone, V., Derwent, R., Atkinson, R. W., Cohen, A., Shonkoff, S. B., Krewski, D., Pope, C. A., Thun, M. J., and Thurston, G.: Public health benefits of strategies to reduce greenhouse-gas emissions:

health implications of short-lived greenhouse pollutants., Lancet, 374, 2091–103, doi:10.1016/S0140-6736(09)61716-5, 2009.

Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., and Petzold, A.: The aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 5, 1125-1156, doi:10.5194/acp-5-1125-2005, 2005.

Stier, P., Seinfeld, J. H., Kinne, S., and Boucher, O.: Aerosol absorption and radiative forcing, Atmos. Chem. Phys., 7, 5237–5261, doi:10.5194/acp-7-5237-2007, 2007.

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

Takemura, T., Okamoto, H., Maruyama, Y., Numaguti, A., Higurashi, A., and Nakajima, T.: Global three-dimensional simulation of aerosol optical thickness distribution of various origins, J. Geophys. Res., 105, 17853, doi:10.1029/2000JD900265, 2000.

Takemura, T., Nakajima, T., Dubovik, O., Holben, B. N., and Kinne, S.: Single-Scattering Albedo and Radiative Forcing of Various Aerosol Species with a Global Three-Dimensional Model, J. Climate, 15, 333–352, 2002.

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

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Feichter, H., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I., Iversen, I., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: 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, 2006.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Feichter, J., Fillmore, D., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I. S. A., Iversen, T., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J. E., Pitari, G., Reddy, M. S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: The effect of harmonized emissions on aerosol properties in global models – an AeroCom experiment, Atmos. Chem. Phys., 7, 4489–4501, doi:10.5194/acp-7-4489-2007, 2007.

Twomey, S.: Pollution and the planetary albedo, Atmos. Environ., 8, 1251–1256, 1974.

Vignati, E., Wilson, J., and Stier, P.: M7: An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models, J. Geophys. Res., 109, D22202, doi:10.1029/2003JD004485, 2004.

Vink, S. and Measures, C.: The role of dust deposition in determining surface water distributions of Al and Fe in the South West Atlantic, Deep Sea Res. Pt. II, 48, 2787–2809, doi:10.1016/S0967-0645(01)00018-2, 2001.

Zhang, J. and Reid, J. S.: MODIS aerosol product analysis for data assimilation: Assessment of over-ocean level 2 aerosol optical thickness retrievals, J. Geophys. Res., 111, D22207, doi:10.1029/2005JD006898, 2006.

Zhang, K., O'Donnell, D., Kazil, J., Stier, P., Kinne, S., Lohmann, U., Ferrachat, S., Croft, B., Quaas, J., Wan, H., Rast, S., and Feichter, J.: The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations, Atmos. Chem. Phys., 12, 8911–8949, doi:10.5194/acp-12-8911-2012, 2012.