An Intercomparison of Cloud-Resolving Models with the ARM Summer 1997 IOP Data

Kuan-Man Xu[1], Richard T. Cederwall[2], Leo J. Donner[3], Wojciech W. Grabowski[4], Francoise Guichard[5], Daniel E. Johnson[6], Marat Khairoutdinov[7], Steven K. Krueger[8], Jon C. Petch[9], David A. Randall[7], Charles J. Seman[3], Wei-Kuo Tao[6], Donghai Wang[10,1], Shao Cheng Xie[2], J. John Yio[2], and Ming-Hua Zhang[11]

[1] NASA Langley Research Center, Hampton, VA, USA
[2] Lawrence Livermore National Laboratory, Livermore, CA, USA
[3] NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[4] National Center for Atmospheric Research, Boulder, CO, USA
[5] Centre National de Recherches Meteorologiques, France
[6] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[7] Colorado State University, Fort Collins, CO, USA
[8] University of Utah, Salt Lake City, UT, USA
[9] The Met Office, Bracknell, UK
[10] Hampton University, Hampton, VA, USA
[11] State University of New York, Stony Brook, NY, USA

Submitted to
Quarterly Journal of the Royal Meteorological Society

January 19, 2001

Corresponding author address:
Dr. Kuan-Man Xu
Mail Stop 420
NASA Langley Research Center
Hampton, VA 23681
E-mail: k.m.xu@larc.nasa.gov

Summary

This paper reports an intercomparison study of midlatitude continental cumulus convection simulated by eight 2-D and two 3-D cloud resolving models (CRMs), driven by observed large-scale advective temperature and moisture tendencies, surface turbulence fluxes, and radiative heating profiles during three subperiods of the Summer 1997 Intensive Observation Period (IOP) of the U.S. Department of Energy's Atmospheric Radiation Measurement (ARM) program. Each subperiod includes two or three precipitation events of various intensities over a span of 4 or 5 days. The results can be summarized as follows.

CRMs can reasonably simulate midlatitude continental summer convection observed at the ARM Cloud and Radiation Testbed (CART) site in terms of the intensity of convective activity, and the temperature and specific humidity evolutions. Their performance is comparable to that of tropical oceanic convection, especially the temperature and specific humidity departures from the observations. The 2-D results are very close to those produced by 3-D version of the same models. The observed cloud properties are useful for identifying some model deficiencies. Delayed occurrences of the initial precipitation events are a common feature for all three subcases among the models. Cloud mass fluxes, condensate mixing ratios and hydrometeor fractions produced by all CRMs are basically consistent with each other. All CRMs produce large downdraft mass fluxes with magnitudes similar to those of updrafts, in contrast with CRM results for tropical convection. Some intermodel differences in cloud properties are likely to be related to those in the parameterizations of microphysical processes.

There is generally good agreement between the CRMs and observations with CRMs being significantly better than single-column models (SCMs), suggesting that current results are suitable for use in improving parameterizations in SCMs. However, improvements can still be made in the CRMs; those include the proper initialization of the CRMs, additional sensitivity tests for addressing some model deficiencies and a more proper method of diagnosing cloud boundaries in model outputs for comparison with satellite and cloud radar observations.

Keywords: Cloud Resolving Models, Intercomparison Study, Continental Cumulus Convection

1. Introduction

Cloud-related processes occur on finer scales than those resolved by large-scale models. A subset of these models are the general circulation models (GCMs) used for weather forecast and climate studies. Therefore, these models have to use parameterizations to represent these subgrid-scale processes, for example, cumulus convection, cloud microphysics and cloud fraction parameterizations. Improvements of GCMs rely heavily on the development of more physically-based parameterizations of cloud processes. It is the objective of the Global Energy and Water-cycle Experiment (GEWEX) Cloud System Study (GCSS) to develop new parameterizations of cloud-related processes for large-scale models (Browning 1994; Randall et al. 2001).

An important tool for achieving the GCSS objective, in addition to observational data, is fine-resolution numerical models such as large-eddy simulation (LES) and cloud resolving models (CRMs). Both allow an explicit treatment of fine-scale cloud dynamics and cloud systems. In fact, observations alone, though extremely valuable, cannot provide all the data needed for a thorough development and evaluation of many aspects of the parameterizations of cloud processes. For instance, cloud mass fluxes, which are central to many existing cumulus parameterizations, are very difficult to retrieve from observations. Therefore, LES and CRMs appear as the well-suited tools that fill the gap between sparse observations and parameterization development (Randall et al. 1996), for an in-depth understanding of cloud-related processes, an essential step to the formulation of any advanced and physically sound parameterization of these processes.

Because LES and CRMs have their own subgrid-scale parameterizations and numerical uncertainties and there are no complete data set to verify the performance of all aspects of numerical simulations by these models, a standard approach that has been widely adopted in the community is the intercomparison study (e.g., Gates, 1992; Moeng et al. 1996; Boyle et al. 2000; Ghan et al. 2000). In order to have a successful intercomparison study, high quality data are needed. In outlining the approach of precipitating convective cloud systems working group (WG)

of GCSS i.e., WG 4, Moncrieff et al. (1997) concluded: “A comprehensive evaluation of state-of-the-art CRMs will require state-of-the-art observations.” In particular, cloud parameter observations should be available for the intercomparison, in addition to the large-scale thermodynamic variables and the radiative fluxes from the surface and the top of the atmosphere. Some recent field campaigns have provided increasingly more comprehensive observations of cloud properties, in particular, the Tropical Ocean and Global Atmosphere (TOGA) Coupled Ocean-Atmosphere Response Experiment (COARE; Webster and Lucas, 1992) and the U.S. Department of Energy’s Atmospheric Radiation Measurement (ARM) program (Stokes and Schwartz 1994).

GCSS WG 4 has conducted two case studies: Case 1, two- and three-dimensional modeling of a squall line on a time scale of a few hours, and Case 2, two-dimensional simulation of the evolution of convection over a week-long period under imposed large-scale conditions observed during TOGA COARE. Case 1 simulations were compared with kinematic and radar reflectivity observations from airborne Doppler radar (Redelsperger et al. 2000). Open or periodic boundary conditions were used without large-scale advective forcings. This study shows: 1) a broad agreement in describing the overall structure and propagation of the squall line, 2) some quantitative agreement in the evolution of surface precipitation, and 3) a qualitative agreement in describing the vertical structure of the apparent heat source and moisture sink [see Yanai et al. (1973) for their definitions]. In addition, ice-phase microphysics is important for a proper simulation of the dynamic structure of the squall line. There are differences in cumulus momentum transports between 2-D and 3-D simulations.

The second case, on the other hand, deals with multiple episodes of deep convection during 20 - 26 December 1992 subperiod of TOGA COARE Intensive Observation Period (IOP; Krueger et al. 2001). One 3-D CRM, seven 2-D CRMs and five single column models (SCMs) participated in the intercomparison. All models are driven by the observed large-scale advective tendencies and sea surface temperatures (SSTs). The main findings are: 1) that the bulk character-

istics of convection are determined (in a diagnostic sense) by the large-scale advective tendencies; 2) that the temperature errors of 2 K in all CRMs and 3K in all SCMs are mostly due to errors in the imposed large-scale advective tendencies of potential temperature and/or water vapor, not to model deficiencies; and 3) that the intermodel differences in most time series and profiles of cloud properties among the CRMs are significantly smaller than those among the SCMs.

There have also been a few “long-term” simulations (i.e., over one week periods) using the same approach as in the Case 2 study (e.g., Grabowski et al. 1996, 1998; Guichard et al. 2000; Johnson et al. 2001; Li et al. 1999; Su et al. 1999; Wu et al. 1998; Xu and Randall 1996, 2000a). These studies focused on tropical convection using either the Global Atmospheric Research Program’s Atlantic Tropical Experiment (GATE; Kuettner and Parker 1976) Phase III or the TOGA COARE data set to conduct 2D and sometimes 3D CRM simulations, except for Xu and Randall (2000a), who simulated midlatitude convection with the ARM July 1995 IOP data. In all these studies, the simulated thermodynamic profiles and characteristics of convective cloud systems can be compared with observations. However, a quantitative evaluation of cloud properties such as cloud mass fluxes with observations could not be performed, due to lack of adequate observations. It is sometimes difficult or impossible to quantitatively evaluate several aspects of a model from an individual study. Therefore, a model intercomparison study is a more proper avenue.

In a related research effort, the ARM single-column model (SCM) WG conducted its intercomparison study of midlatitude summertime convection using the July 1995 IOP data set (Ghan et al. 2000), i.e., ARM SCM Case 1. Only one CRM participated in this study, compared to eleven SCMs. Only two of the SCMs also participated in Case 2 of GCSS WG 4. The major findings of this study are related to the relative magnitudes of model deficiencies to errors due to large-scale advective forcing errors, the analysis methods of large-scale advective forcings, the merit of different methods of imposing the large-scale advective forcings, the sensitivity to prescribed surface fluxes, and the relatively better performance of the CRM compared to the SCMs.

To further advance the goals of both GCSS WG 4 and ARM SCM WG, both groups decided to have a joint intercomparison project to maximize the resources. Advancing the understanding of midlatitude continental convection is the primary goal of this case. This joint project (Case 3) compares two 3-D CRMs, eight 2-D CRMs and thirteen SCMs by simulating continental cumulus convection observed at the Southern Great Plains (SGP) Cloud and Radiation Testbed (CART) site during Summer 1997 IOP of the ARM program. A rich variety of cloud parameter observations are obtained by the ARM Cloud Properties Working Group. This paper reports the results of cloud-resolving simulations and the findings of the intercomparison, while the SCM part of the project is reported elsewhere (Xie et al. 2001).

The goals of the present CRM intercomparison study are:

1) to evaluate CRM simulations with the detailed cloud parameter observations;

2) to compare the performance of CRMs in simulating midlatitude convection with that of tropical convection; and

3) to provide detailed analyses of results, essential to further improvement of the representations of cloud processes in GCMs.

Although the CRMs differ significantly in their treatments of cloud microphysics, radiation, and turbulence closures, they all explicitly (and marginally) resolve cloud-scale dynamics in two or three dimensions of limited-size domains. The majority of them have participated in GCSS Cases 1 and 2 (Krueger et al. 2001; Redelsperger et al. 2000), simulating tropical deep convection. Most of them have, however, not been used to simulate midlatitude continental convection with an observed large-scale data set. This is another unique aspect of this intercomparison study, in addition to the comparison with more comprehensive cloud parameter observations than earlier intercomparison efforts, e.g., GCSS WG 4 Case 2 (Krueger et al. 2001) and ARM SCM Case 1 (Ghan et al. 2000). Furthermore, this work serves as a foundation for further work by contributing groups to address many issues raised in this paper, in particular, some model deficiencies.

2. Design of simulations and descriptions of CRMs

a) Simulations

Eight groups in this intercomparison perform simulations with 2-D (x and z) CRMs (Table 1). All 2-D models orient on the east-west direction. CSULEM and UKLEM (one subcase only) also perform 3-D simulations. The domain sizes are 512 km in the horizontal for most models except for UKLEM (500 km) and EULAG (600 km) and approximately 20 km (25 km for EULAG, 26 km for LaRC and 27 km for CSULEM) in the vertical for 2-D simulations. For 3-D simulations, the domain sizes are 250 km x 250 km x 20 km for UKLEM and 250 km x 250 km x 27 km for CSULEM, respectively. The horizontal grid sizes used in these simulations are 2 km (2-D) or 2 km x 2 km (3-D) except for EULAG (3 km). Periodic (cyclic) lateral boundary conditions are implemented in all models. The vertical resolution varies from model to model with stretched coordinates, with layer thickness in the range of 70 m and 1000 m, or uniform 500 m layer thickness (EULAG and GFDL). The number of vertical layers varies from 34 to over 60 (Table 2).

Table 1: Summary of CRMs used in this intercomparison study.

Model	Model full name	Reference(s)	Modeler(s)
CNRM	Centre National de Recherches Meteorologiques	Redelsperger and Sommeria (1986); Caniaux et al. (1994)	Guichard
CSULEM	Colorado State University LES/CRM	Khairoudinov and Kogan (1999)	Khairoudinov
EULAG	NCAR EUlerian/semi-LAGrangian cloud model	Grabowski and Smolarkiewicz (1996); Smolarkiewicz and Margolin (1997); Grabowski (1998)	Grabowski
GCE	NASA Goddard Cumulus Ensemble model	Tao and Simpson (1993)	Tao, Johnson
GFDL	NOAA Geophysical Fluid Dynamic Laboratory	Lipps and Hemler (1986); Held et al. (1993); Donner et al. (1999)	Donner, Seman
LaRC	NASA Langley Research Center Advanced Regional Prediction System (ARPS)	Xue et al. (2001)	Wang, Xu
UCLA/ CSU	University of California-Los Angeles/Colorado State University	Krueger (1988); Xu and Randall (1995)	Xu
UKLEM	UK Met Office Large-Eddy Model	Shutts and Gray (1994)	Petch

Table 2: The cloud resolving model codes: numerics, cloud microphysics and turbulence closure. NFT stands for nonoscillatory forward-in-time while A-B for Adams-Bashforth.

Model	Dimension	Time differencing	Time step	Momentum advection	Vertical layers	Bulk cloud microphysics	Turbulence closure
CNRM	2-D	Leapfrog	12 s	2nd-order	48	5 category	1.5 order
CSULEM	2-D/3-D	A-B 3d	10 s	2nd-order	64	4 category	1.5 order
EULAG	2-D	NFT	15 s	2nd-order	51*	4 category	1.5 order
GCE	2-D	Leapfrog/NFT	6 s	2nd-order	41	5 category	1.5 order
GFDL	2-D	Leapfrog	2 s	2nd-order	42*	4 category	1st order
LaRC	2-D	Leapfrog	3 s / 6 s	4th-order	53	5-category	1.5 order
UCLA/CSU	2-D	A-B 2d	10 s	2nd-order	34	5 category	3rd order
UKLEM	2-D/3-D	Leapfrog	variable	2nd-order	60	5 category	1st order

The horizontally uniform, large-scale advective cooling and moistening rate profiles are prescribed in all simulations. The observed horizontal wind components are nudged so that the domain-averaged horizontal wind components are approximately equal to the observed, with a nudging time of 1 or 2 h (Grabowski et al. 1996; Xu and Randall 1996). The horizontal inhomogeneity of *u* and *v* components inside the CRM domain is preserved by the nudging procedure.

Also prescribed are the radiative heating rate profiles, based upon the European Center for Medium-range Weather Forecasting (ECMWF) forecast model (not shown). This eliminates the complicated interactions between clouds and radiation and simplifies the interpretation of the intercomparison results (Ghan et al. 2000). The impact of interactive radiation may be a focus of further studies from contributing groups (e.g., Xu and Randall 2000b).

Similarly, observed surface turbulence fluxes from Energy Balance/Bowen Ratio (EBBR) measurements are imposed to all models, so that the complicated land surface processes and their feedback to cloud processes are eliminated, as in Ghan et al. (2000). In CRMs, however, only the domain-averaged fluxes are constrained to the observed values. The horizontal variations of the surface fluxes are retained for most models.

In all models, convection is initiated by introducing small random perturbations in the temperature field (0.5 K maximum magnitude) in the subcloud layer for the first hour or so, as in simulations of tropical convection (e.g., Krueger 1988). This aspect of the design of simulation will be further discussed later.

In summary, major differences in the design of simulations between Case 2 (Krueger et al. 2001) and Case 3 consist of 1) prescribing the radiative heating rate profiles, instead of interactive radiation, and 2) prescribing the domain-averaged surface turbulence fluxes of heat and moisture, instead of computing them from the prescribed land-surface temperature and soil wetness.

b) Numerics

All of the model codes have been developed independently although some parts of the models are rather similar between some models. Most of the CRMs are based upon the anelastic dynamics, which filters out the sound waves, except for GFDL and LaRC CRMs which uses the elastic dynamics. Two time steps are used in LaRC CRM (Xue et al. 2001), with the smaller time step for sound waves. The anelastic dynamics allows for a larger time step (see Table 2) for integration but needs to solve an elliptic equation for pressure. The time differencing is either leapfrog (CNRM, GFDL, LaRC, UKLEM, and GCE for dynamic variables only) or Adams-Bashforth schemes (CSULEM and UCLA/CSU; Table 2). A nonoscillatory forward-in-time scheme is used for all variables in EULAG (Smolarkiewicz and Margolin 1997) but only for scalar variables in GCE. The leapfrog scheme requires a time smoothing scheme to overcome the decoupling of solutions at odd- and even-number steps. The advection of momentum is mostly 2nd-order accuracy except for LaRC CRM with 4th-order accuracy. The advection of scalar variables, especially for the mixing ratios of condensate, is accomplished with some form of positive-definite flux-corrected transport schemes of different accuracies and complexities. Other numerical details, such as numerical diffusion and smoothing, and treatment of sponge layers in the upper portion of the CRMs, vary from one model to another.

c) Cloud microphysics

Cloud microphysics parameterization is one of most complicated and least understood aspects of CRMs. A common feature of all cloud microphysics parameterizations is that a bulk representation is used, with 4 or 5 condensate (cloud water, cloud ice, rainwater, snow and graupel/hail) categories. The majority of CRMs use some variations of Lin et al. (1983) and Rutledge and Hobbs (1984) with 5 condensate categories, for example, CNRM, GCE, LaRC, UCLA/CSU and UKLEM. UKLEM also predicts the number of concentration of cloud ice (Swann 1998). CNRM relaxes the constant slope and intercept parameter assumptions (Caniaux et al. 1994) and includes partial saturation over grid volumes for warm clouds.

Simpler parameterizations are used in CSULEM, EULAG and GFDL. GFDL has prognostic equations for cloud water, snow/ice and rainwater mixing ratios, and a diagnostic partitioning equation between liquid and ice forms of cloud water (Donner et al. 1999). CSULEM predicts two species, total water, which is a sum of water vapor and cloud condensate (water/ice), and precipitating water (rainwater/graupel/snow). The cloud condensate amount is diagnosed by a simple "all-or-nothing" moist-adjustment scheme. The partitioning of different categories of cloud condensate and precipitating water depends upon the ambient temperature. EULAG also uses a simple representation that is an extension of the classical Kessler (1969) scheme (Grabowski 1998). Cloud condensate and precipitating water are predicted, with partitioning between cloud water and cloud ice (between rainwater and snow) on ambient temperature. Graupel/hail is not explicitly considered in EULAG and GFDL, but may be embedded in either rainwater or snow category.

d) Turbulence closure

Turbulence parameterization is another important aspect of CRMs, especially in-cloud turbulence. In all eight CRMs, two use the first-order closure (GFDL and UKLEM; Table 2), five use the one and a half-order closure (CSULEM, EULAG, GCE, LaRC and CNRM), and one uses a third-order closure (UCLA/CSU). In UKLEM, Smagorinsky-Lilly model predicts viscosity and

thermal diffusivity coefficients (Smagorinsky 1963; Lilly 1967), and neutral mixing length is assumed to be 250 m (Brown et al. 1994). GFDL uses a deformation- and Richardson number-dependent subgrid-scale mixing formulation. On the other hand, CSULEM, EULAG, GCE, LaRC and CNRM have a prognostic equation for TKE. The eddy diffusion is computed through the TKE equation (Deardorff 1980). In CNRM, liquid water effect is included in the turbulence flux terms. UCLA/CSU CRM uses a third-moment closure with a diagnostic equation for turbulence length scale. The in-cloud turbulence is treated in great details (Krueger 1988).

Another important aspect of CRMs is the formulation of surface turbulent fluxes of heat, moisture and momentum. Because the domain-averaged fluxes are prescribed in all models, the impact of different formulations on simulated cloud processes cannot be significant. For the sake of brevity, details of these formulations are thus omitted here.

3. Characteristics of Case 3

Three subperiods of 4- to 5-day durations (Fig. 1) are chosen for the simulations, instead of an entire IOP. Each subperiod corresponds to a subcase of Case 3 (Table 3). The Summer 1997 IOP covers a 29-day period, starting from 2330 UTC on 18 June and ending at 2330 UTC on 17 July (Julian Day 170 to 199). Balloon-borne soundings of winds, temperature and dewpoint temperature were obtained every 3 h from the CART central facility located near Lamont, OK (36.61 °N, 97.49 °W) and from four boundary facilities, which form a rectangle of approximate 300 x 370 km^2. The sounding/profiler data, combined with the surface and the top-of-the-atmosphere flux observations, are analyzed over such a horizontal domain, using a constrained variational objective analysis method (Zhang and Lin 1997; Zhang et al. 2001). Details of the ARM IOP observations are also provided by Ghan et al. (2000).

The large-scale advective cooling rates associated with the strong event of Subcase A reach 1.8 K h^{-1} at 7 km on Julian Day 181 (Fig. 2), while the large-scale advective moistening

rates have maxima of 0.4 g kg^{-1} h^{-1} at 2 km and 5 km, respectively (Fig. 3). The zonal wind component has a weak deep shear for most of the subperiod except for strong low-level shear in the last 36 h that is associated with the strong precipitation event (Fig. 4).

Table 3: Summary of characteristics of subcases in this study.

Subcase	Duration	Characteristics of convection
A	2330 UTC 26-30 June 1997 (Julian Day 178 to 182)	A strong precipitation event with a maximum precipitation rate of 3.5 mm h^{-1}on Julian Day 181, and weak precipitation events on Julian Day 179.
B	2330 UTC 7-12 July 1997 (Julian Day 189 to 194)	Three moderate precipitation events with maximum precipitation rates of approximately 1.0 mm h^{-1} (Julian Days 190, 191.5 to 193), with a very short (3 - 5 h) break period between the second and third events.
C	2330 UTC 12-17 July 1997 (Julian Day 194 to 199)	A moderate precipitation event (~ 1 mm h^{-1}) on Julian Day 198 and a few weaker ones (< 0.7 mm h^{-1}) in the middle of the subperiod.

In Subcase B, the large-scale advective cooling maxima (0.5, 0.3 and 0.4 K h^{-1}, respectively) coincide with the observed surface precipitation maxima rather well (Fig. 2). The large-scale advective moistening maxima (0.2 - 0.3 g kg^{-1} h^{-1}) also coincide with the precipitation maxima reasonably well for the last two events (Fig. 3), but not the first event. Two weak moistening maxima (0.1 g kg^{-1} h^{-1} at 5 km, 0.15 g kg^{-1} h^{-1} below 2 km) occur slightly after the first precipitation event. An advective drying maximum of equal strength appears between 2 km and 5 km during this precipitation event, which might not be favorable for the occurrence of cumulus convection. The zonal wind component also shows weak deep shear for most of the subperiod (Fig. 4), except with strong low-level shear in the first 36 h.

In Subcase C, the large-scale advective cooling maxima (0.5 K h^{-1}) correspond to the two relatively strong precipitation events reasonably well, and so do the two advective moistening maxima except that the durations are shorter than those of the precipitation events (Figs. 2 and 3). Advective drying maxima are, however, rather pronounced, occurring just before the two relatively strong precipitation events. The first drying maximum of the two is probably an artifact of the missing soundings on Julian Day 195, which impacts the simulated timing and magnitude of convective activity. The zonal wind shear is also rather weak throughout the subperiod (Fig. 4).

4. Intercomparison results

Two types of the intercomparison results are shown in this paper: the subcase time-mean profiles and time series of surface or integrated variables. Most of the variables shown in this section will be compared with available observations, except for cloud mass fluxes and condensate mixing ratios, for which no data is available. All of them are temporally and spatially averaged, i.e., 3 h in time, and the entire horizontal domain in space.

a. Temporal evolution of surface precipitation and precipitable water

Observations of surface precipitation rates were combined from the rain gauges at the central facility, four boundary facilities and the Oklahoma and Kansas Mesonet stations, as well as NEXRAD radar retrievals. Left panels of Fig. 5 show that all CRMs basically capture the overall temporal evolutions of surface precipitation rates in Subcases A and B and most CRMs have difficulties simulating the observed temporal evolution of Subcase C, in terms of amplitudes and durations of precipitation events. In all three subcases, the degree of agreement with observations is somewhat related to the amplitudes of the large-scale advective tendencies (Subcases A and B) and the accuracy of the large-scale advective tendencies (Subcase C). The high-frequency variations of surface precipitation rates in all subcases, e.g., higher amplitudes and some phase differences from observations, could be also attributed to a single realization of the simulations. An ensemble of simulations with slightly different initial conditions are expected to provide more smoothly varying time series (e.g., Xu and Randall 1996).

Another common feature among the models is that several precipitation events (often the first) are delayed by several hours. The delay is most pronounced in the first precipitation event of subcase B where the onset of precipitation in the models is about 6 hours after the observed precipitation event. As mentioned in Section 2, for this intercomparison and the previous one (Case 2), all models were initialized with horizontally homogeneous soundings apart from small random temperature perturbations applied to the lowest model layers. In the Tropics (Case 2), these

perturbations were able to induce convection because of the small inhibition and large source of moisture. However, in the continental midlatitude (Case 3) these perturbations were not able to readily generate convective circulations due to the large inhibitions at the boundary layer top and the drier environment. Convection was therefore delayed until large amounts of moisture could build up in the boundary layer (see Fig. 5, right panels).

There are several potential reasons why the convection is delayed in Case 3 and it is likely that all are important to some extent. Observations suggest that many convective events in the midlatitudes are initiated by mesoscale circulations but these are not included in the initialization of the CRMs. It is also likely that the CRMs resolution of the boundary layer is important as they need to generate shallow cumulus before the deep convection (i.e., overcoming the large inhibition). The current choice of 2 km is too low for this and sensitivity tests performed by J. Petch using the UKLEM (see Petch and Gray 2001) have shown that the delay in subcase B is completely removed when the horizontal grid length of 2 km is reduced to 250 m. Another suggestion is that inhomogeneous large scale large-scale forcings may improve the simulated timing of precipitation events because the observed cloud systems only occupy a portion of the SGP CART site. The weaker, averaged forcings cannot trigger convection as quickly as in the observations. This conjecture was somewhat supported by sensitivity tests with some idealized partitioning of forcing magnitudes within the domain for the initial 12 h of simulations, performed by the lead author. This approach implicitly imposes some mesoscale circulations, due to horizontal discontinuity of the advective forcings. Finally, the temporal decoupling of large-scale advective moistening from advective cooling maxima, as mentioned in Section 3, is perhaps another reason for the delay in Subcase B (Figs. 2 and 3).

The general characteristics of the observed temporal evolutions of precipitable water (right panels of Fig. 5), which measures the total vapor mass in a vertical column, are basically captured by all CRMs. The observations are based upon radiosondes measurements and estimates

of column water vapor from microwave radiometer (MWR) measurements (Liljegren 1994). The intermodel differences are smaller than the observed temporal variations for most of three subperiods but increase as the model integration time increases in Subcases A and C. For example, the temporal evolution is rather similar among the models for the first two days of Subcase C, but it diversifies greatly in the last two and a half days. Subcase A shows the same behavior more dramatically; i.e., an intermodel difference of 9 kg m^{-2} after the strongest precipitation event (Fig. 5).

Other features noted in the right panels of Fig. 5 are 1) differences in the impact of the delayed initiation of convection on the temporal evolution of precipitable water among the CRMs and 2) some differences between 2-D and 3-D results. Precipitable water is larger for 3-D simulations than for their 2-D counterparts. This may be related to differences in the cloud-scale circulations between 2-D and 3-D models, which impact cloud microphysical processes and their interactions with the environment. Larger cloudy areas in 3-D simulations, as shown later, are associated with more humid environments.

b. Temperature and moisture profiles

In this section, the root-mean-square (RMS) errors of CRM simulations relative to observations and the temporal correlation coefficients between the simulated and observed temperature and water vapor mixing ratio are shown.

First of all, let us examine the impact of the delayed occurrence of the first precipitation event of each subperiod on the temperature and water vapor mixing ratio departures from observations. In Subcase B, the observed maximum precipitation occurs at 30 h, while the simulated maxima appear between 33 and 38 h (Fig. 5). Between 30 and 36 h (early stages of simulated precipitation events), temperature and moisture departures are determined by the large-scale advective effects (Figs. 2 and 3). There are more unstable lapse rates and moister lower/middle troposphere (Fig. 6), because of the imbalance between the large-scale advective effects and the lack of simulated convection. After the maximum precipitation is reached in the models (39 - 45

h), the atmospheres are more stable and the boundary layer is drier, but the middle troposphere is as moist as in the 30 - 36 h period (Fig. 6). The magnitudes of temperature biases are as high as ± 5 K in Subcase B (± 3 K in Subcase A, but ± 5 K in Subcase C) before convection is simulated, but they are about half of the magnitudes after the precipitation event. Large moisture biases in the lower troposphere correspond to larger temperature biases in the middle troposphere, but with opposite signs, in all subcases before the first precipitation events are simulated. There are large differences in the magnitudes of the biases among the models (Fig. 6). However, models with larger temperature biases are generally associated with smaller moisture biases, or vice versa. Finally, the temperature and moisture biases are generally very small when precipitation events are not delayed (not shown).

Figure 7 shows that the typical magnitudes of RMS temperature departures from observations are 1 - 2 K for Subcase A, and 1 - 3 K for Subcases B and C, with largest departures in the upper troposphere. These are rather close to the range of observed variabilities (0.5 - 2.5 K) in these subcases. There are secondary maxima of RMS departures around 7 km and in the planetary boundary layer (PBL). The observed PBL depths exhibit diurnal variations between 0.2 and 2.1 km for these subcases (Krueger et al. 2000). Locations of these maximum RMS departures are generally coincident with the large biases caused by the delayed occurrence of the first precipitation event in each subcase (Fig. 6). If the first precipitation event were adequately simulated, the typical magnitudes of the RMS departures would probably be halved.

Other features appearing in Fig. 7 are 1) different performance of an individual CRM from one subcase to another and 2) the larger intermodel differences above 11 km. The latter may be chiefly related to the weak forcings/small stability and the different treatments of gravity waves reflections from the upper boundary of the models.

The temporal correlation coefficients between the simulated and observed temperature for Subcase B are extremely high below 11 km for all models, i.e., greater than 0.8 (Fig. 8). The lack

of temporal variabilities on the biases after the first precipitation event (not shown) is partially responsible for this. In Subcases A and C, however, the correlation coefficients are generally much lower than in Subcase B. And there are much larger intermodel differences for these two subcases. For example, CNRM and CSULEM perform much better than the other models for the layer between 5 and 10 km in Subcase A. In Subcase C, CSULEM and EULAG performs somewhat poorer than the rest of the models although the RMS temperature errors above the PBL of CSULEM and EULAG are not the largest among the models (Fig. 7). Finally, an ensemble of simulations with slightly different initial conditions may more adequately quantify the differences among the models.

Figure 8 also shows that the intermodel differences in the correlation coefficients can be much larger than those detected by the RMS errors. This is rather obvious for Subcases A and C. The opposite is also true, for example, in Subcase B. Therefore, the combination of Figs. 7 with 8 gives a more comprehensive intercomparison than using either Fig. 7 or 8 alone. This, as shown below, is also true for water vapor mixing ratios (Figs. 9 and 10).

In addition, a comparison between 2-D and 3-D simulations from CSULEM and UKLEM shows that the results are very similar (Figs. 7 and 8). Little improvement is shown with the third dimension, which is also true for the results shown in the rest of the paper. This finding, in agreement with Grabowski et al. (1998), justifies the usage of 2-D models to examine many statistical properties of cumulus convection, at least, the mean fields presented in this paper.

The RMS errors of water vapor mixing ratio are larger in the lower troposphere, i.e., greater than 1 g kg^{-1} in the PBL where the mixing ratio is also larger (Fig. 9), compared to the observed variabilities of up to 1.2 g kg^{-1}. The largest RMS errors occur in Subcases B and C, probably related to the more significant delayed initiation of convection and the larger errors in the advective forcing data than in Subcase A. The large moisture errors in the PBL are caused by the delayed occurrence of the first precipitation event (Fig. 6) and due perhaps to deficiencies of

turbulence parameterizations. The latter is partially indicated by the large differences in the biases of the PBL among the models, especially for Subcase B. In addition, it is difficult to point out which models perform much better, based upon the results shown in Fig. 9. These may suggest that the observed moisture variations are more difficult to simulate.

The temporal correlation coefficients between the simulated and observed water vapor mixing ratios show very high values (> 0.7) above 2 km in Subcase B and above 3 km in Subcase C for most models. The intermodel differences are much greater in Subcase A (Fig. 10), as are those of temperature (Fig. 8). Performance of an individual model varies greatly with height for Subcase A. Inspection of the time-height cross sections of the moisture biases (not shown) suggests that the poor simulations of Julian Day 181 (right panels of Fig. 5) are responsible for the low correlations below 7 km in Subcase A of most models, due possibly to intermodel differences in the formulation of rainwater evaporation.

The small correlation coefficients of water vapor mixing ratio in the PBL of all subcases are possibly caused by 1) the large biases associated with the delayed initiation of convection (Fig. 6), 2) the lack of interactive land-surface processes, 3) the uncertainties in turbulence parameterizations (mentioned earlier), and 4) the time-invariant surface pressure values for each subcase. Sensitivity tests by the lead author show that imposing the time-varying, observed surface pressure values in a CRM also slightly improves the overall results of the simulations. Simulated specific humidities in the PBL exhibit far greater variabilities than the observed (Xu and Randall 2000a) with prescribed surface fluxes. Interactive land-surface physics indeed improves the PBL simulations, according to sensitivity tests performed by the LaRC CRM. In addition, uncertainties of measured surface turbulence fluxes may play a role (Stephen Klein, personal communications).

When compared to the Case 2 study of tropical oceanic convection, the departures from the observations for all three subcases are not much greater. The data quality of Case 2 is perhaps not as high as that of the present case. However, the delayed occurrence of the first precipitation

event of each subcase is largely responsible for the large departures shown in Figs. 7 and 9. Therefore, the performance of CRMs is reasonably acceptable for simulating midlatitude continental convection, relative to that of tropical oceanic convection (Krueger et al. 2001; Xu and Randall 2000a). Moreover, prompter initiation of the first precipitation events in all subcases should significantly reduce the temperature and moisture departures from observations (Fig. 6). Finally, despite the noticed shortcomings of these CRM simulations, shortcomings at least partly explained by oversimplifications in the initiation and forcing methods, the temperature and water vapor mixing ratio simulated by CRMs are much better than those from SCMs (Ghan et al. 2000; Xie et al. 2001). The intermodel differences of the temperature and moisture departures from observations, are much smaller for CRMs, compared to those among the SCMs (Xie et al. 2001), which gives support for the GCSS strategy.

c. Temporal evolution of cloud liquid water path and total cloud amount

A novel aspect of this intercomparison study is that the ARM Cloud Properties Working Group provides observations of several cloud properties such as the total cloud amount, the cloud liquid water path (CLWP), and the hydrometeor fraction profile, which can be extensively used for intercomparison among CRMs for the first time and provide constraints for the simulated cloud properties.

The CLWPs are measured with MWRs at the central and four boundary facilities of the ARM SGP CART site (Liljegren 1994). The uncertainty of the measurements is about 0.03 kg m^{-2} when raindrops are absent from the instrument. Some corrections are also made to eliminate the contamination by raindrops on the instrument. A significant impact of this procedure is that the CLWPs could be severely underestimated during intense precipitation events, for example, on Julian Day 181 of Subcase A (Fig. 11). Also, these “point measurements” might not be fully representative of the domain mean values. Given these shortcomings, left panels of Fig. 11 clearly show that most CRMs produce the CLWP magnitudes comparable to the observed.

The temporal evolutions of the simulated CLWPs (Fig. 11) are generally similar to those of surface precipitation rates (left panels of Fig. 5). Obviously, the magnitudes of CLWPs are not simply related to those of surface precipitation rates; indeed not all clouds produce surface precipitation. Evaporation of raindrops is a significant factor for producing downdrafts in midlatitude (Srivastava 1985). The extent of the agreement between the observed and simulated CLWPs varies among the models. That is, the intermodel differences in CLWPs are much greater than those in surface precipitation rates. It appears that CLWPs from the models with 5-category representations are more similar to each other than with those from 4-category microphysics (EULAG and GFDL), especially in Subcase B. The different magnitudes of temperature and moisture departures from observations among the CRMs (Figs. 7 and 9) may also contribute to the intermodel differences in CLWPs.

In summary, the main point from the left panels of Fig. 11 is that the simulated CLWPs are unambiguously comparable to the observed in magnitudes. The models have difficulties matching the observed temporal evolutions, a shortcoming mostly related to the simulation design, less to the physics of the models, as shown previously. Nevertheless, the consensus of all models agrees with the observed temporal evolutions and magnitudes extremely well except for the first convective event of each subcase and occasionally lack of representativeness of the point measurements.

Two observed column cloud fractions are shown in the right panels of Fig. 11, one from satellite observations (GOES-7 satellite; Minnis et al. 1995) and the other from the retrievals of (single) point measurements of ground-based millimeter-wave cloud radar (MMCR; Moran et al. 1998) at the CART central facility. The latter is just a frequency of retrieved cloudy columns that are sampled at 3-min interval. The definition of a cloudy column is based upon MMCR reflectivity. The satellite procedure uses a threshold method on the brightness temperature. In general, the MMCR cloud fraction is higher than that given by satellite and there are larger temporal variations in the MMCR cloud fractions when satellite observed cloud fractions are low, due to inco-

herent spatial and temporal scales of these data. This comparison suggests that uncertainties of column cloud fraction observations are in the range of 10 - 30%.

The column cloud fractions (right panels of Fig. 11) show some temporal correlations between models and observations, even though the intermodel differences are probably greater than those in CLWP. The simulated column cloud fractions are calculated, based upon the grid-column cloud liquid water + ice path exceeding a threshold of 0.01 kg m^{-2} (Cahalan et al. 1995; Harshvardhan et al. 1994). This threshold could be too high to include many MMCR-observed thin cirrus clouds (Mace et al. 2001). It should also be noted that these clouds are probably not fully resolved by models, with vertical spacings of 500 - 1000 m (Section 2a). Therefore, the magnitudes of observed and simulated column cloud fractions are expected to significantly differ.

After the first day (Subcases A and B) or the first two days (Subcase C), most CRMs produce temporal evolutions of column cloud fractions somewhat similar to the observations. CNRM, GFDL, LaRC and UKLEM reach the observed overcast conditions on a few occasions. However, the column cloud fractions produced by most models are smaller than the observations, especially CSULEM, GCE and UCLA/CSU (right panels of Fig. 11). This is mainly due to the lack of the low-level clouds, according to a comparison between satellite-observed and simulated cloud amounts for different layers (Xu and Randall 2000a). The lack of subgrid saturation parameterizations in most CRMs may be one of the reasons for this discrepancy, because 1- or 2-km grid size is too coarse to resolve many small clouds, especially in the lower troposphere. Finally, it is interesting to notice that the differences between 2-D and 3-D cloud fractions are within the intermodel differences among 2-D CRMs although cloud fractions are higher from 3-D models.

Another common feature among the CRMs is the delayed development of clouds in the first day of each subcase (Fig. 11). The lack of the agreement in the temporal variations after the first day of each subcase is most likely related to the lack of horizontal hydrometeor advection (Petch and Dudhia 1998). This may help overcoming the delayed occurrence of convection for the

first precipitation event of each subperiod if the associated mesoscale circulations could be incorporated into the CRM simulations.

d. Cloud property profiles

For all cloud property and mass flux profiles shown hereafter, the mean profiles averaged over all three subcases are produced, instead of the mean profiles of the individual subcases. The latter were shown in Xu et al. (2000). This procedure does not impact the discussion of the results.

Figure 12 shows the mean profiles of mixing ratios of cloud water, cloud ice, rainwater, snow, graupel/hail, and their sum (total condensate mixing ratio). There are no observations available that can be used to compare with model results. Intermodel differences in cloud water mixing ratios are much smaller than those in cloud ice mixing ratios. As far as the profiles of cloud water mixing ratios are concerned (Fig. 12a), all models agree with each other well except for EULAG. The profiles are significantly different between EULAG and GFDL even both have 4-category microphysics. Among the models with 5-category microphysics, the locations of maximum cloud water mixing ratios are generally similar and the magnitudes are only slightly different (smallest by GCE and largest by CNRM). This result is expected due to their similar representations of warm cloud microphysics.

The intermodel differences in the magnitudes of cloud ice mixing ratios are significant (Fig. 12b). The locations of the maximum values are also different (7.5 km in GFDL to 10 km in 3-D CSULEM, GCE and UKLEM). For example, UKLEM shows the smallest values in the middle troposphere, GFDL has its maximum at 7.5 km and LaRC has the smallest maximum value. Surprisingly, the cloud ice mixing ratio profile of EULAG is rather similar to the other models with the Lin et al./Rutledge and Hobbs ice microphysics schemes, although profiles of other water species are not. The additional dimension significantly impacts the cloud ice/snow mixing ratios, lifting the profiles upwards (CSULEM) or increasing the magnitudes (UKLEM, not shown), which suggests that cloud dynamics are probably different between 2-D and 3-D simulations.

The total condensate mixing ratio does not depend upon the details of conversion processes among water/ice species, but does depend upon their conversions with water vapor and the vertical transport of the hydrometeors both through the in-cloud dynamics and gravitational fall out of the hydrometeors. Apparently, there is consistency for the profiles of total condensate mixing ratios among the CRMs except for EULAG and for the middle troposphere (Fig. 12c). GCE and LaRC CRMs have hail, instead of graupel, as one of the ice-phase categories, for simulating midlatitude convection. Both have smaller total condensate mixing ratios than the other models between 3 and 9 km. EULAG show appreciably large differences in the middle/upper troposphere from other models mainly because snow mixing ratios are much larger than other models (Fig. 12e), due perhaps to its simple parameterization of conversion processes (see section 2). The large rainwater amount in EULAG and GFDL (Fig. 12d) is due perhaps to the omission of graupel/hail as one of the ice-phase categories.

Another feature of Fig. 12c is that there are larger intermodel differences in the middle/upper troposphere, compared to cloud water or cloud ice mixing ratio (Figs. 12a, b). For example, cloud ice/snow mixing ratios are very small in the upper troposphere of GFDL (Figs. 12b, e), and graupel/hail and snow mixing ratios are very small in GCE and LaRC (Figs. 12e, f). On the other hand, rainwater mixing ratio profiles are rather similar among the models except for CNRM, EULAG and GFDL (Fig. 12d) but those of snow and graupel/hail are less similar (Figs. 12e, f). Perhaps it is worth pointing out that there is no reason to believe there would be “rain” throughout the troposphere. The freezing of rain is probably the main process missing in the 4-category microphysics parameterizations. Also, the small cloud ice and graupel mixing ratios in the middle troposphere of UKLEM are partially compensated by the large snow mixing ratios there. These intermodel differences could partly be due to the difficulties of ice-phase microphysics representations in CRMs. However, the definitions of what exactly is meant by ice, snow and graupel may differ from model to model which makes a direct comparison of individual species difficult.

Next, observed hydrometeor fractions are compared with CRM results (Fig. 13). They are based upon the retrievals of MMCR measurements that are averaged on 3-min interval. Whether or not clouds are detected at a height by the MMCR at the central facility of ARM SGP CART site is dependent upon a height-dependent reflectivity threshold (-60 to -45 dBZ; Clothiaux et al. 1999). The reflectivity is contributed not only by cloud water and ice droplets, but also by precipitating hydrometeors. The MMCR measured frequency is called "hydrometeor fraction." Similarly, the CRM hydrometeor fraction is composed cloud occurrence and precipitating fractions.

Cloud occurrence from CRMs is the sum of the "cloudy" grid points at a height, divided by the total number of grid points. A CRM grid point is identified as cloudy if the sum of cloud water and cloud ice mixing ratios exceeds 1% of saturation water vapor mixing ratio with respect to liquid (Xu and Krueger 1991). Precipitation fraction from CRMs is similarly defined with a threshold (10^{-6} kg kg^{-1}) on the sum of precipitating water species. Apparently, the criteria used in CRM diagnoses are not identical to those used in the retrievals of MMCR measurements.

The intermodel consistency for cloud occurrences is rather good among the CRMs, especially in the lower and middle troposphere (Fig. 13a). As expected, cloud occurrences are smaller than the MMCR hydrometeor fractions in the lower and middle troposphere but are very comparable to the MMCR estimates in the upper troposphere (especially in GCE and CNRM).

Most models produce the mean profiles of the hydrometeor fractions similar to the observed, although the simulated fractions are higher than the MMCR estimates below 11 km in EULAG, UKLEM, and UCLA/CSU, and for the middle troposphere in GCE and LaRC, and for the lower troposphere in CNRM and GFDL (Fig. 13b). Both 2-D and 3-D versions of CSULEM have the smallest hydrometeor fractions and smaller than the MMCR estimates. These differences from MMCR estimates are mostly related to those of the snow and graupel/hail mixing ratios (Fig. 12). The small threshold used in the diagnosis of precipitation fractions is another reason. A diagnosis that is consistent with the MMCR retrievals is needed to narrow down the differences.

e. Cloud mass fluxes

The CRM cloud properties shown above are in reasonable agreement with observations (Figs. 11 and 13). Particularly essential to improve cloud parameterizations are variables such as cloud mass fluxes (M_c). There are no corresponding direct observations available. The updraft, downdraft mass fluxes and their sum are compared in this section. Updraft mass flux is defined as:

$$M_u = \int_{\sigma} (\bar{\rho} w d\sigma) \qquad \text{if } w>0,$$

where $\bar{\rho}$ is the density of air, and σ is the updraft area which satisfies the criterion of cloud occurrence described earlier. For downdraft mass fluxes (M_d), two components are considered: saturated downdrafts (*ds*), which satisfy the cloud occurrence criterion, and unsaturated downdrafts (*du*):

$$M_{ds} = \int_{\sigma} (\bar{\rho} w d\sigma) \qquad \text{if } w<0,$$

$$M_{du} = \int_{\sigma_p} (\bar{\rho} w d\sigma_p) \qquad \text{if } w<0,$$

where σ_p is the precipitation area, which is identified using a larger threshold (10^{-4} kg kg^{-1}) than that used for defining the hydrometeor fraction. Because many different scales of motion are present in CRM simulations, the diagnosed mass fluxes include contributions not only from convective-scale (individual strong drafts), mesoscale circulations (weak stratiform precipitation), but also from gravity waves. Other criteria on defining updrafts and downdrafts have also been used in the literature, mainly using the draft intensity (e.g., Tao et al. 1987; Gray 2000).

The consistency of M_c, which are the sums of M_u and M_d, among the models is very good for the mean profiles, as indicated by the small differences from the consensus of all models (Fig. 14a). For comparison, the observed large-scale mass flux, $\overline{M}$, is also shown. Most models produce subsidence in the environment of the middle and upper troposphere, i.e., M_c is greater

than $\overline{M}$, except for UKLEM and the middle troposphere of CNRM and UCLA/CSU (Fig. 12f). That is, downdrafts are relatively strong in these three models.

Another consistent feature among the models is the lack of environmental subsidence in the lower troposphere and the negative M_c in the PBL of all models. The consensus shows the zero-subsidence level at approximately 5 km. This feature is due to the presence of strong precipitating downdrafts and to high cloudbase heights at the midlatitude CART site. A detailed analysis of the downdraft characteristics from CRM simulations is required in order to understand this feature and to improve cumulus parameterizations in SCMs/GCMs.

The mean profiles of M_u and M_d are also quite consistent among the CRMs (Figs. 14b and 14c). A striking feature, however, is that the intermodel differences are larger than those of M_c (Fig. 14a). The intermodel differences in M_u (Fig. 14b) are consistent with those in cloud water mixing ratios (Fig. 12a) and cloud ice mixing ratios (Fig. 12b), except for the large M_u in the upper troposphere of CNRM and UCLA/CSU (perhaps contributed by gravity waves). For example, CNRM and GFDL have the largest cloud water mixing ratios (Fig. 12a) and the largest M_u in the lower troposphere. The smallest cloud ice water mixing ratios correspond to the smallest M_u in the upper troposphere of EULAG and GFDL (Fig. 12b). Significant differences in the profiles of hydrometeor species from model to model also contribute to the intermodel differences in M_u and M_d.

Beyond these intermodel differences, there is a strong consensus among CRMs towards comparable magnitudes in M_u and M_d at most heights. As a result, the net cloud mass flux appears as a relatively small residual of these two mass fluxes (Fig. 14). This feature does not appear in the simulations of tropical oceanic convection (e.g., Xu and Randall 2000a) and might be characteristic of midlatitude convection over land. A change of thresholds used for diagnosis of

updraft and downdraft areas may not significantly impact this result. Clearly, this result stresses the equally important roles of updrafts and downdrafts in midlatitude convection over land. It is probably essential that cloud-related parameterizations capture this feature for a proper representation of these convective systems (Xie et al. 2001).

Further analyses are need to isolate contributions from convective and mesoscale processes, as well as from gravity waves, especially in the upper troposphere. Partitioning of convective and mesoscale processes (e.g., Tao and Simpson 1989; Xu 1995) is a well suited approach to fully understand the physical processes leading to these mass fluxes.

f. Discussion

The agreements between simulations and observations are rather remarkable in many aspects of the Case 3 simulations, for example, intensity of convective events and timing of some events, and temperature and specific humidity evolutions. Some noticeable disagreements are, however, present among the CRMs. Chiefly, the initial convective precipitation events in the CRM simulations of all subcases tend to be delayed relative to observations (Fig. 5). Major probable causes for this are 1) the coarse horizontal resolutions (1-3 km), 2) the lack of initial mesoscale circulations due to initialization from horizontally homogeneous soundings, 3) the initial uniform surface fluxes, and 4) the lack of spatially non-uniform large-scale advective tendencies. Most of these causes are related to oversimplifications in the initiation and forcing methods, not to shortcomings in the models. The delayed occurrence on the simulated thermodynamic profiles leads to significant RMS departures from observations (Figs. 7 and 9), which also impact the simulations of cloud fields and cloud variables in the initial one to two days.

In the present intercomparison study, variable observations of cloud parameters such as cloud liquid water path, column cloud fraction and hydrometeor fraction are available for comparisons with model simulations (Figs. 11 and 13). Some intermodel differences in cloud microphysics parameterizations are readily revealed, for example, 4- vs. 5-category bulk representations and

the ice-phase microphysics. It is, however, difficult to pinpoint the differences between simulations and observations because of some uncertainties in observations, i.e., point measurements vs. spatial averages, and in the best-suited definitions of cloud boundaries (lateral, top and bottom) used in the CRM diagnoses. The definitions of cloud boundaries in the CRM diagnoses are not consistent with those of the different measurements of cloud properties. For example, the column cloud fractions are all severely underestimated, compared with either MMCR or satellite observations (Fig. 11). The hydrometeor fractions show large intermodel differences at all heights (Fig. 13), due perhaps to the small thresholds used in the diagnoses.

Updraft and downdraft mass fluxes also show some intermodel differences among the models though much weaker than those from SCMs (Fig. 14; Xie et al. 2001). Methods of diagnosing M_u and M_d need to be refined because of the presence of multiple-scale processes in the models, as in the real atmosphere. The mass flux profiles are not available from observations but are needed for evaluating cumulus parameterizations, in addition to the diagnoses of cumulus transports of heat, moisture and momentum.

To further understand the differences between simulations and observations and the intermodel differences, further analyses of observations, for example, based upon mesonet measurements, gridded satellite and radar precipitation data, improvements in the variational analysis of the forcing data, e.g., for the horizontal moisture/condensate advection. Furthermore, model sensitivity studies will be helpful to reduce the extent of disagreements between models and observations, e.g., sensitivities to horizontal or vertical resolutions, representations of microphysical processes, and relaxations of oversimplifications in the initiation and forcing methods. In addition, some differences between 2-D and 3-D results also need to be further analyzed. Sensitivity studies by individual modelers would help finding out the causes of some intermodel differences and deficiencies; and addressing some major issues raised in this study. These sensitivity studies are beyond the scope of this intercomparison but should provide very useful findings in the future.

5. Summary and conclusions

In summary, this intercomparison study shows:

1. CRMs can reasonably simulate midlatitude continental summer convection observed at the ARM CART site in terms of the intensity, temperature and specific humidity evolutions;

2. The performance of CRMs for simulating midlatitude continental summer convection is comparable to that of tropical oceanic convection, especially the temperature and specific humidity departures from the observations;

3. Delayed occurrences of the initial precipitation events are a common feature for all three subcases among the CRMs;

4. The observed cloud properties are very useful to identify some model deficiencies;

5. The 2-D results are very close to those produced by 3-D version of the same models; some differences between 2-D and 3-D simulations need to be examined further;

6. Cloud mass fluxes, condensate mixing ratios and hydrometeor fractions produced by all CRMs are consistent with each other. Some intermodel differences in cloud properties are likely to be related to those in the parameterizations of microphysical processes;

7. The magnitudes of the updraft and downdraft mass fluxes are more comparable than those produced by simulations of tropical deep convection.

Acknowledgments: This research was partially supported by the Environmental Sciences Division of the U.S. Department of Energy as part of the Atmospheric Radiation Measurement Program, under grants DE-FG03-95ER61968 (Khairoudinov, Randall and Xu), DE-FG03-94ER61769 (Krueger), and DE-FG02-98ER62570 (Zhang), and Contract W-7405-Eng-48 to LLNL (Cederwall, Xie and Yio). The work at NASA Langley Research Center (Xu and Wang) was partially supported by the NASA Earth Observation System (EOS)/Interdisciplinary Science Program under grant 291-01-91-10. Work at GFDL (Donner and Seman) was partially supported by NASA Contract RR1BNC97. Johnson's and Tao's work is supported by the NASA Headquarters Atmospheric Dynamics and Thermodynamics Program and the NASA Tropical Rainfall Measuring Mission (TRMM). Zhang's research was also partly supported by NSF under grant ATM9701950 to the State University of New York at Stony Brook. Work at NCAR (Grabowski) was supported by NCAR's Clouds and Climate Program. The Met Office (Petch) acknowledges support from the EU contract EVK2 CT199900051 (EUROCS). The simulations by F. Guichard were run on a Cray C90 at NCAR and she was partly funded by the European Program EUROCS during the course of this work.

References

Boyle, J. D., D. R. Durran, C. Chen, B. A. Colle, M. Georgelin, V. Grubisic, W. R. Hsu, C. Y. Huang, D. Landau, Y. L. Lin, G. S. Poulos, W. Y. Sun, D. B. Weber, M. G. Wurtele, and M. Xue, 2000: An intercomparison of model-predicted wave breaking for the 11 January 1972 Boulder windstorm. *Mon. Wea. Rev.*, **128**, 901-914.

Brown, A. R., S. H. Derbyshire, and P. J. Mason, 1994: Large-eddy simulation of stable atmospheric boundary layers with a revised stochastic subgrid model. *Q. J. Roy. Met. Soc*., **120**, 1485-1512.

Browning, K. A., 1994: GEWEX cloud system study (GCSS) science plan. *IGPO Publication Series*, **No. 11**, International GEWEX Project Office, 84 p.

Cahalan, R. F., D. Silberstein, and J. B. Snider, 1995: A validation of a satellite cloud retrieval during ASTEX. *J. Atmos. Sci*., **52**, 3002-3012.

Caniaux, G., J.-L. Redelsperger, and J.-P. Lafore, 1994: A numerical study of the stratiform region of a fast-moving squall line. Part I: General description, and water and heat budgets. *J. Atmos. Sci.*, **51**, 2046-2071.

Clothiaux, E. E., K. P. Moran, B. E. Martner, T. P. Ackerman, G. G. Mace, T. Uttal, J. H. Mather, K. B. Widener, M. A. Miller, and D. J. Rodriguez, 1999: *J. Atmos & Oceanic Tech.*, **16**, 819-827.

Deardorff, J. W., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteor.*, **18**, 495-527.

Donner, L., C. J. Seman, and R. S. Hemler, 1999: Three-dimensional cloud-system modeling of GATE convection. *J. Atmos. Sci*., **56**, 1885-1912.

Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc*., **73**, 1962-1970.

Ghan, S. J., D. A. Randall, K.-M. Xu, R. Cederwall, D. Cripe, J. Hack, S. Iacobellis, S. Klein, S. K. Krueger, U. Lohmann, J. Pedretti, A. Robock, L. Rotstayn, R. Somerville, G. Stenchikov, Y. Sud, G. Walker, S. C. Xie, J. Yio, and M.-H. Zhang, 2000: An intercomparison of single column model simulations of summertime midlatitude continental convection. *J. Geophys. Res*., **105**, 2091-2124.

Grabowski, W. W., 1998: Toward cloud resolving modeling of large-scale tropical circulations: A simple cloud microphysics parameterization. *J. Atmos. Sci*., **55**, 3283-3298.

--------, and P. K. Smolarkiewicz, 1996: On two-time-level semi-Lagrangian modeling of precipitating clouds. *Mon. Wea. Rev.*, **124**, 487-497.

--------, X. Wu, and M. W. Moncrieff, 1996: Cloud-resolving modeling of tropical cloud systems during Phase III of the GATE. Part I: Two-dimensional experiments. *J. Atmos. Sci.*, **53**, 3684-3709.

--------, X. Wu, M. W. Moncrieff, and W. D. Hall, 1998: Modeling of cloud systems during Phase III of GATE. Part II: Effects of resolution and the third spatial dimension. *J. Atmos. Sci*., **55**,Journal of the Atmospheric Sciences: Vol. 57, No. 24, pp. 3953-3970.

Gray, M. E. B., 2000: Characteristics of numerically simulated mesoscale convective systems and their application to parameterization. *J. Atmos. Sci*., **57**, 3264-3282.

Guichard, F., J.-L. Redelsperger, J.-P. Lafore, 2000: Cloud-resolving simulation of convective activity during TOGA-COARE: Sensitivity to external sources of uncertainties. *Q. J. Roy. Meteor. Soc*., **126**, 3067-3096.

Harshvardhan, B. A. Wielicki, and K. M. Ginger, 1994: The interpretation of remotely sensed cloud properties from a model parameterization perspective. *J. Climate*, **7**, 1987-1998.

Held, I. M., R. S. Hemler, and V. Ramaswamy, 1993: Radiative-convective equilibrium with explicit two-dimensional moist convection. *J. Atmos. Sci*., **50**, 3909-3927.

Johnson, D., W.-K. Tao, J. Simpson, and C.-H. Sui, 2001: A study of the response of deep tropical clouds to mesoscale processes, Part I: Modeling strategy and simulation of toga coare convective systems. *J. Atmos. Sci.,* (submitted).

Kessler, E., 1969: *On the Distribution and Continuity of Water Substance in Atmospheric Circulation. Meteor. Monogr.,* **32**, *Amer. Meteor. Soc.*, 84 pp.

Khairoutdinov, M., and Y. L. Kogan, 1999: A large-eddy simulation model with explicit microphysics: Validation against aircraft observations of a stratocumulus-topped boundary layer. *J. Atmos. Sci.,* **56**, 2115-2131.

Krueger, S. K., 1988: Numerical simulation of tropical cumulus clouds and their interaction with the subcloud layer. *J. Atmos. Sci.,* **45**, 2221-2250.

--------, S. M. Lazarus, Y. Luo, and K.-M. Xu, 2000: Interactions of deep cumulus convection and the boundary layer over the Southern Great Plains. *14th Symp. on Boundary Layers and Turbulence, Amer. Meteor. Soc.,* Aspen, CO, 114-117.

Krueger, S. K., S. M. Lazarus, P. Bechtold, S. Chen, D. Cripe, L. Donner, W. W. Grabowski, M. Gray, D. Gregory, F. Guichard, H. Jiang, D. Johnson, R. McAnelly, J. C. Petch, D. A. Randall, J.-L. Redelsperger, C. Seman, H. Su, W.-K. Tao, X. Wu, and K.-M. Xu, 2001: Intercomparison of multi-day simulations of convection during TOGA COARE with several cloud-resolving and single-column models. (*in preparation)*

Kuettner, J. P., and D. E. Parker, 1976: GATE: Report on the field phase. *Bull. Amer. Meteor. Soc.*, **41**, 11-27.

Li, X., C.-H. Sui, K.-M. Lau, and M.-D. Chou, 1999: Large-scale forcing and cloud-radiation interaction in the tropical deep convective regime. *J. Atmos. Sci.*, **56**, 3028-3042.

Liljegren, J. C., 1994: Two-channel microwave radiometer for observations of total column precipitable water vapor and cloud liquid water path. *Proceedings of the Fifth Symposium on*

Global Change, Jan. 23-28, 1994, Nashville, TN, *Amer. Meteor. Soc*., Boston, MA, 266-269.

Lilly, D. K., 1967: The representation of small-scale turbulence in numerical simulation experiments. *Proc. IBM Scientific Computing Symp. on Environmental Science*, Yorktown Heights, NY, 195-210.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22**, 1065-1092.

Lipps, F. B., and R. S. Hemler, 1986: Numerical simulation of deep tropical convection associated with large-scale convergence. *J. Atmos. Sci*., **43**, 1796-181

Mace, G. G., E. E. Clothiaux, and T. P. Ackerman, 2001: The composite characteristics of cirrus clouds; Bulk properties revealed by one year of continuous cloud radar data. *J. Climate*, (revised July 2000).

Minnis, P., W. L. Smith, Jr., D. P. Garber, J. K. Ayers, and D. R. Doelling, 1995: Cloud properties derived from GOES-7 for Spring 1994 ARM Intensive Observing Period using Version 1.0.0 of ARM satellite data analysis program. *NASA Reference Publication 1366*, NASA Langley Research Center, Hampton, VA 23681-0001.

Moeng, C.-H., W. R. Cotton, C. S. Bretherton, A. Chlond, M. Khairoutdinov, S. K. Krueger, W. S. Lewellen, M. K. MacVean, J. R. M. Pasquier, H. A. Rand, A. P. Siebesma, B. Stevens, and R. I. Sykes, 1996: Simulation of a stratocumulus-topped planetary boundary layer: 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.

Moran, K. P., B. E. Martner, M. J. Post, R. A. Kropfli, D. C. Welsh, and K. B. Widener, 1998: An anattended cloud-profiling radar for use in climate research. *Bull. Amer. Meteor. Soc*., **79**, 443-455.

Petch, J. C., and J. Dudhia, 1998: The importance of the horizontal advection of hydrometeors in a single-column model. *J. Climate*, **11**, 2437-2452.

--------, and M. E. B. Gray, 2001: Sensitivity studies using a cloud resolving model simulation of the tropical western Pacific. *Q. J. Roy. Meteor. Soc.*, (submitted).

Randall, D. A., K.-M. Xu, R. C. J. Somerville, and S. Iacobellis, 1996: Single-column models and cloud ensemble models as links between observations and climate models. *J. Climate*, **9**, 1683-1697.

--------, J. Currey, P. Duynkerke, S. K. Krueger, M. W. Moncrieff, B. Ryan, D. O. Starr, M. Miller, W. Rossow, G. Tselioudis, B. A. Wielikci, 2001: The GEWEX cloud system study: A view from 2001. *Bull. Amer. Meteor. Soc.*, (to be submitted).

Redelsperger, J.-L., and G. Sommeria, 1986: Three-dimensional simulation of a convective storm: Sensitivity on subgrid parameterization and spatial resolution. *J. Atmos. Sci*., **43**, 2616-2635.

Redelsperger, J.-L., P. R. A. Brown, F. Guichard, C. Hoff, M. Kawasima, S. Lang, T. Montmerle, K. Nakamura, K. Saito, C. Seman, W.-K. Tao and L. J. Doner, 2000: A GCSS model inter-comparison for a tropical squall line observed during TOGA-COARE. I: Cloud-resolving models. *Quart. J. Roy. Met. Soc*, 115, 222-2121.

Rutledge, S. A., and P. V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cold-frontal rainbands. *J. Atmos. Sci*., **41**, 2949-2972.

Srivastava, R. C., 1985: A simple model of evaporatively driven downdrafts: Application to microburst downdraft. *J. Atmos. Sci*., **42**, 1004-1023.

Shutts, G. J., and M. E. B. Gray, 1994 A numerical modeling study of the geostrophic adjustment process following deep convection. *Q. J. Roy. Meteor. Soc*., **120**, 1145-1178.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Mon. Wea. Rev.*, **91**, 99-165.

Smolarkiewicz, P. K., and L. G. Margolin, 1997: On forward-in-time differencing for fluids: An Eulerian/semi-Lagrangian nonhydrostatic model for stratified flows. *Atmos.-Ocean Special*, **35**, 127-152.

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**, 1202-1221.

Su, H., S. S. Chen, and C. S. Bretherton, 1999: Three-dimensional week-long simulations of TOGA COARE convective systems using the MM5 mesoscale model. *J. Atmos. Sci.*, **56**, 2326-2344.

Swann, H., 1998: Sensitivity to the representation of precipitating ice in CRM simulations of deep convection. *Atmos. Res.*, **48**. 415-435.

Tao, W.-K., and J. Simpson, 1993: The Goddard Cumulus Ensemble Model. Part I: Model description. *Terr. Atmos. Oceanic Sci*., **4**, 35-72.

--------, --------, 1989: Modeling study of a tropical squall-type convective line. *J. Atmos. Sci*., **46**, 177-202.

--------, --------, and and S.-T.Soong, 1987: Statistical properties of a cloud ensemble: A numerical study. *J. Atmos. Sci.*, **44**, 3175-3187.

Webster, P., J., and R. Lukas, 1992: TOGA COARE: The coupled ocean-atmosphere response experiment. *Bull. Amer. Meteor. Soc*., **73**, 1377-1417.

Wu, X., W. W. Grabowski, and M. W. Moncrieff, 1998: Long-term behavior of cloud systems in TOGA COARE and their interactions with radiative and surface processes. Part I: Two-dimensional modeling study. *J. Atmos. Sci*., **55**, 2693-2714.

Xie, S. C., R. T. Cederwall, K.-M. Xu, P. Bechtold, D. G. Cripe, A. D. Del Genio, S. J. Ghan, J. J. Hack, S. F. Iacobellis, S. A. Klein, S. K. Krueger, U. Lohmann, J. C. Petch, D. A. Randall, L. D. Rotstayn, R. J. C. Somerville, Y. C. Sud, K. von Salzen, G. K. Walker, A. Wolf, J. J. Yio, G. Zhang, and M.-H. Zhang, 2001: Intercomparison and evaluation of GCM cumulus parameterizations under summertime midlatitude continental conditions. *Q. J. Roy Meteor. Soc.,* (To be submitted).

Xu, K.-M., 1995: Partitioning mass, heat and moisture budgets of explicitly simulated cumulus ensembles into convective and stratiform components. *J. Atmos. Sci*., **52**, 551-573.

--------, and S. K. Krueger, 1991: Evaluation of cloudiness parameterizations using a cumulus ensemble model. *Mon. Wea. Rev.*, **119**, 342-367.

--------, D. A. Randall, 1995: Impact of interactive radiative transfer on the macroscopic behavior of cumulus ensembles. Part I: Radiation parameterization and sensitivity test. *J. Atmos. Sci*., **52**, 785-799.

--------, --------, 1996: Explicit simulation of cumulus ensembles with the GATE Phase III data: Comparison with observations. *J. Atmos. Sci*., **53**, 3710-3736.

--------, --------, 2000a: Explicit simulation of midlatitude cumulus ensembles: Comparison with ARM data. *J. Atmos. Sci.*, **57**, 2839-2858.

--------, --------, 2000b: Cloud-resolving model simulation of the July 1997 IOP: Comparison with ARM data on short, medium and long subperiod. *Proceedings of the Tenth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, March 13-17, 2000, San Antonio, Texas.

--------, S. K. Krueger, R. T. Cederwall, L. J. Donner, W. W. Grabowski, F. Guichard, D. E. Johnson, M. Khairoutdinov, J. C. Petch, D. A. Randall, C. J. Seman, W.-K. Tao, S. C. Xie, J. J. Yio, and M.-H. Zhang, 2000: Cloud-resolving model intercomparison with the ARM Summer 1997 IOP data. *Proceedings of the Tenth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, March 13-17, 2000, San Antonio, Texas.

Xue, M., K. K. Droegemeier, and V. Wong, 2001: The Advanced Regional Prediction System (ARPS) -- A multi-scale nonhydrostatic atmospheric simulation and prediction tool: Part I: Model dynamics and verification. *Meteor. Atmos. Phy.*, (in press).

Yanai, M., S. Esbensen, and J.-H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci*., **30**, 611-627.

Zhang, M.-H., and J. L. Lin, 1997: Constrained variational analysis of sounding data based on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements. *J. Atmos. Sci*., **54**, 1503-1524.

--------, --------, R. T. Cederwall, J. J. Yio, and S. C. Xie, 2001: Objective analysis of ARM IOP data: Method, feature and sensitivity. *Mon. Wea. Rev.*, **129**, 295-311.

Figure caption

Figure 1: Time series of observed surface precipitation rates during Summer 1997 IOP.

Figure 2: Time-height cross sections of observed large-scale advective cooling rates for Subcases A, B and C. Contour interval is 0.1 K h^{-1}.

Figure 3: Time-height cross sections of observed large-scale advective moistening rates for Subcases A, B and C. Contour interval is 0.05 g kg^{-1} h^{-1}.

Figure 4: Time-height cross sections of observed zonal wind components for Subcases A, B and C. Contour interval is 2 m s^{-1}.

Figure 5:Time series of surface precipitation rates (left panels) and precipitable water (right panels).

Figure 6: Six-hour averaged temperature and water vapor mixing ratio biases for periods before (30-36 h) and after (39-45) the simulated precipitation peaks in Subcase B.

Figure 7: Root-mean-square errors of temperature for Subcases A, B, and C.

Figure 8: Temporal correlation coefficients of temperature for Subcases A, B and C.

Figure 9: Same as Fig. 7 except for water vapor mixing ratio.

Figure 10: Same as Fig. 8 except for water vapor mixing ratio.

Figure 11: Time series of cloud liquid water paths (left panels) and the column cloud fraction (right panels).

Figure 12: The all-subcase mean profiles of the mixing ratios of cloud water mixing ratio, cloud ice mixing ratio, total condensate, rainwater, snow and graupel/hail.

Figure 13: Same as Fig. 12 except for cloud occurrence and hydrometeor fractions. The thick black dashed line shows the consensus of all models.

Figure 14: Same as Fig. 12 except for the net cloud mass flux, updraft and downdraft mass fluxes. The thick black dashed line shows the consensus of all models.

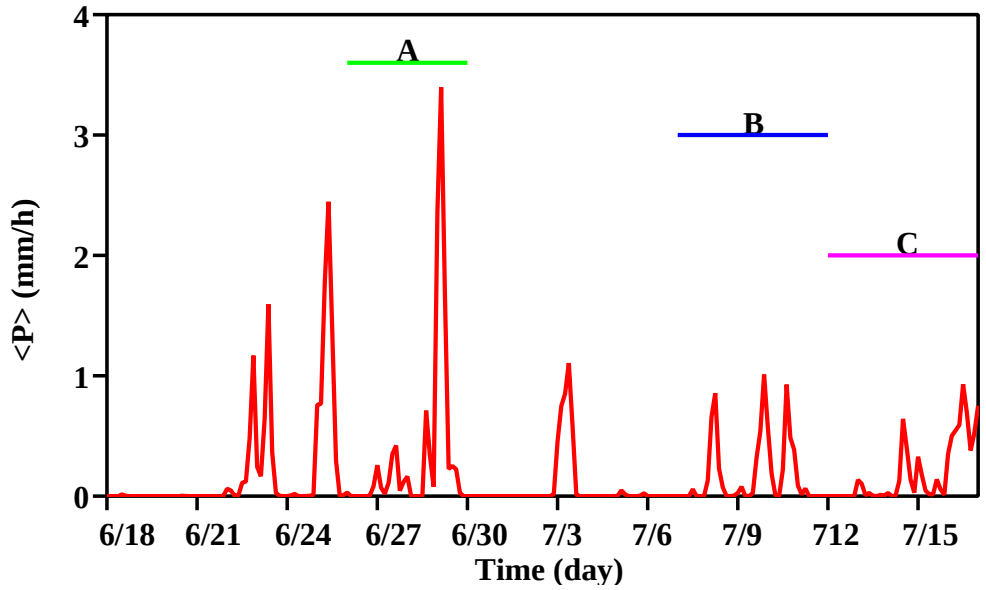

Fig. 1: Time series of observed surface precipitation rates during Summer 1997 IOP.

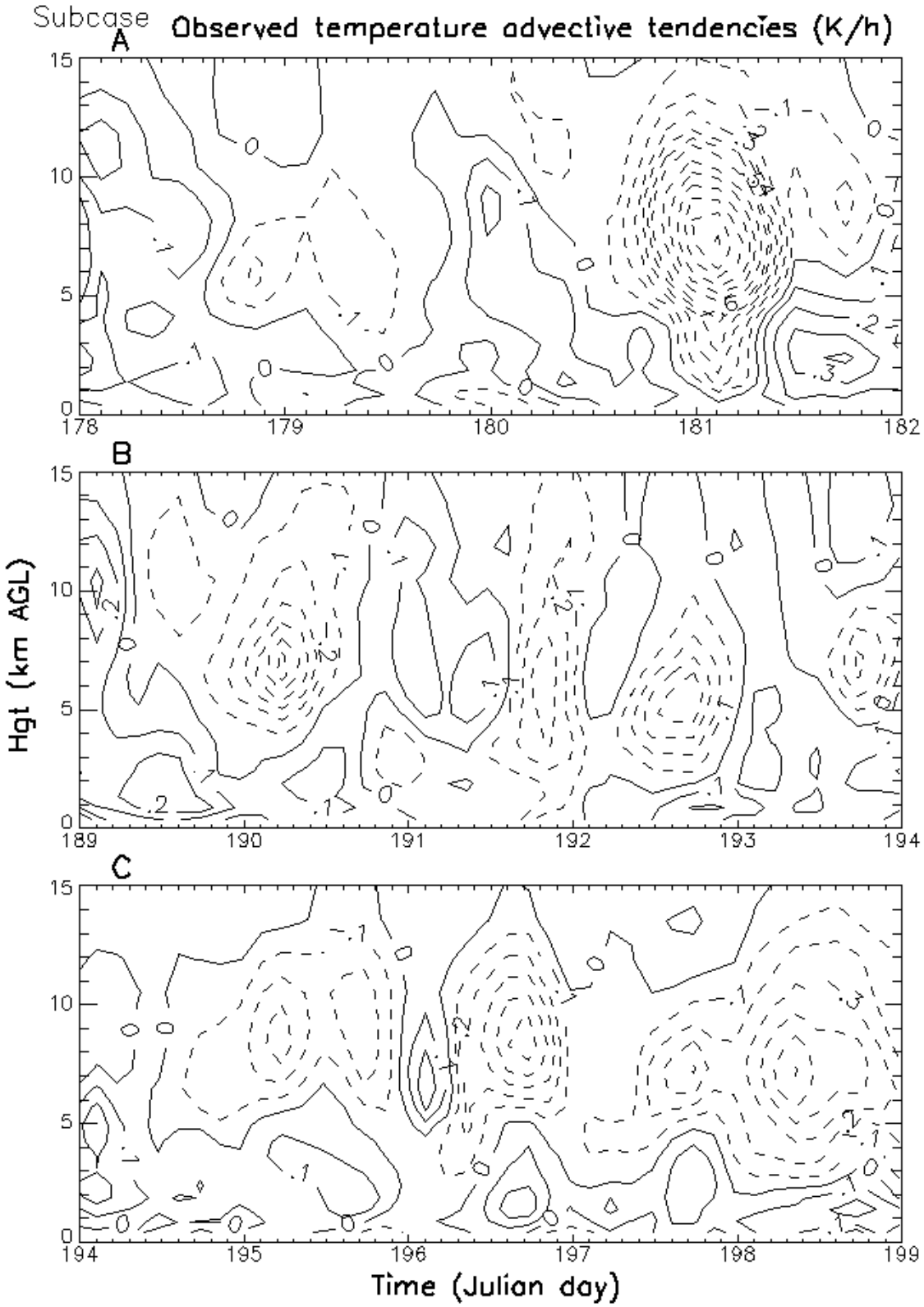

Fig. 2: Time-height cross sections of observed large-scale advective cooling rates for Subcases A, B and C. Contour interval is 0.1 K h^{-1}.

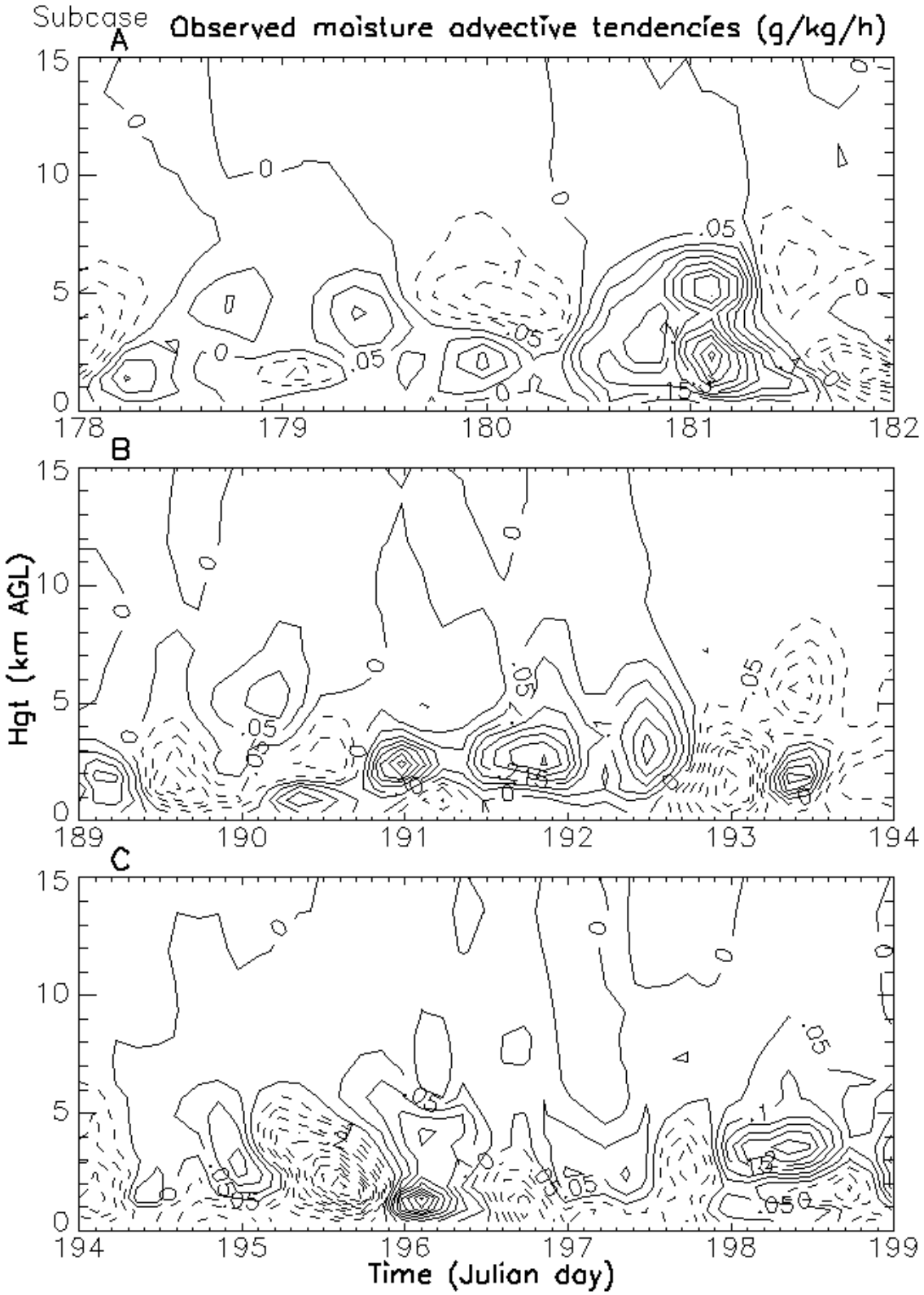

Fig. 3: Time-height cross sections of observed large-scale advective moistening rates for Subcases A, B and C. Contour interval is 0.05 g kg^{-1} h^{-1}.

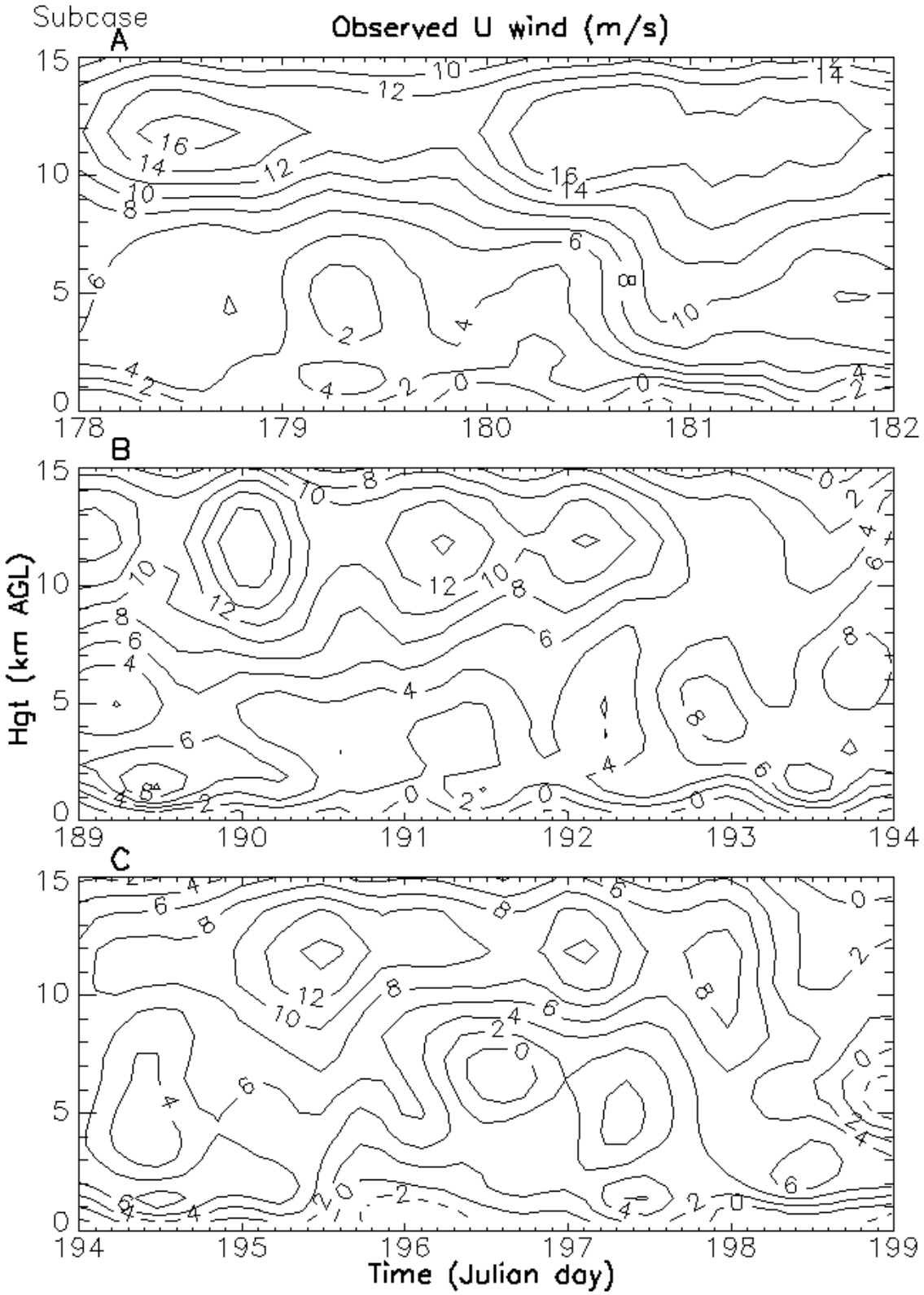

Fig. 4: Time-height cross sections of observed zonal wind components for Subcases A, B and C. Contour interval is 2 m s^{-1}.

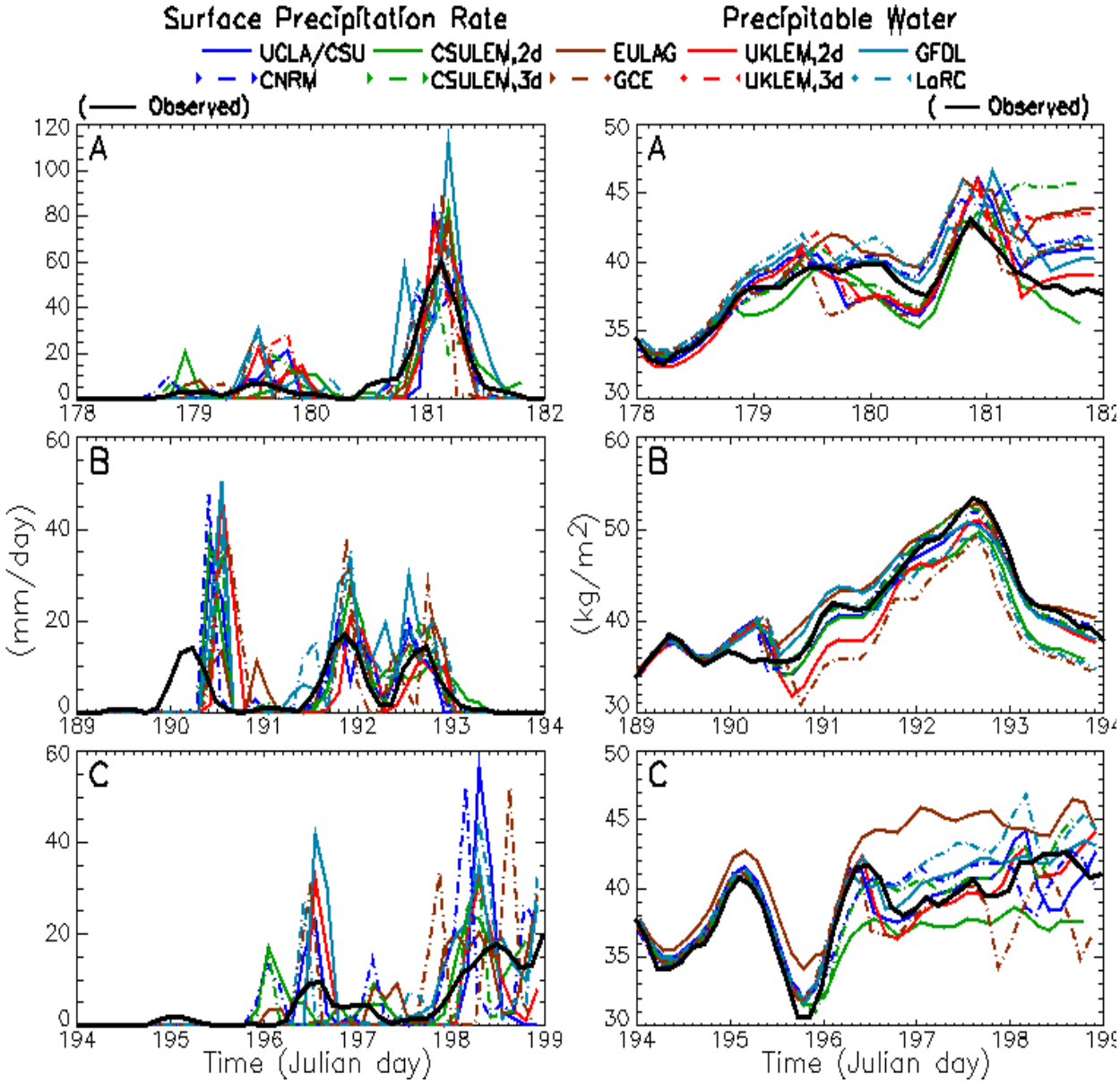

Fig. 5: Time series of surface precipitation rates (left) and precipitable water (right panels).

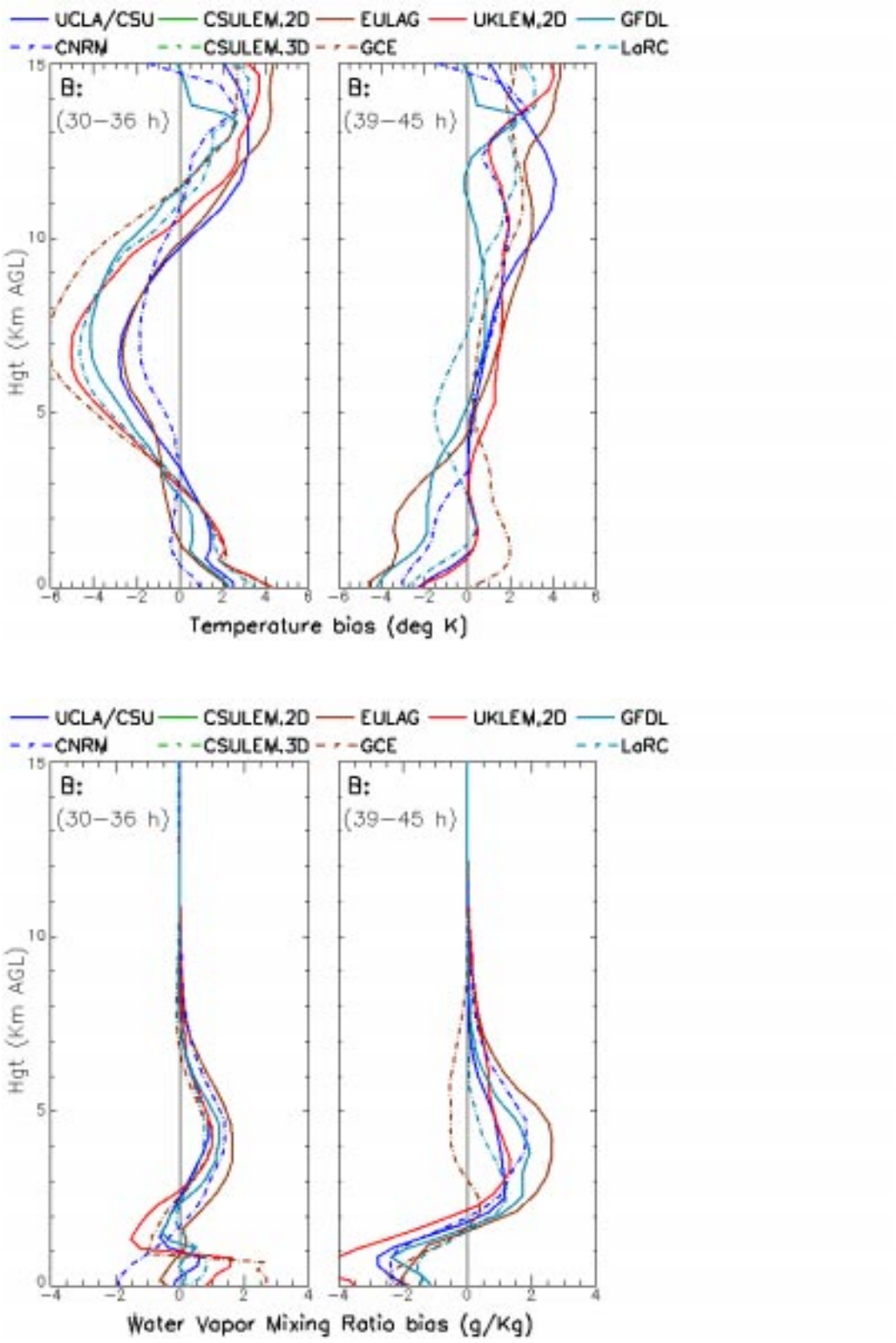

Fig. 6: Six-hour averaged temperature and water vapor mixing ratio biases for periods before (30 - 36 h) and after (39-45) simulated precipitation peaks in Subcase B

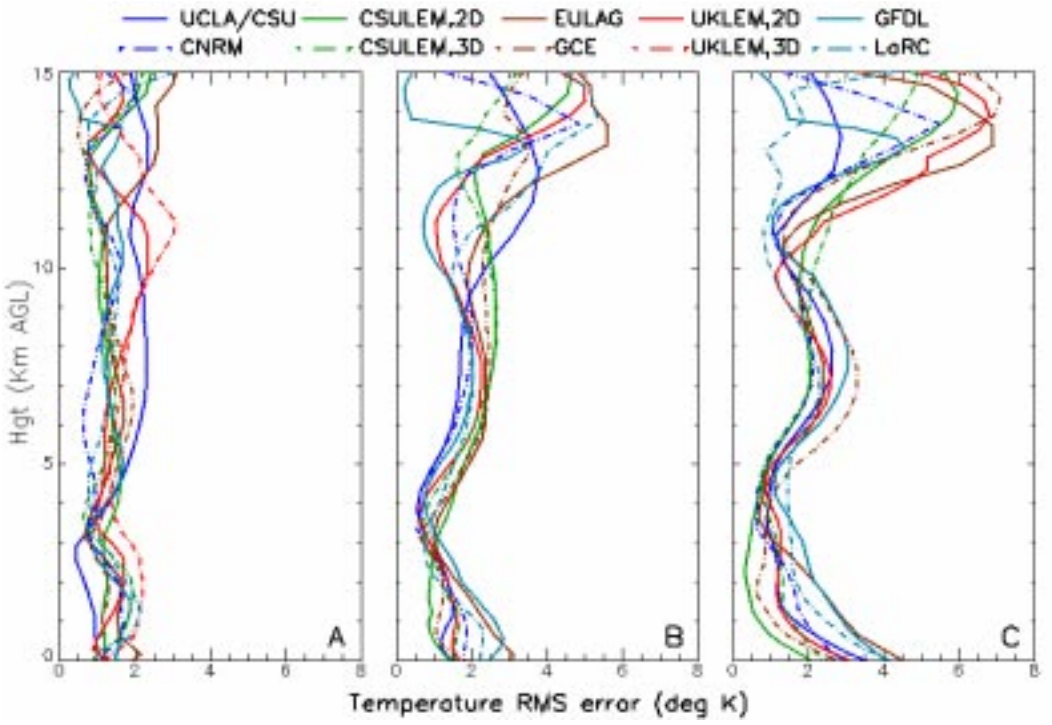

Fig. 7: Root-mean-square errors of temperature for Subcases A, B and C.

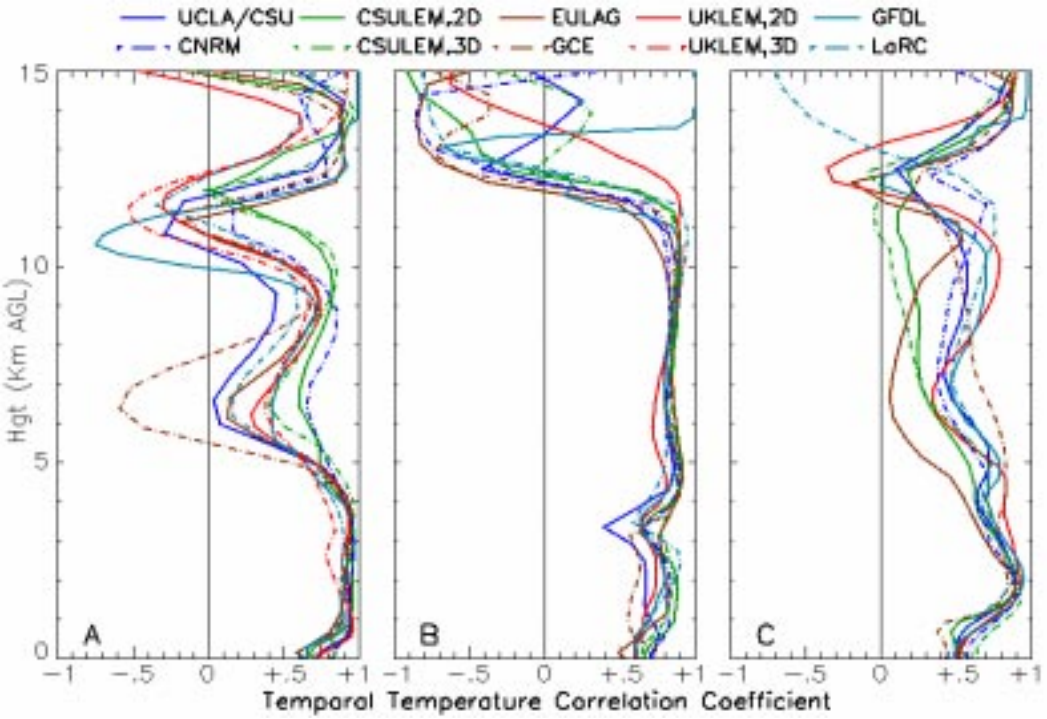

Fig. 8: Temporal correlation coefficients of temperature for Subcases A, B and C.

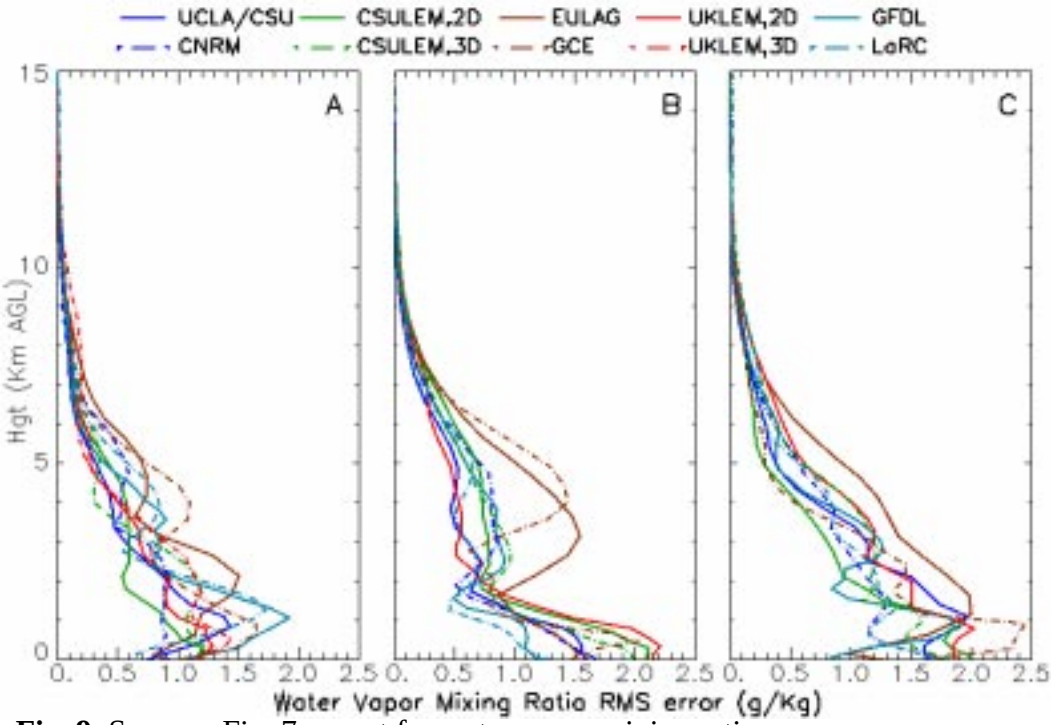

Fig. 9: Same as Fig. 7 except for water vapor mixing ratio.

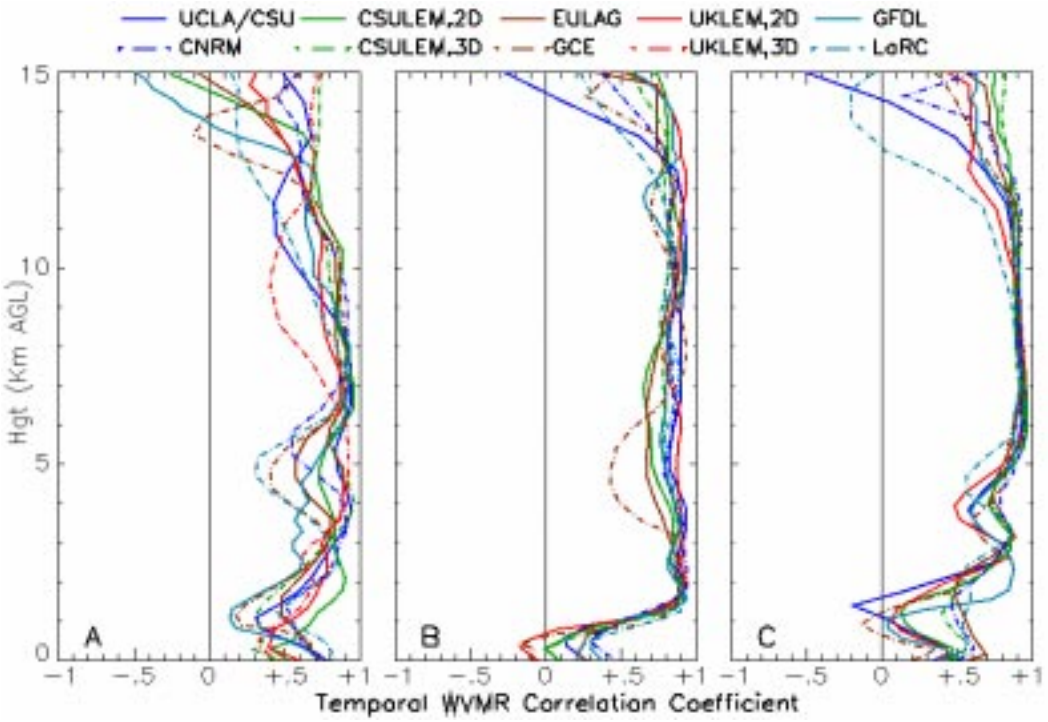

Fig. 10: Same as Fig. 8 except for water vapor mixing ratio.

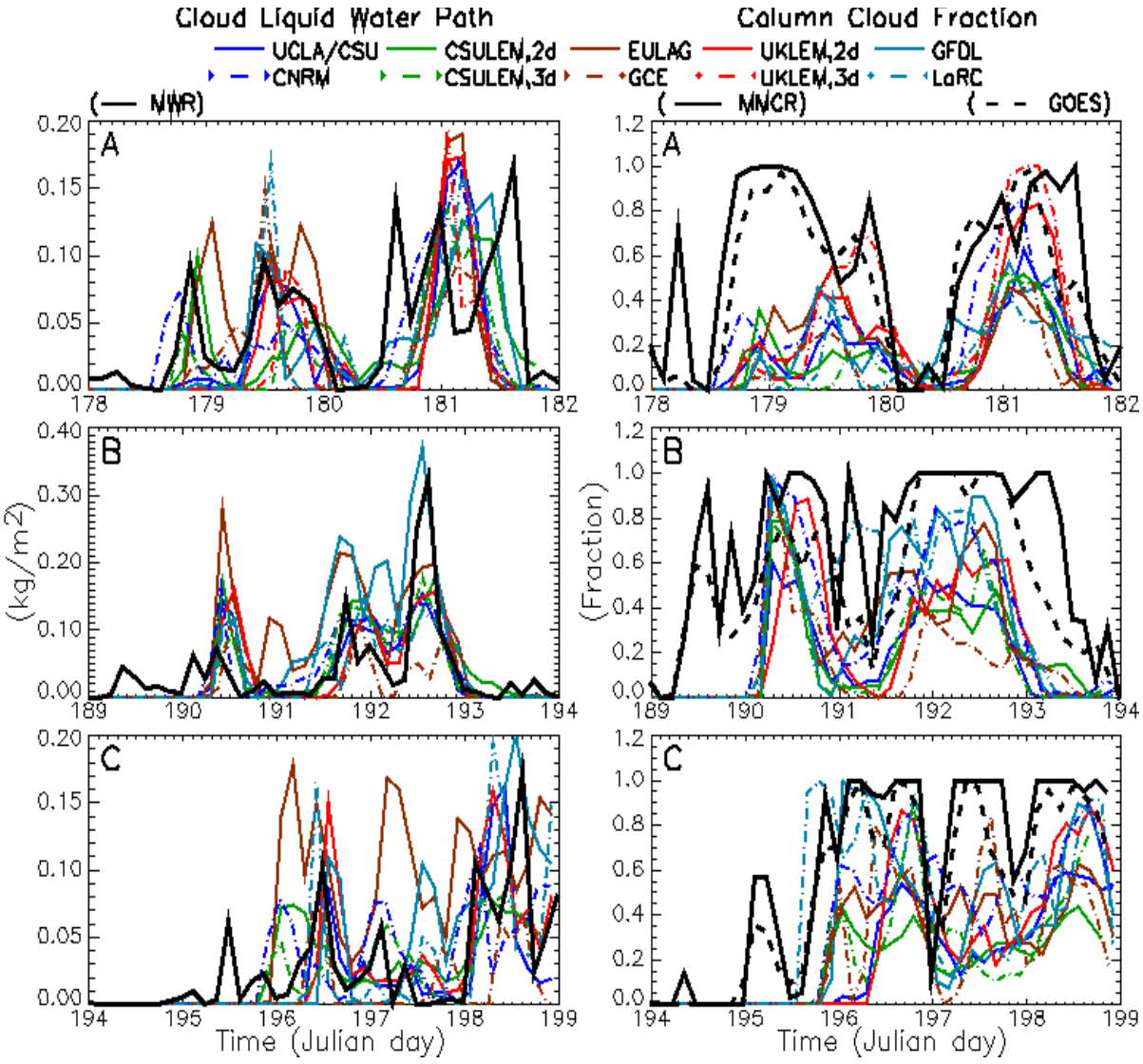

Fig. 11: Time series of liquid water path (left) and the column cloud fraction (right panels).

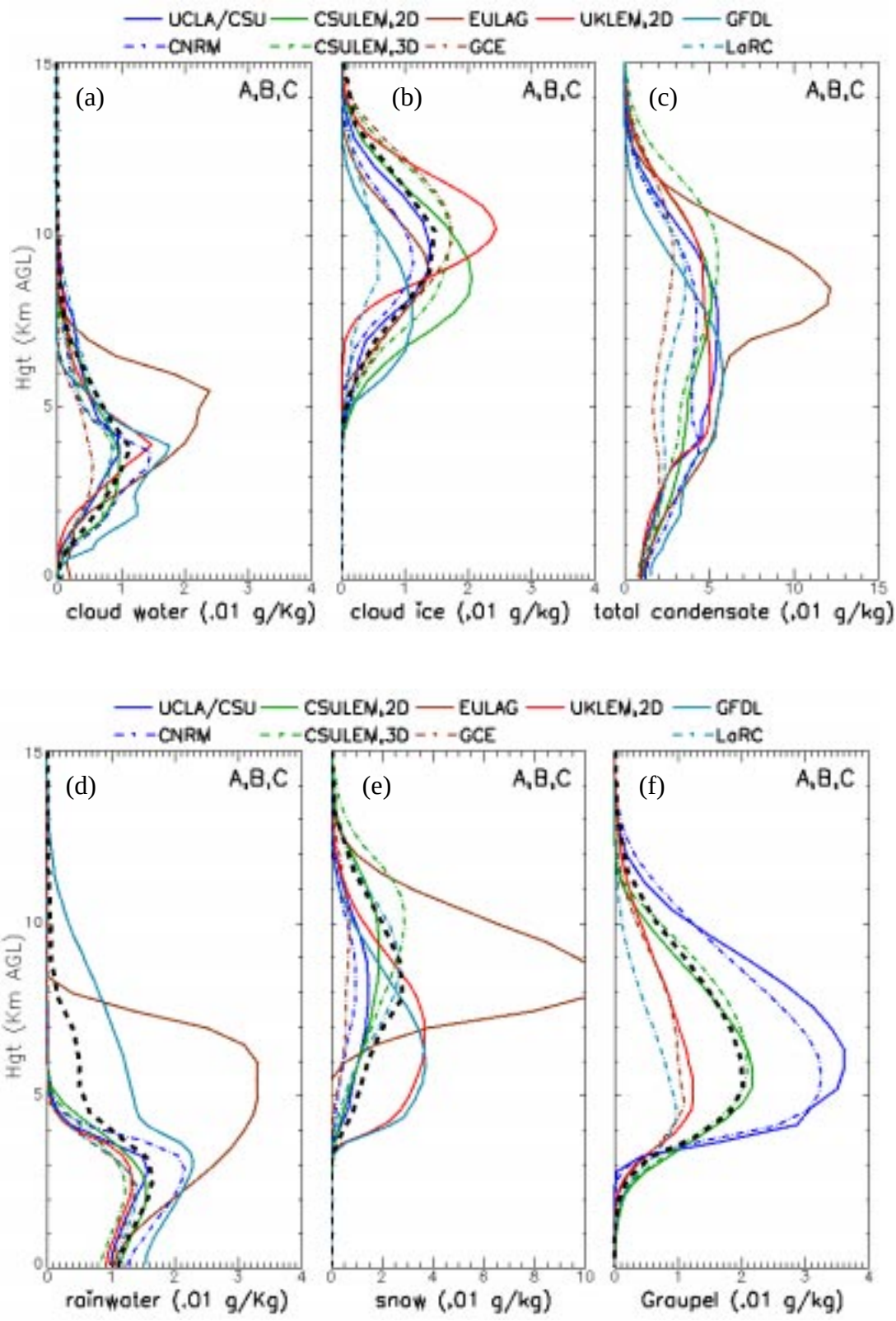

Fig. 12:The all-subcase mean profiles of the mixing ratios of cloud water, cloud ice, total condensate, rainwater, snow and graupel/hail.

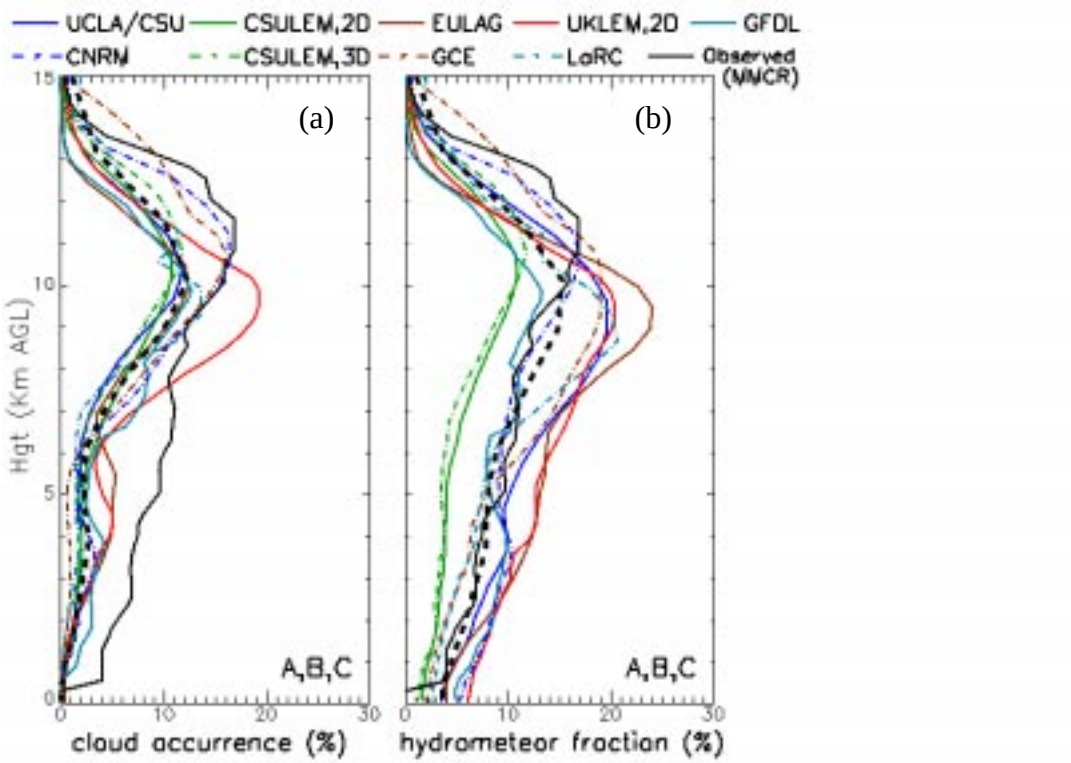

Fig. 13: Same as Fig. 12 except for cloud occurrence and hydrometeor fractions.

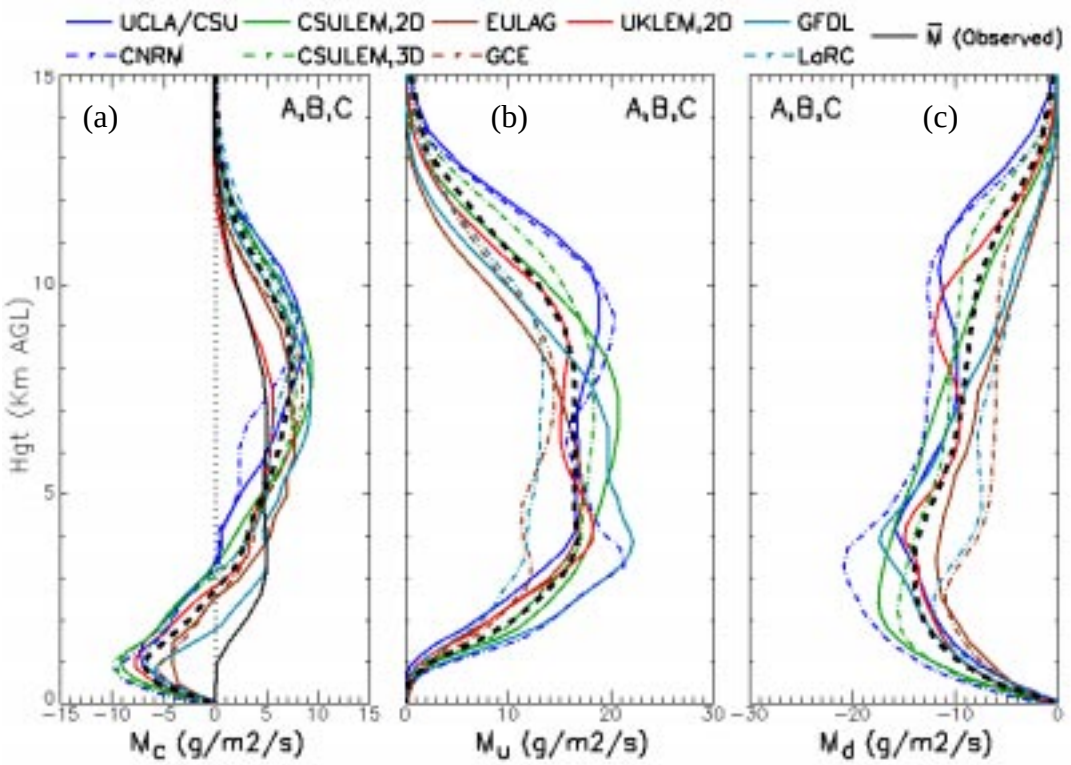

Fig. 14: Same as Fig. 12 except for cloud, updraft and downdraft mass fluxes.