Different Approaches for Constraining Global Climate Models of the Anthropogenic Indirect Aerosol Effect

BY U. LOHMANN, J. QUAAS, S. KINNE, AND J. FEICHTER

Strategies to detect and attribute aerosol global impacts on clouds and climate from synergetic approaches involving modeling and observational evidence at different spatial and temporal scales.

Aerosol particles resulting from human activity such as sulfate and carbonaceous aerosols have substantially increased the global mean aerosol burden since preindustrial times. Aerosol particles can affect the climate system via several mechanisms. The most prominent impacts are 1) the reflection of solar radiation back to space (a "direct" effect), 2) the absorption of solar radiation by soot and mineral dust to warm the atmospheric aerosol layer, which could hinder cloud formation and/or cause cloud droplets to evaporate (a "semi-direct" effect), and 3) the capability to act as condensation nuclei for (water and ice) clouds ("indirect effects"). The last effect, which is expected to increase the solar reflection of (water) clouds, is often distinguished into a cloud albedo and a cloud lifetime effect. The cloud albedo effect captures the process by which polluted clouds with more but smaller droplets appear brighter (Twomey 1959), whereas the lifetime effect considers that polluted clouds with more but smaller droplets reduce the likelihood for cloud droplets to grow to raindrop size, thereby extending the cloud lifetime (Albrecht 1989). Modeling results suggest that these indirect effects are more important than the direct and semi-direct. Still despite many efforts large uncertainties remain for all simulated aerosol indirect effects (Penner et al. 2001; Ramaswamy et al. 2001).

Anderson et al. (2003) identified for the climatic impact of anthropogenic aerosols a discrepancy between climate model simulations and estimates from inverse models. Inverse models are conceptual models that derive the impact of aerosols by subtracting all other better-quantified anthropogenic impacts from observed changes of surface temperatures where also the heat storage in oceans is considered (Knutti et al. 2002; Forest et al. 2002; Crutzen and Ramanathan 2003). Inverse methods constrain the aerosol impact on the energy balance at the top of the atmosphere (TOA) to about \(-1\) W m\(^{-2}\), with an uncertainty range from 0 to \(-1.9\) W m\(^{-2}\) (Anderson et al. 2003). In contrast, climate model simulations suggest a TOA forcing centered around \(-1.5\) W m\(^{-2}\), and an uncertainty range that
extends beyond $-2.5 \text{ W m}^{-2}$ (Anderson et al. 2003; Lohmann and Feichter 2005).

Thorough validation of aerosol–cloud interactions with observational data is missing in all climate model simulations of anthropogenic aerosol effects on clouds. There are some physical arguments, in particular related to the treatment of the indirect effects, as to why estimates of the aerosol impact by current general circulation model (GCM) parameterizations suggest a stronger-than-expected cooling. There are a number of offsetting effect that are poorly, if at all, considered in global modeling: 1) the semi-direct effect as mentioned above (Gräbl 1979; Hansen et al. 1997); 2) the dispersion effect, which considers that the shape of the cloud droplet spectra is broader in polluted conditions [when the growth of the majority of the newly activated smaller aerosols is retarded (Liu and Daum 2002)], is believed to reduce both the albedo effect by 0.2–0.5 W m$^{-2}$ (Peng and Lohmann 2003; Rotstyn and Liu 2003) and the cloud lifetime effect by a similar amount (Rotstyn and Liu 2005); 3) anthropogenic aerosols can impact the microphysics of mixed-phase and ice clouds by acting as ice nuclei thus accelerating ice-crystal-induced precipitation, which would reduce cloud cover and lifetime (Lohmann 2002); 4) the more and smaller cloud droplets can also reduce the collision rate of snow crystals with cloud droplets (the riming process) in mixed-phase clouds (Borys et al. 2000, 2003) but nevertheless increase in the snowfall rate (Lohmann 2004); and 5) more and smaller cloud droplets resulting from increasing aerosol concentrations also suppress low-level rainout and aerosol washout in convective clouds. This allows transport of water and smoke to upper levels, thus elevating the onset of precipitation and the release the latent heat, which would result in more intensive convection (Rosenfeld and Woodley 2000; Andreae et al. 2004).

Most of these aerosol indirect effects have been deduced from the analysis of in situ observations and satellite data on local scales. However, to be useful for global modeling with its coarse horizontal and vertical resolution, parameterizations must be developed that capture the essence of all aerosol–cloud interactions. And it must be demonstrated with available observational data that the developed parameterizations apply globally.

In "Methods to derive aerosol–cloud parameterizations," we outline four different pathways to parameterizations. In "Constraints of parameterizations with large-scale observations," methods to constrain parameterizations in global models are presented, and in "Recommendations," we outline how those methods can be further developed in the future.

**METHODS TO DERIVE AEROSOL–CLOUD PARAMETERIZATIONS.** Processes that act on spatial scales smaller than the horizontal ($200 \times 200 \text{ km}^2$) and vertical resolution ($-20 \text{ tropospheric layers}$) and are shorter than the typical time step in global modeling ($-20 \text{ minutes}$), must be parameterized. It basically means to express variability on smaller scales in terms of large-scale variables. The concepts of the development of parameterizations can be divided into the four different methods outlined below.

**Derivation from first principles.** If a process has an analytical solution, then this method should be used. Even if analytical solutions cannot be obtained directly, they often can be approximated with some simplifications. An example is the cloud droplet activation process, where particle size as a function of relative humidity is described by the Köhler curve for equilibrium conditions. Ghan et al. (1993), Abdul-Razzaq and Ghan (2000), and Nenes and Seinfeld (2003) used the Köhler curve as the starting point for a parameterization of cloud droplet nucleation and assumed that the condensation rate can be related to the dry aerosol number. This way droplet formation can be parameterized as a function of total aerosol number, vertical velocity, and an activation parameter. Likewise, homogeneous ice crystal nucleation can be derived from theory if differences between the supersaturation at which freezing commences and the maximum supersaturation are neglected (Kärcher and Lohmann 2002).

Along the same lines the conversion of cloud droplets to form raindrops, the autoconversion process, can be deduced analytically as discussed by Liu et al. (2004), Liu and Daum (2004), and Liu et al. (2006). The autoconversion parameterizations traditionally require a threshold cloud droplet size, above which the conversion to raindrops takes place. Instead of an arbitrary threshold, Liu et al. (2004) analytically derived an expression for this.

**Derivation from laboratory studies.** Laboratory data have been utilized in particular to study ice crystal formation. For instance, the effectiveness of mineral dust particles and/or black carbon particles to initiate contact or immersion freezing as a function of temperature is derived from a compilation of laboratory data (Diehl and Wurzler 2004; Diehl et al. 2006). These relationships have been applied to GCMs to study the importance of anthropogenic soot aerosols versus mineral dust aerosols to serve as ice nuclei (Lohmann and Diehl 2006). This approach is promising because the relationship among variables of interest is studied...
in isolation. The derived parameterizations should be valid in the atmosphere as long as the heterogeneity and larger scales in real clouds do not influence these relationships. Thus, this method is the preferred method if no analytical solution is available.

**Derivation from focused measurement campaigns.** Field data at various continental and marine sites during cleaner and more polluted events have been summarized to derive robust relationships between (submicrometer size) aerosol number concentration (sometimes via the sulfate mass) and cloud droplet number concentration (Boucher and Lohmann 1995; Gultepe and Isaac 1996; Lin and Leaitch 1997). This concept was later extended for organic carbon and sea salt (Menon et al. 2002). These compiled datasets should be able to represent the spatial and temporal variability of droplet number concentration within a grid box of the model. Thus, they include all influences on the process in question, including those that we do not know. This is an advantage and a risk at the same time. In this regard, this method complements the laboratory method for processes that are more complex than can be studied in a laboratory setting. However, the sample size in a field experiment is normally not large enough to stratify these empirical data according to all influences in question. For example, empirical aerosol mass–cloud droplet number concentration relationships are limited with respect to the number and mixture of aerosol species that they take into account. Moreover, compilations of different datasets are complicated once different instruments are used for observing the same quantity.

**Derivation from models with finer resolution.** The cloud lifetime effect depends critically on the parameterization of the precipitation formation mechanism in warm clouds (Lohmann and Feichter 1997; Menon et al. 2002). These processes occur on much smaller scales. Thus, the growth of cloud droplets to precipitation-sized particles can be described by a stochastic (collection equation) relationship, as long as models keep track of different cloud droplet sizes. Following that idea, simulations of droplet growth with a cloud microphysical model have led to global model-suited parameterizations for the collision–coalescence process, the autoconversion rate, and accretion rate (Khairoutdinov and Kogan 2000). This method is clearly not ideal, because it has no observational database. It is the last resort for a parameterization in the absence of any analytical solution or observations of the process in question. It is only useful if the GCM with different autoconversion schemes is validated against satellite data (Suzuki et al. 2004) so that the best scheme can be identified.

**CONSTRAINTS OF PARAMETERIZATIONS WITH LARGE-SCALE OBSERVATIONS.** All methods to derive parameterizations outlined in the previous chapter have in common that they are based on theory or measurements valid for small scales and/or for specific situations. In a GCM the parameterizations need to be applied for any part of the globe, for any climate state, and for the model’s coarse spatial and temporal resolution. A widely used methodology to adapt parameterizations to changing environments and scales is to adjust (or "tune") parameters to get a more "realistic" match to available local data (e.g., Rotstayn 2000). A better method would be to use observational data at adequate scales to infer these parameters. The use of information from satellite retrievals is particularly appealing, because they provide detailed cloud, aerosol, and radiation measurements at the large horizontal and temporal scales needed to evaluate GCMs. The resolution of the satellite goes down to scales able to distinguish between individual cloud fields and cloud-free regions, which are of subgrid scale for GCMs. There are, however, deficiencies in the use of satellite data in terms of temporal and vertical resolution, at least up to now.

**Constraint with statistical relationships.** Satellite-derived relationships can provide clues about the way specific parameterizations should work on the scales relevant in large-scale modeling. The advantage here is that statistical correlations from satellites are temporally and spatially more robust than individual measurements. This allows derivation of a desired quantity (e.g., cloud droplet number concentration) from predicted quantities (e.g., aerosol mass concentration). Correlations can also identify the overall effect in case that different effects with different signs and magnitudes are involved. Moreover, correlations analyze relative changes and therefore limitations to the absolute accuracy are acceptable. Relationships are supposed to be valid also in a changing climate, whereas absolute values and currently measured distributions are not. For example, based on POLDER satellite data it was found that the negative relationship between cloud droplet size and aerosol concentration is overestimated in global modeling (Lohmann and Lesins 2002). Based on similar findings, satellite-derived relationships were used to revise existing parameterizations of cloud droplet activation from in

---

1 The acronyms are explained in the appendix.
situ measurements used in GCMs (Quaas and Boucher 2005; Quaas et al. 2006).

**Constraint using the data assimilation technique.** The assimilation of observational data in global modeling is one approach to tie modeling to observations. Here, model parameters are relaxed toward observed values. Data assimilation can provide information of model deficiencies and measurement uncertainties especially in data-rich areas. Assimilated climate models that require large model adjustments in certain areas or under certain conditions offer obvious clues to poor representations of the adjusted model quantity. In turn, model output of data assimilations can extend temporally or spatially sparse measurements. There is a European initiative GEMS that will follow this avenue as well as the U.S. initiative PARAGON (Kahn et al. 2004). Moreover, this approach allows testing of parameterizations that were developed from the methods introduced in the previous chapter. Since the meteorology is prescribed for a particular time, testing could be as detailed as in numerical weather prediction. This enables model evaluations around a specific event or at locations of dedicated measurement campaigns. Even then, we still face data issues because measurements 1) are usually unable to match the detail in modeling, 2) suffer from accuracy limitations, and 3) may not be applicable at the temporal and spatial scales of modeling.

**TABLE I.** Physical parameters related to aerosols, clouds, and dynamics needed to evaluate parameterizations of aerosol indirect effects.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Typical value or range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_s )</td>
<td>Aerosol number concentration</td>
<td>( 10^6 ) to ( 10^{10} ) m(^{-3} )</td>
</tr>
<tr>
<td>SD(_s)</td>
<td>Aerosol size distribution</td>
<td>Lognormal</td>
</tr>
<tr>
<td>SF(_s)</td>
<td>Soluble fraction of aerosol population</td>
<td>0–1</td>
</tr>
<tr>
<td>AT</td>
<td>Aerosol type</td>
<td>—</td>
</tr>
<tr>
<td>SSA</td>
<td>Single-scattering albedo</td>
<td>0.6–0.99</td>
</tr>
<tr>
<td>BC(z)</td>
<td>Position of black carbon w.r.t. the cloud</td>
<td>Above/in/below</td>
</tr>
<tr>
<td>CC</td>
<td>Cloud cover</td>
<td>0–1</td>
</tr>
<tr>
<td>( N_d )</td>
<td>Cloud droplet number concentration</td>
<td>( 10^3 ) to ( 10^6 ) m(^{-3} )</td>
</tr>
<tr>
<td>( N_i )</td>
<td>Ice crystal number concentration</td>
<td>( 10^3 ) to ( 10^4 ) m(^{-3} )</td>
</tr>
<tr>
<td>SD(_d)</td>
<td>Cloud droplet size distribution</td>
<td>Lognormal or gamma distribution</td>
</tr>
<tr>
<td>SD(_i)</td>
<td>Ice crystal size distribution</td>
<td>Lognormal or gamma distribution</td>
</tr>
<tr>
<td>LWC</td>
<td>Liquid water content</td>
<td>0–10(^1) kg m(^{-3} )</td>
</tr>
<tr>
<td>IWC</td>
<td>Ice water content</td>
<td>0–10(^3) kg m(^{-3} )</td>
</tr>
<tr>
<td>AU</td>
<td>Autoconversion rate</td>
<td>0–10(^{-4}) kg kg(^{-1}) s(^{-1})</td>
</tr>
<tr>
<td>( \omega )</td>
<td>Vertical wind speed</td>
<td>-10 to 10 m s(^{-1})</td>
</tr>
<tr>
<td>RH</td>
<td>Relative humidity</td>
<td>20%–100% (up to 170% w.r.t. ice)</td>
</tr>
<tr>
<td>TKE</td>
<td>Turbulent kinetic energy</td>
<td>0–10 m(^2) s(^{-1})</td>
</tr>
<tr>
<td>T</td>
<td>Temperature</td>
<td>210–273 K</td>
</tr>
</tbody>
</table>

**RECOMMENDATIONS.** In order to constrain global models of the aerosol indirect effect, accurate and long-time observations for all relevant (aerosol, cloud, and environmental) quantities are required at adequate spatial and temporal resolutions (Tables 1 and 2). We should be able at least to address seasonal and interannual variability. Thus, at a minimum data are needed for one complete annual cycle and if possible on a global or quasi-global (e.g., through intercalibrated ground networks such as AERONET) scale to establish links in space. Even longer records are certainly desirable given the year-to-year variability of aerosol (e.g., biomass burning) and dynamics (e.g., El Niño, North Atlantic Oscillation), especially so if data originate from the same instrument (assuming that no instrument response changes and orbital changes occur).

Observations should discriminate the cloud type (shallow and deep convective, stratiform), cloud phase (water, ice, mixed phase), season, region (Arctic, midlatitude, subtropic, and tropic), and underlying surface (ocean, coastal, and continental). In terms of observational evidence on aerosol–cloud interactions, we currently benefit from the increased capabilities of passive remote sensing (e.g., MODIS, MISR, PARASOL) and the establishment of passive and active remote sensing networks at surface sites (e.g., AERONET, BSRN, EARLINET). All of these measurements need to be combined in order to retrieve the whole suite of aerosol and cloud parameters that are relevant for the indirect aerosol effect. Instead of just comparing individual models with observations, Kinne et al. (2003, 2006) pointed out the strength of investigating an ensemble of models at the same time not only to determine a most likely modeling value (model median) or regional model diversity, but also to identify common biases in large-scale modeling.

More reliable correlations between aerosol, cloud, and environmental atmospheric properties will be possible as available multi-year datasets from the newer generation of satellite sensors are analyzed. Multispectral (MODIS, MERIS), multi-
Table 2. Attribution of the parameters listed in Table 1 to the evaluation of the different aerosol indirect effect mechanisms, including the required data resolution.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Aerosols</th>
<th>Clouds</th>
<th>Large-scale environment</th>
<th>Required resolution (x, z, t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud albedo effect</td>
<td>$N$, $SD$, $SF$, $AT$</td>
<td>$N$, $SD$, $LWC$, $CC$</td>
<td>w, RH</td>
<td>1 km, 100 m, 1 h</td>
</tr>
<tr>
<td>Cloud lifetime effect</td>
<td>—</td>
<td>$N$, $SD$, $LWC$, $AU$, $CC$</td>
<td>TKE</td>
<td>1 km, 100 m, 1 h</td>
</tr>
<tr>
<td>Semi-direct effect</td>
<td>$N$, $SSA$, $BC(z)$</td>
<td>$LWC$, $CC$</td>
<td>RH</td>
<td>10 km, 1 km, 6 h</td>
</tr>
<tr>
<td>Aerosols effects on mixed-phased and ice clouds</td>
<td>$N$, $SD$, $SF$, $AT$</td>
<td>$N$, $SD$, $IWC$, $CC$</td>
<td>w, RH, T</td>
<td>1 km, 100 m, 1 h</td>
</tr>
</tbody>
</table>

Table 3. Advantages and disadvantages of the different satellite observing concepts.

<table>
<thead>
<tr>
<th>Satellite orbits</th>
<th>LEO(^a)</th>
<th>GEO(^b)</th>
<th>L1(^c)/L2(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude above ground</td>
<td>Low</td>
<td>Mid</td>
<td>High</td>
</tr>
<tr>
<td>Global coverage</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Temporal resolution</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Overpass time</td>
<td>Specific time</td>
<td>Always</td>
<td>Day (L1), night (L2)</td>
</tr>
<tr>
<td>High latitudes</td>
<td>Good</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Examples</td>
<td>MODIS</td>
<td>GOES</td>
<td>TRIANA (planned)</td>
</tr>
</tbody>
</table>

\(^a\)LEO: Low-earth orbit (e.g., polar orbiting—crossing up and down the equator).
\(^b\)GEO: Geostationary earth orbit (placed at a fixed longitude over the equator).
\(^c\)L1: Placed in the equal attraction point of earth and sun (between Earth and sun).
\(^d\)L2: Placed in the equal attraction point of Earth and sun (behind the Earth). At the Lagrangian points L1 or L2 a satellite can "move with the sun."
ACKNOWLEDGMENTS. This publication resulted from discussions held during the January 2005 Specialty Conference on the Indirect Effect of Aerosols on Climate in Manchester, United Kingdom, which was jointly sponsored by the International Global Atmospheric Chemistry Project [IGAC; a Core Project of the International Geosphere Biosphere Programme (IGBP)] and by the U.S. NOAA and NASA programs. Financial support was also generously provided by the European Network of Excellence Atmospheric Composition Change (ACCENT).

APPENDIX: LIST OF ACRONYMS.

AATSR: Advanced Along Track Scanning Radiometer
AERONET: Aerosol Robotic Network
BSRN: Baseline Surface Radiation Network
CALIPSO: Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations
EARLINET: European Aerosol Research Lidar Network
GEMS: Global Earth System Monitoring using Space and in situ Data
GOES: Geostationary Operational Environmental Satellite
MERIS: Medium Resolution Imaging Spectrometer
MISR: Multi-angle Imaging Spectro-Radiometer
MODIS: Moderate Resolution Imaging Spectroradiometer
MSG: Meteosat Second Generation
PARAGON: Progressive Aerosol Retrieval and Assimilation Global Observing Network
PARASOL: Polarization and Anisotropy of Reflectances for Atmospheric Sciences Coupled with Observations from a Lidar
POLDER: Polarization and Directionality of the Earth’s Reflectances
SEVIRI: Spinning Enhanced Visible and Infrared Imager

REFERENCES


Copyright of Bulletin of the American Meteorological Society is the property of American Meteorological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.