

Cloud-controlling Factors

A. Pier Siebesma, Rapporteur

Jean-Louis Brenguier, Christopher S. Bretherton,
Wojciech W. Grabowski, Jost Heintzenberg,
Bernd Kärcher, Katrin Lehmann, Jon C. Petch,
Peter Spichtinger, Bjorn Stevens, and Frank Stratmann

Introduction

To address the multitude of issues that contribute to a thorough understanding of clouds, science needs to consider and incorporate a wide variety of biological, physical, and chemical processes that are embedded in flows. A host of processes emerge from the categories of atmospheric dynamics—kinetics, diffusion, and turbulence—and span scales from the millimeter, or smaller, scale to the planetary dimension (see Figure 12.1). The development of a simulation system, or theory, that would incorporate the full range of all relevant cloud-controlling factors does not appear to us to be foreseeable, and it is unlikely that we will ever have a reference model that incorporates all relevant processes and their interactions.

Nonetheless, a variety of models or theories continue to be developed in an attempt to describe specific sets of processes over a subset of relevant scales. These models and theories, as well as the connections between them, form an informal network or hierarchy. We refer to this network as being informal because the connections between models cannot be rigorously justified; we term it hierarchical in the sense that the flow of information through the hierarchy tends to have an aggregative character, wherein one attempts to represent the statistical behavior of clouds on small scales in larger-scale models.

The informal nature of the network raises two important issues: How can formalism (i.e., the connections) be increased among processes and models at different scales? How should the network be justified in the context of each problem one wishes to solve, in the present case perturbed clouds.

By “perturbed clouds,” we mean clouds that are perturbed either by anthropogenic aerosols, which initiate such perturbations at the smallest spatial

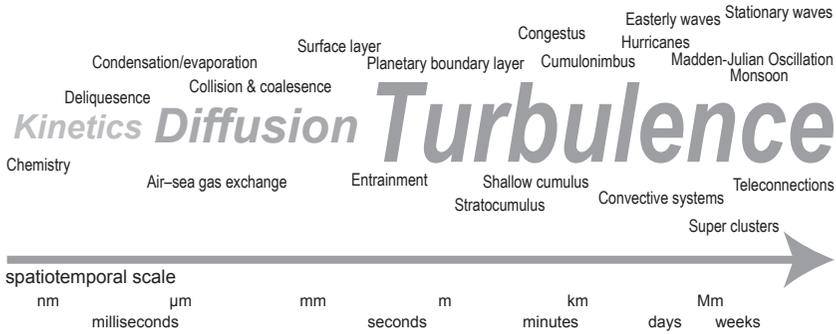


Figure 12.1 Depiction of the continuum of relevant cloud-related processes across the full range of spatiotemporal scales, showing the underlying categorical behavior (gray text) from which processes emerge.

scales (i.e., the microscale), or through global warming attributable to additional greenhouse gases. These perturbed clouds include possibly all tropospheric clouds, including contrails which are a direct consequence of anthropogenic influences, and may include polar stratospheric clouds as well.

What are the practical considerations of this network? For individual models that can encompass at most three orders of magnitude of spatial (horizontal) scales, we need to rethink carefully how to use these in the network. On the smallest scale, direct numerical simulations (DNS) are used, ranging from the millimeter scale to the 1-meter scale. Large eddy simulations (LES) resolve partly the turbulence and the individual clouds up to a scale of 100 km but need to parameterize microphysics and small-scale turbulence. At the next level, cloud-resolving models (CRMs) address cloud clusters up to scales of 1000 km, but these require additional parameterized turbulence and clouds for more coarse CRMs in the boundary layer (see Grabowski and Petch, this volume). On the largest scale, general circulation models (GCMs) are employed. They have the obvious advantage of not requiring lateral boundary conditions; however, as they operate at resolutions of typically 100 km, they fail to represent explicitly most of the cloud-controlling factors and essential cloud processes. Perhaps surprisingly, the effects of perturbed clouds in our climate system have been primarily studied by using these GCMs. One important theme that emerged from our discussions is the need to develop a more optimal use of this hierarchical network of models to answer questions on the representation of clouds in our climate system. This is a difficult task, as we do not yet understand completely how the various interactions propagate across the many scales.

Similar arguments hold for experimental research, which ranges from small-scale laboratory experiments that address microphysical issues to large-scale field experiments and global satellite observations.

Connecting Scales in Global Climate Models

Traditional Pathways and Shortcomings

GCMs have been the major modeling tool used to address the issue of perturbed clouds, despite the fact that many relevant cloud processes are not explicitly resolved by them. In our opinion, such models are useful for specific aspects that they resolve (i.e., the large-scale dynamics), but care must be taken when interpreting the impact of cloud-scale physical, chemical, and dynamic processes in these models, as they are represented imperfectly. Figure 12.2 illustrates how the various scales are connected in most of the state-of-the-art GCMs today: The box on the right represents all of the resolved processes on a mean grid box state (scales of typically 100–500 km); on the left, boxes depict the unresolved processes relevant for clouds, including turbulent processes, moist convection, cloud microphysics, and radiation. Traditionally, impacts of these processes are modeled in a statistical manner in terms of the grid box-averaged fields: temperature, T , specific humidity, q_v , and the velocity, \mathbf{v} (u , v , w). The impact of these statistical models, or parameterizations, are directly communicated with the grid box mean state, as indicated in Figure 12.2.

With this approach, we face at least four categories of errors and uncertainties. The first relates to our inherent lack of knowledge of certain physical processes. A prime example is our lack of understanding of microphysics for mixed-phase and, in part, ice clouds; they are challenging to simulate irrespective of resolution. For these types of processes, further experimental research is especially required (discussed in detail below). In addition to new experimental research, we require new theoretical concepts (e.g., a consistent framework for explicit ice supersaturation in GCMs).

The second category of errors relates to processes that we understand at some level of aggregation, but for which there is no clear way to parameterize these in GCMs, mainly because of the nonlinear character of these processes. Examples include radiative flux transfer functions $R(q_p, q_v, T)$, which describe the vertical radiative fluxes as a function of the atmospheric conditions, and

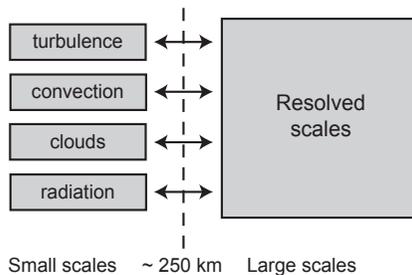


Figure 12.2 Depiction of the interaction between resolved and parameterized unresolved cloud-related processes (convection, turbulence, clouds and radiation) in present-day climate models.

to a certain extent warm-phase microphysical processes. In terms of the latter, consider a cloud-scale adjustment scheme for cloud liquid water, q_l , and cloud fraction, c , in which it is assumed that condensation occurs if supersaturation occurs and cloud droplets coalesce into rain drops at a rate A , where for the sake of argument we take A following Kessler (1969):

$$\begin{aligned} c &= H(q_t - q_s), \quad q_l = (q_t - q_s)H(q_t - q_s), \\ A &= K(q_l - q_{cr})H(q_l - q_{cr}), \end{aligned} \quad (12.1)$$

where H is the Heaviside step function, q_s is the saturation specific humidity, $q_t = q_v + q_l$ is the total water-specific humidity, K an inverse timescale, and q_{cr} a critical threshold of the liquid water below which there is no conversion to rain. In a GCM, these processes must be represented by averaging over a family of clouds; however, typically, the above relationships are simply implemented in terms of grid-scale properties, which causes biased errors because the underlying relationships are nonlinear:

$$\begin{aligned} \overline{c(q_t, q_s)} &\geq c(\overline{q_t}, \overline{q_s}), \\ \overline{q_l(q_t, q_s)} &\geq q_l(\overline{q_t}, \overline{q_s}), \\ \overline{A(q_l)} &\geq A(\overline{q_l}), \end{aligned} \quad (12.2)$$

where overbars denote grid box averages. In a similar fashion, applying the grid box-averaged liquid water in radiative transfer calculations leads to a systematic overestimation of the albedo α (i.e., $\overline{\alpha(q_l)} \leq \alpha(\overline{q_l})$, the so-called plane-parallel cloud albedo bias). For calculations capable of representing the scales at which the physical relationships are valid (in our example, the cloud scale), these biases are negligible, which partly explains the success, if not authority, of fine-scale calculations.

The third category of errors relates to the fact that current practice largely requires various subgrid processes to interact through modifications to the mean state, as depicted in Figure 12.2. In reality, however, virtually all relevant subgrid processes displayed in Figure 12.2 are related to the same joint probability density function (PDF) of vertical velocity, w , q_p , and the liquid water (or ice) potential temperature θ_l . Vertical transport through turbulence, unresolved mesoscale gravity waves, and convection creates subgrid variability in temperature, moisture, and vertical velocity. This subgrid variability determines cloud amount and cloud condensate. It is also these subgrid updraft velocities that influence the cloud particle number concentrations through the cloud condensation nuclei (CCN), droplet freezing, or ice nuclei (IN) activation (ice-forming nuclei constitute the subset of aerosol particles on which ice forms). Conversely, cloud amounts influence the vertical stability and therefore the vertical transport of moisture and heat. Finally, the subgrid variability of cloud amounts influences radiative heating profiles. Thus, in reality, all of

the subgrid processes interact with each other, and this is not usually taken into account by GCMs, which leads to biases and inconsistencies.

A last category of errors is associated with the mean field representation of subgrid processes, whereby there is a one-to-one correspondence between the subgrid processes and the local (in space and time) resolved state. Although such a quasi-equilibrium assumption might be defensible in situations in which the grid box size is much larger than the typical size of the subgrid process of interest, or timescales are long compared to the timescale of the process being represented, the assumption breaks down if the spatiotemporal size of a grid box becomes comparable to the size of the subgrid process of interest. In that case, resolution is still too coarse to resolve the process of interest; however, the grid box is also too small to allow a statistical parameterization that assumes quasi-equilibrium. An interesting example is the convective mass flux, indicating convective activity within a grid box, which can be reasonably represented by a single value for grid box sizes of several hundred kilometers and a range of possible values for a given mean state if the grid box has a typical size of tens of kilometers. This observation motivates a *stochastic* approach rather than a *mean field* representation.

New Pathways

The most direct, but also computationally most demanding alternative is to use global CRMs (see Figure 12.3a). With this class of models, it is now possible to achieve a resolution of 1 to ~5 km for time integrations of a month (Collins and Satoh, this volume). This has the obvious advantage of allowing processes beyond this scale to be represented, in principle, in a more realistic, consistent way. High computational costs, however, prohibit long (years to many decades) or repeated simulations.

A computationally less demanding, but still challenging alternative involves cloud-resolving convective parameterization or super-parameterization (see Grabowski and Petch, this volume), where a two-dimensional CRM is embedded in each grid box of a conventional GCM (see Figure 12.3b). Unlike traditional nesting techniques, this approach uses the fine-scale model to parameterize the fine-scale physics rather than span the space distended by the grid box itself. The computation advantage over the previous approach rests on the degree to which the fine-scale model leaves out some of these intermediate scales. This has the advantage of allowing cloud processes to interact on the more appropriate scale of a CRM grid point, while global circulation feedbacks are still being simulated. The disadvantage is that mesoscale processes are often poorly represented as the intermediate scales are missing. With current computational resources, super-parameterized GCMs are becoming practical for five- to ten-year integrations, thus providing an alternative framework to explore the global effect of greenhouse gas and aerosol perturbations on clouds.

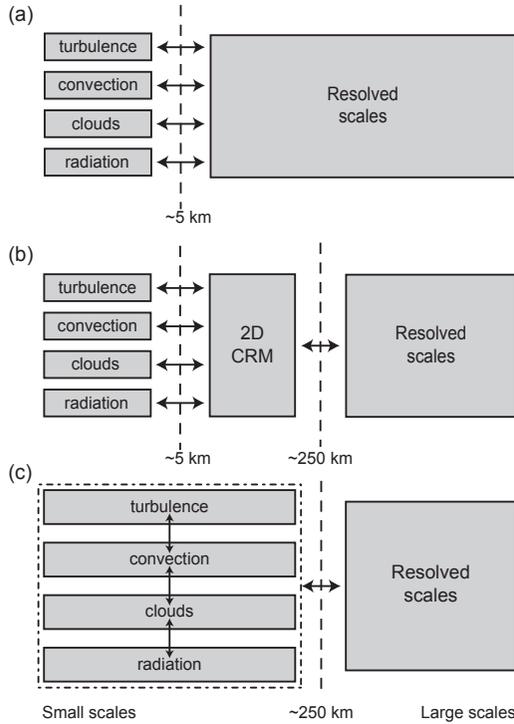


Figure 12.3 New pathways to improve the representation of cloud-related processes for the next generation of climate models: (a) global CRMs, (b) cloud-resolving convective parameterization or super-parameterization and (c) interactive parameterizations that communicate with each other directly rather than through the mean resolved state.

For both methods, clouds must be adequately resolved to be realistic. A 4-km horizontal resolution (as currently used) is still too coarse to resolve deep convection properly, let alone to represent boundary layer clouds, which are believed to be the most sensitive cloud types for cloud–climate feedbacks. Therefore, boundary clouds, turbulence, and microphysics require parameterization, and biases associated with the poor resolution of deep convection should be expected. Hence, we need to invest computational resources in even higher resolution studies (not forgetting the all important role of vertical resolution) to gain understanding of the implications of running these global CRMs and super-parametrizations at these coarse resolutions. Limited area tests should be conducted to compare runs using 4-km grid lengths with reference runs using mesh sizes well below 100 m for different cloud regimes. An analysis of such simulations should focus additionally on the representation of convection and cloud processes.

A third alternative is to develop more consistent and sophisticated parameterizations for clouds, turbulence, convection, and microphysics. If the joint PDF $P(q_p, \theta_1, w)$ of the vertical velocity w , the total water specific humidity, q_p ,

and the liquid (ice) water potential temperature, θ_l , for a given grid box could be estimated, a consistent treatment of convective and turbulent transport, cloud properties and their interaction with radiation can be inferred. Aerosol and microphysical processing could also be addressed in this framework. Various different pathways for these PDF-based schemes can be explored in which extra subgrid information is obtained through extra prognostic equations of higher-order moments or of cloud variables, through multiple updraft equations as well as through the use of an assumed shape of the PDF. This approach is proving successful for cloudy boundary layer processes, as it addresses what we referred to above as the second (nonlinear processes that can generate biased errors) and third category (a consistent treatment of the relevant processes) errors. However, it is computationally demanding because of the large number of prognostic equations and the short timestep needed for numerical stability, and it requires numerous closure assumptions.

Finally, stochastic parameterizations that provide a framework to account for deviations from the mean field prediction may be necessary, if not convenient, especially for GCMs that use resolutions finer than 100 km. In such a stochastic approach, the constraint of having only a one to one correspondence between the mean state and the subgrid response is relaxed. As a result, different subgrid-scale responses can occur that are selected stochastically, leading in general to an increased variability of the climate system.

Use of Process Models to Improve Cloud Representation in Climate Models

For any of the pathways described above, we must strive to improve our understanding of the cloud processes that remain unresolved in GCMs over the coming decades. By evaluating GCMs with global observational datasets, weak points in the representation of certain cloud regimes can be identified. Subsequently, by making use of theory and fine-scale flow solvers (such as LES and CRMs) guided by field studies, more insight can be gained into the cloud processes that require parameterizations in GCMs. Intercomparisons of these process models can advance theoretical understanding, thereby promoting the development of new parameterizations which can be further tested in a single column model environment or through a numerical weather prediction (NWP) model. As a final step, the parameterizations can be evaluated in a full GCM using global observations to assess improvements in our ability to simulate the current climate system. This top-down/bottom-up approach has proven to be a slow but successful way of making use of the network of models. In particular, it has led to demonstrable progress in our representation of top-entrainment in stratocumulus, lateral entrainment and detrainment in shallow cumulus, and closures for convection schemes. By carefully comparing observations and the fine-scale modeling frameworks (LES, CRMs), these models (and best practices in their use) have improved substantially over the

last ten years. This strategy has also been effective in formulating key questions, which in turn serve to direct field and laboratory studies. For example, despite high resolutions of up to 5 m, LES intercomparison studies show great uncertainty in cloud-top entrainment rates in stratocumulus, a crucial dynamic process for this cloud type. As a result, the DYCOMS-II field study (Stevens and Brenguier, this volume) was designed, and it successfully narrowed the uncertainty in cloud-top entrainment rates for several well-observed cases. Similarly, LES results inspired a recent field study, RICO (Rauber et al. 2007), and an ongoing field study, VOCALS¹, designed to quantify the role of precipitation and its interaction to aerosol and microphysical processes in trade-wind cumulus and stratocumulus, respectively.

Finally, these process models have been useful to study and define processes that are not well represented in GCMs. Prime examples are the dependence of tropospheric humidity on moist convection, the transition from shallow to deep convection in relation to the diurnal cycle, and the formation of ice in stratiform cirrus by homogeneous freezing of supercooled aerosols.

Connecting Scales Using Alternative Modeling Techniques

Accepting the limitations of GCMs to study perturbed clouds, let us now proceed to discuss a number of alternatives.

Numerical Weather Prediction

The NWP model is a useful tool in the evaluation of (perturbed) clouds. Since most of the cloud processes, in which we are interested, operate on short timescales, we are able to view the impact of current cloud parameterizations on forecasts. It is therefore useful to run climate models in a NWP mode, as many climate biases are visible within two days. The performance of cloud parameterizations can be critically evaluated using ground-based profiling stations, such as those from ARM and Cloudnet (Illingworth and Bony, this volume). Furthermore, aerosol impacts may be testable in forecast models.

There is a large potential for further exploitation of reanalysis methods, as at present they are not sufficiently constrained by cloud and boundary layer properties. Introducing constraints to reflect these quantities in a better manner would advance the use of a wider variety of data as analysis products are developed. This could be very powerful in helping us to untangle meteorological from aerosol effects. It must be recognized, however, that even if an aerosol is not explicitly included in the analysis, its effects on cloud or thermal structure may be indirectly evident.

¹ <http://www.eol.ucar.edu/projects/vocals/>

Large Eddy Simulation Models, Cloud-resolving Models, and Super-parameterized Single Column Models

We have already mentioned that this class of models offers a way to improve the representation of unresolved processes in climate models. However, one can also imagine ways of using these models more directly to study perturbed clouds. To this end, it would be helpful to have “model problems” that exhibit aspects of the global cloud response of a type of model in a conceptually simpler framework. One potentially useful framework is to understand the response of clouds in an appropriately chosen “single column” framework, in which the effect of the large scale is encapsulated in specified advective forcings of temperature, moisture, and surface boundary conditions.

In reality, at any given location, these forcings constantly change as weather systems move through, thus generating a distribution of clouds in the column. In other words, at any location, the climatological cloud distribution is built up through patterns of weather variability. Thus, to utilize a single column framework to understand the cloud processes, it may be necessary for the clouds in this column to be simulated for months or years with realistic covariability in the forcing to build up a climatologically realistic response. This approach must be repeated in the perturbed climate (with a correspondingly different time series of forcings) to understand how the cloud response changes. To predict the cloud properties and their sensitivity to climate perturbations, the forcings can either be applied to a single column model or, preferably, a small-domain CRM. A similar framework can be used to study clouds perturbed by anthropogenic aerosols.

Model Simplification

Over the last decade, the climate modeling community has tended to increase the complexity of models by adding more processes or components (e.g., aerosols, chemistry, carbon-cycle, vegetation) as well as, but to a lesser extent, by increasing the model resolution. It appears that this tendency has not resulted in a better understanding of cloud–climate feedbacks, as the uncertainty in climate sensitivity has not decreased with time. In fact, the range of global warming estimates for the end of the 21st century, has increased with the inclusion of carbon-cycle feedbacks in coupled ocean–atmosphere models. Although this reflects, in part, the manner in which scientific progress is made, it is nonetheless disappointing, because the aim of climate modeling is not only to provide global warming estimates or global cloud feedback estimates but to learn how the climate works and responds to external perturbations.

A better understanding of the physical mechanisms that underlie the cloud–climate feedbacks produced by climate models would enable us to design a strategy to evaluate these feedbacks using observations. By simplifying models (rather than making them more complex) and by conducting idealized

experiments, we should be able to identify key critical processes, to arrange them hierarchically, and to test resultant ideas or theories. Interpretation frameworks may then be proposed to help us understand the physics of cloud–climate feedbacks and intermodel differences in cloud–climate feedbacks.

With GCMs, this could be done, for instance, by simplifying the large-scale boundary conditions (e.g., aquaplanet versions of the models), by reducing the dimensionality of the system (e.g., 2-D or 1-D model versions derived from the 3-D model), or even by removing some processes (e.g., by replacing complex microphysical schemes with simpler ones). Models of intermediate complexity might also be implemented. Simple conceptual models (e.g., 2-box models) may be viewed as the ultimate step in this simplification process. We need, however, to approach these conceptual models with a sense of caution, as there is the risk of inadvertently neglecting interactions or feedbacks, which could result in biased behavior. For instance, Albrecht’s conceptual model hypothesised how an increase of nucleating aerosol particles could suppress the development of precipitation in boundary clouds and increase cloud fraction (Albrecht 1989). Although this mechanism appears to be relevant in some circumstances, recent fine-scale modeling studies suggest a plethora of circumstances for which the relationship between cloudiness and precipitation is exactly the opposite (see Stevens and Brenguier; Brenguier and Wood, both this volume).

In general, the extent to which simplified models can be useful in reproducing and interpreting complex model results remains to be tested and quantified through comparison with process-resolving fine-scale models. Simplified models are useful in conceptualizing the essential processes, and therefore they can provide guidance if these processes are realistically represented in complex GCMs.

We strongly recommend that simplified climate models and/or experiments be developed with the aim of better identifying and *understanding* the key processes that control cloud–climate feedbacks. Ultimately, our trust in the results of a climate model is leveraged against this understanding. An ideal situation would be for each global climate model to be associated with a suite of simpler or more idealized model versions to support the analysis and understanding of the results obtained. This would, in a sense, formalize the hierarchy.

Global Climate Models: What Level of Complexity Is Necessary, and What Drives Complexity?

Does higher complexity provide better (more reliable) estimates of climate sensitivity? In climate models, the complexity of all modules (e.g., radiation, diffusion, dynamics, cloud microphysics, and convection) should be more or less balanced. It is obviously not an easy task to qualify what “more or less” means. However, one should ask whether a given GCM can cope with the complexity of a proposed new parameterization. It is, for instance, not helpful

to incorporate a full-fledged spectral microphysics parameterization when the GCM is unable to predict the cloud-scale updraft velocities.

The level of complexity required in a GCM will depend on the question being posed. For example, even though contrail cirrus contribute only a small amount to the global-mean radiative forcing, they need to be studied in order to quantify the significant regional effects already in place or their potential effects in a future climate. This task requires that the contrail cirrus microphysical and radiative properties need to be investigated within GCMs at the same level of sophistication that exists in natural cirrus. Similarly, the surface tension of organics should not matter for the present-day simulation of clouds; however, if one is interested in the past or future perturbation of organic aerosols and their effects on clouds and climate, it may well be important.

Because GCMs are used for different applications, and since people prefer to utilize only one model to address all possible questions of interest, it appears that this path toward increasing complexity will continue, as new processes or submodules are added. To increase our understanding of climate, however, it may be more beneficial to reduce complexity and employ a suite of different models in an effort to understand which processes are necessary to reproduce a specific phenomenon.

Tuning and Metrics of Global Climate Models

Models used to project behavior of clouds in the perturbed climate system are generally optimized by adjusting uncertain parameters according to a mixture of developed intuition and empirical experiences. In effect, top-of-the-atmosphere radiative fluxes are adjusted by tuning cloud optical properties. The task becomes more difficult as model complexity increases. The use of formal nonlinear optimization methods offers a way to systematize these procedures and possibly identify solutions to improve desired model metrics. If more complete observational datasets become available on both cloud properties and their radiative properties, this will put strong constraints on the freedom in the tuning of the cloud properties. We would expect that such observational constraints will decrease the large disagreement in cloud properties in existing GCMs and improve the physical basis for representing the climate system.

Fundamental Problems in Cloud Processes

We have discussed problems and errors of the representation of cloud processes that are primarily related to a lack of resolution. These problems would theoretically disappear if only we had enough computational power to resolve them, given that we fully understood the physical mechanisms involved. However, an increase in resolution requires most of the (scale-dependent) parameterizations in GCMs to be revised. These revisions of code physics may turn out to

be very challenging. On a more fundamental level, it must be recognized that we are also coping with processes for which we do not even know the principal processes, let alone their governing equations. These type of problems are predominantly related to microphysics, in general, and further to the ice phase in cold and mixed-phase clouds, in particular.

In addition, hygroscopic growth and activation behavior of atmospheric aerosol particles are not yet fully understood. In this context, possible kinetic effects deserve mention: key words here are surface active substances and their influences on water uptake (“accommodation coefficient”) during droplet activation and growth processes (see Stratmann et al., and Kreidenweis et al., both this volume). Further experiments are thus needed to understand and quantify the effects of (a) slightly soluble, surface active substances and soluble gases on high relative humidity hygroscopic growth and activation (i.e., CCN closure experiments), and (b) the influence of kinetic limitations during droplet activation and growth processes.

Ice processes are significantly more complicated than those associated with warm rain. Below, we highlight several fundamental issues associated with the ice phase in mixed-phase clouds (in particular deep convection) and in cold cirrus clouds (at temperatures below the spontaneous freezing point of supercooled water droplets).

Mixed-phase Clouds

Mixed-phase clouds are loosely defined as clouds which contain liquid and ice in close proximity, so that the liquid and ice particles interact microphysically. They are common in subfreezing conditions (below 0°C) and are associated with a wide variety of systems, including deep convection, fronts, and orographic systems. A critical aspect of mixed-phase clouds is their colloidal instability; that is, ice tends to grow through vapor deposition at the expense of liquid water because saturation vapor pressure is lower with respect to ice than liquid (the Bergeron–Findeisen mechanism). Liquid water is generated through upward motion within the cloud. Thus, the existence of liquid water in mixed-phase clouds results from a balance between its evaporation via the Bergeron–Findeisen mechanism and its generation via vertical motion. The accretion of liquid droplets onto snow (riming) may also play an important role in the depletion of liquid, especially if high concentrations of large ice crystals are present. The amount of liquid water in mixed-phase clouds and its lifetime is important to the climate because small liquid droplets have a much larger impact on the cloud radiative forcing relative to the larger ice crystals. In mixed-phase clouds, supercooled liquid water has also been important for weather forecasting because of its role in aircraft icing.

Assuming saturation with respect to water, the supersaturation with respect to ice, and thus the strength of the Bergeron–Findeisen mechanism, increases with decreasing temperature. The dominant microphysical controls

on the Bergeron–Findeisen mechanism (as well as depletion of liquid water through riming) are the concentration, size, and shape of the ice crystals. These parameters determine how quickly water vapor can be taken up by the ice. However, vapor depositional growth of ice particles remains uncertain from a basic physical standpoint; additional laboratory measurements and theoretical understanding are needed to better constrain this process (Stratmann et al., this volume). Studies have suggested that larger updraft speeds are required to sustain liquid water in mixed-phase clouds as the concentration and/or size of the ice particles increase. For example, in weakly forced mixed-phase stratus, which are endemic to the Arctic and have relatively weak updrafts, models have suggested strong sensitivity of the lifetime of these clouds to small increases in crystal concentration; in the presence of stronger updrafts (e.g., in deep convection), however, the existence of liquid may be less sensitive to crystal concentration. Concentrations of small ice crystals have been difficult to observe in the past, although new techniques and instrumentation have led to better observational constraints (Kärcher and Spichtinger, this volume). Nonetheless, small ice crystals remain difficult to measure.

The concentration of crystals in mixed-phase clouds may be largely dependent on the concentration of IN and, under some conditions, secondary ice formation processes (i.e., ice multiplication), such as the production of ice splinters during riming via the Hallet–Mossop mechanism. Still, ice nucleation processes remain poorly understood, and concentrations of ice crystals often far exceed concentrations of IN, even in conditions that do not support any known ice multiplication processes. Thus, additional laboratory studies are required to expand our understanding of ice particle formation, along with field-deployable instruments to study ice initiation processes in real clouds, which cannot be measured by current methods (Stratmann et al., this volume).

Cold Cirrus Clouds

Cold cirrus clouds are composed of ice crystals. The interplay between dynamic and aerosol impacts on stratiform (non-convective) cirrus is less complex than for mixed-phase clouds. The dynamic forcing is driven by small-scale vertical wind variability (gravity waves, turbulence) on scales of tens of kilometers. Those regions of small-scale variability are embedded in larger-scale ice-supersaturated regions, where cirrus formation takes place *in situ* through ice nucleation. These meteorological factors led to the development of a relatively simple mechanistic understanding of the ice formation process in rising air parcels. In anvil clouds and the lower parts of frontal cirrus, ice is formed within a mixed-phase cloud environment and transported aloft. In such cases, *in-situ* nucleation can occur after most of the preexisting ice mass has been removed by sedimentation.

The relative role of meteorological factors and aerosol impact on cirrus properties and their variability (e.g., PDF of ice crystal concentration) has

been addressed by Kärcher and Spichtinger (this volume), based on *in-situ* data and emphasizing the key role of meteorological factors. The balance in a rising air parcel between increasing supersaturation attributable to cooling and decreasing supersaturation by forming ice condensate implies a predominantly dynamic control of cirrus formation, with the total initial ice crystal concentration being a strong function of cooling rates or vertical velocities. Heterogeneous IN may modulate cirrus formation in regions with relatively small cooling rates. Surprisingly few LES or CRM studies are available that address the links between dynamic and aerosol control of cirrus.

Several field measurements in the tropical and midlatitude upper troposphere have suggested a ubiquitous background of mesoscale (scale of tens of km) temperature fluctuations, leading to typical mean cooling rates of the order 10 K h^{-1} . The origin of these background fluctuations is not well known, but is thought to be generated by gravity waves caused by flow over terrain, and amplifying with height. A predictive capability of geographical and seasonal dependences of this type of forcing is lacking. In such a dynamic environment, the effect of an IN population on cirrus properties depends primarily on the IN number concentration, the ice nucleation threshold (as determined by size, chemical composition and other factors), and the local cooling rate. Theory shows that the likely effect of adding IN to the ubiquitous liquid aerosol background is to reduce actually the total number of ice crystals formed in most cases, keeping other factors fixed. Numerical simulations show that this effect should be robust (though smaller in magnitude) when variability in the dynamic forcing conditions consistent with observations is explained. Adding IN tends to weaken the homogeneous freezing process, leading to slightly larger effective ice crystal sizes and less bright cirrus, and perhaps more significant changes in cloud frequency of occurrence. However, although these theoretical and numerical results are entirely plausible, no closure field studies are available to demonstrate that these processes are actually at work in nature. The current limitations in measuring relative humidity in upper tropospheric conditions accurately (e.g., to within a few percent in absolute terms) limit our ability to discriminate between ice nucleation behavior of various IN in the field. After ice initiation, large ice crystals are generated by sedimentation and aggregation processes. The largest ice particle dimensions that have been observed rarely exceed a few millimeters, and precipitation formation (graupel, hail, snow) commences only after ice crystals have fallen into rather warm cloud layers at low altitudes. The longevity of stratiform cirrus permits radiative heating to feed back to cloud dynamics: absorption of thermal emission by large ice crystals may induce internal convective instability (in particular in a less stable thermodynamic cloud environment), prolonging cirrus lifetime by additional cooling and possibly triggering turbulence-induced ice nucleation.

Further Issues for Ice Clouds

In addition to the issues noted above, there are additional characteristic differences between ice and warm clouds that are common to all ice-containing clouds. These include the coupling between the cloud phase and the moisture field, radiative/microphysical effects of particle shape, and the lack of observational data to constrain cloud models.

In general, clouds that contain liquid water droplets are always close to water saturation. By contrast, if ice forms in a warm cloud at very low concentrations, it can turn into a mixed-phase cloud with the ice phase growing by the diffusion of water vapor and accretion of supercooled water. If the ice concentration is very high, growth occurs primarily by vapor diffusion. This difference is crucially important for local cloud characteristics and aerosol effects, but they also affect significantly the bulk measures in mixed-phase clouds, such as the latent heating or precipitation fallout. Low-temperature ice clouds are often far away from equilibrium; that is, they are less strongly coupled to the ice-supersaturated moisture field because of relatively long supersaturation relaxation timescales. The evolution of their local characteristics depends more sensitively on the history of individual air parcels. Therefore, relative humidity-based, diagnostic GCM cloud schemes fail generally to capture the observed behavior of relative humidity and cirrus in the upper troposphere.

The various forms, shapes (“habits”), and surface characteristics of ice crystals compared to spherical cloud droplets affect scattering and absorption of shortwave radiation and thus the albedo of an ice cloud. Inhomogeneities in the cloud structure affect radiative properties as well, rendering radiative transfer calculations as very challenging. For the same reason, retrieval algorithms used to interpret ground- or satellite-based remote-sensing data of ice clouds are far more uncertain than for low-level water clouds. Moreover, the sedimentation velocities of ice crystals vary significantly, depending on the particle mass and shape (e.g., the habit for pristine crystals and the degree of riming or aggregation for snowflakes), and affect the simulation of cirrus cloud life cycles. Finally, ice initiation has been intensively studied in the laboratory, but it is only poorly understood on the theoretical level. There is a substantial lack of atmospheric IN measurements in all (particularly in the highest and coldest) regions of the troposphere; based on discussion at a recent workshop² on IN instruments, this situation will hopefully improve in the near future (Stratmann et al., this volume).

All of the above issues render the treatment of the ice phase in process models (LES, CRM models) difficult, let alone their representations in models that rely on cloud parameterizations (e.g., mesoscale models and especially in GCMs). Therefore, validation studies of LES and CRM simulations (e.g.,

² <http://lamar.colostate.edu/~pdemott/IN/INWorkshop2007.htm>, Intl. Workshop on Comparing Ice Nucleation Measuring Systems.

using the GCSS strategy) are an important step. How to transfer the knowledge from small-scale models into GCMs via physically based parameterizations poses, however, another challenge.

The difficulty in providing a robust, quantitative answer to how aerosols affect the distribution of ice in cirrus lies in the fact that our understanding of how small changes in formation conditions (i.e., initial ice crystal size distribution) affect cloud ice distribution is very limited. The further development of clouds is strongly tied to dynamic forcing and the evolution of the moisture field. It is unclear how long an initial perturbation of a cloud attributable to IN may last, given the potentially long lifetime of cirrus. We also do not know whether changes in aerosols significantly alter the net cloud radiative forcing. Exploring this parameter space poses a challenge for emerging LES studies of cirrus, including detailed aerosol and ice microphysics. On the GCM level, we are beginning to explore this issue as tools become available to parameterize the relevant subgrid-scale processes and their mutual interactions. Closure measurements that are successful in disentangling aerosol and dynamic effects on cirrus may be possible in the near future given the substantial progress that is currently being made in airborne instrumentation.

Entrainment

To measure entrainment of stratocumulus-topped boundary layers, tracer measurements are employed, which enable an estimate of the entrainment velocity. Depending on the variability of the tracers within and above the boundary layer, they can provide very accurate estimates of this entrainment velocity, particularly when performed with multiple tracers. However, they tell us nothing about the physical mechanism by which air is entrained into the boundary layer, and this information is critical if we wish to relate our parameterizations to physical principles and processes (e.g., sedimentation). Methods using near-stationary platforms, such as ACTOS, and laboratory studies under well-defined turbulence conditions (Stratmann et al., this volume), could provide valuable information about the details of the entrainment interface and nature of the entrainment process itself, and thus merit support. Such methods could be similarly useful for constraining the entrainment interface along cloud boundaries.

Overlap and Heterogeneity Assumptions

Currently, GCMs predict cloud fraction in each vertical level. Uncertainty on how to distribute this fractional cloud, both as a function of horizontal resolution and as a function of the distribution of clouds in the other layers of the model, gives model developers considerable freedom in distributing the modeled cloud amount so as to satisfy the radiative constraints at the top of the atmosphere. Experience with different GCMs show that while there may be

large layer-by-layer differences in cloud fraction and cloud amount, suitable choices for the cloud overlap assumption enables each to match the top-of-the-atmosphere radiative constraints. Thus, providing measurements of cloud overlap along with cloud horizontal inhomogeneity (e.g., as from the Cloudnet and ARM remote-sensing sites) would provide further and valuable constraints on the models. In addition, such an approach should encourage representations of clouds that encapsulate the entirety of the vertical column, rather than processes on a grid by grid-level basis.

Cloud Lifetime

Satellite measurements with high spatial resolution lack typically temporal resolution. Being able to measure the temporal evolution or even life cycle of clouds would provide a valuable constraint on process models of clouds and their environmental interactions. For instance, a geostationary observatory, such as Meteosat Second Generation, could provide invaluable inputs. Going a step further, one could imagine and promote the deployment of a geostationary observatory that could steer and focus a large telescope on particular cloud scenes (to complement field measurements or *in-situ* studies).

Observational Strategies and Proposals

Observational Strategies to Quantify Aerosol-perturbed Clouds

To a first order, the main cloud-controlling factor is relative humidity, which combines two atmospheric state parameters: specific humidity, q_v , and temperature, T . Clouds form if supersaturation occurs in a clear-sky region, i.e., if $q_v > q_s(T)$. Since both temperature and specific humidity vary strongly in both time and space, as a result of meteorology, the performance of the existing observation, forecast, and reanalysis systems is far below what is necessary to discriminate an aerosol impact from changes attributable to the meteorology (see Stevens and Brenguier, this volume). This holds both for the cloud-controlling factors and for the cloud macrophysical properties that are supposed to be impacted by aerosol particles (radiative properties and precipitation at the surface), and is valid for all cloud types. This makes it extremely hard to demonstrate the effect of aerosols on clouds on an observational basis (see Kärcher and Spichtinger, Stevens and Brenguier, and Brenguier and Wood, all this volume).

In attempting to detect an aerosol impact on cloud macrophysical properties (e.g., radiative properties or precipitation), field studies face the same obstacle that weather modification has experienced for over fifty years. Indeed, the sensitivity of these macrophysical properties to meteorology is so high, and our capability of measuring cloud-controlling factors so limited, that the number of

case studies is generally too small to reduce statistically the variability caused by cloud-controlling factors below the level of expected aerosol impacts. One way to optimize the chances of detecting an aerosol signal in the meteorological noise is to choose cases in which the perturbation of the aerosol particle is large, the variability of the meteorology small, and the covariance between the two negligible. Fires, particularly those that result from human activity, offer a chance to examine the effects of large-scale changes in the atmospheric aerosol independent of changes in the meteorological environment. Indeed, fires are preferable under certain meteorological conditions. Whether or not fires actually develop, however, often depends on mere chance. For instance, during periods of Santa Ana Winds (offshore flow over the southwestern United States and northwestern Mexico) fires are common, but they are invariably triggered by chance human activity. By sampling aerosol and cloud properties, well downstream of the fire regions, during periods of the Santa Anna Winds, with and without fires, it may be possible to develop a dataset within which meteorological and aerosol effects are independent. The utility of such an approach could first be explored using the satellite record; then, if warranted, it could be realized by targeted deployments of suitably instrumented aircraft over several fire seasons. We recognize that even this may be difficult to achieve. Apart from the rigor of obtaining sufficient samples, one must sample well downwind of the fires because conditions that favor such winds do not favor clouds. Doing so, however, yields measurements in a more disperse aerosol, thereby lessening its impact. Perhaps more importantly, such a strategy allows sufficient time for the development of meteorological biases because of the direct effect of the smoke on radiative fluxes in the evolving cloud-free air.

Another clear example of anthropogenic aerosols that perturb clouds is the existence of ship tracks in stratocumulus. These were the subject of the 1994 MAST experiment (Durkee et al. 2000) as well as numerous remote-sensing studies. These studies suggest that when cloud droplet concentrations exceed 100 cm^{-3} , the cloud in the ship track has a similar or lower liquid water content than that of the ship track itself, suggesting a negative secondary indirect effect on climate (Platnick et al. 2000). When cloud droplet concentrations are much lower, however, the ship track tends to be much thicker than the surrounding stratocumulus. There are almost no realistic modeling studies of ship tracks that quantitatively reproduce these results, including the possibly large effects of circulations between the perturbed cloud in the ship track and the unperturbed surrounding clouds. Ship track measurements provide a good opportunity to test whether LES models can simulate the response of a stratocumulus cloud to a large aerosol perturbation, but as yet this opportunity has been underutilized.

Finally, one can also follow the opposite approach: a situation could be chosen where aerosol properties are similar but cloud-controlling factors vary significantly. Such situations occur every day, at all locations, over the diurnal cycle, with a significant and precisely reproducible change in sun radiative

forcing. Over such a period and if air masses are sufficiently homogeneous in terms of aerosol properties, this reverse approach might actually be more successful in detecting an aerosol impact on cloud macrophysical properties than doing so directly. The chance of success of the two approaches should be evaluated following the same methodology as recommended by the WMO in the design of weather modification experiments (NRC 2003). Since aerosols have a diurnal cycle attributable to photochemistry, vertical mixing, and diurnal variation of human and natural aerosol/chemical sources, concern was expressed about the applicability of this approach.

Strategies Using Satellite Measurements

Satellites provide wonderful global observations of cloud properties, but their ability to provide sensible inputs on cause and effect relationships is often questioned. It is interesting to note that this same criticism is not applied to ground remote-sensing or *in-situ* observations from instrumented aircraft, although they too provide just snapshots of various physical parameters. There are two differences which may explain such contrasting views:

1. Time evolution: Observations of clouds help demonstrate, for instance, that clouds are responsible for generating precipitation and not the opposite, because time series show that clouds come before precipitation. This simplistic example shows that time evolution is a crucial ingredient when establishing cause and effect relationships. Present cloud satellite instruments, mainly polar orbiting, are not suited for short timescale evolution experiments, but the increasing performance of geostationary satellites, in terms of cloud observations and time resolution, opens new opportunities.
2. Spatial distributions: Most satellite observations, based on passive sensors, are two-dimensional and cannot document the vertical distribution of the components. For instance, statistics on horizontal correlations of aerosol and cloud liquid water path cannot discriminate aerosols located higher than cloud layers, hence limited to radiative interactions, from those entering cloud base that can control cloud microphysics. The availability of the A-Train observations (Anderson et al. 2005), some of which are vertically resolved, opens new opportunities.

Lacking these two capabilities, satellite observations leave numerous possible explanations open for the observed correlations between aerosol and cloud properties, in terms of cause and effect relationships (cf. Nakajima and Schulz, this volume).

Experimental Strategies for Cloud Ensembles

The last decade has witnessed a number of field studies (DYCOMS-II, RICO, EPIC-2001) designed to measure processes related to the cloud ensemble. Whereas earlier attempts to make such measurements often required many platforms (BOMEX, ATEX, GATE), the emergence of a wide variety of satellite products as well as dropsondes and aircraft-based remote sensing make it possible for a single, long-endurance aircraft instrumented for *in-situ* measurements of dynamics, radiation, gas and aerosol chemistry, and microphysics both to constrain important remaining aspects of the cloud environment and to measure the statistical properties of clouds in this environment. These studies appear most promising for understanding the behavior of cloud ensembles, which is critical in advancing parameterizations as statistical representations of fields of clouds. Combining such measurements with long-term space and ground-based remote sensing offers further opportunity for progress. One measurement that is, however, conspicuous in its absence is that of cloud fraction, such as might be measured by a highly sensitive scanning cloud (K-band) radar, or a spaceborne observatory capable of being focused on a given experimental area (for discussion on the satellite techniques CloudSat/CALIPSO, see Anderson et al., this volume).

Conclusions

Processes that impact clouds span scales from the molecular to the planetary level, a fact that makes their description in a single reference model unimaginable. This implies that large-scale models will continue to hinge on parametric (or statistical) representations of cloud processes on at least some scale for the foreseeable future. As we endeavor to represent present-day and perturbed climates, we must recognize that this introduces a source of uncertainty at best, and likely bias. The origins of uncertainty in our representation of clouds in present-day GCMs are:

1. a fundamental lack of knowledge of some key cloud processes (e.g., upper tropospheric ice supersaturation and role of ice nucleating particles in mixed-phase and cold ice clouds),
2. a lack of knowledge about how to represent the aggregate properties (statistics) of processes that are well understood at their native scale,
3. a lack of interaction among subgrid processes,
4. a breakdown of quasi-equilibrium assumptions at the intermediate scales.

While some of these issues may play less of a role when increasing the resolution of GCMs (category 2 and 3), others (category 1) remain or might even play a larger role (category 4). We propose a number of new pathways for

GCMs, including global GCRMs and cloud-resolving convective (super) parameterization (Grabowski and Petch, this volume). However, even with such approaches, it will remain important to have consistent and interactive treatments of the remaining unresolved scales, such as boundary-layer cloud processes. When moving to higher resolutions, stochastic parameterizations might become more relevant as quasi-equilibrium assumptions break down. We have discussed the more fundamental problems (category 1) pertaining to cloud microphysics (e.g., primary and secondary ice formation, droplet/ice particle growth kinetics, entrainment, cloud overlap) and have made recommendations for both observational studies (laboratory and field experiments) and high-resolution process models.

GCMs should not be used as the sole instrument to quantify the effect of perturbed clouds in our climate system. Instead, a better use of the full hierarchy of models could improve our understanding of the key processes that control cloud–climate feedbacks. For instance, process-resolving, fine-scale models can be used under perturbed conditions to understand and quantify the cloud dynamic response to specified perturbations. Simplified frameworks (aquaplanets, 2-D models, column models with large-scale feedbacks) are useful for conceptualizing these physical cloud processes and provide guidance of how they could be operating in complex GCMs. They can also advance our understanding upon which our confidence in our climate predictions ultimately rests.

Finally, we have proposed appropriate experiments to study perturbed clouds by aerosols. As it is extremely hard to detect an aerosol signal in the meteorological noise, we suggest that cases be chosen in which the perturbation of an aerosol is large compared to the meteorological variability and the covariance between the two is negligible. Proposed candidates are fires triggered by human activity and perturbed stratocumulus clouds in ship tracks. In addition, we recommend that the reverse approach be considered; namely, aerosol properties remain similar but meteorological factors vary significantly (e.g., the diurnal cycle).

References

- Albrecht, B. A. 1989. Aerosols, cloud microphysics, and fractional cloudiness. *Science* **245**:1227–1230.
- Anderson, T. L., R. J. Charlson, N. Bellouin et al. 2005. An “A-train” strategy for quantifying direct climate forcing by anthropogenic aerosols. *Bull. Amer. Meteor. Soc.* **86**:1795–1809.
- Durkee, P. A., K. J. Noone, and R. T. Bluth. 2000. The Monterey area ship track experiment. *J. Atmos. Sci.* **57**:2523–2553.
- Kessler, E. 1969. On the Distribution and Continuity of Water Substance in Atmospheric Circulations. Meteorol. Monogr., vol. 32. Boston: Amer. Meteorol. Soc.

- NRC. 2003. Critical Issues in Weather Modification Research. Board on Atmospheric Sciences and Climate, Division on Earth and Life Studies. Washington, DC: National Academies Press.
- Platnick, S., P. A. Durkee, K. Nielsen et al. 2000. The role of background cloud microphysics in the radiative formation of ship tracks. *J. Atmos. Sci.* **57**:2607–2624.
- Rauber, R. M., B. Stevens, H. T. Ochs et al. 2007. Rain in shallow cumulus over the ocean: The RICO campaign. *Bull. Amer. Meteor. Soc.* **88**:1912–1928.