Chapter 25

Cloud-Resolving Modeling: ARM and the GCSS Story

STEVEN K. KRUEGER
University of Utah, Salt Lake City, Utah

HUGH MORRISON
National Center for Atmospheric Research, Boulder, Colorado

ANN M. FRIDLIND
NASA Goddard Institute for Space Sciences, New York, New York

1. The GEWEX Cloud System Study

The Global Energy and Water Cycle Experiment (GEWEX) Cloud System Study (GCSS) was created in 1992. As described by Browning et al. (1993, p. 387), "The focus of GCSS is on cloud systems spanning the mesoscale rather than on individual clouds. Observations from field programs will be used to develop and validate the cloud-resolving models, which in turn will be used as test-beds to develop the parameterizations for the large-scale models." The most important activities that GCSS promoted were the following:

- Identify key questions about cloud systems relating to parameterization issues and suggest approaches to address them, and
- Organize model intercomparison studies relevant to cloud parameterization.

Four different cloud system types were chosen for GCSS to study: boundary layer, cirrus, frontal, and deep precipitating convective. A working group (WG) was formed for each of the cloud system types. The WGs organized model intercomparison studies and meetings to present results of the intercomparisons. The first such intercomparison study took place in 1994 (Moeng et al. 1996; Bechtold et al. 1996).

The GCSS approach uses cloud-resolving model (CRM) simulations to estimate cloud system structure, vertical fluxes, and other characteristics from large-scale atmospheric conditions. A CRM is a 2D or 3D nonhydrostatic numerical model that resolves cloud-scale motions while simulating a cloud system. For example, to simulate a convective cloud system that contains both cumulus-scale and mesoscale circulations, a CRM would typically have a horizontal grid size of about 2 km and a horizontal domain size of about 400 km.1 The GCSS approach is an example of dynamical downscaling, which uses larger-scale conditions to drive a smaller-scale numerical model at higher spatial resolution, which in turn is able to simulate atmospheric phenomena in greater detail.

It is interesting to compare the time and space scales of CRMs used to simulate convective cloud systems to those of global climate models (GCMs; see Table 25-1). The scales are set by the characteristics of the dominant resolved eddies: cumulus clouds in CRMs and baroclinic eddies in GCMs. The scales of cumulus clouds are about a hundredth of those of baroclinic eddies. Thus, a 30-day CRM simulation is equivalent to a 10-yr GCM simulation. CRMs have appropriately been called "cloud GCMs."

1. Krueger (2000) described CRMs in greater detail and provided examples of using them to study several types of cloud systems.
The results of the CRM simulations are synthetic or virtual datasets of cloud system realizations that can be used to test, improve, and develop parameterizations for large-scale models. However, because CRMs also contain parameterizations of unresolved processes (including turbulence, microphysics, and radiation), the results of such CRM simulations should be evaluated before being used for large-scale-model parameterization development.

The GCSS approach thus requires observations for both the downscaling and for the evaluation (Randall et al. 1996, 2003). Downscaling requires observations in order to provide a CRM with the same large-scale “forcing” that an atmospheric column in a GCM would have, while evaluation requires observations of macrophysical and microphysical cloud system properties. Large-scale forcing (Zhang et al. 2016, chapter 24) includes the profiles of the large-scale 1) horizontal advective tendencies of temperature and water vapor, 2) vertical velocity, 3) advective tendencies of the horizontal wind components, and 4) horizontal pressure gradient, as well as the surface and top-of-atmosphere boundary conditions, all varying in time. Because of the difficulties of determining the large-scale horizontal wind’s advective tendency and the large-scale horizontal pressure gradient, the large-scale forcing for the horizontal wind is usually specified as a relaxation of the CRM’s horizontally averaged wind towards the observed wind, with a relaxation time scale of about 2 h. This relaxation term also represents the Coriolis acceleration.

Obviously, accurate large-scale forcing will minimize simulation errors due to the forcing and thereby facilitate identification of errors due to CRM physics. Forcing errors can be detected through model intercomparison studies because such errors typically produce similar impacts in all of the simulations. Once they are detected, the only recourse is to improve the accuracy of the forcing. ARM contributed greatly to producing accurate large-scale forcing datasets by collecting high-frequency vertical profiles of temperature, water vapor, and wind with an array of balloonborne sounding systems (rawinsondes) during intensive observation periods (IOPs) and by using the constrained variational analysis method to process the sounding array profiles.

2. ARM data and single-column modeling

The ARM Program was established to improve the understanding of atmospheric radiation and its interaction with clouds and cloud processes. The goals were to make measurements of atmospheric radiation and the atmospheric properties that affect it at five (later reduced to three) fixed sites for up to 10 years and to use the measurements to test cloud and radiation parameterizations of varying complexity. Randall et al. (1996) proposed to use field data such as those collected by ARM together with single-column models (SCMs) and CRMs to test physical parameterizations used or to be used in GCMs, as shown in Fig. 25.1.

a. Single-column model IOPs

The first ARM single-column model IOP took place in winter 1994 and was led by David Randall. Starting with this IOP, seasonal SCM IOPs were conducted at the Southern Great Plains (SGP) site (Zhang et al. 2016, chapter 24) to enhance the frequency of observations for SCM uses, particularly vertical soundings of temperature, water vapor, and winds. These SCM IOPs (listed in Table 25.2) were conducted for a period of 21 days each. During each IOP, radiosondes were launched at the Central Facility and the four boundary facilities eight times per day, seven days per week. The data were required for quantifying boundary forcing and column response.

b. Variational analysis

The constrained variational analysis method was developed by Zhang and Lin (1997) for deriving large-scale vertical velocity and advective tendencies from
sounding arrays. A history of the development of this technique is provided in Zhang et al. (2016, chapter 24). It is used to process atmospheric soundings of winds, temperature, and water vapor mixing ratio over a network of a small number of stations. Given the inevitable uncertainties in the original data, the basic idea in this objective analysis approach is to adjust these atmospheric state variables by the smallest possible amounts to conserve column-integrated mass, moisture, dry static energy, and momentum. The analysis products include both the large-scale forcing terms and the evaluation fields, which can be used for driving SCMs and CRMs and for evaluating model simulations. The first variational analysis dataset produced for ARM was based on the Summer 1995 SCM IOP. This was followed by ones for the Spring, Summer, and Fall 1997 SCM IOPs. The Summer 1995 dataset was used for the first ARM SCM intercomparison study (Cederwall et al. 1998), which was used to evaluate the adequacy of the forcing dataset and to investigate various prescriptions of advective forcing.

c. Active Remote Sensing of Clouds

Once ARM began producing accurate large-scale forcing data for the SGP site, the next task was to evaluate the results of CRM (and SCM) simulations that were based on these forcing data. This was accomplished with the extensive array of ARM instruments at the SGP site, particularly those designed to remotely sense cloud properties from the ground. The Active Remote Sensing of Clouds (ARSCL; Clothiaux et al. 2000; Kollias et al. 2016, chapter 17) value-added product combines data from millimeter cloud radars, laser ceilometers, microwave radiometers, and micropulse lidars to produce a time series of vertical distributions of cloud hydrometeors over the ARM sites. A preliminary version of ARSCL first became available for the SGP site in November 1996, and a stable, general version became available on 1 April 1997. Other measurements used to directly evaluate CRM and SCM results included surface broadband radiative fluxes, surface turbulent fluxes, surface precipitation rate, and profiles of temperature and water vapor.

3. Continental deep convection: Bringing GCSS and ARM together

The inaugural model intercomparison meeting of GCSS Working Group 4 (precipitating convective cloud systems) was held 21–23 October 1996 in Annapolis, Maryland. Moncrieff et al. (1997) described the first two model intercomparison projects organized by Working Group 4, which were designed in accord with Randall et al.'s (1996) proposed method of using SCMs and CRMs to test and develop physically based parameterizations. The main actions and recommendations from the GCSS Science Panel Sixth Session (GCSS Science Team 1998) included this action item for Working Group 4:

Investigate the possibility of developing a continental case study built on data from the USA, Department of Energy, Atmospheric Radiation Measurement (ARM) Program. Include the possible options for such a case in the Working Group 4 report at the 1998 GCSS meeting.

The GCSS Science Panel Seventh Session (GCSS Science Team 1999) recommended that:

WG 4 should proceed with a continental deep convection case drawn from data taken at the ARM Southern Great Plains experimental site during July 1997. Specialized instrumentation including a millimeter cloud radar and an extensive array of other meteorological instruments were operational at that time. This case will be done in collaboration with the ARM Single-Column Modeling Group as a means of involving more of the SCM community to participate in the process and to gain support of the ARM Data and Science Integration Team in the provision of forcing data and the compilation of results submitted by the modeling groups.

The first joint GCSS and ARM model intercomparison case was WG 4 Case 3: Summer 1997 ARM SCM IOP. GCSS and ARM collaborated on several major studies of cloud systems during the next 10 years.

4. GCSS and ARM: Confronting models with data

a. Continental deep convection

Ric Cederwall (Lawrence Livermore National Laboratory), Steve Krueger (University of Utah), and Dave Randall (Colorado State University) initiated a collaborative intercomparison of SCMs and CRMs that

---

focused on the Summer 1997 SCM IOP at the SGP site (Cederwall et al. 1999). The study involved the ARM Single-Column Model Working Group, the ARM Cloud Working Group, GCSS WG 4 (precipitating convective cloud systems), and the National Centers for Environmental Prediction (NCEP) Environmental Modeling Center. The Summer 1997 IOP provided a 30-day dataset that includes a range of meteorological conditions but was dominated by deep convection.

Large-scale modelers at NCEP and the European Centre for Medium-Range Weather Forecasts (ECMWF) identified the following processes as two of those most in need of improved representation in their parameterizations of precipitating convective cloud systems:

- The occurrence (frequency and intensity) of deep convection, including the diurnal cycle over land and other interactions with the boundary layer
- The production of upper-tropospheric stratiform clouds by deep convection and a related issue: How much microphysical complexity is required in GCMs?

Continental deep convection is strongly coupled to the surface and boundary layer processes. The methodology applied by WG 4 was to start with idealized approaches (for instance, by prescribing surface fluxes) and then to proceed toward more realistic approaches after building physical understanding and developing/improving parameterizations of relevant physical processes. For instance, the diurnal cycle of continental convection, with shallow convection early in the day giving way to the deep convection in the afternoon, could be studied initially without considering the mesoscale variability of surface fluxes, without considering the impact of surface topography on convection development, and without coupling to a land surface model. Moreover, the impact on convection development of larger-scale features (e.g., synoptic-scale fronts and dry lines) that are abundant over midlatitude continents and challenging for the single-column modeling framework underlying the GCSS approach had to be addressed as well.

This model intercomparison project evaluated 8 2D CRMs, 2 3D CRMs, and 11 SCMs by testing their abilities to determine the large-scale statistics of precipitating convective cloud systems for three multiday periods during the Summer 1997 SCM IOP at the SGP site (Xu et al. 2002; Xie et al. 2002). The time-averaged CRM results for these periods of deep convection showed consistently smaller biases of time-averaged temperature and water vapor than did the SCM results. The time-averaged CRM cloud fraction profiles were in reasonable agreement with the observations from the cloud radar, while many of the SCM profiles were not. The CRM and SCM convective mass flux profiles differed significantly in the lower troposphere, and the CRM surface precipitation rates were better correlated with the observations than those from the SCMs. The results confirmed that, as expected, the CRMs are better able to reproduce the observed cloud system properties than are the SCMs. The CRM results for cloud system properties such as convective mass flux that are difficult to observe can therefore be used to evaluate and improve the parameterizations (of convective mass flux, for example) used in the SCMs.

Khairoutdinov and Randall (2003) used a new 3D CRM to simulate the 28-day evolution of clouds over the ARM SGP site during the Summer 1997 IOP. The sensitivity of the results to the domain dimensionality and size, horizontal grid resolution, and parameterization of microphysics was tested. In addition, the sensitivity to perturbed initial conditions was also tested using a 20-member ensemble of runs. The ensemble runs revealed that the uncertainty of the simulated precipitable water due to the fundamental uncertainty of the initial conditions can be as large as 25% of the mean value. Even though the precipitation rates averaged over the whole simulation period were virtually identical among the ensemble members, the timing uncertainty of the onset and the precipitation maximum were as long as one full day. Despite these predictability limitations, the simulation statistics were found to be almost insensitive to the uncertainty of the initial conditions.

The overall effects of the third spatial (second horizontal) dimension were found to be minor for the simulated mean fields and scalar fluxes but were considerable for the velocity and scalar variances. Neither changes over a rather wide range of domain sizes nor in the horizontal resolution had any significant impacts on the simulations. Although a rather strong sensitivity of the mean hydrometeor profiles and, consequently, cloud fraction to the microphysics parameters was found, the effects on the predicted mean temperature and humidity profiles were modest. The spreads among the time series of the simulated cloud fraction, precipitable water, and surface precipitation rate due to changes in the microphysics parameters were within the uncertainty of the ensemble runs.

b. Continental shallow convection

The second joint GCSS and ARM model intercomparison case was referred to as WG 1 ARM, which was an intercomparison study for CRMs and SCMs of the diurnal cycle of shallow cumulus over land based on an idealization of measurements at the ARM SGP site on 21 June 1997. The case coordinator was...
Andy Brown from the Met Office. Brown et al. (2002) described the large-eddy simulation (LES)³ results from this case, while Lenderink et al. (2004) described SCM results from the same case.

Many characteristics of the cumulus layer previously found in simulations of quasi-steady convection over the sea were reproduced in this more strongly surface-forced, unsteady case. Furthermore, the results were encouragingly robust, with similar results obtained with eight independent models and also across a range of numerical resolutions. The LES results for the subcloud layer were consistent with well-established convective boundary layer scalings. Direct validation of the cloud-layer results was difficult, although the tendency of the models to have cloud cover decreasing with time is at least broadly consistent with the observations. In the cloud layer, many of the results previously found in overseas cases are still applicable. For example, cloud cover and convective mass flux decrease with height. The similarity of the cloud-layer structures in the overseas and ARM cases was encouraging, because it suggested that there is no fundamental reason why a parameterization that performs well for one case should not also do so for the other. However, an intercomparison of single-column model results for the present ARM case (Lenderink et al. 2004) revealed the following general deficiencies: values of cloud cover and cloud liquid water that are too large, unrealistic thermodynamic profiles, and significant numerical noise. The results were also strongly dependent on vertical resolution. In part motivated by these results, several of the SCMs have been updated successfully with new parameterization schemes, and/or their present schemes have been modified successfully.

c. Midlatitude frontal clouds

Xie et al. (2005) and Xu et al. (2005) described the CRM and SCM intercomparison study of midlatitude frontal clouds based on the ARM March 2000 Cloud IOP. Xie et al. evaluated the overall performance of nine SCMs and four 2D CRMs in simulating a strong midlatitude frontal cloud system during the IOP, while Xu et al. focused on a 27-h period when only shallow frontal clouds were observed. Xie et al. (2005) found that all models captured the bulk characteristics of the frontal system and the frontal precipitation. However, there were significant differences in the detailed structures of the frontal clouds. All models overestimated high thin cirrus cloud amounts before the main frontal passage. During the frontal passage with strong upward motion, the CRMs underestimated middle and low cloud amounts, while the SCMs overestimated cloud amounts at levels above 765 hPa. All CRMs and some SCMs also underestimated the middle cloud amounts after the frontal passage. In general, the CRM-simulated cloud water and ice contents were comparable to the observations, while most SCMs underestimated cloud water content. SCMs showed very large biases of cloud ice content. Many of the model biases could be traced to the lack of subgrid-scale dynamical structure in the applied forcing fields and to the lack of organized mesoscale hydrometeor advection.

Xu et al. found that all CRMs and SCMs correctly simulated clouds in the observed shallow frontal cloud layer. Most SCMs also produced clouds in the middle and upper troposphere while only shallow clouds were observed. Intermodel differences in the SCM-simulated cloud properties and their profiles were as large as those for summertime continental convection (Xie et al. 2002; Xu et al. 2002), but the intermodel differences among the thermodynamic profiles were comparable between the CRMs and SCMs for this case study. Overall, some CRMs did not perform better than the SCMs did.

d. Midlatitude cirrus clouds

The first two WG 2 (cirrus cloud) projects were the Idealized Cirrus Model Comparison (ICMC) Project, developed and led by David Starr (NASA GSFC), and the Cirrus Parcel Model Comparison (CPMC) Project, developed and led by Ruei-Fong Lin (NASA GSFC). The ICMC Project involved the comparison of simulations of cirrus development and dissipation in two idealized baseline environments: a “warm” cirrus case and a “cold” cirrus case (Starr et al. 2000). The cloud was generated in an ice supersaturated layer about 1 km in depth. Continuing cloud formation was forced via an imposed diabatic cooling representing a 3 cm s⁻¹ uplift over a 4-h time span. The simulations then proceeded for an additional 2 h to enable assessment of the cloud dissipation phase of the cloud life cycle among the models. This is a critical issue in that cirrus clouds are commonly observed to be long lasting.

The disagreements between the results from the 16 models submitted for the ICMC Project were substantial. The spread in ice water path (IWP) among the models revealed that at least some of the models had previously unexposed major deficiencies in their representations of ice water precipitation. The differences among the CRMs exceeded what had been expected based on previous studies. The SCMs exhibited a similar

³ An LES is a 3D simulation that resolves the large, energy-containing turbulent eddies and thus does not rely on a turbulence parameterization to represent the effects of turbulence, unlike a CRM.
range. It was concluded that the models, which spanned a considerable range in terms of their design and heritage, should be able to be tested under much better circumstances as new information becomes available. Advances in observational capabilities, including those in measurement of small ice crystal populations, should be able to adequately resolve the shape of the ice water content profile and the overall ice water path. The result that internal cloud circulation intensity is highly correlated with the ice crystal size distribution should allow for an additional confirming test. At the time of this study, observations of bulk ice water fall speed were just starting to be derived from millimeter-wavelength Doppler radar.

The purpose of the CPMC Project was to compare cirrus simulations by parcel models as well as by parcel-model versions of multidimensional models (Lin et al. 2002). The CPMC Project directly addressed the primary source of dispersion found in the results of the ICMC Project: the development of the ice crystal size distribution. The models used microphysical schemes in which the size distribution of each class of particles (aerosols and ice crystals) is represented by bins or the evolution of each individual particle is calculated. The concept was to look at relatively simple calculations and then to investigate more complex issues, such as effects of gross changes in the aerosol particle size distribution, effects of aerosol particle composition, and direct radiation effects on ice crystal growth rate. The initial focus was on any and all ice nucleation mechanisms included in cirrus parcel models. A second focus was on the homogeneous nucleation process operating in isolation.

There was qualitative agreement for the homogeneous-nucleation-only simulations. For example, the number density of nucleated ice crystals increased with the strength of the prescribed updraft. However, significant quantitative differences were found. The differences in cloud microphysical properties between state-of-the-art models of ice crystal initiation were significant. Inter-model differences in the case of all-nucleation-mode simulations were even greater than in the case of homogeneous nucleation acting alone. It was concluded that definitive laboratory and atmospheric benchmark data are needed to improve the treatment of heterogeneous nucleation processes.

GCSS did not organize any modeling studies of cirrus clouds after 2002 because the cirrus community became heavily involved with convectively generated cirrus in the tropics and the ice microphysics community with Arctic mixed-phase clouds (see section 3). However, two modeling studies of midlatitude cirrus clouds observed at the ARM SGP site were made independently of GCSS under ARM support and are described next.

![Fig. 25-2. Time–height cloud fraction at ARM SGP, 19 Jun–17 Jul 1997, surface to 16 km: (top) observed by millimeter cloud radar (3-h averages), (middle) simulated by the UCLA–CSU CRM (1-h averages), and (bottom) simulated by the NCEP SCM (3-h averages). Color indicates cloud fraction, which ranges from 0 (violet) to 1 (red).](image-url)
based on intensive observations made at the ARM SGP site from 19 June to 17 July 1997. During this period, cirrus clouds, many generated by deep convection, were observed about 30% of the time by the cloud radar. There are significant differences between the NCEP SCM and observed cloud fraction profiles, most notably in the SCM’s underestimate of cloud fraction at high levels (Fig. 25-2).

To produce cirrus statistics from the SCM results that are comparable to the cloud radar observations, Luo et al. (2005) employed a method described by Klein and Jakob (1999) that uses the SCM cloud fraction profile and the SCM’s overlap assumption (random or maximum/random) to create a synthetic cloud field. Luo et al. sampled the synthetic cloud fields as would a cloud radar to determine the statistical properties of cirrus. They then compared the SCM’s cirrus cloud properties to those obtained by Mace et al. (2001) using the MMCR and to a corresponding CRM simulation. Luo et al. found that cirrus properties simulated by the SCM significantly differed from cloud radar observations, while the CRM simulation reproduced most of the cirrus properties as revealed by the observations.

Luo et al. (2006) used the CRM results to evaluate the SCM’s representation of the physical processes determining the simulated cirrus: specifically, cumulus detrainment and ice microphysics. They found that in the SCM (i) detrainment occurs too infrequently at a single level at a time, though the detrainment rate profile averaged over the entire IOP is comparable to that in the CRM simulation; (ii) too much detrained ice is sublimated when first detrained; and (iii) snow sublimates over too deep of a layer.

Solch and Karcher (2011) applied a cloud-resolving model with explicit aerosol and ice microphysics and Lagrangian ice particle tracking to simulate the evolution of a cirrus cloud field observed during the ARM IOP in March 2000. This dataset includes remote sensing, radiosonde, and aircraft measurements of a midlatitude cirrus cloud system, supported by estimates of the dynamical cloud forcing. The dataset allowed Solch and Karcher to evaluate and study in great detail the process-oriented representation of the microphysical processes relevant to the formation and evolution of deep, stratiform cirrus (in particular, ice crystal sedimentation and aggregation). The suite of explicitly resolved physical processes in their model enabled them to better understand the sensitivity of the simulated cirrus properties to a large number of microphysical and environmental parameters.
Yang et al. (2012) developed an observationally based case study that is suitable for a Global Atmospheric System Studies (GASS) cirrus modeling intercomparison study. The case is based on measurements made on 9 March 2000 at the ARM SGP during an IOP. The retrievals of ice water content, ice number concentration, and fall velocity provide several constraints to evaluate model performances. Initial testing of the case using the Met Office Large Eddy Model suggests that the case is appropriate for a model intercomparison study.

A midlatitude cirrus case based on ARM measurements was developed in 2012 and included as part of the joint 8th International Cloud Modeling Workshop (Muhlbauer et al. 2013) and GASS Cirrus Model Intercomparison Project led by Andreas Muhlbauer and Thomas Ackerman (both at the University of Washington). The case describes a synoptically driven thick cirrus developing over the ARM SGP site on 1 April 2010 during the Small Particles in Cirrus (SPARTICUS) field campaign. The goal of the project was to investigate the microphysical and macrophysical evolution and life cycle of a synoptically driven cirrus and to compare simulated cirrus cloud properties and radiative effects among models. Special emphasis was placed on the contribution of small ice crystals in cirrus and the role of homogeneous and heterogeneous ice nucleation. Simulations were compared and evaluated with in situ aircraft observations and with various ground-based and spaceborne remote sensors. This project specifically targeted CRM, LES models, and SCMs with advanced cloud microphysics schemes, such as multimoment bulk microphysics parameterizations or bin microphysics schemes.

Preliminary analysis of the modeling results (Muhlbauer et al. 2013) suggests that models have difficulty in predicting the observed ice number concentrations and representing the vertical structure of the cirrus. Ice number concentrations are overestimated in the homogeneous freezing regime and at cold temperatures but are underestimated in the warmer temperature regime in which ice is initiated by heterogeneous ice nucleation mechanisms. The modeling results also indicate deficiencies in correctly representing the observed vertical profiles of ice water content and radar reflectivity and show an underestimation of the ice water path in the mesoscale cirrus cloud field.

5. Long-term cloud-resolving simulations

Wu et al. (2007) performed a CRM simulation for a period of 26 days (22 June–17 July) during the Summer 1997 IOP at the ARM SGP site. A 2D version of the CRM was aligned east to west and used a domain 600 km long and 40 km deep. The model employed a 3-km horizontal grid size and a stretched grid in the vertical (100 m at the surface, increasing to 1500 m at the model top). The CRM differences from the ARM measurements, when averaged over the entire period, were less than 5 W m^-2 in both longwave and shortwave radiative fluxes at the top of the atmosphere and surface. Using the CRM simulation as a benchmark, it was demonstrated that the conventional GCM radiation calculation greatly underestimates the shortwave and longwave fluxes at the top of atmosphere and surface because of the use of homogeneous cloud condensates and unrealistic random overlap assumptions.

Xie et al. (2004) developed long-term single-column model/cloud-system-resolving model forcing datasets using numerical weather prediction products constrained by surface and top-of-atmosphere observations. Wu et al. (2008) performed a year-long 2D CRM simulation using the variationally constrained large-scale forcing produced by Xie et al. for the year 2000 at the ARM SGP site. The CRM-simulated year-long cloud liquid water path and cloud (liquid and ice) optical depth are in good agreement with the ARM retrievals over the SGP site. The cloudy-sky total (shortwave and longwave) radiative heating profile shows a dipole pattern (cooling above and warming below) during spring and summer, while a second peak of cloud radiative cooling appears near the surface during winter and fall.

Zeng et al. (2007) simulated two 20-day continental midlatitude cases with a 3D CRM and compared the results to ARM data. The first case was the ARM Spring 2000 IOP (Xie et al. 2005; Xu et al. 2005). The second case covered the period from 25 May to 14 June 2002 for the same ARM domain. The two cases represent springtime and summertime midlatitude clouds, respectively. Large-scale forcing was based on the variational analysis approach. The surface fluxes were specified using sitewide averages of observed fluxes from the ARM Energy Balance Bowen Ratio (EBBR) stations or estimated from a land data assimilation system. Evaluation of the simulations shows that the model overpredicts cloud amounts, especially in the upper troposphere. Comparisons of 2D and 3D simulations showed that the 2D CRM not only had unrealistic rapid

\footnote{In 2011, GCSS became a part of the GEWEX GASS Panel. It also abandoned its working group structure and began operating instead through projects that could be initiated by any member of the community.}
fluctuations in surface precipitation but also a spurious dehumidification (or a decrease in cloud amount). ARM surface fluxes were obtained from the EBBR stations, which use the Bowen ratio to partition the fluxes. When Land Information System (LIS) surface fluxes replaced ARM data in the CRM simulations, similar results were obtained, but LIS fluxes produced a better simulation of diurnal cloud variation in the lower troposphere.

6. Convective and cloud processes during TWP-ICE: A multimodel evaluation project

Accurate representation of the characteristics and effects of tropical deep convection remains a leading challenge in global modeling. Cirrus are also ubiquitous in the tropics and have major radiative impacts on their environment, but the properties of these clouds and their connections to deep convection processes are poorly understood. The Tropical Warm Pool–International Cloud Experiment (TWP-ICE) was aimed at describing the properties of tropical cirrus and the convection that leads to their formation (May et al. 2008). TWP-ICE took place 21 January–13 February 2006 in the region near the ARM fixed site at Darwin, northern Australia. The experiment was a collaborative effort between ARM, the Australian Bureau of Meteorology, NASA, the European Commission Directorate-General for Research and Innovation, and several U.S., Australian, Canadian, and European universities. Measurements included a five-station, 3-hourly sounding array, with one station aboard ship offshore. The sounding array enclosed a coastal area roughly the size of a coarse-resolution GCM grid box. Multiple surface energy budget sites characterized maritime and continental surface sensible, latent, and radiative heat fluxes. Scanning precipitation radars provided domainwide polarimetric retrievals of near-surface rain rate and dual-Doppler retrievals of vertical winds over a smaller interior region. Xie et al. (2010) applied the variational analysis approach of Zhang and Lin (1997) to this extensive set of measurements to perform an objective analysis from which the large-scale vertical velocity and the advective tendencies of temperature and water vapor were derived. These were used as the large-scale forcing for the CRM and SCM simulations. TWP-ICE overlapped with the Aerosol and Chemical Transport in Tropical Convection (ACTIVE) program at Darwin, which provided extensive collocated airborne aerosol measurements during TWP-ICE (Vaughan et al. 2008). A chief outcome from TWP-ICE was a dataset suitable to provide the forcing and testing required by modern cloud-resolving models and parameterizations in GCMs. More than 50 observational and modeling papers have been based primarily on TWP-ICE observations. Here, we briefly discuss papers that included cloud-resolving simulations and associated model intercomparison studies.

Because the outstanding questions regarding tropical deep convection (e.g., how do individual cirrus cloud particles form and evolve, and what is the structure and evolution of ascent and descent regions in mesoscale convective systems?) range from the microscale to the mesoscale and larger, the TWP-ICE experimental data motivated a series of interlocking model intercomparison studies that were coordinated by the GCSS Precipitating Cloud System (PCS) group led by Jon Petch (Met Office), following an approach previously tested at the Met Office (Petch et al. 2007). The TWP-ICE model intercomparison studies included one for CRMs with periodic boundary conditions led by Ann Fridlind (NASA Goddard Institute for Space Studies), one for SCMs led by Laura Davies (Monash University), one for limited-area models (LAMs) with open boundary conditions and nested grids led by Ping Zhu (Florida International University), and one for global atmospheric models led by Yanluan Lin (NOAA/Geophysical Fluid Dynamics Laboratory). By coordinating the days simulated to insure spatiotemporal overlap, the four studies allowed direct comparison of results. Each study optimized its approach by requesting diagnostics that focused on its individual primary research questions, which were selected from the wide range of possible relevant issues. Thus, four separate studies emerged, each with some modeling aspects that could be considered cloud resolving (albeit coarse in the GCM case).

The SCM study utilized a 100-member ensemble of large-scale forcings that was produced by considering experimental uncertainty of the retrieved surface rainfall rate that was an input to the variational analysis code (Davies et al. 2013). Since computational expense was not an obstacle, the SCM study spanned the full experiment period, including sequential active monsoon, suppressed monsoon, and break periods. Based on results submitted from 11 SCMs plus a subset of ensemble members from two CRM models (one 2D and one 3D), the largest range of model sensitivity occurred under weak forcing conditions. Overall, SCMs and CRMs differed substantially in predicted surface evaporation. In addition, the vertical structure of cloud variables was relatively insensitive across ensemble members for any given model but demonstrated pronounced differences from one model to the next.

The CRM study emphasized rigorous comparison of simulation results with observations, including forward simulation of Rayleigh radar reflectivity from the 3D...
model output fields for evaluation of convective and stratiform rain areas (Varble et al. 2011; Fridlind et al. 2012a). Owing to computational expense, simulations were limited to the first 16 days of active and suppressed monsoon. A single optional sensitivity test included domain-mean relaxation to observed conditions, which was intended to facilitate comparison with observations by limiting drift of the simulated moist static energy from that observed. In 10 submitted 3D simulations, predicted convective rain area fractions were highly correlated with observations. However, convective areas were systematically overestimated, and stratiform rain areas varied widely around observed values. Considering the handful of domainwide observables during active and suppressed periods (obtained via satellite and scanning radar), the strongest rank correlation was found between large stratiform-area and small top-of-atmosphere outgoing longwave radiation during the active period. Fridlind et al. (2012a) concluded that CRMs require closer observational evaluation of the often poorly predicted and radiatively important stratiform properties and the factors controlling them. Focusing on the source of differences between simulations, Varble et al. (2011) concluded that varying degrees of overestimation among the models of ice radar reflectivity could be attributed more to different assumptions about hydrometeor size distributions among the models than to differences in predicted mass mixing ratios among the models.

The LAM study focused on six simulations from LAMs that used three different large-scale input fields (i.e., different boundary conditions) and compared them with one another and with an ensemble from the CRM intercomparison results (Zhu et al. 2012; Varble et al. 2014a,b). All of the LAM simulations reproduced the observed large-scale wind, temperature, and water vapor fields quite successfully, but the predicted rain rates and cloud properties varied widely, especially in the stratiform rain regions. In a detailed comparison of simulated stratiform rain properties to disdrometer and profiling radar datasets, Varble et al. (2014b) found that biases in rain properties, precipitation radar reflectivities, and mean Doppler fall speeds relative to the observations were similar in LAM and CRM simulations that used similar microphysics schemes. Comparison with observations revealed errors associated with specific one- or two-moment microphysics scheme components, but also indicated the presence of a systematic underestimation of stratiform ice mixing ratio at the melting level in the LAM and CRM simulations. In a companion paper that described a detailed comparison of convective updraft properties with dual-Doppler retrievals, Varble et al. (2014a) indicated that there was a systematic overestimation of updraft speeds in the LAM and CRM simulations. Excessive updraft speeds were most pronounced in the upper troposphere but likely also significant near the surface, where excessive radar reflectivity was collocated with excessive lofting of rain above the freezing level. Together, the results supported a previously reported connection between decreased stratiform detrainment and increased detrainment height.

Petch et al. (2014) compared the LAM, CRM, and SCM results with an intercomparison of nine global atmospheric model simulations of the TWP-ICE monsoon and break periods (Lin et al. 2012). The global model study, which included two simulations with roughly 20-km horizontal resolution, found that predicted cloud properties varied widely across models but were not strongly sensitive to increased horizontal resolution. Petch et al. (2014) speculated that the cloud-resolving nature of the CRM and LAM simulations served to constrain predicted liquid water contents, in contrast to the highly parameterized schemes in SCM and global models, but that ice water contents remained strongly dependent on the representation of the microphysics across all model types.

Although the four independent TWP-ICE intercomparison studies resulted in a wide range of progress and conclusions, a disadvantage of this approach was that substantial differences in large-scale forcing methods and diagnostic definitions complicated direct comparisons across the various simulations and observations. By weighing the benefits of loose coordination against the benefits and costs of tighter coordination, in future studies it may be possible to better optimize the intercomparison approach by identifying a subset of well-constructed diagnostics and a limited set of goals that span the component studies without compromising structural differences that serve independent goals.

Several follow-on studies used the CRM intercomparison simulations: Rio et al. (2013) refined a cold pool parameterization; Mrowiec et al. (2012) reported roughly linear relationships between simulated updraft and downdraft mass fluxes and convective area fractions; and Mrowiec et al. (2015) applied an isentropic analysis approach to remove substantial gravity wave contributions to calculated convective mass fluxes for parameterization development.

Other studies that used independent 2D or 3D CRM simulations with TWP-ICE observations reported a variety of results, including the following:

- Substantial effects of cloud condensation nuclei concentrations, ice nuclei concentrations, freezing parameterizations, and wind shear on simulated anvil area
and water budget terms (Fan et al. 2009b, 2010a,b; Zeng et al. 2009a,b, 2013);

- Increased cloud condensation nuclei concentrations leading to weaker convection because of changes in anvil ice characteristics and subsequent upper-tropospheric radiative heating and weaker tropospheric destabilization in one study (Morrison and Grabowski 2011);

- Weak enhancement of surface precipitation but substantial effects on its spatiotemporal distribution (Lee and Feingold 2010, 2013);

- Relaxation of water vapor toward saturation in overshooting convection (Hassim and Lane 2010);

- Forward simulation of multidirectional polarized reflectance to evaluate anvil-top ice properties (van Diedenhoven et al. 2012).

Other studies using independent LAM simulations reported a strong sensitivity to the choice of bulk parameterization of ice microphysics (Wang et al. 2009), identified the triggering mechanisms of Hector convection\(^5\) (Ferretti and Gentile 2009; Zhu et al. 2013), and identified a relatively shallow entrainment-dominated layer in simulated deep updrafts in contrast to convective parameterization assumptions (Wang and Liu 2009). The studies also tested entrainment parameterization schemes against Weather Research and Forecast (WRF) Model simulations (Wu et al. 2009; Del Genio and Wu 2010), demonstrated the skill of WRF simulations to reproduce observed rainfall statistics (Wapler et al. 2010), and reported improvement of WRF simulations using a 3D variational data assimilation (3DVar) system or another observation incorporation technique (Yeh and Fu 2011; Zhu et al. 2012). Del Genio et al. (2012) used cloud-resolving simulations of TWP-ICE to characterize mesoscale organization processes and to provide parameterization guidance.

Taken as a group, these TWP-ICE cloud-resolving simulation studies emphasized understanding the mechanistic behavior of tropical convection, its sensitivity to environmental conditions, and the degree to which various modeling choices give divergent results or reproduce observations. The need to improve the simulation of ice microphysical processes—including the role of dynamics in determining ice properties—is evident.

7. Arctic clouds

Accelerated warming and rapid environmental change highlight the Arctic as a region particularly sensitive to climate change. Studies have linked this sensitivity to various feedbacks operating in the region, with changes in cloud properties central to these feedbacks. Mixed-phase clouds comprise a large fraction of clouds occurring in the Arctic and are a critical component of the regional climate, but the relevant physical processes and their parameterization in models remain uncertain. Several field experiments (led by ARM or with significant ARM contributions) conducted over the last 15 years have provided measurements to test and improve models. These have included the 1997–98 Surface Heat Budget of the Arctic Ocean (SHEBA) experiment ( Uttal et al. 2002), the 2004 Mixed-Phase Arctic Cloud Experiment (MPACE; Verlinde et al. 2007), and the 2008 Indirect and Semi-Direct Aerosol Campaign (ISDAC; McFarquhar et al. 2011). Observations from these three field experiments have been used to develop case studies for ARM/GCSS/GASS model intercomparison projects.

The first intercomparison project was led by Stephen Klein (Lawrence Livermore National Laboratory) and Hugh Morrison (National Center for Atmospheric Research). It comprised two cases derived from MPACE observations gathered over northern Alaska and the adjacent Beaufort Sea: 1) a single-layer, boundary layer mixed-phase cloud system associated with cold air outflow from the Arctic pack ice southward across open ocean (Klein et al. 2009) and 2) a deeper multilayered mixed-phase cloud system driven by midlevel mesoscale and synoptic-scale forcing as well as cold air outflow over open ocean (Morrison et al. 2009). Xie et al. (2006) developed SCM and CRM large-scale forcing data for MPACE from sounding array data collected by ARM using an objective variational analysis approach. Multiple LES, cloud-resolving, and single-column models participated in the project. The single-layer case included 28 submissions from 17 SCMs and 9 CRMs, the most of any GCSS/GASS intercomparison to date. There was a wide spread of the simulated liquid and ice water paths among the models, leading to large differences in the cloud radiative forcings at the surface. A majority of the models underpredicted liquid water path for the single-layer case, with the opposite for the multilayer case. Models with more sophisticated microphysics schemes tended to produce more realistic results, although there was considerable scatter. There was also a range of several orders of magnitude in the modeled ice crystal concentrations that likely contributed to the large spread in the results. These cases also served as the basis for several additional publications focusing on LES, cloud-resolving, or mesoscale modeling, including Fridlind et al. (2007), Prenni et al. (2007), Luo et al. (2008a,b), Morrison et al. (2008), Fan et al.

\(^5\) Thunderstorms that form regularly over the Tiwi Islands just to the north of Darwin, Australia.
The next GCSS intercomparison project was based on observations of a single-layer mixed-phase case from SHEBA over the central Arctic Ocean and led by Hugh Morrison (Morrison et al. 2011). This case differed from the single-layer MPACE case in several ways: colder cloud temperatures, a sea-ice-covered surface instead of open ocean, and relatively polluted aerosol characteristics. This work built upon the MPACE intercomparison project with the goal of further exploring differences in model results and relationships between the dynamics, radiation, and microphysics driving this spread. To further simplify the model setup and analysis, ice nucleation was constrained in the simulations in a way that held the ice crystal concentrations approximately fixed, with two sets of sensitivity runs in addition to the baseline simulations utilizing different specified ice nucleus concentrations. Simulations clustered into two distinct quasi-steady states consisting of persistent mixed-phase clouds or all-ice clouds. Transitions from the mixed-phase to the all-ice state were accelerated by feedbacks between the dynamics, microphysics, and radiation when the bulk deposition rate of water vapor onto ice exceeded a threshold value. Additional publications based on modeling results from this case include Morrison and Pinto (2005, 2006), de Boer et al. (2010, 2013), van Diedenhoven et al. (2009, 2011), and Fridlind et al. (2012b).

The third and most recent intercomparison project was developed from observations of a single-layer, mixed-phase cloud system over northern Alaska and the adjacent Beaufort Sea observed during ISDAC (McFarquhar et al. 2011) and led by Mikhail Ovchinnikov (Pacific Northwest National Laboratory). This case was similar to the SHEBA single-layer case, but cloud temperatures were warmer, and the cloud-topped mixed layer was decoupled instead of coupled with the surface. Measurements of this cloud system were also much more comprehensive compared to the SHEBA case. The goal of the ISDAC intercomparison was to build upon the MPACE and SHEBA intercomparisons and further examine causes of differences in large-eddy simulations of mixed-phase clouds. To this end, simulations were further constrained compared to the previous intercomparisons. This was done by using the same ice microphysical characteristics (bulk density, capacitance, and fall speed relationships), simplified radiation scheme, and horizontal and vertical grid spacings in all models. Despite these constraints, there was still a fairly large spread of simulated cloud characteristics, including liquid water path and surface precipitation. This was attributed to differences in model dynamics, numerics, and subgrid-scale mixing schemes as well as in representations of the ice particle size distribution shape. Additional LES and cloud-resolving modeling publications based on ISDAC observations include Solomon et al. (2011, 2014), Ovchinnikov et al. (2011), Avramov et al. (2011), and Fan et al. (2011).

8. Summary

In 1993, GCSS promoted a two-stage methodology by which field observations can be combined with simulations using CRMs to develop parameterizations of cloud and precipitation processes for use in global models (Browning et al. 1993). The first stage is to evaluate and improve the CRMs with the help of observational field experiments. The second stage is to use CRMs to develop parameterizations for large-scale models. In 1994, while GCSS was organizing its first model intercomparison, ARM was carrying out its first SCM IOPs. The first variational analysis dataset produced by ARM was based on the Summer 1995 SCM IOP. The first stable version of ARSCL, based primarily on the MMCR, became available in 1997. The first joint GCSS and ARM model intercomparison case was WG 4 Case 3 based on the Summer 1997 ARM SCM IOP. As described in this chapter, GCSS and ARM collaborated on several major studies of cloud systems during the next 10 years. These included studies of the diurnal cycle of shallow cumulus based on an idealization of measurements at the ARM SGP site, mid-latitude frontal clouds based on the ARM March 2000 Cloud IOP, and idealized cirrus clouds via the Idealized Cirrus Model Comparison (ICMC) and Cirrus Parcel Model Comparison (CPMC) Projects. They also included synoptically driven thick cirrus observed over the ARM SGP site on 1 April 2010 during the Small Particles in Cirrus (SPARTICUS) field campaign and deep tropical convection during the Tropical Warm Pool–International Cloud Experiment (TWP-ICE) that took place during 2006 at the ARM site in Darwin, Australia. Arctic cloud studies were based on cases from three field experiments: the 1997–98 Surface Heat Budget of the Arctic Ocean (SHEBA) experiment, the 2004 Mixed-Phase Arctic Cloud Experiment (MPACE), and the 2008 Indirect and Semi-Direct Aerosol Campaign (ISDAC).

Other GCSS-type modeling projects described in this chapter were carried out independently by ARM scientists and included studies of midlatitude cirrus generated by deep convection during the Summer 1997 ARM SCM IOP, the evolution of a cirrus cloud field observed during the ARM IOP in March 2000, the seasonal variation of cloud systems over ARM SGP for the year 2000, and midlatitude clouds during the ARM Spring 2000 IOP and the May–June 2002 IOP.
When GCSS and ARM began collaborating, CRMs were 2D and used single-moment microphysics. Now, 20 years later, 3D CRMs have become standard, and two-moment microphysical schemes are commonly used. Another change has been a shift from the single-column modeling approach to a nested modeling approach, in which the highest-resolution domain is embedded within one or more larger and lower-resolution domains that provide lateral boundary conditions. ARM's observational capabilities also have increased with the deployment of improved cloud and precipitation radars.

There will always be a need for parameterizations in atmospheric models. The need for cumulus parameterization has been removed from some global models by embedding CRMs within each grid column or by increasing the horizontal resolution to that of a CRM. However, these approaches are computationally expensive, so improving conventional cumulus parameterizations remains a priority. Furthermore, the need for turbulence parameterization and especially for microphysics parameterization in global models and cloud-resolving models will not go away any time soon. Improving these parameterizations will remain challenging tasks for GASS and Atmospheric System Research (ASR; Mather et al. 2016, chapter 4) in the years ahead.

The impacts of GCSS-style modeling projects are wide but diffuse and difficult to summarize strictly in terms of scientific results in papers. GCSS promoted the use of CRMs (and LES models) to better understand the cloud-scale processes that must be parameterized in GCMs. This methodology (the SCM approach) requires high-quality datasets for forcing and evaluating the SCMs and the CRMs, and it requires CRMs that reproduce the observations to an adequate degree, in order to convince parameterization developers to use the results of the CRMs for parameterization development and improvement. What may not be so obvious is that observationally based model intercomparison projects do not necessarily lead directly to parameterization improvements. More often, they establish the capabilities or lack thereof of the CRMs, and are followed by idealized and simplified model intercomparison cases, which have two important advantages over strictly observational based ones: by simplifying a case (by limiting the physical processes involved), the CRMs could reach a consensus; and by idealizing a case, a particular aspect of a cloud system could be studied without extraneous complications. There were many examples of first trying a realistic case, then simplifying and idealizing it. Insights that led most rapidly to parameterization improvements were almost always obtained from the idealized cases.

An indirect but ultimately perhaps very important impact of GCSS on parameterization development was and remains that of building a community of cloud-resolving modelers who are interested in parameterization development and know how to use CRMs and observations for that purpose.

Acknowledgments. The authors acknowledge the support and efforts of the ARM Climate Research Facility site and field campaign teams whose measurements are the basis of the modeling studies described in this chapter. In addition, the first author in particular acknowledges David A. Randall, Minghua Zhang, Richard T. Cederwall, and Shaocheng Xie for instituting the first SCM IOPs and for developing the methods used to create high-quality large-scale forcing data sets from SCM IOP measurements. 

REFERENCES


—, —, —, and —, 2011: Influence of humidified aerosol on lidar depolarization measurements below ice-precipitating Arctic


