

## Comparison of SCM and CSRМ forcing data derived from the ECMWF model and from objective analysis at the ARM SGP site

Shaocheng Xie and Richard T. Cederwall

Atmospheric Science Division, Lawrence Livermore National Laboratory, Livermore, California, USA

Minghua Zhang

Marine Sciences Research Center, State University of New York, Stony Brook, New York, USA

J. John Yio

Atmospheric Science Division, Lawrence Livermore National Laboratory, Livermore, California, USA

Received 26 February 2003; revised 6 May 2003; accepted 22 May 2003; published 20 August 2003.

[1] The large-scale forcing data diagnosed from the European Center for Medium Range Weather Forecast (ECMWF) model for driving Single-Column Models (SCMs) and Cloud System Resolving Models (CSRMs) are compared with forcing data derived using the objective variational analysis constrained by observations collected at the Atmospheric Radiation Measurement program (ARM) Southern Great Plains (SGP) site. The comparison covers the following three different synoptic conditions: a strong precipitation period dominated by subgrid scale processes during the ARM summer 1997 Intensive Operational Period (IOP), a moderate precipitation period dominated by synoptic scale processes during the spring 2000 IOP, and a nonprecipitation period during the fall 2000 IOP. In the study we demonstrate that the differences between the two forcing data sets are considerably large during the strong convective precipitation period, while they are much less during the moderate and nonprecipitation periods. By analyzing the column-integrated heat and moisture budgets we show that errors in the ECMWF-model-derived forcing are closely associated with errors in the model-predicted surface precipitation, which largely reflect deficiencies of model parameterizations. In SCM tests we show that SCM simulations are sensitive to the prescribed large-scale forcing data. The simulation errors are not well correlated between the SCM runs with the two different forcing data sets for all the three cases. Some important SCM simulated fields, such as surface precipitation, tend to follow the ECMWF model simulations rather than the observations when it is forced with the ECMWF forcing, especially for the summer case. *INDEX TERMS*: 0910 Exploration Geophysics: Data processing; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; 3399 Meteorology and Atmospheric Dynamics: General or miscellaneous; *KEYWORDS*: large-scale forcing, objective analysis, single-column model, cloud system resolving model

**Citation:** Xie, S., R. T. Cederwall, M. Zhang, and J. J. Yio, Comparison of SCM and CSRМ forcing data derived from the ECMWF model and from objective analysis at the ARM SGP site, *J. Geophys. Res.*, 108(D16), 4499, doi:10.1029/2003JD003541, 2003.

### 1. Introduction

[2] The Single-Column Model (SCM) and Cloud System Resolving Model (CSRМ) are useful tools to test and evaluate physical parameterizations used in climate models [Randall *et al.*, 1996]. A successful SCM or CSRМ test requires highly accurate large-scale forcing data, such as the large-scale advective tendencies of temperature and moisture and the vertical velocity. These forcing data can be derived from the data collected in major field programs (e.g., Atmospheric Radiation Measurement program (ARM) and Tropical Ocean-Global Atmosphere Coupled Ocean-

Atmosphere Response Experiment (TOGA-COARE)) through objective analysis. However, the observations are often available only over a limited time periods and regions. Over regions and periods where observations are not available or data density is low, the large-scale forcing data are usually obtained from output of operational numerical weather prediction (NWP) models. A potential use of the NWP products is to develop long-term continuous forcing data sets for statistical studies of SCM and CSRМ results that are not possible with limited observations. The NWP forcing has been used in some recent SCM studies [e.g., Iacobellis *et al.*, 2002]. A problem in using the NWP data is that the forcing data themselves are affected by deficiencies of the model physical parameterizations used in generating the data. However, how much model physical parameter-

izations influence these forcing fields and how the forcing data affect SCM results have not been discussed previously in the literature.

[3] In this paper, we attempt to address the above issues through assessment of the forcing data diagnosed from the European Center for Medium Range Weather Forecast (ECMWF) model by using data collected from the ARM Intensive Operational Periods (IOPs) and processed by the ARM objective variational analysis [Zhang *et al.*, 2001]. We will also present results from SCM tests to demonstrate impacts of using NWP forcing on SCM simulations.

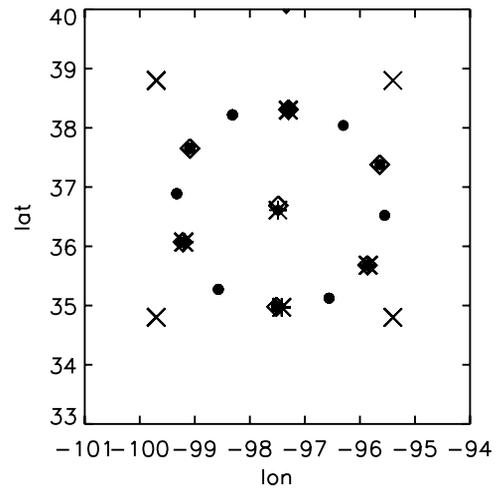
## 2. Large-Scale Forcing

### 2.1. Observed Forcing

[4] The large-scale forcing data, such as the large-scale advective tendencies and vertical motion, cannot be observed directly from field measurements. They are derived from the field observations by using objective analysis methods. The objective analysis scheme used in this study is the constrained variational analysis approach developed by Zhang and Lin [1997]. The variational analysis approach uses the domain-averaged surface precipitation, latent and sensible heat fluxes, and radiative fluxes at the surface (SRF) and the top of the atmosphere (TOA) as the constraints, to constrain the atmospheric state variables to satisfy the conservation of mass, heat, moisture, and momentum. Therefore the derived data set from this approach is dynamically and thermodynamically consistent. Zhang *et al.* [2001] showed that the constrained variational analysis could significantly reduce the sensitivity of the final analyzed products to uncertainty in the upper air input observations. Other studies have also shown that this approach significantly improves the accuracy of the large-scale forcing, and therefore its derived forcing data sets have been used in several SCM and CSRM studies [e.g., Ghan *et al.*, 2000; Xie *et al.*, 2002; Xu *et al.*, 2002]. Nevertheless, it should be noted that uncertainties still exist in the variational analysis forcing data because of the inevitable measurement error and arbitrary parameters used in the analysis procedure. In fact, perfect forcing data never exist in the real world. In the following discussion we use the objective variational analysis forcing as the “truth” to evaluate the forcing data derived from ECMWF model analysis. The uncertainty of the variational analysis forcing will be briefly discussed in section 5. Figure 1 displays the variational analysis domain that is circled by the analysis grids (solid circles in Figure 1), which includes the five ARM sounding stations (asterisks in Figure 1) and seven wind profiler stations (diamonds in Figure 1) near the ARM Southern Great Plains (SGP) site.

### 2.2. ECMWF Forcing

[5] ECMWF has been providing ARM with continuous data sets including the large-scale forcing data, covering all three ARM field research sites, North Slope of Alaska (NSA), SGP, and Tropical Western Pacific (TWP), since 1995. The model forcing is specifically extracted from the ECMWF model runs to force SCMs. These data are averaged over an area that is close to the ARM variational analysis domain (see Figure 1). The data set is a composite of 12 to 36-hour forecasts. The model used to generate the



**Figure 1.** Boundary locations of the Atmospheric Radiation Measurement program (ARM) Single-Column Model (SCM) variational analysis domain (solid circles) and the ECMWF analysis domain (crosses). Asterisks and diamonds represent the locations of the balloons and the wind profilers, respectively.

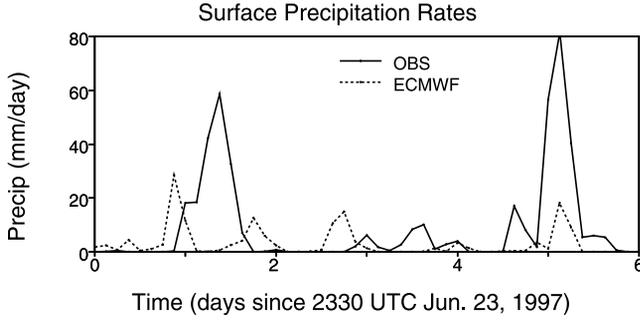
data set is the ECMWF global spectral model. Detailed information can be found in the release notice for the SGP ECMWF data sets (available at [www.arm.gov/docs/xds/static/ecmwf.html](http://www.arm.gov/docs/xds/static/ecmwf.html)). Information about the model physical parameterizations can be seen in the work of Gregory *et al.* [2000].

## 3. Analysis

[6] In this study, a strong precipitation period from 23 June (2330 UTC) to 29 June (2330 UTC) during the summer 1997 IOP, a moderate precipitation period from 8 March (1730 UTC) to 18 March (1730 UTC) during the spring 2000 IOP, and a nonprecipitation period from 27 November (1730 UTC) to 3 December (1730 UTC) during the fall 2000 IOP are selected to assess the ECMWF-derived forcing under different weather conditions. All observations are collected at the ARM SGP site.

### 3.1. Summer Strong Precipitation Case

[7] The summer strong precipitation case contained two strong precipitation events on day 2 and day 6 and a weak precipitation event on day 4 (solid line in Figure 2). Note that “day  $n$ ” here refers to the day between  $n - 1$  and  $n$  in the plots. For example, day 2 refers to the day between 1 and 2 in Figure 2. This convention is used throughout the paper. These precipitation events were associated with mesoscale convective systems and were dominated by subgrid scale precipitation [Xie *et al.* 2002]. It is seen that the ECMWF model fails to correctly simulate the strong summertime continental precipitation events (dotted line in Figure 2). It largely underestimates the observed precipitation and tends to trigger convection earlier than the observations. Note that strong convective events are generally associated with large-scale dynamic processes of upward motion and low-level moisture convergence. The failure to correctly reproduce the observed precipitation events, which is likely related to deficiencies of the model cumulus



**Figure 2.** Time series of the observed (solid line) and ECMWF-model-produced (dotted line) surface precipitation rates (millimeters per day) during the selected strong precipitation period in the 1997 summer Intensive Operational Period (IOP).

parameterization, could have a large impact on the diagnosed vertical velocity and advective tendencies as discussed later.

[8] To assess the ECMWF diagnosed forcings, we first examine the column-integrated heat and moisture budgets as follows:

$$C_p \frac{\partial \langle T \rangle}{\partial t} + C_p \langle \nabla \cdot \vec{\nabla} T \rangle - \left\langle \frac{\omega}{\rho} \right\rangle = R_{\text{TOA}} - R_{\text{SRF}} + LP_{\text{rec}} + SH + L \frac{\partial \langle q_l \rangle}{\partial t} \quad (1)$$

$$\frac{\partial \langle q \rangle}{\partial t} + \langle \nabla \cdot \vec{\nabla} q \rangle = E_s - P_{\text{rec}} - \frac{\partial \langle q_l \rangle}{\partial t}, \quad (2)$$

where

$$\langle X \rangle = \frac{1}{g} \int_{p_t}^{p_s} (X) dp.$$

In the above equations,  $\vec{\nabla}$  is the wind,  $T$  is the temperature,  $\omega$  is the vertical  $p$  velocity,  $\rho$  is the air density,  $q$  is the mixing ratio of water vapor,  $p_s$  is the surface pressure,  $p_t$  is the tropopause pressure,  $q_l$  is the cloud liquid water content,  $R$  is the net downward radiative flux at TOA and SRF,  $P_{\text{rec}}$  is precipitation,  $L$  is the latent heat of vaporization,  $C_p$  is the heat capacity,  $SH$  is the sensible heat flux, and  $E_s$  is the surface evaporation. Note that the terms on the right-hand side of the equations are the constraints used in the variational analysis. These constraints are not changed in the analysis.

[9] Table 1 lists the statistics of the observed (values in parentheses in Table 1) and the ECMWF-model-calculated column heat and moisture budget components during the strong summer convective period. The observed values are obtained from the variational analysis. In Table 1, “MEAN” represents an a time average over the selected period, “STD” represents standard deviation, “RMSE” denotes root-mean-square (RMS) error, and “CORR” denotes correlation coefficient. “RMSE/STDO” is used to measure RMSE relative to the variability in the observed field itself, where “STDO” represents standard deviation in

observations. Since cloud liquid term is very small compared to other terms, it is not shown in Table 1. Note that in Table 1 we use “Col\_Tadv” to represent the column integral of large-scale advective tendency of temperature and the adiabatic expansion term, i.e.,

$$-C_p \langle \nabla (\vec{\nabla} T) \rangle + \left\langle \frac{\omega}{\rho} \right\rangle,$$

and “Col\_qadv” to represent the column integral of large-scale advective tendency of moisture, i.e.,  $-L^* \langle \nabla (\vec{\nabla} q) \rangle$ . The budget check shows that the column-integrated energy and moisture budgets are balanced in the constrained variational analysis. The area-averaged ECMWF model data also conserve well the column energy and moisture budgets with a time-averaged budget imbalance of  $\sim 1.45 \text{ W m}^{-2}$  and  $-1.28 \text{ W m}^{-2}$ , respectively (not shown in Table 1).

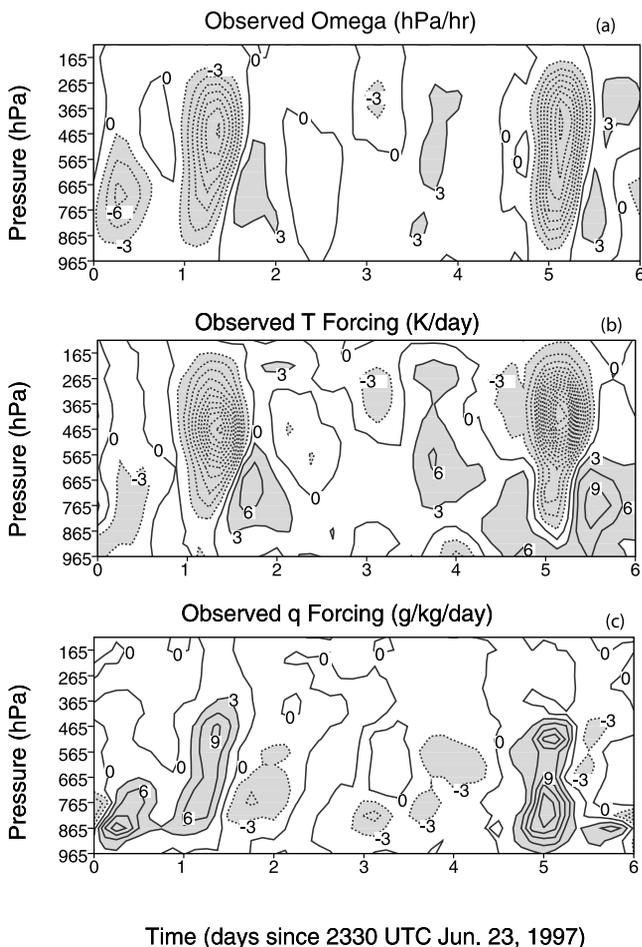
[10] It is seen from Table 1 that significant disagreements exist in all column budget components between the model data and the observations. The ECMWF model produces much stronger column radiative cooling and larger surface evaporation while it dramatically underestimates the surface precipitation ( $LP_{\text{rec}}$ ) and sensible heat flux, compared to the observations. The calculated latent heating associated with surface precipitation shows a bias of  $\sim 65\%$  of the observed mean. The RMS error in this term is significantly large judged by RMSE/STDO, which is 1.03; that is, it is larger than the observed temporal variations. Note that any RMS error close to the temporal variability of the observed field itself is a serious concern. The ECMWF-model-produced surface precipitation shows very weak correlation (0.2) with the observed value. It should be noted that these constraint variables used in the variational analysis are obtained directly from the observations and are not changed in the variational analysis procedure. The uncertainty in these column budget variables can be considered smaller than that in the derived forcing fields. Therefore the discrepancies presented above are largely related to model deficiencies.

[11] Given the large disagreements in the surface fields, it is not surprising to see the considerable differences in the

**Table 1.** Comparison of ECMWF Model Data to the ARM Observations for the Column-Integrated Heat and Moisture Budget Components During the Strong Convective Period<sup>a</sup>

	MEAN	STD	RMSE	RMSE/STDO	CORR
$R_{\text{TOA}} - R_{\text{SRF}}$	-97.7 (-57.1)	127.9 (162.1)	63.8	0.39	0.97
$LP_{\text{rec}}$	90.8 (261.8)	163.6 (517.4)	533.3	1.03	0.20
$SH$	27.8 (40.8)	76.9 (46.6)	42.4	0.91	0.90
$L(E_s)$	146.5 (109.6)	151.3 (111.9)	64.1	0.57	0.96
$C_p \langle \partial(T)/\partial t \rangle$	13.7 (33.6)	284.1 (247.6)	165.9	0.67	0.81
$\text{Col\_Tadv}$	-8.8 (-210.7)	254.5 (519.4)	608.7	1.05	0.23
$L \langle \partial(q)/\partial t \rangle$	11.8 (26.5)	271.1 (225.6)	329.7	1.46	0.11
$\text{Col\_qadv}$	-42.6 (178.6)	355.5 (519.4)	558.6	1.08	0.34

<sup>a</sup>Values in parentheses are ARM observations. Values given in  $\text{W m}^{-2}$ . ECMWF, European Center for Medium Range Weather Forecast; ARM, Atmospheric Radiation Measurement program; MEAN, time average; STD, standard deviation; RMSE, root-mean-square errors; STDO, standard deviation in observations; CORR, correlation coefficient;  $R_{\text{TOA}}$ , net downward radiation flux of the top of the atmosphere;  $R_{\text{SRF}}$ , net downward radiation flux of the surface;  $LP_{\text{rec}}$ , surface precipitation;  $SH$ , sensible heat flux;  $L(E_s)$ , latent surface evaporation;  $C_p \langle \partial(T)/\partial t \rangle$ , heat term;  $\text{Col\_Tadv}$ , column integral of large-scale advective tendency of time;  $L \langle \partial(q)/\partial t \rangle$ , moisture term;  $\text{Col\_qadv}$ , column integral of large-scale advective tendency of moisture.



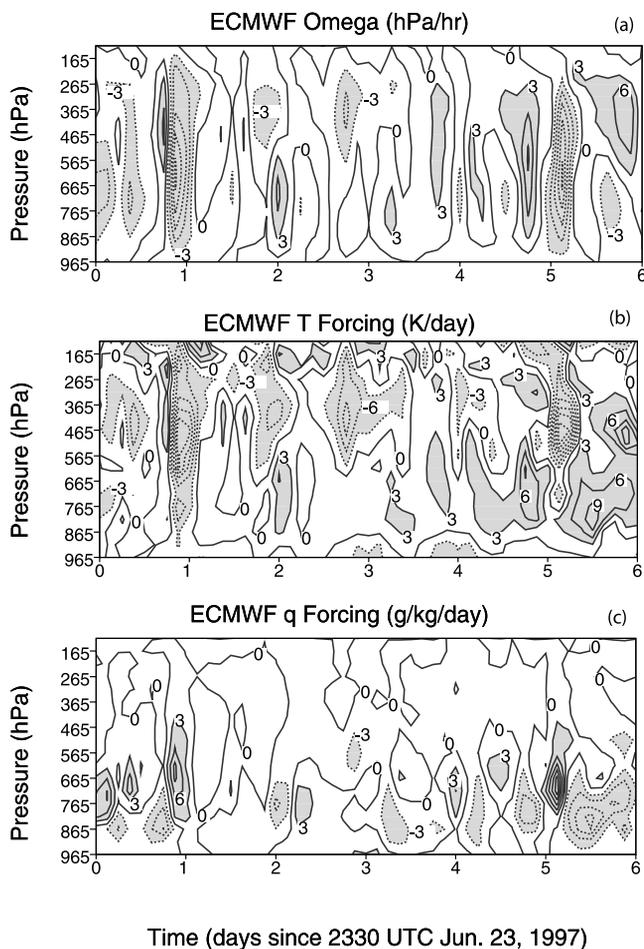
**Figure 3.** Time-height distributions of the derived (a) vertical velocity, (b) total advective tendency of temperature, and (c) total advective tendency of moisture during the selected strong precipitation period in 1997 summer IOP. Contour interval is 3. Contours  $>3$  or less than  $-3$  are shaded. In Figures 3a–3c, solid lines denote contours  $\geq 0$ , and dotted lines denote contours  $< 0$ .

derived forcing fields (i.e.,  $Col\_Tadv$  and  $Col\_qadv$ ) between the variational analysis and the ECMWF model. It is seen that the variational analysis shows very strong advective and adiabatic cooling and large moisture convergence during the strong convective period. These values are consistent with other previous studies [e.g., Hudson, 1971; Cho and Ogura, 1974; Houze and Rappaport, 1984]. In contrast, the model exhibits rather weak advective cooling and weak divergence. The model-derived temporal variations in these two terms are much weaker than the observations. Similar to the surface precipitation rates, the column integrated forcing fields show RMS errors that are larger than the observed temporal variations. The correlations between these two types of forcing data are fairly weak (less than 0.35). Note also that the heat and moisture storage terms  $C_p * \partial(T)/\partial t$  and  $L * \partial(q)/\partial t$  show noticeable differences between the model and the observations, even though the model-time-averaged temperature and moisture agree well with the observations with errors of  $< 0.5$  K in temperature and  $0.3 \text{ g kg}^{-1}$  in moisture, respectively.

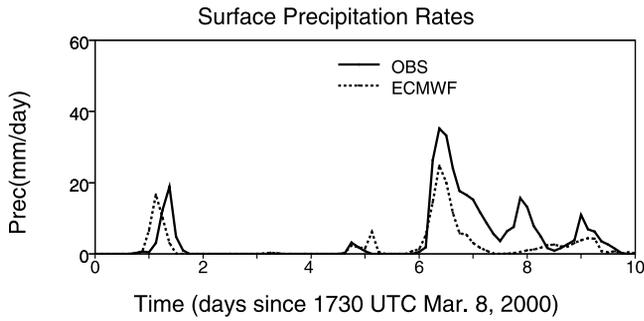
[12] Another noteworthy feature is that the observed energy budget is dominated by the latent heating associated with surface precipitation and advective/adiabatic cooling, and the observed moisture budget is largely balanced by surface precipitation and moisture convergence, for the strong convective case. In the model, however, the latent heating is balanced by the column net radiative cooling; the surface precipitation and the moisture divergence are balanced by the surface evaporation. The relationship presented in the model-calculated column-integrated budget of heat is usually found in the radiation-induced convective atmosphere rather than the strong summer convective atmosphere where the large-scale dynamic forcing plays the most important role in destabilizing the atmosphere.

[13] Although there are large disagreements between the model and the observations, it is interesting to see that the model-calculated column net radiation, sensible heat flux, and evaporation terms show rather high correlation (above 0.9) with the observations. This is mainly because these processes are largely dominated by the strong solar diurnal variations over the midlatitude land in the summer.

[14] The time-height distributions of the derived vertical velocity (omega) and the total advective tendencies of temperature and moisture from the variational analysis are shown in Figures 3a–3c for the strong convective period. Note that the total advection of temperature



**Figure 4.** Same as Figure 3 except for the ECMWF-derived forcing fields.



**Figure 5.** Same as Figure 2 except for the selected moderate precipitation period in 2000 spring IOP.

includes the adiabatic expansion term. Corresponding to the observed surface precipitation events (Figure 2), the derived forcings show strong large-scale advective cooling (associated with strong upward motion) in the middle and upper troposphere and strong moisture convergence in the lower troposphere.

[15] Figures 4a–4c are the same as Figures 3a–3c except for the model-derived forcing fields. It is seen that the model-derived forcing fields are closely associated with its calculated precipitation. As we showed earlier, however, the calculated precipitation events are much weaker than the observations and also are triggered too early. Associated with these problems, the model-derived forcing fields are much weaker compared to those derived from the variational analysis. For some periods, such as on day 2, in which a strong convective event was observed, the two different forcings are even out of phase. On this day the objectively analyzed data show very strong upward motion and advective cooling in the middle and upper troposphere and large lower-level moisture convergence, while the ECMWF data displays weak downward motion and small advective heating in the middle and upper troposphere and weak lower-level moisture divergence.

[16] In summary, the above discussions show that the forcing data derived from the ECMWF model differ dramatically from those derived from the objective variational analysis method. The biases are closely reflected in the errors of the model-simulated surface precipitation, which partially relate to potential deficiencies in the model parameterizations. In addition, the lack of sufficient mesoscale observations available for the analysis of the mesoscale structures, which are important for generating mesoscale convection, in the ECMWF data assimilation system or deficiencies in the ECMWF data assimilation itself may also contribute to the errors in the ECMWF forcing data.

**Table 2.** Same as Table 1 Except for the Spring Precipitation Case

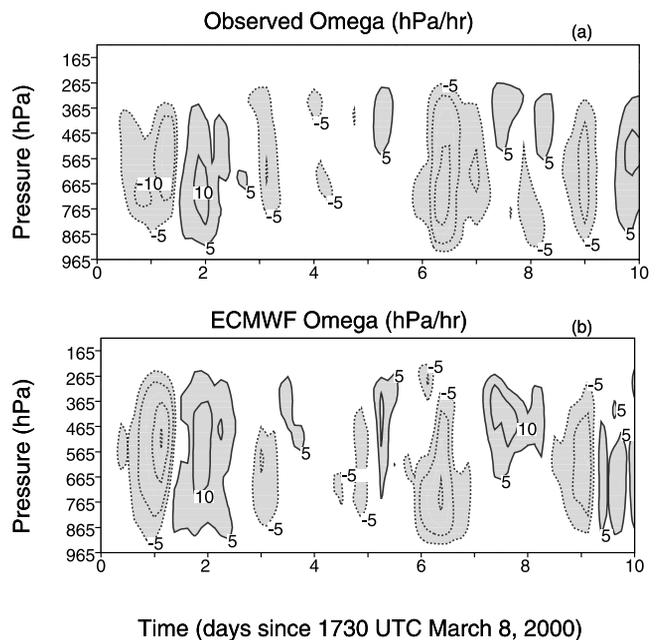
	MEAN	STD	RMSE	RMSE/STDO	CORR
$R_{TOA} - R_{SRF}$	-87.7 (-71.5)	352.5 (423.5)	40.9	0.33	0.96
$L(\text{PREC})$	64.0 (124.3)	131.9 (221.3)	156.7	0.71	0.77
SH	29.9 (40.3)	89.0 (76.6)	38.2	0.50	0.91
$L(E_s)$	48.3 (30.9)	49.0 (36.3)	30.5	0.84	0.87
$C_p(\partial(T)/\partial t)$	-21.9 (-24.5)	282.1 (263.9)	116.1	0.44	0.91
$\text{Col}_{\text{Tadv}}$	-39.6 (-116.7)	352.5 (423.5)	204.3	0.48	0.90
$L(\partial(q)/\partial t)$	28.0 (3.1)	243.4 (177.1)	154.3	0.87	0.78
$\text{Col}_{\text{qadv}}$	54.5 (96.6)	272.4 (260.4)	222.9	0.86	0.66

### 3.2. Spring Moderate Precipitation Case

[17] The spring precipitation period contained one single-day precipitation on day 2 and one multiday precipitation event on days 7–10 (solid line in Figure 5). In comparison with the summer case the spring precipitation events were relatively weaker and were mainly dominated by large-scale frontal systems. It is seen that the ECMWF model is able to generally capture well most of the precipitation events while it underestimates the multiday precipitation event and triggers the first precipitation event a little earlier (dotted line in Figure 5).

[18] Table 2 is analogous to Table 1 but for the spring case. In comparison with Table 1, much better agreement in the column budget terms between the ECMWF simulations and the observations is seen in the spring case. The temporal correlations between the model data and the observations are rather high (close to or larger than 0.9) for most of the budget terms. Yet relatively weak correlations are seen in those precipitation-related terms, such as the latent heating and moisture convergence, where the correlation coefficients are below 0.80. Both the model and the variational analysis show moisture convergence and advective/adiabatic cooling during these precipitation periods. The RMS errors in those constraint terms are less than the magnitude of the temporal variability in the observations. In terms of the mean value, however, the model overestimates the observed column radiative cooling and surface evaporation and underestimates the surface precipitation and sensible heat flux. Discrepancies in the mean diagnosed heat and moisture convergences are also quite large.

[19] Figures 6a and 6b display the derived large-scale vertical motions from the variational analysis and the



**Figure 6.** Time-height distributions of the derived vertical velocity from (a) variational analysis and (b) the ECMWF model for the selected moderate precipitation period in 2000 spring IOP. Contour interval is 5. Contours  $>5$  or less than  $-5$  are shaded. In these figures, solid lines denote contours  $\geq 0$ , and dotted lines denote contours  $< 0$ .

**Table 3.** Same as Table 1 Except for the Fall Nonprecipitation Period

	MEAN	STD	RMSE	RMSE/STDO	CORR
$R_{TOA} - R_{SRF}$	-113.6 (-106.3)	65.4 (93.2)	35.8	0.38	0.96
$L(\text{PREC})$	0.83 (0)	2.9 (0)	3.1	N/A	N/A
SH	8.2 (12.5)	66.2 (68.8)	37.6	0.55	0.84
$L(E_s)$	26.9 (14.8)	38.6 (18.3)	26.8	2.61	0.88
$C_p(\partial(T)/\partial t)$	-32.6(-37.9)	256.7 (294.5)	104.8	0.36	0.94
Col_Tadv	71.6 (55.9)	224.7 (295.4)	117.9	0.40	0.93
$L(\partial(q)/\partial t)$	-24.9(-16.7)	171.0 (131.2)	80.2	0.61	0.89
Col_qadv	-51.1 (-31.7)	163.2 (126.5)	81.2	0.64	0.88

ECMWF model, respectively, for the spring case. The model-derived vertical motions agree well overall with those derived from the variational analysis. Larger disagreements are seen on days 7 and 8 where the model produces weaker upward motions and stronger downward motions. This is clearly related to the model’s underestimation of the observed precipitation that occurred on these days (see Figure 5). Another noticeable feature in the figure is that model upward/downward motions in the upper troposphere are stronger than observed. This might be because the TOA is set to 10 hPa in the ECMWF model while it is set to 100 hPa in the variational analysis. Therefore, omega in the model is allowed to extend higher than that in the variational analysis. Since there are no accurate measurements available above 100 hPa, it is not possible to judge if the ECMWF upper troposphere omega is more realistic. Similar results can be seen in the derived temperature and moisture forcing fields (not shown).

**3.3. Fall Nonprecipitation Case**

[20] Table 3 analyzes the statistics of the model-calculated and observed column heat and moisture budget components for the non-precipitation period of the fall 2000 IOP. As expected, the disagreement between the ECMWF model data and the observations is further reduced in the fall nonprecipitation case. Most budget terms show RMS errors considerably less than the observed temporal variations. The only exception is the surface evaporation term, which shows significantly large error in terms of RMSE/STDO. The calculated column net radiative cooling agrees well with the observation, with the mean bias <7% of the observed value. The model-produced spurious precipitation is very small ( $0.83 \text{ W m}^{-2} - 0.03 \text{ mm d}^{-1}$ ) and can be neglected.

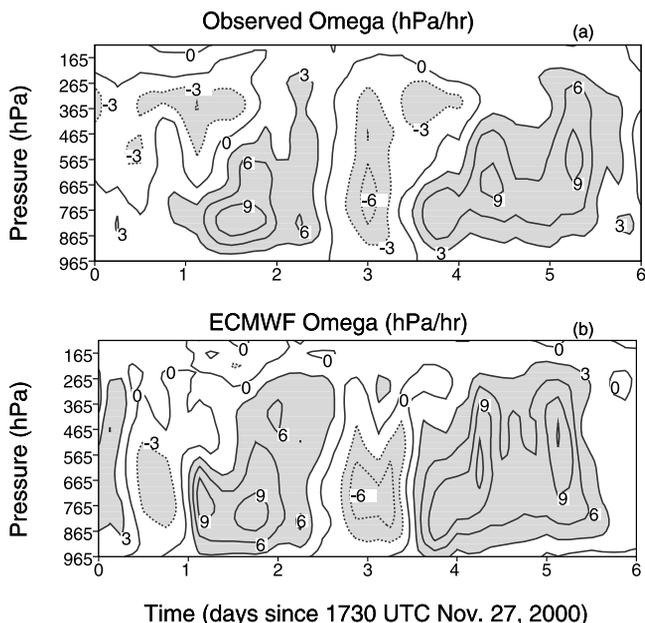
[21] For the derived forcing fields, the EMWCF model produces similar temporal variabilities in the heat and moisture convergence terms as those in the variational analysis data. The correlations for these two fields between these two data sets are 0.94 and 0.88, respectively. Yet discrepancies in the mean diagnosed heat and moisture convergences are still noticeably large. Another noticeable feature in Table 3 is that both the model and the variational analysis show that the decrease in the heat storage is balanced by the advective heating and column radiative cooling, and the decrease in the moisture storage is balanced by the moisture divergence and surface evaporation, in absence of precipitation.

[22] Figures 7a and 7b are the same as Figures 6a and 6b except for the nonprecipitation period. Both the observations and the model show that large-scale downward motion dominates this nonprecipitation period. In general, the model captures well the observed vertical velocity field

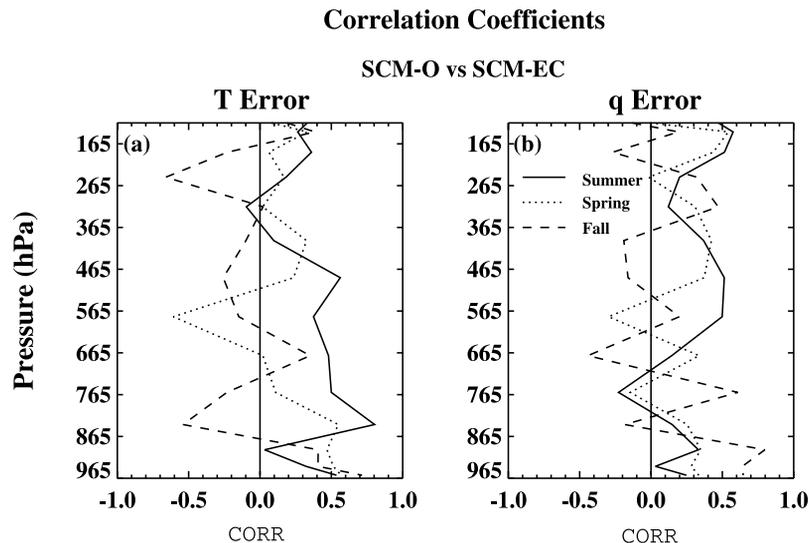
even though the downward motions are somewhat overestimated.

**4. Single-Column Model Simulations**

[23] In section 3 we discussed the large-scale forcing data derived from the ECWMF model and from the objective variational analysis for three selected cases. One important question is whether or not a bias or model error can be detected regardless of the forcing data set. To address this question, the National Center for Atmospheric Research (NCAR) Community Climate Model version 3 (CCM3) SCM [Hack et al., 1998] with a modified cumulus convection scheme [Xie and Zhang, 2000] is used to investigate the impact of the different large-scale forcings derived from the ARM objective variational analysis and the ECMWF model on SCM simulations. In the SCM runs the large-scale total advective tendencies of temperature (including the adiabatic expansion term) and moisture are specified from the two different forcing data sets. The surface forcing is calculated by the model surface parameterizations. Since CCM3 uses a diagnostic cloud scheme, no initial cloud condition and cloud forcing are required to drive the SCM. To prevent the problem that SCM simulations could drift away from observations over long time integration, a 36-hour forecast of the SCM is launched every day. For each forecast the temperature and moisture are initialized with the observations. The prognostic land variables used in the CCM3 land surface model are initialized using climatological values if there are no observations. A composite of 12- to 36-hour forecasts from the series of 36-hour runs is analyzed. For the convenience of discussion we use SCM-O to represent the SCM run with the objective variational analysis forcing and SCM-EC to represent the run with the ECMWF forcing in the following discussions.



**Figure 7.** Same as Figure 6 except for the selected nonprecipitation period in 2000 fall IOP. Contour interval is 3.



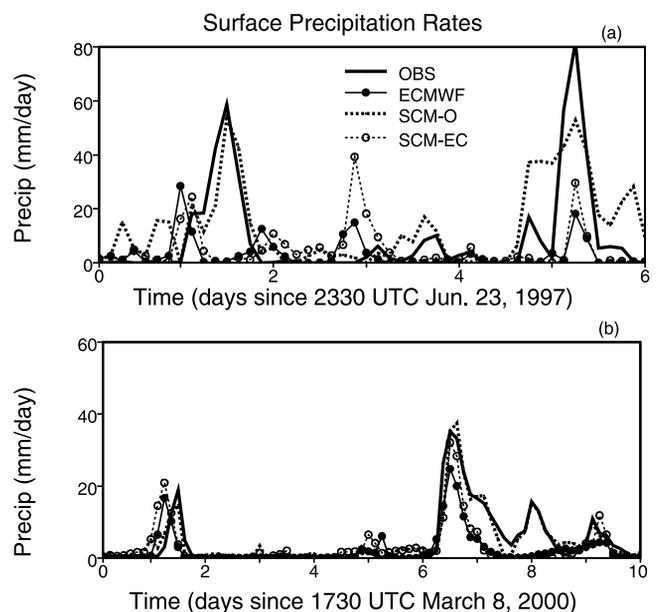
**Figure 8.** Correlations as a function of height between the simulation errors produced by the SCM with the different forcing data sets. (a) Temperature bias. (b) Moisture bias.

[24] Both runs generally produce much smaller simulation errors ( $<6$  K in the temperature and  $<4$  g kg $^{-1}$  in the moisture) than those found in other SCM studies (e.g., Ghan *et al.*, 2000; Xie *et al.*, 2002) due to the short-range (36-hour) SCM runs conducted in this study. However, the temporal correlation between the simulation errors produced from the two runs is rather weak or sometimes shows negative correlations for the three cases (Figure 8). It is quite surprising that the temperature errors produced in the two runs are less correlated in the spring and fall cases than the summer case, although the large-scale forcing data are less divergent during the spring and fall periods, as shown before. It should be noted that Figure 8 just gives the relationship between the model errors produced from the two runs.

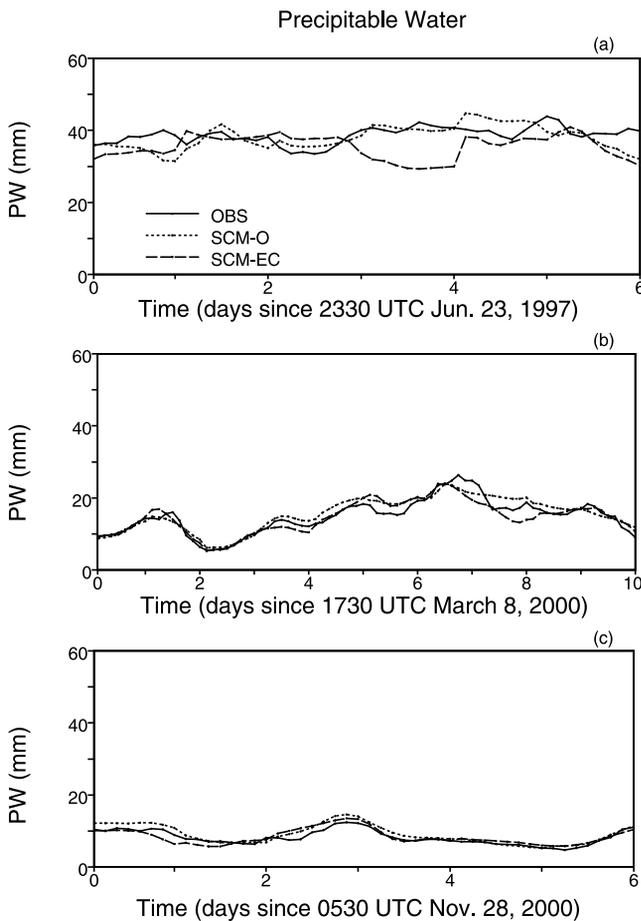
[25] Figures 9a and 9b show the simulated surface precipitation rates of the two runs during the summer and spring periods, respectively. For the summer case, SCM-O generally reproduces well the observed precipitation. In contrast, SCM-EC largely underestimates the observed precipitation. In fact, it regenerates well the ECMWF-model-predicted precipitation. Its temporal correlation with the ECMWF precipitation is much higher than that with the observed precipitation (0.72 versus 0.21). This is also true for the spring case where SCM-O is successful in reproducing the observed precipitation in both magnitude and phase with the correlation around 0.95 while SCM-EC captures well the ECMWF model precipitation with the correlation around 0.94, which is much higher than its correlation with the observations (0.70). Similar results can be found in other simulated fields, such as the precipitable water, the cloud liquid water, the TOA longwave and shortwave radiative fluxes, and the surface sensible heat flux. For all these fields, SCM-EC shows higher correlation with the corresponding ECMWF simulations than the observations.

[26] Figure 10 shows the time series of the observed and model-simulated precipitable water (PW) for the three cases. The observed PW data are from the ARM microwave radiometer measurements. The simulated PW results of the two runs are in much better agreement with the observations for the spring and fall cases than the summer case, where

SCM-EC produced atmosphere is too dry on days 3 and 4 and both runs show problems in correctly simulating the observed PW temporal variations. In comparison with the observations, SCM-O produces more realistic PW results than SCM-EC for the summer case while both runs show comparable results for the spring and fall cases. Nevertheless, noticeable differences are evident in the simulated PW results between these two runs. Similar to the surface precipitation rates, it is noted that SCM-EC has negative



**Figure 9.** Time series of the observed (solid line), ECMWF-produced (solid line with solid circles), and SCM-simulated surface precipitation rates (millimeters per day). (a) Over the strong convective period in 1997 summer IOP. (b) Over the moderate precipitation period in 2000 spring IOP. Dotted line is for using the ARM variational forcing, and dotted line with open circles is for using the ECMWF forcing.



**Figure 10.** Time series of the observed and the SCM-simulated precipitable water (mm) (a) Over the strong convective period in 1997 summer IOP. (b) Over the moderate precipitation period in 2000 spring IOP. (c) Over the nonprecipitation period in 2000 fall IOP. Solid line is for the observations. Dotted line is for using the ARM variational forcing, and dashed line is for using the ECMWF forcing.

correlation ( $-0.35$ ) with the observations while it shows fairly high correlation ( $0.61$ ) with the ECMWF simulated PW (not shown) for the summer case.

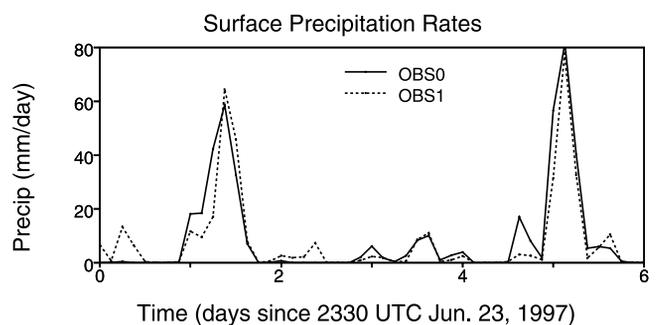
[27] The above discussions demonstrate the sensitivity nature of the SCM to the prescribed forcing data. This highlights a problem when using the forcing derived from NWP models to run SCMs. Because of uncertainties in the NWP-derived forcing and the model sensitivity to these uncertainties, it is difficult to correctly interpret SCM simulation results. Problems explored from such SCM tests may not really reflect problems in the tested parameterizations. It may be just because the enforced NWP-derived forcing does not correctly capture the large-scale dynamic features in the observations due to impacts from imperfect model parameterizations that are used to generate the forcing data, such as during the strong summer convection periods.

### 5. Discussion of the Comparisons

[28] This study used the large-scale forcing data derived from the ARM field measurements by the constrained

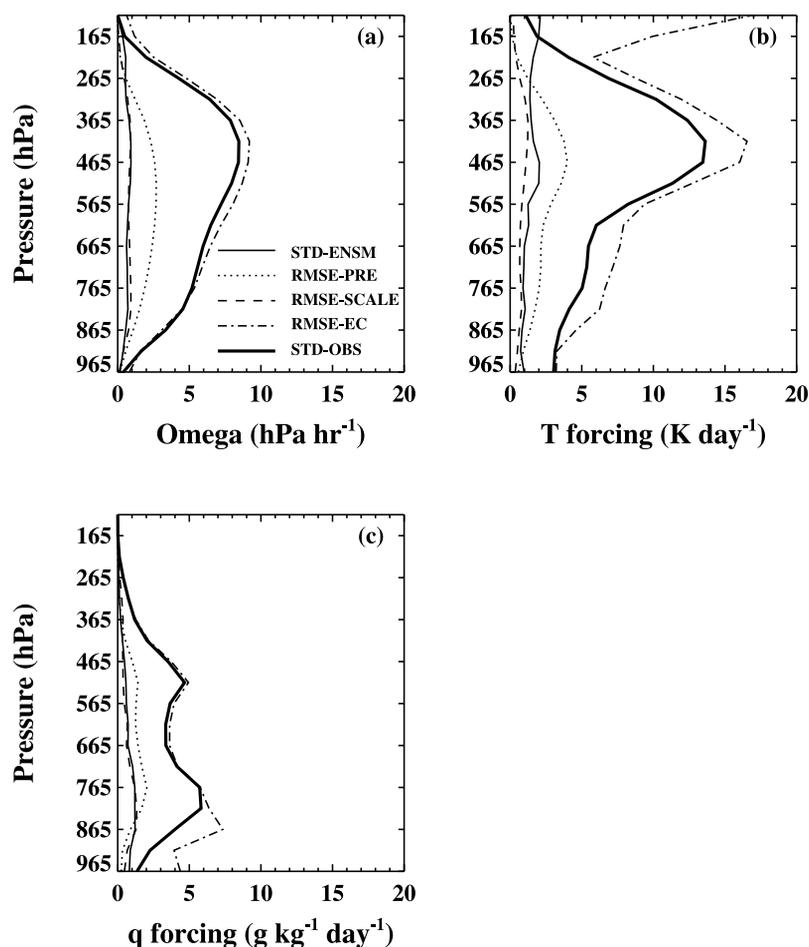
objective analysis method to assess those diagnosed from the ECMWF model. One concern for this comparison is the accuracy of the objective analysis forcing data. Because of the inevitable error in the measurements, it is impossible to obtain perfect forcing data in reality. However, sensitivity tests of the analysis forcing data to uncertainties in the original measurements and some arbitrary parameters used in the variational analysis procedure can help to provide a crude estimate of the accuracy limit in the derived large-scale forcing fields.

[29] Three sets of sensitivity tests are conducted in this study to assess the uncertainty in the variational analysis forcing data using the summer 1997 IOP data. We first examined the sensitivity of the forcing data to uncertainties in the upper air input data. By perturbing the upper air input data, we produced a set of 20 ensembles of forcing data sets constrained with the same column budgets. The perturbation errors added to horizontal winds, temperature, and water vapor mixing ratio fields are randomly generated and are bounded by an RMS error of  $1.5 \text{ m s}^{-1}$  for the wind components,  $0.5 \text{ K}$  for the temperature, and  $5\%$  of the locally observed water vapor mixing ratio for the moisture. These numbers are close to the typical uncertainties in the measurements. The standard deviations of the ensemble mean of the forcing data can be considered as a rough estimate of the uncertainty in the derived forcing data due to errors in the upper air input data. The second test is to examine how the derived forcing data are sensitive to uncertainties in the domain-averaged surface precipitation rates. Zhang *et al.* [2001] showed that the accuracy of the domain-averaged precipitation had the largest impact on derived forcing fields, compared to other constraints. So results obtained from the second test can be considered as a “worst case scenario,” relative to those that could be obtained due to uncertainties in other budget constraint variables. In the current analysis the domain-averaged precipitation is obtained from the Arkansan Basin Red River Forecast Center (ABRFC) 4-km hourly WSR-88D radar precipitation estimates. Over the ARM SCM domain, surface precipitation data are also available from the surface meteorological observational stations (SMOS) and the Oklahoma and Kansas mesonet stations. The locations for these measurement stations can be seen in the work of Zhang *et al.* [2001, Figure 5a]. Figure 11 compares the



**Figure 11.** Time series of the surface precipitation rates (millimeters per day) from the Arkansan Basin Red River Forecast Center radar measurements (solid line) and the surface meteorological observational stations and Oklahoma and Kansas mesonet measurements (dotted line).

## RMS Errors



**Figure 12.** Comparison of uncertainties in the large-scale forcing fields for the summer 1997 case. (a) Vertical velocity. (b) Temperature forcing. (c) Moisture forcing. See text for description of standard deviations of the current analysis forcing data (STD-OBS), upper-air input data (STD-ENSM), root-mean-square error precipitation data (RMSE-PRE), RMS errors larger length scale (RMSE-SCALE), and ECMWF-derived forcing (RMSE-EC).

domain-averaged surface precipitation rates from the ABRFC radar measurements (OBS0) and from the SMOS and mesonet measurements (OBS1). Differences between these two data are considered as a crude estimate of the uncertainty in the observed surface precipitation field. The third test is to examine the sensitivity of the derived forcing data to the length scale parameters that need to be prescribed in determining the weighting coefficients in the interpolation technique used in the analysis procedure [see Zhang *et al.*, 2001, for details]. In this study, we changed the length scale parameters ( $L_x$ ,  $L_y$ ,  $L_p$ ,  $L_t$ ) from current values (50 km, 50 km, 50 hPa, 3 hours, respectively) to (100 km, 100 km, 50 hPa, 6 hours, respectively). Again, the current analysis results are used as the truth to assess results from these sensitivity tests.

[30] Figures 12a–12c, respectively, show the maximum standard deviations of the ensemble forcing data sets by perturbing the upper-air input data (STD-ENSM), the RMS errors caused by using the SMOS and surface mesonet precipitation data (RMSE-PRE), the RMS errors from using a larger length scale (RMSE-SCALE), and the RMS errors

from the ECMWF-derived forcing (RMSE-EC) in the vertical velocity, the temperature forcing, and the moisture forcing fields, during the selected period in the summer 1997 IOP. To judge these errors large or small, the standard deviations of the current analysis forcing data (STD-OBS) are also shown in these figures. It is seen that the range of the uncertainties due to the imperfect input data (STD-ENSM) is within  $1 \text{ hPa h}^{-1}$  for the omega field,  $2 \text{ K d}^{-1}$  for the temperature forcing, and  $1 \text{ g kg}^{-1} \text{ d}^{-1}$  for the moisture forcing, for this summer case. The vertical distribution of STD-ENSM shows larger errors in the middle and upper troposphere for the temperature forcing and larger errors in the lower troposphere for the moisture forcing. Consistent with Zhang *et al.* [2001], smoother and weaker large-scale forcing data fields are produced with using a larger length scale. The difference in the derived forcing fields due to the change in the length scale parameter (RMSE-SCALE) is very close to the uncertainty in these forcing data due to the imperfect upper air input data (STD-ENSM). The derived forcing data from the objective analysis constrained by the SMOS and mesonet precipitation show the RMS errors

(RMSE-PRE)  $< 3 \text{ hPa h}^{-1}$  for the vertical velocity,  $4 \text{ k d}^{-1}$  for the temperature forcing, and  $2 \text{ g kg}^{-1} \text{ d}^{-1}$  for the moisture forcing, compared to those derived from the current analysis that is constrained by the ABRFC radar precipitation. It is noted that all these errors shown in the sensitivity tests are significantly smaller than the “observed” temporal variability of these forcing fields (STD-OBS). In contrast, the ECMWF-derived forcing data exhibit much larger RMS errors, which are generally comparable or slightly larger than the observed standard deviations. The exception is for the ECMWF-derived temperature forcing, which shows dramatically large RMS error in the upper troposphere.

[31] We note that these sensitivity tests just give a rough estimate of the uncertainty in the objective analysis forcing data. Other factors, such as the scale aliases due to insufficient sampling of measurements, the various interpolation methods, the weighting coefficients used in the variational cost function, and uncertainties in other constraint variables, can also affect the analysis results. More complete characterization of the uncertainty in these derived forcing fields requires further detailed studies in the future.

[32] Another concern for this comparison is that the objective analysis derived large-scale forcing fields may contain subgrid scale information. This concern can be somewhat alleviated in this study because the variational analysis approach is intended to dealias small-scale features from the instantaneous soundings by using the domain-averaged constraints to diagnose the desired large-scale forcing fields. However, this approach cannot dealias data in time and in the vertical direction. In the variational analysis we have implemented vertical smoothing and time filtering techniques to reduce impacts of the small-scale noise on the derived large-scale forcing variables, which are somewhat subjective.

[33] It is also noticed from Figure 1 that the ECMWF domain is slightly larger than the variational domain. So one cannot expect the domain-averaged forcing fields derived from ECMWF to be exactly the same as those from the variational analysis. However, the significant disagreements between these two types of forcing data shown during the convective period in this study cannot be easily explained by the differences in the size of averaging domains. In fact, an additional check for the ECMWF-diagnosed forcing data averaged over a smaller domain shows very similar results.

## 6. Summary

[34] The large-scale forcing data set diagnosed from the ECMWF model has been assessed under different weather conditions using data collected at the ARM SGP site during three IOPs. Over the strong convective period during the summer 1997 IOP we have shown that the ECMWF-diagnosed forcing fields are much weaker than those derived from the ARM objective variational analysis. The correlation between these two different forcing data sets is very small. We have shown that the differences are closely related to the errors in the ECMWF-model-predicted surface precipitation. The errors in the model-predicted surface latent and sensible heat fluxes and surface and TOA radiative fluxes also reflect those in the diagnosed forcing fields because they are the important components in the

column-integrated budgets of heat and moisture. Over the moderate precipitation period during the spring 2000 IOP and the nonprecipitation period during the fall 2000 IOP, the disagreements between these two forcing data sets are significantly smaller compared to those over the strong summer convective period. The two forcing data sets display high correlation, although differences between the ECMWF diagnosed data and the variational analysis data are still noticeable.

[35] Using the summer 1997 IOP data, the uncertainty in the variational analysis forcing data has been briefly discussed by examining the sensitivity of the forcing to uncertainties in the upper air input data, uncertainties in the surface precipitation constraint, and uncertainties in selecting length scale used in determining the weighting coefficients in the analysis procedure. We have shown that the uncertainty of the variational analysis forcing due to these changes is much smaller than the RMS error shown in the ECMWF-derived forcing, which is close to or larger than the observed temporal variability. This is a serious concern when using the ECMWF forcing to drive SCMs/CSRMs.

[36] It has been shown that SCM simulations are sensitive to differences in these two different forcing data sets. The SCM simulation errors are not well correlated between the SCM runs with the different forcing data sets for all the three cases. The SCM with the ECMWF forcing tends to reproduce some important aspects of the ECMWF-model-simulated atmosphere (e.g., the surface precipitation rates and the precipitable water), rather than aspects of the observed atmosphere. This is a worrisome problem for SCM tests because such test results may be misleading.

[37] It should be noted that SCMs and CSRMs have stringent requirements for the large-scale forcing data. This study shows that the forcing data diagnosed from ECMWF model and the objective variational analysis are less divergent over the periods that are dominated by large-scale processes (e.g., the spring case and the fall case) than over strong convective periods where there are considerably large disagreements between the forcing data and therefore the use of the ECMWF-model-derived forcing data should be avoided. The ECMWF model nevertheless provides unique long-term continuous data set, including comprehensive information about the dynamical and physical fields, and there is no doubt that they are very useful for evaluation and development of parameterizations in climate models and understanding the structure of large-scale systems and budgets.

[38] To reduce the impact of model physical parameterizations on the NWP forcing data, we applied the variational analysis method to the National Oceanic and Atmospheric Administration mesoscale model Rapid Updated Cycle (RUC) analysis with the ARM-observed column mass, heat, moisture, and momentum constraints in a separate study. In this approach the state variables from the RUC analysis are adjusted to balance the new column budgets by using the ARM surface and TOA observations rather than the RUC model simulations. This is an important difference between the forcing directly derived from RUC analysis and that derived from the new approach. Preliminary results show that this approach can significantly improve the quality of the derived forcing data based on the RUC analysis. More

detailed description about this study will be given in a follow-up paper.

[39] **Acknowledgments.** The authors would like to thank the anonymous reviewers for several valuable comments that helped to improve the presentation of the paper. This research was performed under the auspices of the U. S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract W-7405-Eng-48. Work at SUNY Stony Brook was supported by ARM grant DE-FG02-98ER62570 and was also supported by NSF under grant ATM9701950. We thank Christian Jakob, formerly of ECWMF and now at the Australian Bureau of Meteorology Research Centre, for valuable discussions and comments on the early version of the manuscript. The ECMWF forcing data are obtained from the ARM data archive and provided by Christian Jakob.

## References

- Cho, H.-R., and Y. Ogura, A relationship between the cloud activity and the low-level convergence as observed in Reed-Recker's composite easterly waves, *J. Atmos. Sci.*, 31, 2058–2065, 1974.
- Ghan, S. J., et al., A comparison of single column model simulations of summertime midlatitude continental convection, *J. Geophys. Res.*, 105, 2091–2124, 2000.
- Gregory, D., J.-J. Morcrette, C. Jakob, A. M. Beljaars, and T. Stockdale, Revision of convection, radiation and cloud schemes in the ECMWF Integrated Forecasting System, *Q. J. R. Meteorol. Soc.*, 126, 1686–1710, 2000.
- Hack, J. J., J. A. Pedretti, and J. C. Petch, *SCCM User's Guide*, Natl. Cent. for Atmos. Res., Boulder, Colo., 1998.
- Houze, R. A., and E. N. Rappaport, Air motions and precipitation structure of an early summer squall line over the eastern tropical Atlantic, *J. Atmos. Sci.*, 41, 553–574, 1984.
- Hudson, H. R., On the relationship between horizontal moisture convergence and convective cloud formation, *J. Appl. Meteorol.*, 10, 755–762, 1971.
- Iacobellis, S. F., R. C. J. Somerville, G. M. McFarquhor, and D. L. Mitchell, Sensitivity of radiation fluxes and heating rates to cloud microphysics, paper presented at Twelfth ARM Science Team Meeting Proceedings, U. S. Dep. of Energy, St. Petersburg, Florida, 2002.
- Randall, D. A., K.-M. Xu, R. J. C. Somerville, and S. Iacobellis, Single-column models and cloud ensemble models as links between observations and climate models, *J. Clim.*, 9, 1683–1697, 1996.
- Xie, S. C., and M. H. Zhang, Impact of the convective triggering function on single-column model simulations, *J. Geophys. Res.*, 105, 14,983–14,996, 2000.
- Xie, S. C., et al., Intercomparison and evaluation of cumulus parameterizations under summertime midlatitude continental conditions, *Q. J. R. Meteorol. Soc.*, 128, 1095–1135, 2002.
- Xu, K.-M., et al., An intercomparison of cloud-resolving models with the Atmospheric Radiation Measurement summer 1997 Intensive Observation Period data, *Q. J. R. Meteorol. Soc.*, 128, 593–624, 2002.
- Zhang, M. H., and J. L. Lin, Constrained variational analysis of sounding data bases on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements, *J. Atmos. Sci.*, 54, 1503–1524, 1997.
- Zhang, M. H., J. L. Lin, R. T. Cederwall, J. J. Yio, and S. C. Xie, Objective analysis of ARM IOP Data: Method and sensitivity, *Mon. Weather Rev.*, 129, 295–311, 2001.

---

R. T. Cederwall, S. Xie, and J. J. Yio, Atmospheric Science Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. (rcederwall@llnl.gov; xie2@llnl.gov; yio@llnl.gov)

M. Zhang, Marine Sciences Research Center, State University of New York, Stony Brook, NY 11794-5000, USA. (mzhang@atmsci.msrc.sunysb.edu)