The Evaluation of Cloud Parametrizations in GCMs - Are We Doing The Best We Can?

Christian Jakob

European Centre for Medium-Range Weather Forecasts

1 Introduction

A realistic simulation of clouds and their effects on the atmosphere and the surface in global atmospheric models as they are used for Numerical Weather Prediction (NWP) and climate simulations has been shown to be of major importance. Therefore large efforts are being made to improve the parametrization of clouds in these models. The last few decades have seen major improvements in the physical realism of cloud parametrizations. Almost all GCMs now use a prognostic equation to represent the time evolution of cloud condensate (e.g., Sundqvist, 1978; Del Genio et al., 1996; Fowler et al., 1996). The treatment of a second crucial parameter, namely cloud cover, is more diverse and ranges from simple diagnostic relations (e.g., Slingo, 1987; Sundqvist, 1978; Smith, 1990) to a fully prognostic treatment (e.g., Tiedtke, 1993).

A crucial aspect when modelling a physical system, or parts of it, is the evaluation of the realism of the model results using observations of the phenomenon in question. With increasing complexity of the model the evaluation of specific components becomes increasingly difficult. Current GCMs can probably be considered as one of the most complex models of a physical system, not only because of the large number of processes described in them, but also because of the non-linear character of their interactions. Hence, evaluating the "model clouds" against their real-life counterparts is a complex problem in itself.

Before embarking on any kind of evaluation study it is important to define its purpose, which must depend on the "end-user" of the results of the study. A forecaster charged with predicting clouds will be interested in evaluating the usefulness of cloud products from NWP systems, but very likely without any interest in the reasons for the level of performance that the evaluation reveals. In contrast the developer of a cloud parametrization wants to devise evaluation studies that expose flaws in the parametrization and, if possible, reveal reasons for their occurence. It is the purpose of this paper to assess current evaluation techniques from a modeller's viewpoint. In doing so it cannot be the intention to provide a comprehensive review of all cloud validation studies ever carried out for GCMs. Instead the ECMWF global forecast model will be used to illustrate various methods used for the evaluation of its cloud simulations.

A fairly large number of studies have dealt with the problem of evaluating the representation of clouds and their radiative effects in GCMs. Despite the large variety in their approaches most of them can be placed in one of two main categories: the evaluation of the model climate and case studies. The paper will give an overview over evaluation techniques for both these categories using the ECMWF global forecast model as an example to highlight the various issues and problems related to the evaluation methods. Section 2 briefly introduces the ECMWF cloud parametrization. Section 3 will show examples of the evaluation clouds and radiation in the ECMWF model climate. Section 4 then provides an overview over different techniques used in evaluating cloud parametrizations using case studies. It will be shown that there exists an unfortunate gap between model climate and case evaluations. Section 5 will introduce the technique of compositing by dynamical regime, which has only recently been used for cloud parametrization evaluation. Recent results of studies carried out at ECMWF will serve as examples to illustrate how this technique might bridge the gap between model climate and case study and thereby provide new insight into the workings of a particular cloud parametrization. The paper will close by proposing a strategy for cloud evaluation in GCMs in Section 6 that integrates most of the current techniques into a coherent procedure.

2 The ECMWF cloud parametrization

The ECMWF prognostic cloud scheme developed by Tiedtke (1993) has been implemented in the operational system in April 1995. This scheme includes prognostic equations for cloud fraction and liquid/ice water content and has been shown to provide a much more realistic description of the cloud cover than the previous diagnostic scheme described by Slingo (1987). The inclusion of cloud fraction as a prognostic variable is an original feature of this scheme since most other prognostic cloud schemes treat cloud fraction by means of diagnostic relations (e.g. Sundqvist, 1988, Smith,1990).

The prognostic scheme consists of two equations, one for liquid/ice water l and one for cloud fraction a:

$$\frac{\partial l}{\partial t} = A(l) + S^l_{conv} + S^l_{bl} + C^l - E - P - D_{entr}$$

$$\frac{\partial a}{\partial t} = A(a) + S^a_{conv} + S^a_{bl} + C^a - D_{evap}$$

where $A(l)$ and $A(a)$ represent the advection of liquid/ice water and cloud fraction, S^l_{conv} and S^a_{conv} represent the source terms of liquid/ice water and cloud fraction from moist convection processes, S^l_{bl} and S^a_{bl} represent the source terms of liquid/ice water and cloud fraction due to boundary layer turbulence, C^l and C^a are the formation of liquid/ice water and cloud fraction by stratiform condensation processes, E is the evaporation rate, P is the precipitation rate, D_{entr} is the destruction of liquid/ice water due to cloud top entrainment and D_{evap} is the reduction of cloud fraction by evaporation. The scheme allows the ice and liquid water to co-exist (mixed phase) base on a simple quadratic temperature dependence between 0°C and -23°C (see Jakob, 1995), therefore requiring only one prognostic equation for condensate.

The scheme accounts explicitely for the physical sources and sinks leading to the production and dissipation of clouds. It is therefore strongly coupled to the other physical parametrizations in the ECMWF model (radiation, turbulence, deep and shallow convection). Clouds formed by convective processes are parametrized by considering them to be condensates produced by cumulus updrafts and detrained in the environmental air. Clouds are also assumed to be formed by non-convective processes (e.g. large-scale lifting of moist air, radiative cooling, etc). Evaporation of clouds is described by two processes in connection with large-scale and cumulus-induced descent and adiabatic heating and by turbulent mixing of cloud air with unsaturated environmental air. Precipitation processes are represented differently for pure ice clouds and for mixed phase and pure water clouds. The rain and snow formed is removed from the atmospheric column immediately but can evaporate and interact with the cloud water in the layers in passes through. The precipitation process in ice clouds is treated as a sedimentation of the ice particles. Ice settling into cloudy areas is treated as source for cloud ice in the layer below, wheras ice settling into clear-sky is converted into snow. For mixed phase and pure water clouds a parametrization describing the conversion of cloud droplets into raindrops is taken from Sundqvist (1978). It represents both the collision process and the Bergeron-Findeisen mechanism. Evaporation of rain/snow is described by a scheme following Kessler (1969) and only takes place when the grid mean relative humidity is below a threshold value.

3 Evaluating the model climate

3.1 Broadband radiative fluxes

The main impetus for the inclusion of clouds into GCMs is their interaction with radiation. It is therefore not surprising that one of the most common techniques to evaluate cloud parametrizations is to compare radiative fluxes produced by the model to those observed by satellites at the top of the atmosphere. One of the most frequently used data set for such comparisons is the broadband flux measurements gathered during the Earth Radiation Budget Experiment (ERBE; Barkstrom and Smith, 1986).

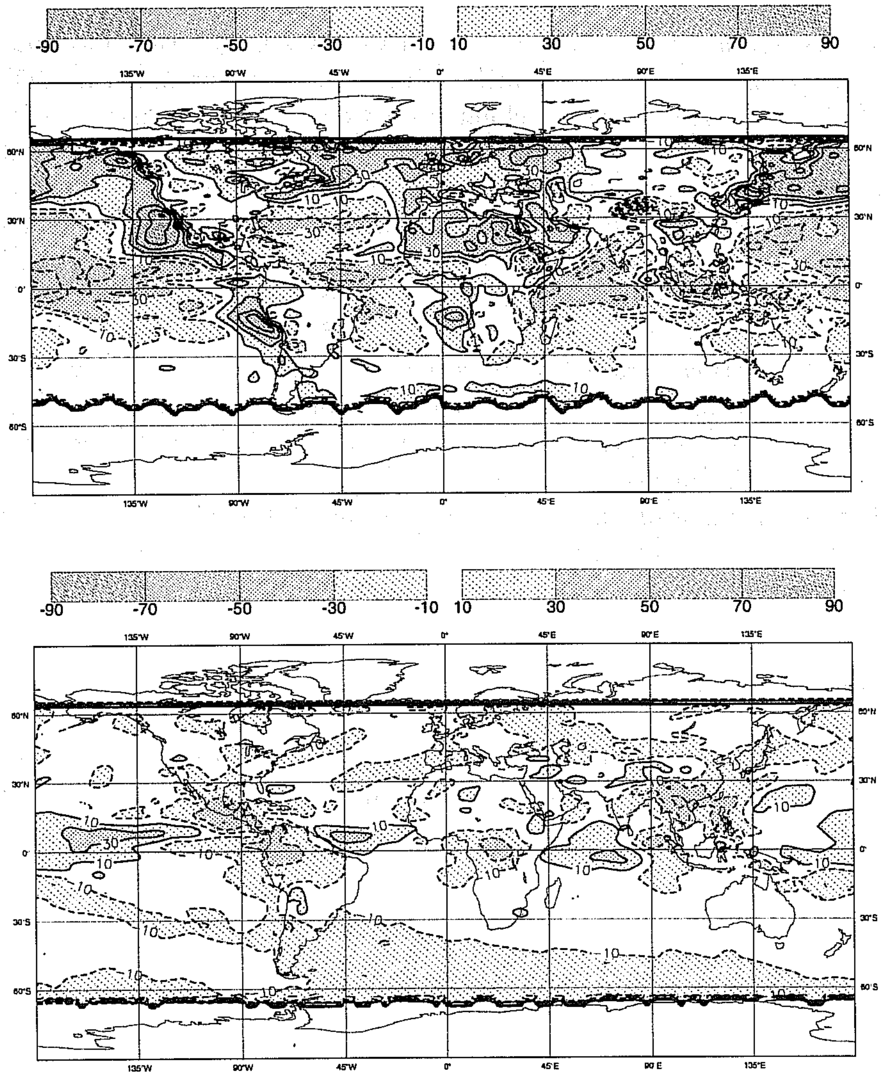

Figure 1: Difference in shortwave (top) and longwave (bottom) radiation at the top of the atmosphere between a model integration and ERBE observations for June/July/August 1987. Positive differences are shown as solid contours, negative differences as dashed. The contour interval is 20 Wm^{-2} with shading starting at ± 10 Wm^{-2}. The model integration is carried out with CY18R6 of the ECMWF model at T63L31 resolution. The initial date is 1 May 1987. SSTs are time varying and prescribed.

Figure 1 shows the difference in TOA shortwave (top panel) and longwave (bottom panel) between simulations for June/July/August (JJA) 1987 carried out with the ECMWF model as used operationally in most of 1998, but at lower horizontal resolution, and the ERBE observations. The model integrations start on 1 May 1987 at a resolution of T63L31 and use prescribed, time-varying sea surface temperatures (SSTs). Upward fluxes are taken as negative so that a negative difference indicates a too strong upward flux, i.e., too much reflection in the shortwave case and too high emission in the longwave part of the spectrum.

Several regions of erroneous TOA radiation emerge for both spectral regions. TOA shortwave radiation is overestimated (pointing to too little reflection) in the extratropics, predominantly over the oceans but also over land, over the eastern part of the sub-tropical oceans and over the Sahara region. It is underestimated over most of the deep tropics and over the western parts of the subtropical oceans. The outgoing longwave radiation (OLR) is underestimated (positive difference) over much of the ocean in the tropical belt, strongly overestimated over the tropical continents and overestimated to a lesser extent in the extratropics.

One of the advantages of evaluating broadband radiative fluxes is that they form part of the planet's energy balance and are therefore a key quantity to be accurately simulated in GCMs. The biggest drawback of the technique when applied to cloud evaluation is that information on the overall model radiative fluxes per sé does not deliver any information about the radiative effect of the model clouds. The errors seen could result not only from erroneous model clouds, but also from the wrong description of the surface albedo in case of the shortwave radiation or from wrong surface temperatures or poor simulations of the water vapour distribution in the longwave part of the spectrum. The coincidence of some of the error patterns with regions dominated by particular cloud types does however raise the suspicion of problems in the description of the radiative effect of these types of clouds. The strong overestimation of TOA shortwave radiation in regions of extensive coverage with stratocumulus off the West coasts of the subtropical continents could well point to a problem in their representation. The introduction of more a priori information, e.g., the use of the relatively good knowledge of the albedo of the sea surface, also makes it more likely that the errors identified in Figure 1 are related to errors in the radiative behaviour of clouds. On the other hand, the large error in shortwave radiation identified over the Sahara, a region of almost no cloud occurrence, probably indicates a problem in the description of surface albedo.

3.2 Cloud radiative forcing

In order to better understand which of the errors identified above are due to cloud radiative effects, a better variable to compare is the cloud radiative forcing as derived from satellite observations. The two components of cloud radiative forcings can be defined as (e.g.,Ellis, 1978)

$$SWCRF = S - S_{clear}, \tag{1}$$

and

$$LWCRF = L - L_{clear}, \tag{2}$$

where $SWCRF$ and $LWCRF$ stand for shortwave cloud radiative forcing and longwave cloud radiative forcing respectively. L and S are the longwave and shortwave radiative fluxes and L_{clear} and S_{clear} their respective clear-sky values.

The CRF at the TOA can easily be derived in a model by storing the clear-sky radiative fluxes at each gridpoint and compare them to the all-sky fluxes. Note that this procedure differs from the way the cloud radiative forcing is derived from data in that the clear-sky radiation in the model is calculated for cloudy columns by just ignoring the cloud variables, but still using the water vapour and temperature profiles of a cloudy column. In contrast, the cloud radiative forcing in the data is derived by comparing cloudy columns with neighbouring (both in space and time) true clear-sky columns. The difference introduced this way can amount to a few Wm^{-2} (Cess and Potter, 1987; Cess et al., 1992) and needs to be considered in cases of small model errors. As will be shown below, the errors of the ECMWF model in many regions are in excess of 10 Wm^{-2} so that this effect should not affect the conclusions drawn here.

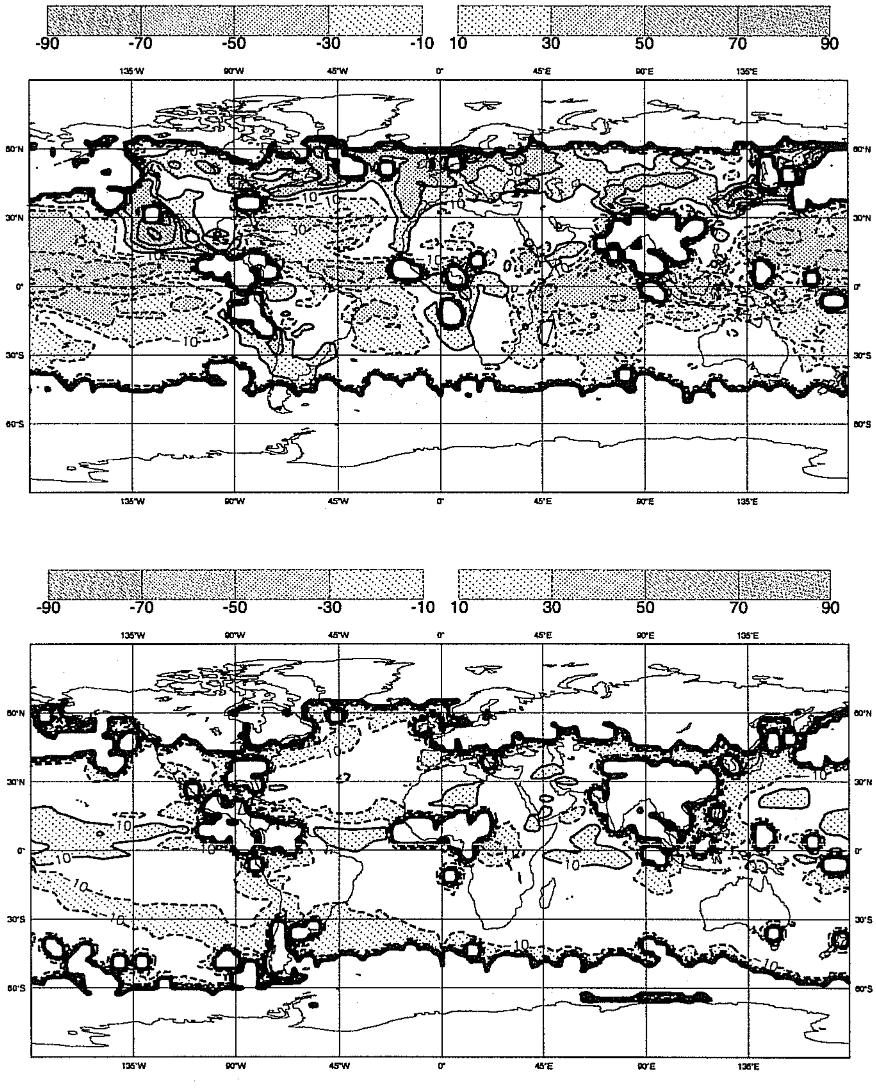

Figure 2: Difference in shortwave (top) and longwave (bottom) cloud radiative forcing at the top of the atmosphere between a model integration and ERBE observations for June/July/August 1987. The model integration is carried out with CY18R6 of the ECMWF model at T63L31 resolution. The initial date is 1 May 1987. SSTs are time varying and prescribed. White areas surrounded by heavy contouring denote missing data.

Figure 2 shows the difference in shortwave (top panel) and longwave (bottom panel) cloud radiative forcing for the same model simulation as above. It is obvious that most of the pattern and size of the errors for both spectral regions are very similar to those in the full radiative fluxes shown in Figure 1. This confirms the suspicions raised in the previous subsection and identifies clouds as the major source for error in the radiative fluxes at the top of the model atmosphere. This is not too surprising, since the knowledge about clear-sky radiative transfer is much further advanced than that of the representation of clouds and their interaction with radiative fluxes. The errors identified in the previous two figures for the ECMWF model can therefore be summarized as

- too reflective clouds (solar) over the deep tropical oceans and the western parts of the subtropical oceans
- too little reflection (solar) from clouds over the eastern part of the subtropical oceans and in the extratropics
- excessively large cloud effects in the longwave over ocean areas in the deep tropics

The advantage of using model climate comparisons to cloud radiative forcing observations is that it is possible to study the overall radiative effects that the clouds exert on the model atmosphere and thereby to assess whether the net effect model clouds have is properly represented. One of the major drawbacks of the technique is that this net effect is the result of many parameters. The overestimated shortwave cloud radiative forcing in the trade cumulus regions described above can therefore be the result of too high cloud fraction, too large cloud liquid water contents, too small assumed particle sizes, the misrepresentation of broken cloud effects in the radiation parametrization or a combination of any of those. From the perspective of cloud parametrization this is an extremely dissatisfying result since it provides no guidance to where the emphasis for future development should be. All one can learn is where the general problem areas (in the geographic sense) are. One step to improving this situation is to try and evaluate the parameters predicted by the cloud parametrization, such as the model cloud fraction.

3.3 Cloud fraction

Figure 3 shows a difference between the mean total cloud fraction for JJA 1987 from the same model simulations as above and that derived from the ISCCP-C2 data set (Rossow and Schiffer, 1983). The main model errors in these extended integrations are

- an underestimation of cloud fraction in the stratocumulus regions off the west coast of the subtropical continents
- an underestimation of cloud fraction over the extratropical oceans
- an overestimation of cloud fraction over the ocean areas of the deep tropics

These errors are consistent with the errors detailed in the cloud radiative forcing. There is however no obvious error in cloud fraction over the trade-cumulus areas that could explain the large errors in cloud radiative forcing there.

The evaluation of cloud-related parameters in the model climate as highlighted in this and the previous subsections provides a first insight into the model errors for those parameters, and is therefore a useful first step in the model evaluation. The use of broadband radiation at the top of the atmosphere in the form of net fluxes or cloud radiative forcing is appealing, since ultimately it is one of the major targets for a successful simulation by GCMs. A significant drawback, however, results from the fact that an error in broadband radiation can never be unambiguously linked to a specific shortcoming in the cloud parametrization, since cloud radiative properties depend on a number of macroscopic and microscopic cloud properties (see e.g., Arking, 1991) and since the environment above and below a given cloud layer may also significantly influence the radiative fluxes. This drawback can be overcome

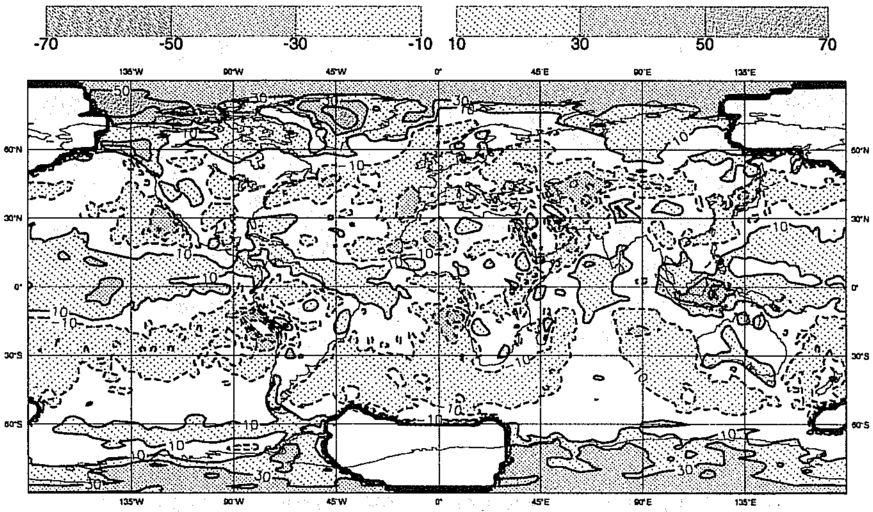

Figure 3: Difference in total cloud cover between a model integration and ISCCP observations for June/July/August 1987. The model integration is carried out with CY18R6 of the ECMWF model at T63L31 resolution. The initial date is 1 May 1987. SSTs are time varying and prescribed. White areas surrounded by heavy contouring denote missing data.

by comparing parameters that are a direct output of the cloud parametrization, such as cloud cover, as highlighted above. The problem here is that direct observations of cloud cover as carried out by surface observers are mostly limited to land areas while cloud cover products from satellites, such as the ISCCP product, do not constitute direct measurements, but involve complex retrieval algorithms. The most serious drawback of a model climate evaluation is however, that in a long model integration feedbacks between different model errors can and will occur, so that an apparently poor representation of clouds in a certain area may be caused not by a failing of the cloud parametrization but for instance by errors in the model's large-scale circulation that may have built up for a variety of reasons. This problem can be overcome to a large extent by using short-range forecasts as produced by NWP systems on a routine basis to build the model "climate". This approach is illustrated using two examples in the following subsection.

3.4 The climate of short-range NWP

As discussed above, the use of short-range NWP for an evaluation of the cloud representation in GCMs avoids the problem of building up large systematic model errors, which can influence the simulation of clouds. This is so, because the large-scale circulation in such forecasts is strongly controlled by the model's initial conditions, which are generated with data assimilation systems taking into account all available observations to correct the model simulation in an statistically optimal way.

An extensive study using this approach for the evaluation of the ECMWF cloud parametrization was carried out by Jakob (1999), who compared monthly cloud cover simulated in short-range forecasts as part of the ECMWF re-analysis (ERA; Gibson et al., 1997) with observations from ISCCP. An example of the findings of this study is shown in Figure 4.

Here the mean total cloud cover from July 1983 to December 1990 from ERA is compared to that derived by ISCCP. Although the model version used in ERA differs from that shown in Figure 3 and although the averaging periods are significantly different, the main error patterns (and in some areas even the magnitude of the errors) in total cloud cover as seen in Figures 3 and 4 are very similar. This suggests that many of the errors in total cloud cover exhibited in the model climate are likely due to

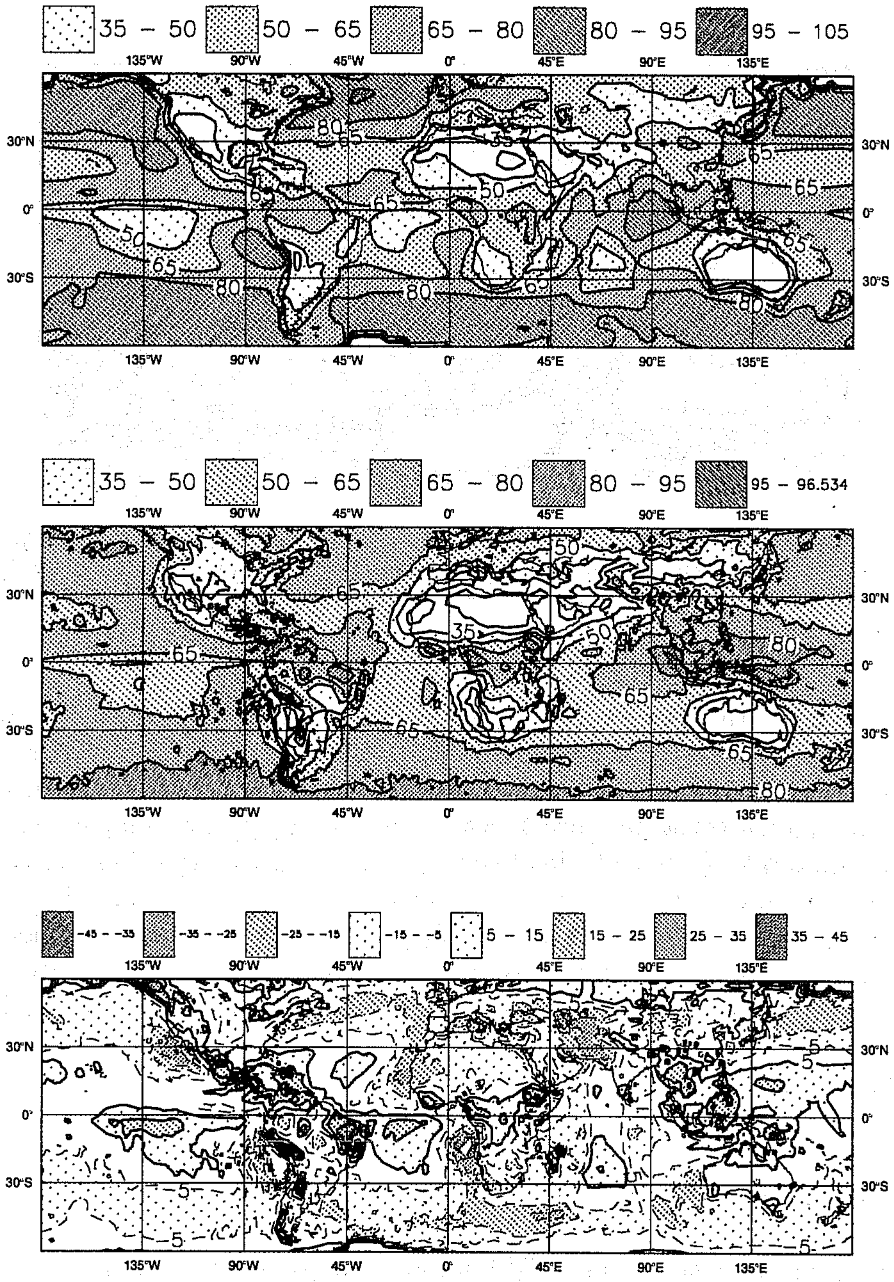

Figure 4: Annual mean of total cloud cover averaged from July 1983 to December 1990 for ISCCP (top), ERA (middle) and ERA minus ISCCP (bottom). Positive differences are depicted by thick solid lines negative by thin dashed lines.

problems in the cloud or related (e.g., convection) parametrizations themselves, an important finding for the improvement of the parametrization.

Another obvious comparison using short-range forecasts is that to cloud observations routinely collected by observers on the ground and included in the reports that are distributed regularly via the Global Telecommunications System (GTS). This type of comparison is carried out on a daily basis at many NWP centres and is used mainly for monitoring forecast performance. One of the immediate benefits of this routine monitoring is that the effects of changes in the cloud parametrization should be reflected in the monitoring statistics. An example for this is shown in Figure 5. This figure shows the evolution of the mean (thick lines) and standard deviation (thin lines) of the forecast error in total cloud cover averaged over many meteorological stations in Europe. The time series shows monthly averages of 60-hour (dashed) and 72-hour (solid) operational forecasts from January 1988 to February 2000. There is a significant reduction in both mean error and standard deviation in April 1995. Not surprisingly, this can be traced back to a change in the model's cloud parametrization, namely the introduction of the parametrization by Tiedtke(1993) into the operational model.

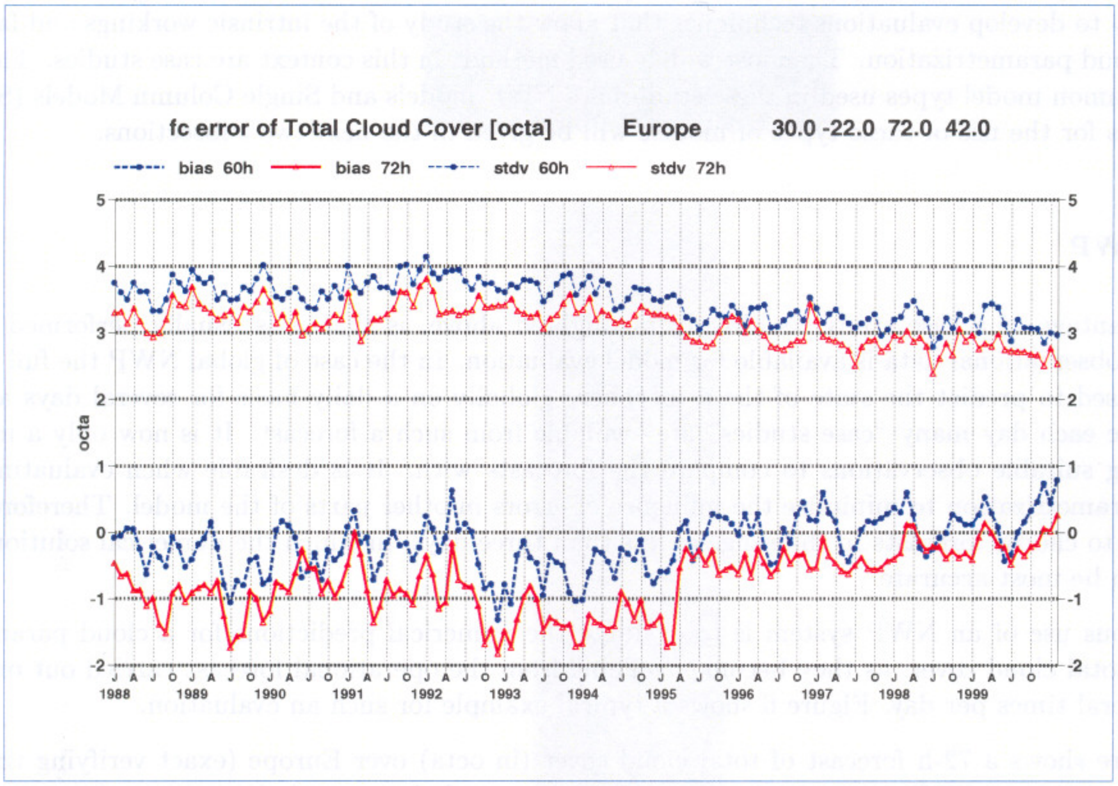

Figure 5: Time series of short-range forecast model errors in total cloud cover (in octa) for Europe (the exact averaging area is indicated in the top right corner) when compared to SYNOP observations. Shown are monthly means of the mean error (thick lines) and the standard deviation (thin lines) for all operational 60-hour forecasts (valid at 00 UTC; dashed) and 72-hour forecasts (valid at 12 UTC; solid).

Although mostly very general in nature, the results of routine monitoring of the forecast results do provide some guidance to the development of cloud parametrization by exposing general problem areas, similar to climate simulations. The big advantage over climate simulations is that, through the use of short-range forecasts, model errors can be more easily ascribed to the parametrization itself since the large-scale flow is captured more realistically.

4 The use of case studies

The previous section has highlighted examples for the evaluation of clouds and cloud-related quantities (e.g., radiative fluxes) on long time-scales. Both long model integrations and averages of short-range forecasts have been shown to be useful for that purpose. Comparisons of this kind have been shown to provide valuable information about major problem areas (often in the geographic sense) in a GCM's cloud representation. It is, however, obvious from the previous discussion, that these kind of studies are not capable of providing crucial insight into the reasons for the model failures. One can of course speculate why certain model climate features exist and one can use intuition and trial and error approaches to correct the model shortcomings. However, without a clear understanding of the causes of the errors this may lead to the undesirable introduction of even more errors that just happen to compensate the already existing ones and hence give a better end result. In the extreme case this approach might lead to an extensive "tuning" exercise in which adjustable model parameters are modified until a satisfactory end result is achieved.

This is clearly not a desirable (even if necessary at times) model development strategy. It is therefore necessary to develop evaluations techniques that allow the study of the intrinsic workings and failings of the cloud parametrization. The most widely used methods in this context are case studies. The two most common model types used in these studies are NWP models and Single Column Models (SCM). Examples for the use of these types of models will be given in the next two subsections.

4.1 NWP

The advantage of a case study, as already highlighted above, is that it is usually performed when detailed observational data is available for model evaluation. In the case of global NWP the full GCM can be used to predict the state of the atmosphere globally on a daily basis for several days ahead. Therefore each day many "case studies" are available from such a forecast. It is now only a matter of finding suitable observations to compare the forecasts with. It is desirable when evaluating the cloud parametrization to minimize the influence of errors in other parts of the model. Therefore it is common to choose forecasts in the range of less than three days, in which the numerical solutions are known to be most accurate.

An obvious use of an NWP system is to evaluate the numerical predictions for a cloud parameter, such as total cloud cover, as they become available from the operational forecast carried out once or even several times per day. Figure 6 shows a typical example for such an evaluation.

The figure shows a 72-h forecast of total cloud cover (in octa) over Europe (exact verifying time: 6 Decmeber 2000, 12 UTC) in comparison to observations made by ground observers at the same time. At first glance the agreement between model and observations is quite striking with the major cloud features well captured. Closer inspection reveals several shortcomings especially over southeastern Europe. Results of a more rigorous statistical evaluation of the figure are indicated at its top and show a negative model bias of about half an octa with a standard deviation of almost 2.5 octa. It is obvious that an evaluation of the kind shown in Figure 6 is far from a comprehensive case study, but it can serve a monitoring purpose if regularly applied to operational NWP forecasts as is done at ECMWF and other NWP centres.

The use of NWP for cloud parametrization evaluation is not restricted, of course, to the use of operational products. If there are dedicated observational campaigns, it is always possible to carry out model simulations covering the period of the observations. This is even desirable, since it allows the evaluation of several versions of a parametrization or even completely different sets of parametrizations. The use of NWP models in this context can provide a natural extension to the SCM approach described below. If data are available at a single point for a given period of time, it is of course feasible to extract the forecast model results at the same point for comparison with the observations. The advantage of using the full GCM for this purpose is twofold. First, there is no need to prescribe the boundary conditions as in the case of an SCM (see below), since those are the result of the full three-dimensional model itself. Second, the parametrization under investigation is working in the environment it is designed for. The disadvantages of the method are first of all the cost (a full GCM needs to be integrated to extract information at a few grid points), and secondly the uncertainties in other model components, which might lead to errors in the parametrization under study. The first problem can be

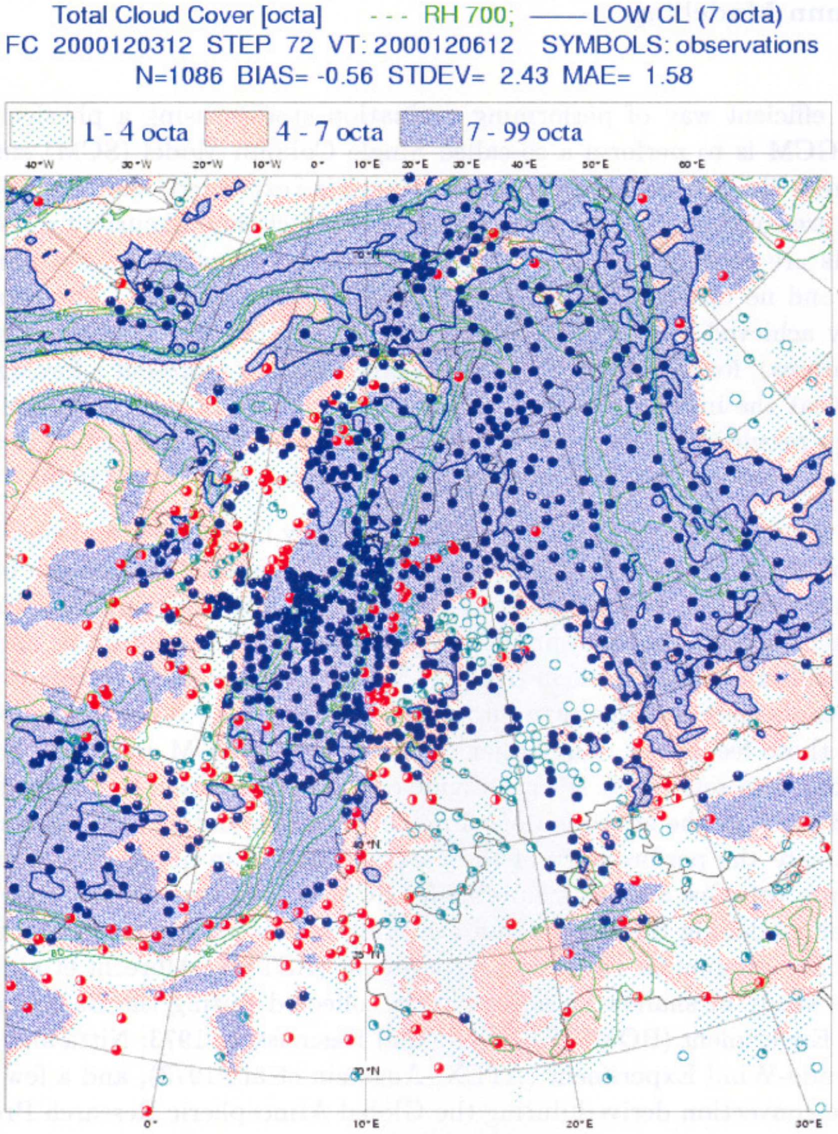

Figure 6: Comparison of the ECMWF 72 hour forecast of total cloud cover (shaded) with obersvations from the SYNOP network (station symbols). The verifying time of the forecast is 6 December 2000, 12 UTC.

overcome by again using products from operational forecasts, which are performed "anyway", while the second problem is minimized by the use of short-range forecasts. ECMWF has been regularly providing data to several field experiments such as the various sites of the Atmospheric Radiation Measurement (ARM; Stokes and Schwartz, 1994) program and the Surface Heat and Energy Budget of the Arctic (SHEBA) experiment (Curry et al., 2000). A number of studies using this kind of data have been carried out (Mace et al., 1998a; Miller et al., 1999; Beesley et al., 2000; Bretherton et al., 2000; Hogan et al., 2000) and provide insight into the ECMWF model's performance.

As highlighted here, the use of a full NWP system for case studies is expensive and it is still cumbersome to store and retrieve all the necessary information to gain insight into possible errors in the parametrization in question. It is therefore desirable to further simplify the approaches to case studies. One such simplification is the use of so-called Single Column Models, which is the subject of the following subsection.

4.2 Single Column Modelling

A computationally efficient way of performing evaluation studies using a physical parametrization package used in a GCM is to perform a so-called Single Column Model (SCM) simulation. Rather than using a full three-dimensional GCM a single column is "extracted" and the results of the model in only this column are considered. This is facilitated by the fact that all current physical parametrizations used in GCMs are assumed to be locally applicable and therefore only require information at a single grid point and no direct interaction between model grid-columns. The three-dimensionality of the GCM is only achieved through the model "dynamics" (i.e., the solution of the grid averaged hydrodynamic equations), for which the physical parametrizations represent a forcing term. Since this is so, it is obvious that the information from neighbouring grid-cells, such as advection terms, which in the full model are provided by the model "dynamics", needs to be prescribed in an SCM. Various techniques have been developed and are summarized in Randall and Cripe (1999).

By design the SCM approach has the advantage to be relatively inexpensive and since only information in a single model column is generated the intricate details of a parametrization can be easily explored. Another potential advantage of the technique is that through the prescription of the boundary conditions, errors created through feedback processes in the full GCM cannot occur and if the boundary conditions were perfect, all errors visible in the SCM would be solely due to errors in the parametrization. However this can sometimes also turn into a disadvantage, e.g., if one wants to study the errors that are caused by those feedbacks. The major difficulty for the SCM approach, however, is to find suitable observational data sets which i) can provide enough information to derive the boundary conditions for the SCM, such as the advection of all model variables and vertical motion, and ii) provide observations to evaluate the performance of the parametrization in question. Because of these high demands on data only very few data sets suitable for SCM studies exist today. Examples of such dat sets include data for modelling the Lagrangian evolution of the marine cloud-topped PBL gathered during the Atlantic Stratocumulus Transition Experiment (ASTEX; Albrecht et al., 1995), a number of data sets for the study of shallow cumulus clouds collected during the Barbados Oceanographic and Meteorological Experiment (BOMEX; Holland and Rasmusson, 1973; Nitta and Esbensen, 1974), and the Atlantic Trade-Wind Experiment (ATEX; Augstein et al., 1973), and a few data sets for the study of penetrative convection derived during the Global Atmospheric Research Program's (GARP) Atlantic Tropical Experiment (GATE; e.g., Houze and Betts, 1981), during TOGA COARE (Webster and Lukas, 1992) and more recently ARM. Designing parametrizations using only a few cases carries the high risk of the nonrepresentativeness of the cases chosen and although SCM simulations are successful, the same parametrizations might not perform well once implemented in the full GCM.

Two major activities have been taking place over the last five to ten years to improve the usefulness of SCMs in parametrization development, in particular that of cloud and convection parametrizations. The first is to gather more observations for the use in case studies. At the forefront of this activity is the ARM programme. The aim of this programme is to collect quasi-continuous data sets related to clouds and cloud-radiation interaction at various locations distributed over the globe. To date ARM is collecting data at sites located in the Southern Great Plains (SGP) of the North American Continent, at Barrow (Alaska) and at two locations in the Tropical Western Pacific (Manus and Nauru island). Since the observations are by design Single Column Observations (SCO), their use in SCM studies is an obvious target. Several studies of this kind using mainly data from the SGP site have already been carried out (e.g., Randall and Cripe, 1999).

A second major activity aimed at the improved use of SCMs in cloud parametrization development is carried out by the World Climate Research Program's (WCRP) Global Energy and Water Cycle Experiment (GEWEX) Cloud System Study (GCSS) (Browning et al., 1993). The basic idea in this programme is to increase the number and therefore representativeness of available SCM case studies through the use of high-resolution cloud models which enable a detailed simulation of cloud processes that need to be parametrized in GCMs. These models are normally referred to as Cloud Resolving Models (CRM). Depending on the cloud type under study the spatial resolution of such models varies from several meters in the horizontal and vertical for boundary layer cloud studies using Large-Eddy Simulation Models (LES; see Mason, 1994, for a review), to several hundred meters in case of Cumulus Ensemble Models (CEM; e.g., Moncrieff et al., 1997, and references therein) used for studies of deep convective systems. The GCSS strategy is to first use these models for the simulation of observed situations. Through the comparison of the model results to observations, the CRMs can be improved and confidence in their ability to accurately simulate cloud processes can be gained. Having established this ability, the CRMs can be used in any kind of simulation, even idealized, to provide the "truth" against which an SCM can be evaluated. GCSS has undertaken a large number of model intercomparisons involving both CRMs and SCMs in order to achieve this goal (e.g., Bechthold et al., 1996; Moeng et al., 1996; Bretherton et al., 1999; Bechthold et al., 2000; Redelsperger et al., 2000; Ryan et al., 2000).

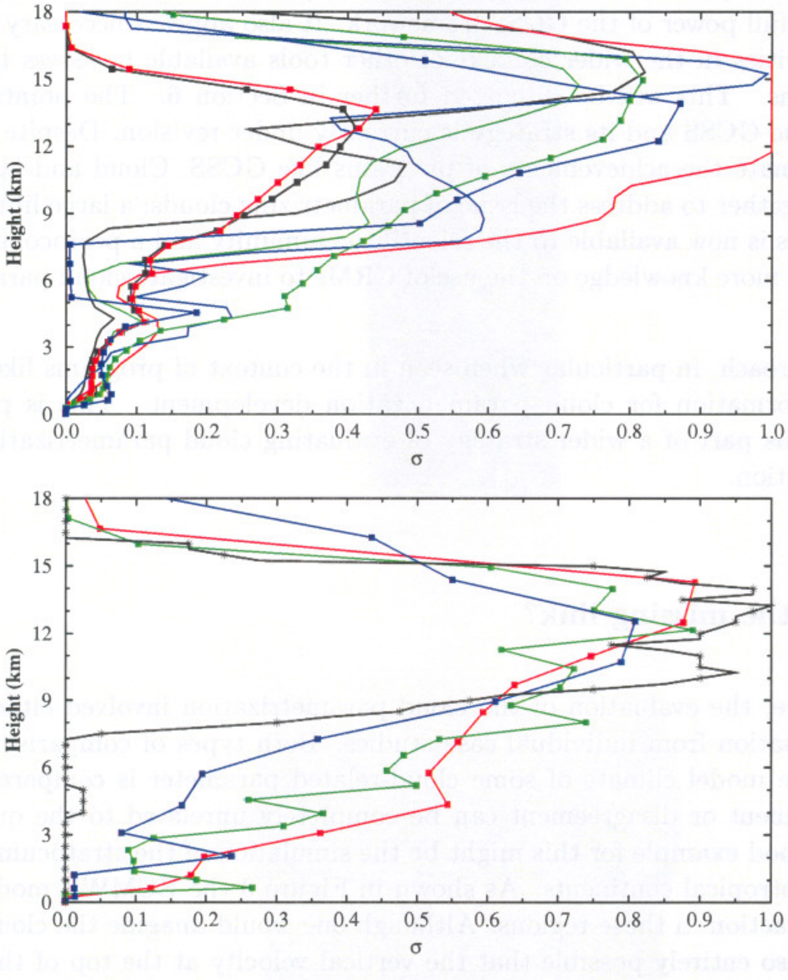

Figure 7: Six-day average of cloud fraction as a function of height from simulations for a period of TOGA COARE (20 to 25 December 1992) with a number of CRMs (top) and SCMs (bottom) (Krueger et al. 2000, pers. communication, http://www.met.utah.edu/skrueger/gcss/case2.html)

Figure 7 shows an example of the kind of results that have become available from GCSS studies for cloud parametrization (Krueger, personal communication; Bechthold et al., 2000; Redelsperger et al., 2000). The figure shows a six-day average of the vertical profile of cloud fraction, one of the crucial results of a cloud parametrization, for a period in late December 1992 of the TOGA COARE field experiment. The top panel shows results from simulations using CRMs whereas the bottom panel depicts the results from SCM simulations. Note, that the external forcing was prescribed identically for all models. The figure highlights both the strength and the intrinsic difficulties of the GCSS approach so far. It is undoubtedly useful to use a variety of state-of-the-art CRMs (instead of just one) to carry out simulations as the above. That way, by assessing the spread in the results, one gains some confidence or, as in this case, one is cautioned on the direct use of the CRM results as a surrogate for observations. From Figure 7 it is difficult to argue that the SCMs perform considerably worse in simulating the vertical distribution of cloud cover, which, given the large spread in the CRM results and the absence of observations, should not necessarily be interpreted as an indication that the SCMs are correct.

The above figure also highlights another danger, the so-called "intercomparison trap", into which the cloud parametrization community is sometimes in danger of falling. Just simulating various cases with a large number of models and comparing the results will not automatically lead to the improvement of any of them. A more useful approach would be to formulate ideas and hypotheses on how the various cloud processes should be included in parametrizations, and then assess and test those ideas using the full power of the GCSS framework. It also appears necessary to view the model intercomparison activities in the wider context of other tools available to assess the performance of cloud parametrizations. This will be discussed further in Section 6. The points raised here have been recognized by the GCSS and its strategy is currently under revision. Despite this criticism, one should not underestimate the achievements of programs like GCSS. Cloud and large-scale modelers have been brought together to address the issue of parametrizing clouds; a large library of case studies for various cloud types is now available to the scientific community and a protocol exists for their use. Futhermore more and more knowledge on the use of CRMs to investigate cloud parametrization issues is emerging.

Overall the SCM approach, in particular when seen in the context of programs like ARM and GCSS provides valuable information for cloud parametrization development. This is particularly true if the approach is seen as part of a wider strategy of evaluating cloud parametrizations, which will be outlined in a later section.

5 Composites - the missing link?

In the examples above, the evaluation of the cloud parametrization involved either highly averaged information or information from individual case studies. Both types of comparisons have some serious drawbacks. If the model climate of some cloud-related parameter is compared to observations, the reasons for agreement or disagreement can be completely unrelated to the quality of the cloud parametrization. A good example for this might be the simulation of the stratocumulus clouds off the west coasts of the subtropical continents. As shown in Figure 3 the ECMWF model strongly underestimates the cloud fraction in those regions. Although one would imagine the cloud parametrization being at fault, it is also entirely possible that the vertical velocity at the top of the PBL, which is a significant component in the equilibrium of many influences that lead to the existence of stratocumulus clouds, is in error. Since there are no observations of vertical velocity (although some comparisons of the ECMWF model subsidence rates in the subtropics have been carried out by Betts et al., 1995), it is difficult to assess whether this is really the case. In turn the possibly wrong subsidence can exist for many reasons, such as errors in the strength of the Hadley circulation due to errors in cumulus parametrization. From this kind of argument, the details of which are not particularly relevant, it is obvious that comparisons based on the model climate alone can shed little light on what might be wrong with the cloud parametrization scheme.

In regions predominantly covered by one cloud type, such as the above quoted stratocumulus areas, the problems mentioned can partly be overcome by using averages of short-range forecasts to build up the "model climate". As already mentioned the advantage of the use of short-range forecasts is that the factors influencing the cloud parametrization can be assumed to be reasonably close to observations, since the latter were used in the generation of the model's initial state. But what about regions of large variance in cloud amount and type, such as over the extratropical oceans. Here one is inclined to prefer case studies of some kind to investigate the performance of the cloud parametrization. This approach can provide substantial amounts of detail on the workings of the parametrization. However, there is one major drawback as well. The choice of the cases for study is far from trivial and is often forced upon the investigation by the limited amount of observations available. This might lead to the choice of interesting, but unrepresentative cases. If a cloud parametrization was made to work for such a case the majority of cases might still be poorly simulated. And even if the case is representative, success in simulating an individual case does not guarantee a good performance of the model and the cloud parametrization in general.

A possible way to reconcile the model climate and case study approaches is to find "more intelligent" ways of averaging the data, so that the general characteristics of certain cloud systems remain intact even when a large number of cases is included in the average. Two attempts of such an averaging approach will be highlighted here.

5.1 A survey of clouds over the North Atlantic and North Pacific

The first composite study outlined here follows an idea of Tselioudis et al. (2000) and is based on data products provided by ISCCP. Up to this point only monthly mean values of total cloud fraction derived in this project have been considered here for the use in model evaluation. ISCCP provides a much larger set of data than that (Rossow and Schiffer, 1991; Rossow and Schiffer, 1999). For each satellite pixel an attempt to derive cloud top pressure and cloud optical thickness is made and the joint statistical distribution of both parameters over larger areas (roughly 2.5 x 2.5 degrees) is provided on a three-hourly basis. Tselioudis et al. (2000) use this data to survey the shape of the above distribution functions for the ocean area of the Northern Hemisphere extratropics (NHE, 30° to $60°N$) as a function of dynamical regime. They use the simplest indicator for different "weather" situations, namely the surface pressure, to define three "dynamical" regimes as anomalously low, normal and anomalously high pressure. Even with this extremely simple classification they find remarkably different cloud distributions. An example is shown in the left panels of Figure 8. Shown are two-dimensional "histograms" of cloud top pressure versus optical thickness for the NHE in April 1992 when the local surface pressure is more than 5 hPa below (top) and 5 hPa above (bottom) its monthly average value.

The difference in the "cloud" distribution between the two chosen regimes is quite marked. In the below-average pressure regime three predominant types of cloud appear, thick high top clouds, very likely associated with frontal systems, as well as medium-high top thin clouds (most likely altostratus and altocumulus), and low clouds of medium optical thickness associated with cloudiness at the top of the PBL. In the above-average regime the latter become the predominant cloud type while the thick high-top clouds are virtually absent. This type of cloud distribution is what one would expect in the subsidence regions of high-pressure systems over the oceans and ahead of and behind extratropical cyclones (see below).

Tselioudis et al. (1998) proposed to use this technique to evaluate the performance of GCMs in simulating not only the mean cloud properties but also the observed difference in cloud structure between the different dynamical regimes. The middle and right panels of Figure 8 show the result of such a comparison using short-range forecasts from the ECMWF model for April 1992. The model

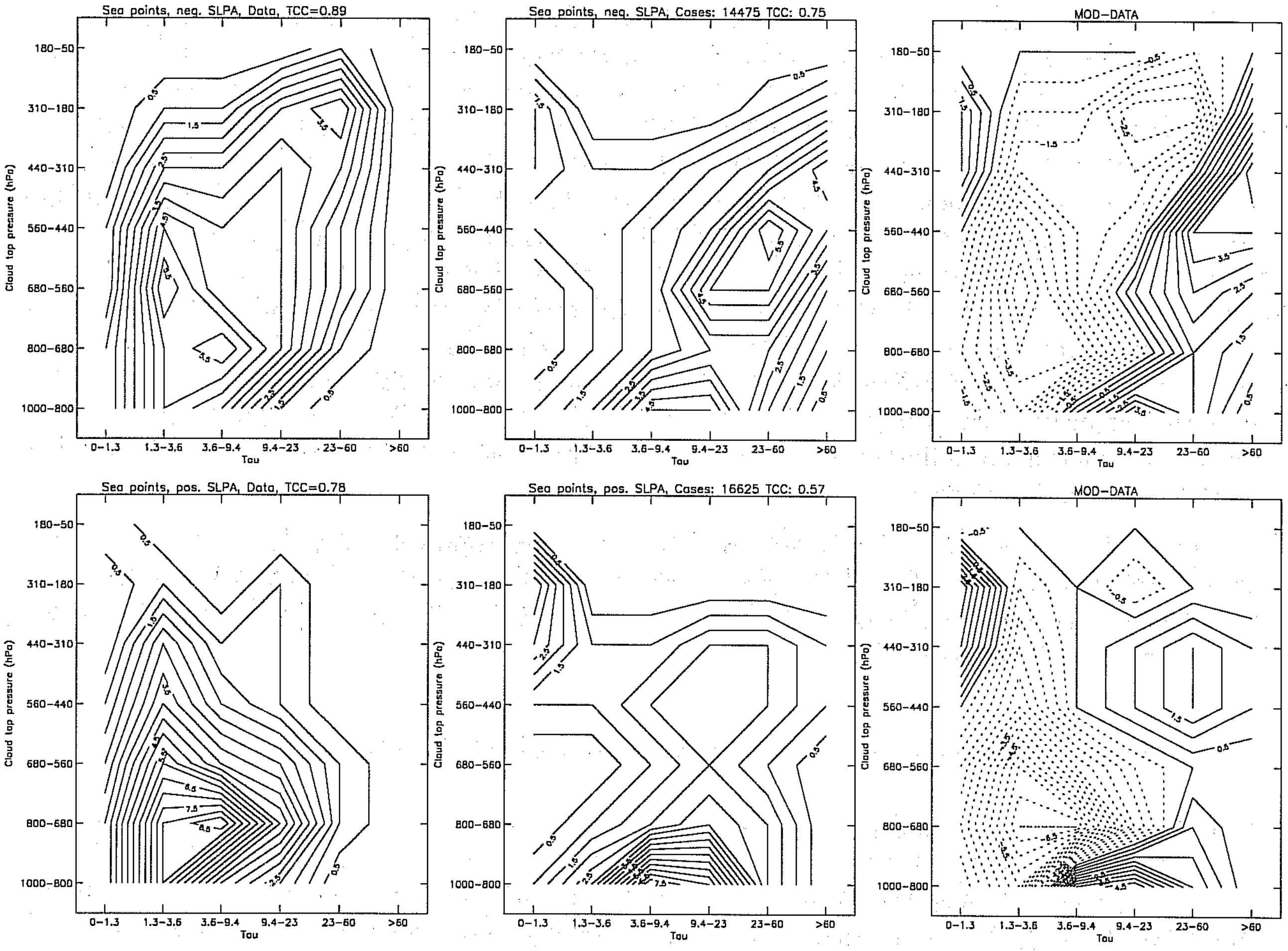

Figure 8: Distribution of cloud top pressure vs. cloud optical thickness for the ocean areas of $30°N$ to $60°N$ in the month of April 1992. The data is stratified into local negative sea level pressure anomalies (from the monthly mean) of more than -5 hPa (top panels) and positive seal level pressure anomalies of more than +5 hPa (bottom panels). Distributions are derived from ISCCP D1 data (left), and ECMWF T106L31 12- and 24-hour forecasts (middle). The right panels show the difference between model and data.

is probed in exactly the same way as the data, i.e., local pressure anomalies based on the model results are calculated and the cloud top pressure vs. optical thickness distributions are derived. The technique to derive these distributions from the model cloud fields attempts to find the radiative cloud top (instead of the physical one) by taking into account that the cloud emissivity of thin cirrus layers can be well below one. The technique is described in detail in Klein and Jakob (1999).

A number of important differences between model and data are evident. First of all the total cloud cover, indicated by "TCC" at the top of each panel, is underestimated for both regimes, by about 15 % in the negative anomaly regime and by 20 % for positive pressure anomalies. For negative pressure anomalies, the clouds are, when present, too optically thick and their tops are too low when compared to the data. When repeating the model analysis using the physical instead of the radiative cloud top (not shown), the latter effect disappears, indicating that the model is producing cloud tops at the right height but that the top parts of the clouds are optically too thin, pointing to a possible deficiency in the ice water content. In the positive pressure anomaly regime the model cloud tops appear to be too low and the clouds are too optically thick. Also the model produces both optically thick mid-level to high-top clouds and thin high-top clouds that are not observed. It is well possible, that the cloud top pressure error for low clouds in this regime is due to the difficulty to determine the exact cloud top in the data for clouds at the top of PBL's capped by an inversion, which can lead to a misinterpretation of the height associated with the measured brightness temperature. However, the overestimation in optical thickness is most likely a true model problem.

It is intriguing that the model has a tendency to underestimate total cloud cover and seems to "compensate" for that by producing too thick clouds. The errors identified here, although for a different period, are nevertheless consistent with the underestimation of the reflection of solar radiation and shortwave cloud radiative forcing that was pointed out in Figures 1 and 2.

It is worthwhile stressing again, that the comparison carried out here is far from comprehensive and is only meant to illustrate a new technique to try and analyse model errors in the representation of clouds. The technique is currently being refined and other "dynamical" compositing criteria are being investigated. Although simple, the approach proves to provide very useful information on the cloud representation. First, it combines cloud fraction information with radiative effects of the clouds when present by studying the optical thickness. Second, by splitting the data set by using some "dynamic" criterion, regimes in which model errors are particularly large can be identified and investigated further using other tools, such as SCM studies. Furthermore, although short-range forecasts have been used here this is not a requirement of the method, since the dynamical criterion (pressure anomaly of a certain size in this case) can be, and is in the example above, entirely defined by model data. This makes this technique useful not only for NWP models but for any GCM.

5.2 Validating clouds associated with extratropical cyclones

More complex compositing techniques, that reveal cloud parametrization problems in even greater detail can of course be applied. One such study is that of Klein and Jakob (1999, hereafter KJ99), which is based on an idea of Lau and Crane (1995, hereafter LC95). Here the cloud structure in extratropical cyclones over the North Atlantic is studied. For a given set of locations LC95 identify from ISCCP data the optically thickest clouds occurring over a number of years and use those events as their compositing criterion. Each maximum-optical-thickness point (in time and space) is then considered the centre of a relative coordinate system and all occurrences (roughly 1200 in LC95) are overlaid using this coordinate system. The result is a composite of the spatial distribution of the observed cloud and other meteorological fields, as shown in the top panel of Figure 9.

Shown are the mean anomalies from a five-day average (taken for each individual case) of 1000-hPa horizontal wind (arrows) and geopotential height, as well as the occurrence of various cloud types. High-top clouds are depicted in red, mid-level top clouds in yellow and low top clouds in blue. The darkness of each colour indicates optical thickness. Note, that the center of the plot contains the optically thickest clouds by construction. The 1000-hPa height field reveals the relative location of the low pressure center, which is positioned to the southwest of the clouds with maximum optical thickness. A large "shield" of high top thick clouds, as they are normally associated with warm fronts, is evident to the northeast of the low pressure centre. Middle-top thick clouds extend out of the low-pressure region to the southwest. Ahead and behind the composite cyclone the cloud fields are dominated by low-top medium-thick clouds.

KJ99 have used short-range forecasts from ERA to generate the same picture of the cloud distribution around extratropical cyclones from the ECMWF model. The exact dates and locations of LC95 were used to position the cyclones and the same analysis techniques were applied. As above the cloud top pressure used to define clouds as high-, mid- and low-top can be derived using the physical or the radiative cloud top (see KJ99 for details). The middle panel of Figure 9 shows the model results when using the physical cloud top while the bottom panel contains the results for the radiative cloud top. It is evident from this figure that the model is able to reproduce the overall distribution of cloudiness around the cyclone quite well, perhaps with the exception of the cloud band extending southwestward

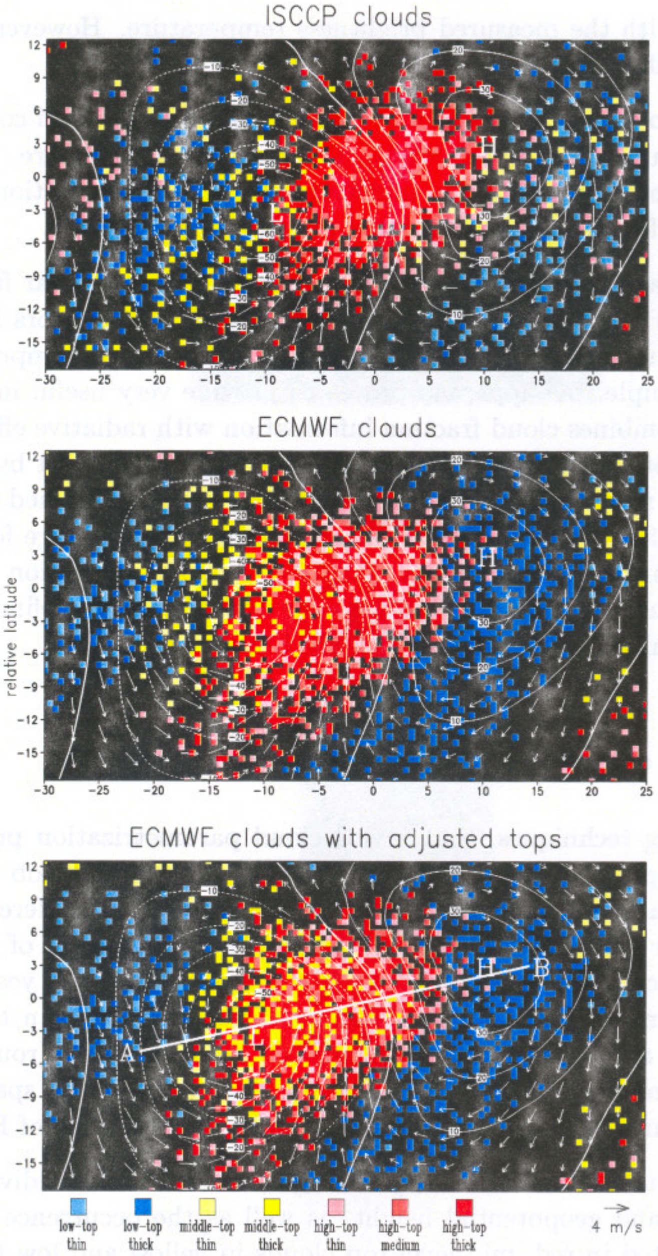

Figure 9: (a) Distributions of 1000 hPa horizontal wind (arrows, see scale at bottom right) and geopotential height (contours, interval 10 m) from ERA analyses, and various cloud types (colour pixels) from ISCCP observations as shown in LC95. The ordinate (abscissa) of the coordinate system corresponds to displacements in degrees from the reference location. Inside each 2.5° × 2.5° the presence and abundance of a certain cloud type is indicated by plotting a number of randomly scattered pixels with the colour designated to the cloud type (see legend). Each pixel represents a 1 % positive deviation from the background field as estimated as the five-day average centered on the key dates. (b) as 1(a) but from 24h ERA forecasts using physical cloud top pressure, (c) as 1(b) but using emissivity-adjusted cloud top pressure. From Klein and Jakob, 1999.

from the low-pressure centre. On closer inspection, however, similar errors to those identified above in the pressure anomaly composites appear. The high-top clouds are optically too thin, leading to large errors in cloud top height when using the radiative cloud top (which is what a satellite would most likely identify). The low-top clouds are optically too thick in particular in the regions ahead of the cyclone. KJ99 carried out a number of sensitivity studies and identified the microphysical assumptions for ice settling as one of the major sensitivities for the simulation of the high-top cloud optical thickness. None of their studies was able to reduce the error in low-top cloud optical thickness.

The two studies briefly summarized here demonstrate the usefulness of the idea of composite-averaging in the evaluation of cloud parametrizations. Other more recent examples using this apporach can be found in Webb et al. (2000) and Bony et al. (2000).

By averaging over a large number of cases in such a way that key dynamical and hence cloud structures remain intact, it is possible to identify not only the deficiencies of the model cloud representation but also the dynamic environment in which they occur. This provides first clues for possible model errors, which can then be investigated further. The next section proposes a new strategy for the evaluation of cloud parametrizations in which "compositing" will play a central role.

6 A strategy for cloud parametrization evaluation

In the previous sections various commonly applied methods for the evaluation of cloud parametrizations have been outlined. Although numerous studies using one or more of the outlined techniques have been published over the last few years, a lack of coherence in the application of several techniques to the same model is clearly visible. Most likely this is because each of the techniques as such requires substantial resources. However, this apparent lack of a strategy when evaluating cloud parametrizations has led to a considerable dilution of efforts. In this final short section a coherent strategy that relies mainly on the known techniques described above will be proposed. It is schematically outlined in Figure 10.

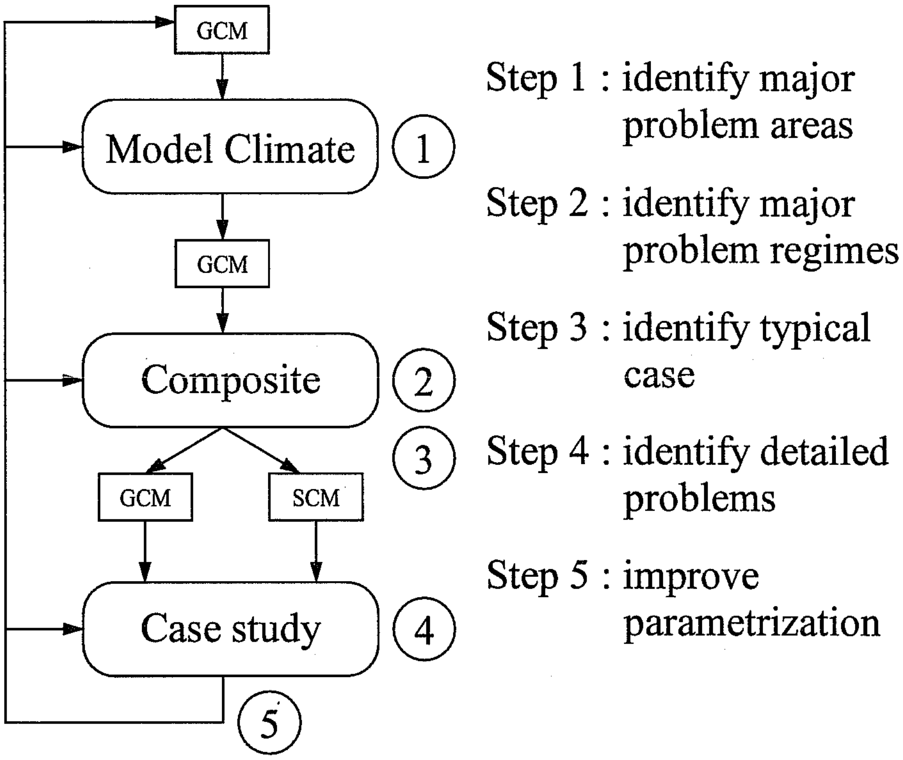

Figure 10: A strategy for the evaluation of cloud parametrizations in GCMs.

At the core of this strategy is the attempt to link the evaluation of the model climate to the selection of cases for case studies through the use of compositing techniques such as those described in the previous section. The evaluation of the model climate normally reveals areas (in the geographic sense) in which the clouds and/or their effects are not correctly represented. As pointed out in Section 3 it is virtually impossible to infer reasons for the observed errors from such studies. Those can normally only be discovered in detailed case studies. But how should a case be chosen, such that it is typical for the model error? This is where compositing observations and model results using some criterion that describes the main mechanisms in the cloud generation and/or maintenance should prove useful. By applying techniques similar to those highlighted in Section 5 not only is a first link to the possible causes for model problems established but also the typical model error is revealed. From the (hopefully) considerable number of cases entering each composite average, one can then select those for which the model error is close to the mean error in the composite. That way it is ensured that the now following case study represents a typical model behaviour, rather than an extreme one. The case study can be carried out either with the full GCM, e.g., in an NWP environment, or with a corresponding SCM. After improving the parametrization it is of course necessary to repeat the entire validation process to i) test the performance of the new parametrization in all aspects of the model and ii) identify the next target for improvement.

One of the caveats of this strategy is of course the availability of the necessary observational data sets. These need to be fairly comprehensive to enable the use of either an NWP model or an SCM and to facilitate the validation of the results. With the availability of long term data sets such as those collected in the context of ARM program and through the combination of analyses produced by operational NWP centres and the vast amount of satellite data available, these problems appear soluble to a large extent.

Finally it is worthwhile pointing out that the real improvements of a cloud parametrization will not result from evaluation studies of any kind, with or without a strategy. They can only result from the knowledge of the model errors, gained from evaluation studies, and first and foremost from ideas of individuals or groups, which can then be put to the test using the strategy outlined above.

References

Albrecht, B. A., C. S. Bretherton, D. W. Johnson, W. H. Schubert and A. Shelby Frisch, 1995: The Atlantic Stratocumulus Transition Experiment – ASTEX. *Bull. Amer. Meteor. Soc.*, **76**, 889-904.

Arking, A., 1991: The radiative effects of clouds and their impact on climate. *Bull. Amer. Meteor. Soc.*, **71**, 795-813.

Augstein, E., H. Riehl, F. Ostapoff, V. Wagner, 1973: Mass and energy transports in an undisturbed Atlantic trade-wind flow. *Mon. Wea. Rev.*, **101**, 101-111.

Barkstrom, B. R., and G. L. Smith, 1986: The Earth Radiation Budget Experiment: Science and implementation. *J. Geophys. Res.*, **24**, 379-390.

Bechthold, P., S.K. Krueger, W.S. Lewellen, E. van Meijgaard, C.-H. Moeng, D.A. Randall, A. van Ulden, and S. Wang, 1996: Modelling a stratocumulus-topped PBL: Intercomparison among different one- dimensional codes and with large eddy simulation. *Bull. Amer. Meteor. Soc.*, **77**, 2033-2042.

Bechtold, P., J. L. Redelsperger, I. Beau, M. Blackburn, S. Brinkop, J. Y. Grandpeix, A. Grant, D. Gregory, F. Guichard, C. Hoff, and E. Ioannidou., 2000: A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE. Part II: Intercomparison of SCMs and with CRM. *Quart. J. Roy. Met. Soc.*, **126**, 865-888.

Beesley, J. A., C. S. Bretherton, C. Jakob, E. Andreas, J. Intrieri, and T. Uttal, 2000: A comparison of ECMWF model output with observations at SHEBA. *J. Geophys. Res.*, **105**, 12337-12349.

Betts, A. K., C. S. Bretherton, and E. Klinker, 1995: Relation between mean boundary layer structure and cloudiness at the R/V Valdivia during ASTEX. *J. Atmos. Sci.*, **52**, 2752-2762.

Bony, S., H. Le Treut, J.-J. Morcrette, and C.A. Senior, 2000: Sensitivity of tropical clouds to sea surface temperature: Observation and climate model simulations. Submitted to *Clim. Dyn.*.

Bretherton, C. S., S. K. Krueger, M. C. Wyant, P. Bechtold, E. van Meijgaard, B. Stevens, and J. Teixeira, 1999: A GCSS boundary layer model intercomparison study of the first ASTEX Lagrangian experiment. *Boundary-Layer Meteorol.*, **93**, 341-380.

Bretherton, C. S., S. R. de Roode, C. Jakob, E. L. Andreas, J. Intrieri, and R. E. Moritz, 2000: A comparison of the ECMWF forecast model with observations over the annual cycle at SHEBA. submitted to *J. Geophys. Res.*

Browning, K. A., 1993: The GEWEX Cloud System Study (GCSS). *Bull. Amer. Meteor. Soc.*, **74**, 387-399.

Cess, R.D., and G.L. Potter, 1987: Exploratory studies of cloud radiative forcing with a general circulation model. *Tellus*, **39A**, 460-473.

Cess, R.D., G.L. Potter, W.L. Gates, J.J. Morcrette, and L. Corsetti, 1992: Comparison of general circulation models to Earth Radiation Budget Experiment data: Computation of clear-sky fluxes. *J. Geophys. Res.*, **97D**, 20421-20426.

Curry, J. A., P. V. Hobbs, M. D. King, D. A. Randall, P. Minnis, G. A. Isaac, J. O. Pinto, T. Uttal, A. Bucholtz, D. G. Cripe, H. Gerber, C. W. Fairall, T. J. Garrett, J. Hudson, J. M. Intrieri, C. Jakob, T. Jensen, P. Lawson, D. Marcotte, L. Nguyen, P. Pielewski, A. Rangno, D. C. Rogers, K. B. Strawbridge, F. P. J. Valero, A. G. Williams, D. Wylie, 2000: FIRE Arctic clouds experiment. *Bull. Amer. Meteor. Soc.*, **81**, 5-29.

Del Genio, A.D., M.-S. Yao, W. Kovari, and K.K.-W. Lo, 1996: A prognostic cloud water parametrization for global climate models. *J. Clim.*, **9**, 270-304.

Ellis, J. S., 1978: Cloudiness, the planetary radiation budget and climate. Ph.D. thesis, Colorado State University, Fort Collins, 129 pp.

Fowler, L.D., D.A. Randall, and S.A. Rutledge, 1996: Liquid and ice cloud microphysics in the CSU general circulation model. Part I: Model description and simulated microphysical processes. *J. Clim.*, **9**, 489-529.

Gibson, J.K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ECMWF Re-Analysis Project Report Series: 1. ERA description. *ECMWF 1997*, 72 pp.

Hogan, R. J., C. Jakob, A. J. Illingworth, 2000: Comparison of ECMWF cloud fraction with radar derived values. *J. Appl. Meteor.*, **40**, 513-525.

Holland, J. Z., and E. M. Rasmusson, 1973: Measurements of atmospheric mass, energy and momentum budgets over a 500-kilometer square of tropical ocean. *Mon. Wea. Rev.*, **101**, 44-55.

Houze jr., R. A. and A. K. Betts, 1981: Convection in GATE. *Rev. Geophys. Space Phys.*, **19**, 541-576.

Jakob, C., 1999: Cloud cover in the ECMWF reanalysis. *J. Clim.*, **12**, 947-959.

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

Lau, N.-C., and M.W. Crane, 1995: A satellite view of the synoptic- scale organization of cloud properties in Midlatitude and Tropical circulation systems. *Mon. Wea. Rev.*, **7**, 1984-2006.

Mace, G.G., C. Jakob, and K. P. Moran, 1998a: Validation of hydrometeor occurrence predicted by the ECMWF model using millimeter wave radar data. *Geophys. Res. Letters*, **25**, 1645-1648.

Mason, P. J., 1994: Large-eddy simulation: A critical review of the technique. *Quart. J. Roy. Met. Soc.*, **120**, 1-26.

Miller, S. D., G. L. Stephens, and A. C. M. Beljaars, 1999: A validation survey of the ECMWF prognostic cloud scheme using LITE. *Geophys. Res. Letters*, **26**, 1417-1420.

Moeng, C.-H., W. R. Cotton, C. S. Bretherton, A. Chlond, M. Khairoutdinov, S. Krueger, W. S. Lewellen, M. K. MacVean, J. R. M. Pasquier, H. A. Rand, A. P. Siebesma, R. I. Sykes, and B. Stevens, 1996: Simulation of a stratocumulus-topped PBL: Intercomparison among different numerical codes. *Bull. Amer. Meteor. Soc.*, **77**, 261-278.

Moncrieff, M.W., S.K. Krueger, D. Gregory, J.-L. Redelsperger, and W.-K. Tao, 1997: GEWEX Cloud System Study (GCSS) working group 4: Precipitating convective cloud systems. *Bull. Amer. Meteor. Soc.*, **78**, 831-845.

Nitta, T., and S. Esbensen, 1974: Heat and moisture budget analyses using BOMEX data. *Mon. Wea. Rev.*, **102**, 17-28.

Randall, D. A., and D. G. Cripe, 1999: Alternative methods for specification of observed forcing in single-column models and cloud system models. *J. Geophys. Res.*, **104**, 24527-24546.

Redelsperger,J. L., P. R. A. Brown, F. Guichard, C. Hoff, M. Kawasima, S. Lang, Th. Montmerle, K. Nakamura, K. Saito, C. Seman, W. K. Tao, and L.J. Donner., 2000: A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE. Part I: Cloud-Resolving Models. *Quart. J. Roy. Met. Soc.*, **126**, 823-863.

Rossow, W.B., and R.A. Schiffer, 1983: The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Program. *Bull. Amer. Meteor. Soc.*, **64**, 779-784.

Rossow, W.B., and R.A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.*, **72**, 2-20.

Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261-2287.

Ryan, B. F., J. J. Katzfey, D. J. Abbs, C. Jakob, U. Lohmann, B. Rockel, L. D. Rotstayn, R. E. Stewart, K. K. Szeto, G. Tselioudis, and M. K. Yau, 2000: Simulation of a cold front using cloud-resolving, limited-area, and large-scale models. *Mon. Wea. Rev.*, **128**, 3218-3235.

Slingo, J.M., 1987: The development and verification of a cloud prediction scheme for the ECMWF model. *Quart. J. Roy. Met. Soc.*, **113**, 899-927.

Smith, R.N.B., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Quart. J. Roy. Met. Soc.*, **116**, 435-460.

Stokes, G.M., and S.E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the Cloud and Radiation Test Bed, *Bull. Amer. Meteor. Soc.*, **75**, 1201-1221.

Sundqvist, H., 1978: A parameterization scheme for non-convective condensation including prediction of cloud water content, *Quart. J. Roy. Met. Soc.*, **104**, 677-690.

Tiedtke, M., 1993: Representation of Clouds in Large-Scale Models, *Mon. Wea. Rev.*, **121**, 3040-3061.

Tselioudis, G., C. Jakob, and U. Lohmann, 1998: A methodology to construct survey cloud datasets for climate model validation. *Report of GCSS Working Group 3 to the GCSS Science Panel*, Hawaii, 1998.

Tselioudis, G., Y. Zhang, and W. B. Rossow, 2000: Cloud and radiation variations associated with northern midlatitude low and high sea level pressure regimes. *J. Clim.*, **13**, 312-327.

Webb, M., C. Senior, S. Bony, J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in three climate models. *Clim. Dyn.*, **17**, 905-922.

Webster, P. J., and R. Lukas, 1992: TOGA COARE: The Coupled Ocean-Atmosphere Response Experiment. *Bull. Amer. Meteor. Soc.*, **73**, 1377-1416.