From the Strüngmann Forum Report, *Clouds in the Perturbed Climate System: Their Relationship to Energy Balance, Atmospheric Dynamics, and Precipitation*
Edited by Jost Heintzenberg and Robert J. Charlson. 2009.
 ISBN 978-0-262-01287-4

22

Observational Strategies at Meso- and Large Scales to Reduce Critical Uncertainties in Future Cloud Changes

Anthony Illingworth[1] and Sandrine Bony[2]

[1]Department of Meteorology, University of Reading, Reading, U.K.
[2]Laboratoire de Météorologie Dynamique, CNRS/UPMC, Paris, France

Abstract

The response of clouds to climate change remains very uncertain. This is attributable to both an incomplete knowledge of cloud physics and to the difficulties that large-scale models have in simulating the different properties of clouds. An observational strategy is proposed to improve the representation of clouds in large-scale models and to reduce uncertainties in the future change of cloud properties. This consists of determining first what key aspects of the simulation of clouds are the most critical, with respect to future climate changes, and then of using specific methodologies and new datasets to improve the simulation of these aspects in large-scale models.

Introduction

After decades of research, the response of clouds to a change in climate, and in particular to a global climate warming induced by anthropogenic activities, remains poorly understood and is still identified as a key source of uncertainty for climate sensitivity estimates. Given the slow progress in this area over the last fifteen years, one may wonder what strategy might help to reduce this uncertainty. There are so many physical processes and cloud properties that need to be better understood, and so many weaknesses in the representation of clouds in climate models.

When proposing an observational strategy, it would be helpful to know whether there is a hierarchy among the different problems; whether there are some priorities among the different processes that need to be better understood, better observed, or better simulated in climate models. Therefore, we think that an observational strategy to reduce uncertainties in cloud–climate feedbacks should be composed of two steps: (a) determine what are the most critical uncertainties, (b) determine how observations might be used to reduce some of these uncertainties.

Critical Uncertainties in the Large-scale Modeling of Clouds and Their Impact on Climate

Key Aspects of the Simulation of Clouds in Large-scale Models

Part of the reason why progress in the representation of clouds in large-scale models has been so slow is that major aspects of the simulated cloud distribution could not be assessed observationally. For instance, the vertical structure of cloud layers, their overlap, the cloud water content, and the cloud water phase are known to play a key role in the radiation budget at the top of the atmosphere (TOA), at the surface, and in the troposphere. Given the lack of reliable and global observations of these quantities, a good agreement between models and observations of TOA radiative fluxes or of the total cloud cover could be obtained with compensating errors. However, differences in the way the radiative balance is achieved can affect the sensitivity of radiative fluxes to a change in climate.

The vertical structure of clouds constitutes one well-known example of a key factor critical for climate studies for which observations have long been lacking. Another example is the cloud phase. Many papers have drawn attention to the vastly different amounts of cloud ice held in various climate models, all of which satisfy the TOA radiation constraint. In large-scale models, the cloud water phase is still commonly diagnosed as a simple function of temperature. As climate warms, the fraction of cloud water may increase at the expense of cloud ice, and owing to differences in the microphysical and radiative properties of liquid and ice clouds, the change in cloud phase contributes to cloud feedbacks. Uncertainties in the diagnostic of the cloud water phase in the current climate thus translate into uncertainties in cloud feedbacks (e.g., Tsushima et al. 2006).

The simplicity of the cloud phase diagnostic in large-scale models has long been justified by the fact that the factors influencing glaciation processes and the presence of supercooled droplets are still poorly understood and loosely constrained by observations. It is thus essential to use new data from satellite or ground-based measurements to provide a better explanation of the factors that influence the cloud phase, and to develop more reliable parameterizations for large-scale models.

In addition to the lack of key measurements, the representation of clouds in large-scale models is complicated by our poor understanding of the large-scale controls of (measured) cloud properties and of the physical processes through which the different cloud properties interact with each other (cf. Bretherton and Hartmann; Grabowski and Petch, both this volume). Such an understanding would require that measurements of cloud properties are done simultaneously for several variables, over a wide range of meteorological situations. Numerical weather prediction (NWP) models may also be very helpful in that regard. If an NWP model is producing the observed cloud amount in the right place at the right time (which has been shown to be the case in Illingworth et al. 2007), then we can be reasonably sure that the meteorological processes which produce the clouds are being well represented. NWP model outputs can then be used to understand how clouds are controlled by meteorological processes.

With the open availability of model simulations to the international climate science community (Meehl et al. 2007), the number of biases of climate models reported in the literature has dramatically increased. However, compared to the hundreds of scientists involved in the analysis of climate simulations, the number of scientists actually working on the development and continuous improvement of physical parameterizations in climate models is fairly small. This may be caused by the fact that institutional structures do not reward this activity enough, especially since it may take a long time and considerable effort to get real improvements, while it is much easier to demonstrate errors in models. In this situation, how should the effort of parameterization improvement be concentrated on the most critical processes?

The same question arises regarding the reduction of uncertainties in the models' projections of the future climate. The Fourth Assessment Report (AR4) (IPCC 2007) reports a large range of global climate sensitivity estimates among climate models. The largest contribution to this spread arises from intermodel differences in cloud feedbacks (Soden and Held 2006). Given the very large number of factors or processes potentially involved in these differences, which ones should we concentrate on to reduce, as efficiently as possible, the uncertainties in future climate change?

Based on different methodologies, recent studies suggest that the response of marine low-level clouds to climate change was the root cause of a large part of intermodel differences in global cloud feedbacks (Bony and Dufresne 2005; Webb et al. 2006; Wyant, Bretherton et al. 2006; Williams and Tselioudis 2007). By using an analysis method based on the stratification of the large-scale tropical circulation into dynamic regimes (Bony et al. 2004) or by analyzing simulations performed with idealized and simplified (aquaplanet) versions of climate models (Medeiros et al. 2008), it was shown that intermodel differences in the tropical cloud response were dominated by the response of clouds in the trade-wind regions. This suggests that an improvement in the representation of shallow convection and trade cumulus clouds is crucial for climate sensitivity. These studies shed light on the "silent majority" of tropical clouds that has locally a less spectacular impact on radiation than deep convective clouds or stratus clouds, but plays a major role for climate sensitivity. This finding will foster further studies focused on the understanding, the simulation, and the evaluation of shallow clouds, and will thus help to reduce this critical uncertainty.

This example shows that by carrying out idealized studies of climate change and by decomposing the global cloud feedback problem into components related to specific physical processes, the problem becomes more tractable and can suggest targeted diagnostics of model-data comparison or of data analysis. Such approaches have the great potential of identifying the processes that are critical for climate change projections. This provides guidance to establish a hierarchy of necessary model developments, helps to fill the gap between climate studies and process studies, and contributes to a better understanding (and thus a better assessment of our confidence) in the models' results. Therefore, such studies should be considered as a key step in the strategy to reduce critical uncertainties associated with the response of clouds to a changing climate.

One caveat associated with this strategy, however, is that processes or cloud regimes that may be missing in all the models or that may be represented equally badly in all the models may not be identified as a key source of uncertainty, as they might actually play an important role in nature. It is thus important to complement this strategy with comparisons of models with observations, and with idealized studies investigating the potential impact of model weaknesses on the simulation of climate.

Studies of this kind are necessary to assess the extent to which some processes contribute more than others to uncertainties in climate change projections. For instance, the inability of large-scale models (NWP or climate models) to simulate accurately the diurnal cycle of convection over tropical land areas is often cited as a concern for the credibility of climate change projections (whatever they are). This bias, which assuredly reveals some weaknesses in the models' representation of physical processes, is likely to be a concern for representing realistically the interactions between local precipitation and vegetation processes for instance. However, for other issues, such as the magnitude of global climate change or the change in some monsoon characteristics, the extent to which this bias actually affects climate model projections has yet to be demonstrated or assessed.

Recent studies show that climate models still exhibit substantial biases in their simulation of the water vapor and temperature distributions in the current climate. However, John and Soden (2007) found no relationship between the biases exhibited by models in the current climate and the magnitude of the water vapor–lapse rate feedback produced in climate change. They interpret this result by the fact that the water vapor feedback depends on the fractional change of humidity, and that this quantity is insensitive to biases in the mean state. Although we cannot exclude the possibility that biases in humidity are associated with biases in cloudiness (such an association may even be expected), this example illustrates the fact that biases in the mean state do not necessarily affect climate change feedbacks.

Similarly, aerosol effects are known to play a key role in the evolution of the 20th century climate through their direct effect on radiation. They have an influence on the formation and radiative properties of clouds at the small or regional scale. It has been proposed that the radiative effects of aerosols at the surface and in the troposphere affect surface temperature and large-scale atmospheric dynamics, and then affect climate phenomena such as the south Asian monsoon (e.g., Ramanathan et al. 2005). Some studies have indicated that the indirect effect of aerosols on liquid water clouds has the potential to affect global-scale climate change. Thus, evaluating and refining the representation of these processes in climate models certainly constitutes an important area of model development. In addition, we know very little about the ability of a very small fraction of the aerosol particles to act as ice nuclei, thus influencing the glaciation of clouds and so affecting the cloud lifetime and precipitation efficiency. Potentially, such processes are very sensitive to small amounts of anthropogenic aerosols and have led to the suggestion that they might affect the development of deep convective clouds.

However, in terms of the understanding and simulation of the tropical or global cloud response to climate change, it is still unclear whether the interaction between aerosols and clouds is of primary importance, compared to the influence of other small-scale (e.g., boundary-layer turbulence and shallow convection) or large-scale (e.g., changes in the large-scale circulation) processes. In the remote trade-wind regions of the Pacific Ocean, for instance, it is unlikely that the physical and radiative properties of trade wind cumuli will be strongly affected by anthropogenic changes in aerosol properties. We can therefore consider, for the moment, that improving the representation of cloud–aerosol interactions in climate models is of lower priority, to reduce the uncertainty in simulated large-scale climate changes, than the improvement of physical processes (e.g., boundary-layer turbulence, atmospheric convection, and radiative transfer).

Observational Strategies that Address These Uncertainties

Cloud Schemes in Large-scale Models

Clouds often form at scales much smaller than the typical size of a grid box in general circulation models (GCMs) and cannot therefore be explicitly predicted in these models. To predict the cloud fraction (together with other cloud properties), many large-scale models diagnose the cloud fraction by using statistical cloud parameterizations. In this approach, subgrid-scale fluctuations of variables (e.g., total water, potential temperature, or vertical velocity) are described by a probability distribution function (PDF) whose statistical moments (mean, variance, and skewness) must be diagnosed based on large-scale prognostic variables plus eventually some subgrid-scale variables predicted by turbulence or convection schemes (e.g., Bony and Emanuel 2001; Tompkins 2002). Here, the cloud fraction and water content are related to the fraction of the PDF above saturation and its first moment, respectively. This PDF can also be used to predict cloud overlap and consequent radiative properties as well as to provide a better description of the development of precipitation. Such an approach is promising to fill the gap between the different cloud scales and to strengthen the physical coupling between the different cloud processes.

A better documentation and understanding of the influence of subgrid-scale processes (e.g., turbulence, convection, gravity waves) on the PDF of large-scale variables for different cloud regimes (e.g., deep convective clouds, shallow clouds, cirrus) is required to guide development or improve statistical cloud parameterizations. For this purpose, modelers often use simulations from cloud-resolving models (CRMs) to get some guidance. This approach might now be developed with the arrival of global CRM simulations and super-parameterizations. This would allow, for instance, the examination and better understanding of how the PDF of different variables relates to small-scale physical processes and interacts with the large-scale environment. One might then investigate why large-scale models fail to simulate middle-level clouds while some CRMs do a better job (e.g., Liu et al. 2001). An important prerequisite for the success of this approach, however, is that CRMs (or large eddy simulations) simulate the subgrid-scale fluctuations accurately. To assess whether it is actually the case, comparisons between observed and simulated fluctuations and cloud distributions produced by high-resolution models on the 100 m to 2 km scale are required. This emphasizes the need to observe the humidity structure of the atmosphere in three dimensions with a high enough resolution.

Evaluating Cloud Properties Simulated by Large-scale Models with Ground-based Data

Ground-based observations, such as those derived from Cloudnet (Illingworth et al. 2007), or ARM-instrumented sites (Mather et al. 1998) have proved useful in evaluating models. Although they lack the global coverage of satellites, they have the advantage of greater spatial and temporal resolution with a more powerful array of remote-sensing instruments. Provided that analysis is restricted to times when winds are high enough to ensure that sufficient amounts of clouds advect past the sensor and a reasonable cross section of the model grid box is sampled, valid comparisons with the model can be made every hour. The Cloudnet study of seven operational models over one year showed that the representation of clouds in a given grid box over the observing site was surprisingly good, but particular biases could be identified. The vertical profile of mean cloud fraction revealed that all models underestimated the occurrence of mid-level cloud. Mean ice water content profiles in the models showed good agreement with observations, and more recent versions of the models captured the observed mean liquid water content well. It is interesting to note that the performance of a mesoscale (12 km resolution) model was not notably better than the same model when run at a global scale with 60 km resolution, although a fairer test would be to carry out the comparisons at the same scale by aggregating the 12 km model data up to the 60 km resolution. It is important to note that the models carry the correct mean values, but this is not the whole story. The PDF of cloud fraction showed that models have fewer completely filled grid boxes than observed and more partially filled grid boxes. Model PDFs of liquid water content were more peaked than observations. The PDFs of ice water content revealed that the one model that had the worst mean value below 7 km had actually the best PDF below 0.1 g m^{-3}; however, because any higher ice water content was considered to be falling snow rather than cloud, the result was a mean value of ice water content that was far too low. Most models have low-level water clouds that drizzle all the time, with the drizzle reaching the ground; observations show the same mean drizzle rate, but a completely different PDF, with occasional bursts of heavier drizzle reaching the ground but usually much lighter drizzle, which evaporates 100 m or so below cloud base. Ground-based studies have also shown (Hogan et al. 2000) that the common assumption of maximum random overlap of clouds is appropriate for clouds shallower than 2 km, but should be modified so that as the clouds become deeper the overlap tends towards maximum. These Cloudnet results, which show that the NWP models have considerable skill in producing clouds at the right time in the right place, suggest that the models are capturing the fundamental meteorological processes that produce the clouds and are correctly locating the regions of ascent. These encouraging results suggest that the assimilation of clouds within NWP models may be feasible, and that such studies could lead to improvements in parameterization schemes. In addition, the success of NWP models give us confidence that climate models should also be representing clouds reasonably well, since they are using essentially the same cloud parameterization schemes.

This example shows how analyzing observations, both in terms of mean values and PDFs, helps to elucidate the processes responsible and should greatly help to improve the physical basis of statistical cloud schemes and to reduce the degree of empiricism in them. Further analysis should be undertaken. This could, for example, investigate if the implicit ice particle sizes in the models are correct, examine if the lack of mid-level clouds in the models is important, and establish the scale and relevance of the errors in the representation of drizzle in low-level clouds. Given the critical uncertainties associated with trade-cumulus clouds (see above), the analysis of long-time series of ground-based data collected in regions covered by such clouds would be very beneficial. Aspects to be investigated would be the values of liquid water content and liquid water path, cloud base, cloud top as a function of the depth of the boundary layer, the formation of any precipitation including small drizzle droplets, and their subsequent fate as they fall below cloud to evaporate or on occasion to reach the ground. A sensitive high-resolution (e.g., 6 m/30 s) ground-based Raman lidar should provide detailed observations of the PDF of water vapor within the boundary layer, which can be combined with the liquid water content within cloud, to provide the observed PDF of total water content.

This PDF, which is the fundamental basis for statistical cloud schemes, can then be compared with that predicted by models. The deployment of the mobile ARM facility in the Azores, from April to December 2009, to observe the springtime overcast stratocumulus regime and the summertime broken trade cumulus, should be particularly fruitful. It may well be that the observations made in recent field projects, such as RICO and BOMEX, are also able to furnish some of the data required to see if climate models are simulating such clouds realistically.

Recent observational campaigns with advanced multiple wavelength and depolarization lidars may provide more information on the ability of aerosol particles to influence cloud properties. In particular, it seems that the size of aerosol particles can be inferred from the lidar backscatter and/or extinction spectrum (Müller et al. 2000, 2001) and the shape from the depolarization ratio, and that large non-spherical particles may act as efficient ice nuclei and promote glaciation. There is some evidence that Saharan dust may be a source of ice nuclei and that when the dust is lofted to high altitudes, such ice nuclei can cross the Atlantic (e.g., DeMott et al. 2003; Ansmann et al. 2008). The degree to which such dust particles promote the glaciation of supercooled clouds and how often this occurs is still questionable, but this is an area of active research which should yield quantitative results. Incorporating such phenomena into climate models will be difficult; the degree to which Saharan dust is lofted is dependent upon the performance of the transport model and the gustiness of the surface winds—a local effect which will be difficult to capture reliably in large-scale models.

Evaluating Cloud Properties Simulated by Large-scale Models with Satellite Data

The evaluation of the clouds in GCMs has long been hampered by the lack of global observations of the vertical structure of clouds. The situation is now radically changing with the arrival of new observations from the A-Train constellation of satellites, including the spaceborne radar (CloudSat) and lidar (CALIOP/CALIPSO) instruments. The observational definition and detection of clouds, however, depends strongly on the type of measurements and sensitivity of sensors, as well as the vertical overlap of cloud layers in the atmosphere. This definition also differs from the definition of a cloud layer in large-scale models or in high-resolution models (e.g., CRMs). Therefore, a raw and direct comparison of cloud products derived from observations with model simulations does not guarantee that apples are not compared with oranges.

To make more meaningful comparisons between models and observations, it is better to use a *simulator* to diagnose from the model outputs some quantities that are directly comparable with observations. Such an approach has been widely used to compare model cloud covers with ISCCP data (e.g., Klein and Jakob 1999; Webb et al. 2001; Zhang et al. 2005). New simulators, which compare the observed radar and lidar backscatter profiles with those profiles calculated from the model parameters, are now under development. First studies using a CloudSat simulator (Bodas-Salcedo et al. 2008) and an ICESAT (Wilkinson et al. 2008) or a CALIPSO (Chepfer et al. 2008) lidar simulator show already how promising the approach is to evaluate the cloudiness simulated by climate models. Biases can now be identified much more clearly and in more detail (in particular, the vertical structure can be documented) than with previous comparisons using passive measurements.

Global comparisons of histograms of CloudSat radar reflectivity (Bodas-Salcedo et al. 2008) as a function of height computed over several months from the Met Office model, which has an implicit exponential ice particle size distribution with an intercept parameter that is a function of temperature and an ice particle density which is inversely proportional to size, appear perhaps initially discouraging. The observed histograms are much smoother than those observed, but they do show that the model is underestimating mid-level clouds. However, when the comparisons are subdivided into geographical regions, they are much more revealing. Over the North Atlantic, the model performance for the ice cloud is quite good, indicating that parameterization of the intercept parameter as a function of temperature performs well. Problems are evident for low cloud: the model has two separate drizzle regimes rather than one. Comparisons over the California stratocumulus region and the tropical Pacific also reveal specific errors in the model. As noted with the ground-based measurements, the lack of mid-level clouds seems to be ubiquitous, and the occurrence of drizzle in low-level clouds seems to be overestimated in the model. Clearly, such an approach has powerful implications, although care is needed to distinguish between the relative influence of cloud and precipitation biases in errors of the simulated radar reflectivities. As with the ground-based observations, comparing the statistics of the mean values of the reflectivity histograms with the observations are just the first step. The next step is to classify the data in terms of different weather regimes; this is accomplished by separating the data according to, for example, large-scale vertical motion and surface stability. Thereafter, those processes which are being poorly represented must be identified to establish if, for example, the lack of mid-level clouds in the models is important for some aspects of the simulated climate.

The computation and interpretation of a radar simulator is reasonably straightforward in that the model holds an implicit size distribution of the ice particles, and for water droplets there is a prescribed droplet size over the ocean and over land, and that in general the attenuation of the radar signal is rather small. Lidar measurements are very sensitive to the presence of cloud particles, and the horizontal and vertical resolutions of the measurements are very high (330 m and 30 m, respectively, for CALIPSO). The analysis of lidar measurements thus constitutes a powerful means of diagnosing the vertical distribution of cloud layers and their overlap. However, the attenuation of lidar signals is much larger than that of radar reflectivities, so that the signal from the satellite can be totally extinguished at low altitudes when thick upper-level clouds are present. The attenuation is related to the observed lidar backscatter through the "lidar ratio," or the ratio of backscatter to extinction, but this lidar ratio is very sensitive to the (unknown) ice particle shape and size. In addition, the penetration of the lidar beam through multiple levels of broken cloud is very sensitive to the degree of cloud overlap; this could be considered as an advantage in that the very sensitivity to the cloud overlap could be regarded as an excellent method of diagnosing if the cloud overlap implied in the model is in fact realistic. In the presence of upper clouds, as in the tropics, the simulated and observed backscatter from lower-level clouds depends on how well the thicker higher-level cirrus clouds are represented. However, a large fraction of tropical oceans are associated with large-scale subsidence in the free troposphere, so boundary-layer clouds are not overlapped by upper-level clouds. In these situations, attenuation problems are minimal and the lidar is able to provide unambiguous returns from cloud top. The lidar simulator is thus particularly useful for studying those (ubiquitous) clouds which are important for the Earth's radiation budget but are often below the sensitivity of radar, such as stratus, stratocumulus, and fair weather cumulus clouds, which we identified earlier as being of crucial importance. Lidar can observe cloud top to 30 m, and thus, for an ensemble of clouds, it should be possible to identify the cloud base of these clouds. Cloud water droplets will generally yield a radar reflectivity too low to be detected from space, so any observed radar reflectivity will indicate the presence of small drizzle droplets or precipitation. If this is combined with inferred values of liquid water path in the cloud and effective radius from passive "MODIS"-type instruments, then the properties of the fair-weather cumulus clouds can be compared in detail with their representation in models. Evaluating the ability of climate models to simulate accurately the geometrical thickness and the precipitation efficiency of shallow-level clouds, together with their variation with natural climate fluctuations, is of paramount importance if we are to have confidence in the simulated response of these cloud properties to climate change and then in the model cloud feedbacks. At high latitudes, the persistent low-level polar clouds should also be well detected by lidar measurements.

Note that the use of simulators is also a way to fill the gap between the different cloud scales since the comparison of the cloud covers predicted at the large scale can be compared to observations derived at a much smaller scale. For example the lidar signals are, in principle, available for each lidar pulse, with a horizontal resolution of 330 m or so for the highly reflecting water clouds and a vertical resolution of 30 m; the radar reflectivity has a horizontal resolution of just 1.1 km and 500 m in the vertical. Unfortunately, when we are considering the representation of tropical broken cumulus clouds, high-resolution observations of the PDF of humidity via Raman lidar do not seem to be possible from space, and, at present, only values of water vapor path integrated over the vertical are available with a horizontal resolution of some 20 km.

Turning to ice clouds, the ice particle size can be derived from the ratio of the radar return (which varies as the sixth power of the particle diameter) to the lidar backscatter signal (which, when corrected for attenuation, depends on the square of the particle diameter). The first stage would be to compare the inferred ice particle size and its variation globally with location and temperature, and then compare it with the particle size, which in most models is prescribed in terms of the temperature alone. Results of this analysis should be available very soon; Delanoë and Hogan (2008) have demonstrated how the errors of the derived products from a combination of active and passive sensors can be obtained using a variational technique. Depending upon the results, one can envisage having a prescribed ice particle size in the models which varies not only with temperature but also with other environmental conditions. The next stage could be to have a double moment scheme to represent the ice particles, as is done in CRMs, provided the particle size in such a scheme could be constrained to agree with the size inferred from the active radar and lidar onboard the satellites.

Supercooled layer clouds can be identified relatively easily from space by their very high lidar backscatter and sharp backscatter gradient at cloud top. Using data from the LITE mission on the space shuttle Hogan et al. (2004) found that around 20% of all clouds between –10°C and –15°C contained supercooled layers. Such thin layer clouds have a much larger radiative impact than ice clouds of the same water content because of their smaller particle size, yet they are scarcely represented in climate models. Quantification of the radiative impact of such clouds on a global scale will soon be possible using CALIPSO data.

In addition, CALIPSO lidar data provides us with high-resolution observations of aerosol backscatter, with the "color ratio" of backscatter at the two wavelengths and depolarization ratio giving us aerosol size and shape information, respectively. The origin of these particles may be desert dust lofted by convection. Clearly, the lidar observations have the potential to quantify the global occurrence of both anthropogenic and natural aerosol but cannot by themselves distinguish the two types. It should be possible to establish just how widespread is the modification by man of the sizes and concentrations of the droplets within liquid water clouds when such clouds are embedded within haze. For ice clouds, lidar returns should reveal the frequency with which dust aerosols (natural or anthropogenic) are being lofted and transported large distances, and whether they are significantly modifying the glaciation rates and ice particle sizes of these high-level clouds. An essential first step is to quantify the magnitude of these effects on a global scale.

Thus far we have discussed the evaluation of NWP models using satellite or ground observations to see if they are producing clouds with the correct average properties and the correct PDF of these properties. This requires several months of data. On a global scale, observations over a few years should be sufficient to establish the characteristics associated with the Madden-Julian Oscillation, interannual variability, and possibly El Niño. The ground-based studies have shown that in regions where there are abundant observations, NWP mesoscale models have skill in producing the right cloud at the right time, and this skill can be evaluated on a monthly basis. In data-sparse regions, this skill is much lower so only an evaluation of the correct statistical properties of the clouds can be achieved. Evaluating the fidelity of clouds in climate models run for many years is more difficult; only the global statistics of mean cloud properties, their PDFs, and the temporal fluctuations of these metrics can be determined. However, experiments in which climate models have been run in a forecast mode indicate that some systematic biases (e.g., in the cloud and humidity fields), noticed in the climate mode, appear in a few days in the model. This suggests that the evaluation of clouds in climate models may be done in part based on short-term experiments and high-frequency observations (e.g., data from field experiments if the model is initialized with large-scale forcings from this experiment). One word of warning is in order: currently, active satellites with active radars and lidars are in sun-synchronous orbits and thus information on the diurnal cycle is limited.

Cloud Feedbacks in a Changing Climate

Cloud and radiative observations are only available for a short period (at best for about 25 years, more generally for just a few years), and no climate variation occurring at this timescale may be considered as an analog of long-term climate change. Until long time series (three decades or more) of cloud and radiation data become available, it is hopeless to assess directly the response of clouds to global climate changes using observations and to compare it with model simulations. This is even more true since establishing long-term trends based on satellite or surface-based measurements is made very difficult by problems such as changes in instrument calibration, or satellite drift in altitude, etc. Once reliable and long time series of observations (of clouds but not only) become available, it might become easier to assess cloud feedback processes directly from observations and then to evaluate cloud feedbacks in climate models. In the meantime, available observational records can be useful in assessing the natural climate variability on various time scales, and also in investigating the physics that controls cloud changes and variability. For this purpose, some approaches have been developed over the last few years that take advantage of the available observations to assess climate model simulations in a way that may be relevant for assessing cloud–climate feedbacks.

Cloud feedbacks are related to the response of clouds to changing climate conditions. To have confidence in the feedbacks produced by climate models, it is therefore not sufficient to evaluate mean cloud properties. Assessing the *sensitivity* of clouds to changing environmental conditions is more likely to be relevant for assessing the realism of the simulated feedbacks.

For this purpose, one approach is to use compositing techniques to assess, in models and in observations, how clouds change in association with dynamic or thermodynamic conditions (e.g., with changes in lower tropospheric stability, in the intensity of large-scale rising or sinking motions in the free troposphere, in humidity and temperature). For this purpose, observations and model simulations are not only compared in terms of geographical distributions but also in terms of covariations between several variables. For instance, recognizing that many cloud properties (in particular, the prominent cloud type) are controlled to a large degree by the large-scale atmospheric circulation, several studies have stratified cloud observations as a function of dynamic regimes (cf. Bretherton and Hartmann, this volume) and then investigated how, for specified dynamic conditions, these cloud properties varied with other environmental conditions, such as surface temperature, static stability, or horizontal advections (Bony et al. 1997; Williams et al. 2003, Bony et al. 2004, Norris and Iacobellis 2005). Other studies have decomposed global cloudiness into a small number of prominent cloud regimes and used this decomposition to understand and assess the response of clouds to long-term climate changes (e.g., Williams and Tselioudis 2007).

Such an approach makes it possible to evaluate simulations from idealized simulations having different geographical distributions of the dynamic features (e.g., aquaplanets) by using observations. Decomposing the large-scale feedback mechanisms or cloud changes in terms of a series of composites also makes it possible to bridge more easily climate studies with process studies. Once a cloud process or a sensitivity is identified as a key component of cloud–climate feedbacks, more detailed investigations using uni-dimensional models, cloud resolving models, or more detailed observations such as those collected during campaigns such as RICO or BOMEX may be performed to explore more deeply the underlying physics.

Cloud–Climate Metrics for Assessing the Relative Reliability of Climate Change–Cloud Feedbacks Produced by Climate Models

With the realization and the open availability of a large coordinated set of climate simulations performed by a large number of climate models (Meehl et al. 2007), the question now arises whether some model results are more reliable than others. Giving more importance (or more weight) to models that seem to perform better in simulating the current climate is sometimes presented as a way to reduce uncertainties in climate projections (Murphy et al. 2004). To address this question, the climate modeling community is currently developing efforts to define a basket of "metrics" to assess the relative merits of the different climate models in reproducing observed features (always remembering that the models may have common errors that need to be identified and corrected, because such errors may offset one another so that some of our present crude criteria for assessing models are satisfied whereas in truth the models are flawed). This effort, in some way, extends to climate models a procedure that has been routinely applied to NWP models for thirty years. However, it raises many questions and concerns.

As provocatively asked during the presentation of the IPCC AR4: "Might the 5th Assessment Report of the IPCC be the end of models democracy?" (IPCC 2007). Indeed, thus far, different climate models have all been treated equally, in terms of their ability to simulate climate change projections. However, we feel that there is a growing desire (and pressure) to rank the different models and to give them different weights depending on their relative ability to reproduce the observed climate.

Certainly, the climate community welcomes the possibility of quantifying, for a large ensemble of climate models and for a wide range of diagnostics, the resemblance between simulations and observations. For example, one may imagine developing metrics focused on the ability of climate models to simulate a realistic diurnal cycle or realistic tropical intraseasonal oscillations. The scrutiny by a very large community of analysts of the simulations performed in support of the IPCC AR4 is already contributing to this very constructively. However, concerns might be expressed regarding the meaning and the future use of these metrics.

As explained above, we still do not know whether some model biases matter more than others for climate change prediction. Common sense suggests that the answer to this question depends on the climate question to be addressed. To assess the reliability of climate model projections in regions dominated by monsoon or ENSO phenomena, one might find useful metrics focused on the simulation of these processes. However, there is some danger in the use of nonspecific metrics based on mean climate features (e.g., mean cloudiness or mean radiative fluxes) to assess the relative reliability of different model estimates of global climate sensitivity (besides, it turns out that climate models producing very different cloud responses to climate change may not be distinguishable in their simulation of mean cloud properties in the current climate). To address this question, we need instead to encourage the development and use of some process-based metrics assessing the ability of climate models to simulate cloud relationships, processes, or composites shown to play a critical role in climate change–cloud feedbacks. Again, analyses and idealized studies of the kind described earlier in this chapter provide some guidance about the processes to be considered in such metrics.

Ways to Reduce Critical Uncertainties in the Prediction of Clouds in a Changing Climate

Comparison of climate simulations with observations reveals a large number of systematic biases in current models. Faced with the long-standing biases of climate models and uncertainties in climate change projections, the optimal way to improve models is still open to question.

Resolution

The increase of the (horizontal and vertical) resolution of large-scale models is often cited as a way to improve model simulations. The experience of many modeling centers indicates that increasing the resolution does reduce some biases, such as the occurrence and strength of midlatitude storms, the simulation of extreme precipitation, or of orographic precipitation. However, it is far from solving all the problems. In particular, the simulation of continental precipitation, of the diurnal cycle, or of the Madden-Julian Oscillations does not improve substantially with resolution. This is the case for many other errors of large-scale models, including the difficulties of representing the cloud processes themselves, such as condensation on aerosol particles, glaciation, the size distribution of the cloud particles and their interaction with radiation, the degree of cloud overlap in the vertical, and the conversion of cloud water into precipitation. We know these are inadequately parameterized and lead to modeled clouds with different characteristics from those indicated by our limited database of cloud observations, but it is unclear if increased resolution will ameliorate the situation. Current models have a higher vertical resolution in the boundary layer and thus can resolve some of the vertical structure of stratocumulus and, to a lesser extent, fair weather cumulus. We know that mid-level clouds are underestimated in nearly all models. Why is this? Is it because we cannot resolve the position of cloud top and base? Is it because the radiation scheme is not called often enough? Is it because the diagnosed phase is incorrect? Is it because there is no turbulent mixing scheme outside the boundary layer? It is also becoming clear that supercooled clouds commonly form and are widespread and persistent and have potentially important radiative effects, but are scarcely represented in the models. Is this also a resolution problem?

Complexity

Another avenue of model development is the increase of complexity. Coupled ocean–atmosphere models are now coupled to complex land-surface schemes, aerosol modules, chemistry, carbon cycle, etc. to form so-called Earth System Models. This allows us to investigate new climate feedbacks (such as carbon–climate feedbacks) but does not reduce the uncertainty in climate change projections. On the contrary, intermodel differences in regional precipitation changes and in climate sensitivity are often amplified by carbon-cycle feedbacks (which are very sensitive to precipitation and climate sensitivity changes) or aerosol feedbacks.

Physical Parameterizations

Improving the physical parameterizations used in large-scale models (in particular, the representation of turbulent, convective cloud processes, and the interaction with radiation) seems to be the most efficient way to reduce uncertainties in model projections of the future climate. However, improving parameterizations is difficult, the number of people actively involved in this work is fairly small at the present, and the progress is slow. National and international funding agencies might play a role in encouraging these activities. Nevertheless, with the arrival of new cloud observations and with the increasing interactions and collaborations between meso- and large-scale modelers, one may expect more progress over the next few years than there has been in the past. From the observations, can we demonstrate that we really need dual moment schemes to represent ice and liquid water, and can we show that such schemes are adequately constrained to lead to improvements? Can observations reliably confirm the existence of large cloud-free regions (see Kärcher and Spichtinger, this volume, and references therein), which are very highly supersaturated with respect to ice, and do we need to adjust our parameterization schemes to take this into account? What level of sophistication is needed in the treatment of aerosols?

Using Cloud-resolving Models instead of Cloud Parameterizations in Climate Models

Now that "super-parameterizations" and "global CRMs" have become available (cf. Grabowski and Petch; Collins and Satoh, both this volume), using CRMs instead of cloud parameterizations might constitute an option to reduce the uncertainty in cloud feedbacks associated with cloud parameterizations. Although these new approaches are promising, they are unlikely, however, to solve the cloud–climate problem issue in the near future for at least for two reasons. First, these approaches are computationally very expensive. Thus, it seems unlikely that ensembles of century-scale simulations can be performed with such models to study changes in the global climate or in climate extremes. Second, the resolution of CRMs is insufficient to resolve boundary-layer turbulence or cloud microphysics and, therefore, parameterizations are still required. The results obtained with these models are likely to depend on these parameterizations and at least on some poorly constrained parameters. This dependence should be explored and quantified before the cloud feedbacks produced by these models can be considered less uncertain than those derived from large-scale models.

On the other hand, sensitivity experiments performed with global CRMs or super-parameterizations can be very instructive to explore the physics of cloud feedbacks and climate sensitivity. It would be very valuable, for example, to understand why an aquaplanet global CRM (Miura et al. 2005) and a GCM embedding a two-dimensional CRM within each grid box instead of a cloud parameterization (Wyant, Khairoutdinov et al. 2006) both predict a climate sensitivity weaker than estimated by most global climate models. It would also be valuable (and computationally cheaper) to perform climate simulations by embedding a CRM or a LES over a limited domain of the Earth instead of globally (e.g., a LES in subtropical regions predominantly covered by boundary-layer clouds). A complementary and constructive (rather than competitive) interaction between large- and mesoscale modeling approaches to study the cloud–climate problem is strongly required.

Conclusion

With the arrival of new and powerful observational datasets, particularly the new space-based active radar and lidar sensors in the "A-Train," we are entering a new era for the evaluation of clouds in large-scale models. Observations with active sensors have already demonstrated that NWP models have skill in representing clouds in the right place and the right time, and have also identified some shortcomings. It will soon be possible to assess key aspects of the simulation of clouds, such as the three-dimensional distribution of cloud layers, the cloud water phase, the cloud precipitation efficiency, and the physical and radiative properties of shallow-level clouds. As these aspects have the potential to affect the response of these crucial shallow tropical clouds to climate change, their evaluation in the current climate under a large variety of environmental conditions will allow us to assess better the realism of their change in the future. Moreover, as high-resolution models are increasingly used to assess and develop physical parameterizations, as well as to investigate cloud-in-climate issues, we recommend that new satellite data be used to evaluate the cloud distributions produced by high-resolution models, including operational NWP models.

We emphasize that the better our physical understanding is of the response of clouds to climate change, the more efficient the strategy for evaluating this response will be. Thus, developing a strategy of evaluation of climate change–cloud feedbacks requires efforts in analyzing and unraveling the physical mechanisms underlying these feedbacks. For this purpose, a promising approach consists of conducting idealized studies using a hierarchy of climate models of different complexities. However, to reduce the uncertainties in cloud–climate feedback processes and improve climate models, it is not sufficient to point out deficiencies in a particular process; physical parameterizations must be improved if we want these deficiencies to be remedied. For this, it is important to keep developing collaborations between the large- and mesoscale cloud communities, as well as between the modeling and observational communities.

The radars and lidars now in space should enable us to observe the global vertical distribution of clouds and aerosols and aid in ascertaining the degree to which aerosols are modifying both warm and cold clouds, as well as the geographic extent of any modification. These measurements should allow us to quantify, for the first time, the effect aerosols are having on the present climate, and hence reduce the large uncertainties in the effects of aerosols on the future climate.

Turning to future satellite-observing systems, we have some concern that the long time series of global monitoring of TOA radiation, which is now being carried out with the CERES sensors, may not continue, although Mega-Tropique may fill the gap but only at low latitudes. We look forward to the launch of the ESA/JAXA EarthCARE mission (in 2013), which will embark a cloud radar and lidar on the same platform. The high spectral resolution lidar should provide direct observations of the optical depths of thin cirrus and aerosols and characterize the aerosol and ice cloud particles. The radar will have improved sensitivity and so should detect more of the high thin ice high clouds as well as the lower-level water clouds, while the Doppler capability should help to characterize the vertical cloud motions and thus contribute to the evaluation of convective parameterization schemes, provide information on ice sedimentation velocities within extensive cirrus decks to inform the model ice schemes, and quantify the drizzling rates in low-level water clouds.

Many small-scale cloud processes remain which must be parameterized in the models; a better understanding of them is needed but cannot be provided from space. Examples include the entrainment and detrainment for both layer and convective clouds; the growth of ice particles from the vapor, their aggregation, riming, and subsequent evaporation; the warm rain coalescence process and the mechanisms that lead to the production and persistence of supercooled layer clouds. Progress can best be provided through detailed observation, whether *in situ* or remotely from the ground, of the evolving physical and dynamic variables. One particularly glaring gap remains: We still have no technique to observe the humidity structure of the atmosphere in three dimensions with a high enough resolution to characterize its PDF within the model grid box even in clear air. To achieve this, when clouds are present, is an even greater challenge.

Acknowledgments

Thoughtful comments from Tony Slingo, Jean-Louis Brenguier, Johannes Quaas, Jon Petch, Graham Feingold, Jost Heintzenberg, Patrick Chuang and an anonymous reviewer are gratefully acknowledged. They were very helpful to clarify some of the issues discussed in this paper.

References

Ansmann, A., M. Tesche, D. Althausen et al. 2008. Ice formation in Saharan dust over northern Africa observed during SAMUM. *J. Geophys. Res.* **113**:D04210.

Bodas-Salcedo, A., M. J. Webb, M. E. Brooks et al. 2008. Evaluation of cloud systems in the Met Office global forecast model using CloudSat data. *Geophy. Res. Abst.* **10**:EGU2008-A-10141.

Bony, S., and J.-L. Dufresne. 2005. Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.* **32**:L20806.

Bony, S., J.-L. Dufresne, H. Le Treut, J.-J. Morcrette, and C. Senior. 2004. On dynamic and thermodynamic components of cloud changes. *Climate Dyn.* **22**:71–86.

Bony, S., and K. A. Emanuel. 2001. A parameterization of the cloudiness associated with cumulus convection: Evaluation using TOGA COARE data. *J. Atmos. Sci.* **58**:3158–3183.

Bony, S., K. M. Lau, and Y. C. Sud. 1997. Sea surface temperature and large-scale circulation influences on tropical greenhouse effect and cloud radiative forcing. *J. Climate* **10**:2055–2077.

Chepfer, H., S. Bony, M. Chiriaco et al. 2008. Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. *Geophys. Res. Lett.* **35**:L15804.

Delanoë, J., and R. J. Hogan. 2008. A variational scheme for retrieving ice cloud properties from combined radar, lidar and infrared radiometer. *J. Geophys. Res.* **113**:D07204.

DeMott, P. J., K. Sassen, M. R. Poellot et al. 2003. African dust aerosols as atmospheric ice nuclei. *J. Geophys. Res.* **30(14)**:1732.

Hogan, R. J., M. D. Behera, E. J. O'Connor, and A. J. Illingworth. 2004. Estimate of the global distribution of stratiform supercooled liquid water clouds using the LITE lidar. *Geophys. Res. Lett.* **31**:L05106.

Hogan, R. J., and A. J. Illingworth. 2000. Cloud overlap statistics from long-term radar observations. *Q. J. Roy. Meteor. Soc.* **126**:2903–2909.

Illingworth, A. J., R. J. Hogan, E. J. O'Connor et al. 2007. Cloudnet, continuous evaluation of cloud profiles in seven operational models using ground-based observations. *Bull. Amer. Meteor. Soc.* **88(6)**:883–898.

IPCC. 2007. Climate Change. 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, ed. S. Solomon, D. Qin, M. Manning et al. New York: Cambridge Univ. Press.

John, V. O., and B. J. Soden. 2007. Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophys. Res. Lett.* **34**:L18704.

Klein, S. A., and C. Jakob. 1999. Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.* **127**:2514–2531.

Liu, C., M. W. Moncrieff, and W. W. Grabowski. 2001. Explicit and parameterized realizations of convective cloud systems in TOGA COARE. *Mon. Wea. Rev.* **129**:1689–1703.

Mather, J. H., T. P. Ackerman, W. E. Clemens et al. 1998. An atmospheric radiation and cloud station in the tropical western Pacific. *Bull. Amer. Meteor. Soc.* **79**:627–642.

Medeiros, B., B. Stevens, I. M. Held et al. 2008. Aquaplanets, climate sensitivity, and low clouds. *J. Climate*, in press.

Meehl, G. A., C. Covey, T. Delworth et al. 2007. The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.* **88**:1383–1394.

Miura, H., H. Tomita, T. Nasuno et al. 2005. A climate sensitivity test using a global cloud resolving model under an aqua planet condition. *Geophys. Res. Lett.* **32**:L19717.
Müller, D., F. Wagner, U. Wandinger et al. 2000. Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization. *Exp. Appl. Opt.* **39(12)**:1879–1892.
Müller, D., U. Wandinger, D. Althausen, and M. Fiebig. 2001. Comprehensive particle characterization from 3-wavelength Raman-lidar observations: Case study. *Appl. Opt.* **40(27)**:4863–4869.
Murphy, J. M., D. M. Sexton, D. N. Barnett et al. 2004. Quantification of modeling uncertainties in a large ensemble of climate change simulations. *Nature* **429**:768–772.
Norris, J. R., and S. F. Iacobellis. 2005. North Pacific cloud feedbacks inferred from synoptic-scale dynamic and thermodynamic relationships. *J. Climate* **18:**4862–4878.
Ramanathan, V., C. Chung, D. Kim et al. 2005. Atmospheric brown clouds: Impacts on South Asian climate and hydrological cycle. *PNAS* **102**:5326–5333.
Soden, B. J., and I. M. Held. 2006. An assessment of climate feedbacks in coupled ocean–atmosphere models. *J. Climate* **19**:3354–3360.
Tompkins, A. M. 2002. A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.* **59**:1917–1942.
Tsushima Y., S. Emori, T. Ogura et al. 2006. Importance of the mixed-phase cloud distribution in the control climate for assessing the response of clouds to carbon dioxide increase: a multi-model study. *Climate Dyn.* **27(2–3)**:113–126.
Webb, M., C. Senior, S. Bony, and J.-J. Morcrette. 2001. Combining ERBE and ISCCP data to assess clouds in three climate models. *Climate Dyn.* **17**:905–922.
Webb, M., C. A. Senior, D. M. H. Sexton et al. 2006. On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dyn.* **27 (1)**:17–38.
Wilkinson, J. M., R. J. Hogan, and A. J. Illingworth. 2008. Use of a lidar forward model for global comparisons of cloud fraction between the ICESat lidar and the ECMWF model. *Mon. Wea. Rev.*, in press.
Williams, K. D., M. A. Ringer, and C. A. Senior. 2003. Evaluating the cloud response to climate change and current climate variability. *Climate Dyn.* **20**:705–721.
Williams, K. D., and G. Tselioudis. 2007. GCM intercomparison of global cloud regimes: Present-day evaluation and climate change response. *Climate Dyn.* **29**:231–250.
Wyant, M. C., C. S. Bretherton, J. T. Bacmeister et al. 2006. A comparison of tropical cloud properties and responses in GCMs using mid-tropospheric vertical velocity. *Climate Dyn.* **27**:261–279.
Wyant, M. C., M. Khairoutdinov, and C. S. Bretherton. 2006. Climate sensitivity and cloud response of a GCM with a superparameterization. *Geophys. Res. Lett.* **33**:L06714.
Zhang, M. H., W. Y. Lin, S. A. Klein et al. 2005. Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *J. Geophys. Res.* **110**:D15S02.