

Developing long-term single-column model/cloud system–resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations

Shaocheng Xie and Richard T. Cederwall

Lawrence Livermore National Laboratory, Livermore, California, USA

Minghua Zhang

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

Received 4 August 2003; revised 15 October 2003; accepted 29 October 2003; published 9 January 2004.

[1] This study proposes another approach to develop multiyear single-column model (SCM) and cloud system–resolving model (CSRМ) forcing data from numerical weather prediction (NWP) model analyses constrained with the observed surface and top of the atmosphere measurements by using a variational analysis approach. In the approach the atmospheric state variables from NWP analyses are adjusted to balance the observed column budgets of mass, heat, moisture, and momentum rather than the NWP model-produced budgets. The derived constrained NWP forcing data from the National Oceanic and Atmospheric Administration rapid update cycle (RUC) analyses are evaluated by the “observed” forcing data derived from radiosonde and wind profiler data collected at the Atmospheric Radiation Measurement (ARM) Program Southern Great Plains site under three selected cases: A strong convective case in the ARM summer 1997 intensive operational period (IOP), a moderate synoptic-scale process-dominated precipitation period in the spring 2000 IOP, and a nonprecipitation period in the late fall 2000 IOP. We show that the forcing data derived from the RUC analyses using ARM column constraints agree with the observed forcing reasonably well. The largest improvements are seen during precipitation periods since precipitation is a strong constraint used in the proposed approach. During the nonprecipitation period the improvements are moderate because the constraints are weak in the absence of precipitation. The constrained NWP forcing and the observed forcing, however, show better agreement during the moderate precipitation period and the nonprecipitation period than during the strong convective period. In SCM tests we show that most model errors revealed by the SCM driven by the observed forcing can be seen in the SCM driven by the constrained NWP forcing. These results suggest the feasibility of using the derived constrained NWP forcing data from RUC analyses for statistical studies of SCM/CSRМ results over long time periods. *INDEX*

TERMS: 0320 Atmospheric Composition and Structure: Cloud physics and chemistry; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; 3399 Meteorology and Atmospheric Dynamics: General or miscellaneous; *KEYWORDS:* large-scale continuous forcing, objective analysis, single-column model and cloud system–resolving model

Citation: Xie, S., R. T. Cederwall, and M. Zhang (2004), Developing long-term single-column model/cloud system–resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations, *J. Geophys. Res.*, 109, D01104, doi:10.1029/2003JD004045.

1. Introduction

[2] Single-column models (SCM) and cloud system–resolving models (CSRМ) have been widely used in recent years to develop and test physical parameterizations in climate models. This is largely because of the ability of these models to isolate targeted parameterizations from the rest of the large-scale model, and the feasibility to validate model

performance directly against available field data. To drive SCMs and CSRMs, however, one needs to specify time-varying vertical profiles of the large-scale vertical velocity and horizontal advective tendencies of atmospheric state variables (i.e., the large-scale forcing). These forcing fields can be derived either from a sounding array in major field programs (e.g., the Atmospheric Radiation Measurement (ARM) Program and the Tropical Ocean-Global Atmosphere Coupled Ocean-Atmosphere Response Experiment (TOGA-COARE)) through objective analysis or from products obtained through data assimilation at operational numerical

weather prediction (NWP) centers. *Randall et al.* [1996] gave a detailed description of the SCM and CSRM approaches and data requirements for running such models.

[3] Most SCM/CSRM studies have focused on evaluating the model's ability to reproduce the detailed time evolution of important meteorological fields (e.g., temperature, moisture, surface precipitation rates, clouds, and radiative fluxes) observed during intensive operational periods (IOPs) in major field programs [e.g., *Ghan et al.*, 2000; *Xie and Zhang*, 2000; *Bechtold et al.*, 2000; *Redelsperger et al.*, 2000; *Xie et al.*, 2002; *Xu et al.*, 2002]. Model errors are then linked to deficiencies in specific physical parameterizations. These studies provide useful insights on how to further improve the model parameterizations.

[4] Other efforts have been recently advocated statistical study of SCM/CSRM results [*Del Genio and Wolf*, 2003]. This is partly due to the sensitive nature of SCMs to uncertainties in the initial conditions and the specified large-scale forcing [*Cripe*, 1998; *Hack and Pedretti*, 2000]. Statistical studies can help smooth out those random errors related to uncertainties in the initial conditions and the specified large-scale forcing so that one can focus on the physically important systematic errors from SCM/CSRM simulations.

[5] Statistical study of SCM/CSRM simulations requires a long-term (preferably, multiyear) forcing data set. This is usually not available from observations. A conventional way to obtain the long-term forcing is to use NWP model analyses. However, such forcing is largely affected by deficiencies in physical parameterizations the forecast model uses. *Xie et al.* [2003] compared the large-scale forcing derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) model and that from an objective analysis of observed data and found that these two forcing data exhibit significant differences over periods where surface precipitation rates were not well simulated by the ECMWF model. They also showed that such differences can have large impacts on SCM simulations. These results are illustrated in Figure 1, which compares the vertical velocity (ω) derived from the constrained objective variational analysis [*Zhang and Lin*, 1997; *Zhang et al.*, 2001] and the ECMWF model during a selected strong convective period in the summer 1997 IOP at the ARM Southern Great Plain (SGP) site. The thick black lines in Figures 1a and 1b are the observed and model-produced surface precipitation rates, respectively. The impact of the two different forcing data on the SCM simulations is shown in Figure 1c. The observed precipitation data are obtained by averaging the hourly Arkansas Basin Red River Forecast Center (ABRFC) 4-km rain gauge adjusted WSR-88D radar measurements over the variational analysis domain (see Figure 3) using a simple arithmetic averaging approach. During this period, two strong convective precipitation events on day 2 and day 6 and a weak precipitation event on day 4 were observed (Figure 1a). 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 1a. This convention is used throughout the paper. Associated with the two strong precipitation events, strong upward motions are seen in the variational analysis derived omega field. In comparison with the observations, the ECMWF model

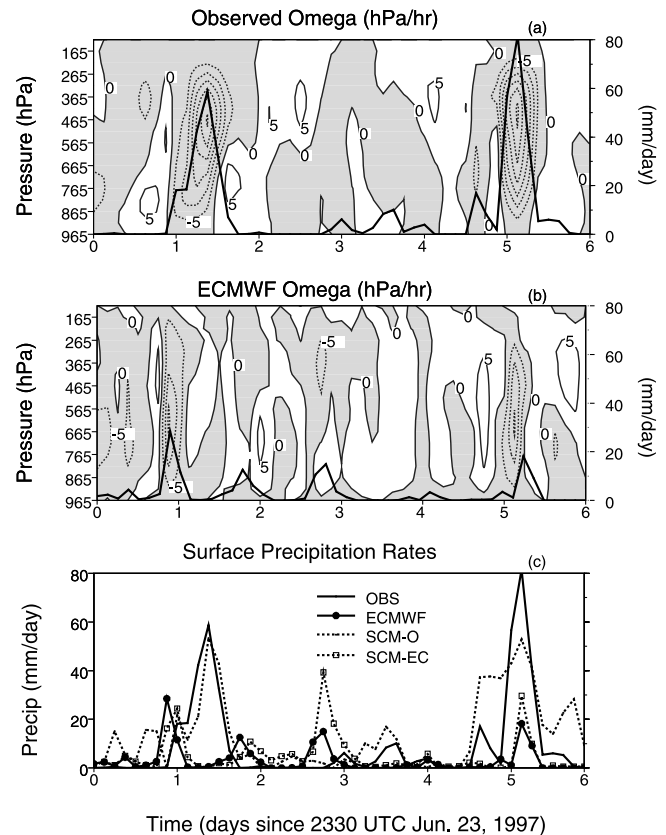


Figure 1. (a) The derived vertical velocity (ω) from the objective variational analysis (OBS) for the selected precipitation period during the summer 1997 IOP. Contour interval is 5. Contours less than 0 are shaded. In Figure 1, solid lines are for contours greater than or equal to zero and dotted lines for contours less than zero. (b) Same as Figure 1a except for the omega derived from the ECMWF model. Thick solid lines in Figure 1a and 1b are the observed and ECMWF model-produced surface precipitation rates (mm d^{-1}), respectively. (c) Time series of the observed (solid), ECMWF-produced (solid line with solid circles), and the SCM-simulated surface precipitation rates (mm d^{-1}). Dotted line shows the ARM variational forcing, and dotted line with open squares shows the ECMWF forcing.

predicts much weaker precipitation and tends to trigger convection more often. This is partially related to potential deficiencies in the model cumulus parameterization. In addition, uncertainties in the observed precipitation field and the lack of sufficient small-scale and mesoscale observations available for the ECMWF data assimilation system may also contribute to the differences between the model and the observations. Consistent with the problems in its produced precipitation, the model shows much weaker upward motions and much higher temporal variability in its derived omega field than the observed. Using these two forcing data sets to drive the National Center for Atmospheric Research (NCAR) Community Climate Model, Version 3 (CCM3) SCM with a modified cumulus convection scheme [*Xie and Zhang*, 2000], Figure 1c shows that the SCM driven by the objective analysis forcing generally captures well the observed precipitation, while

the SCM with the ECMWF forcing largely underestimates the observed values. It actually reproduces well the ECMWF model-calculated precipitation. It should be noted that the forcing data derived from the objective variational analysis can also contain errors due to uncertainties in the upper air data and the column constraints. However, the uncertainty of the variational forcing is much smaller than the differences between the model-derived forcing and the variational forcing as shown in *Xie et al.* [2003], which gave detailed discussions about these two forcing data. The above results indicate that the NWP-derived forcing data should be used cautiously.

[6] In this study, we propose another approach to develop the required long-term forcing data set. This approach is based on the constrained objective variational analysis method that was developed by *Zhang and Lin* [1997] and was applied to process the ARM IOP data by *Zhang et al.* [2001]. The variational analysis uses the domain-averaged surface and top of the atmosphere (TOA) fluxes as the constraints to adjust the balloon soundings and profiler data to satisfy the conservation of mass, heat, moisture, and momentum. Note that previous objective analysis approaches [e.g., *Barnes*, 1964; *Lin and Johnson*, 1996] typically only utilize the mass constraint. A desirable feature of the variational analysis approach is that it can significantly reduce the sensitivity of the resulting derived forcing fields to uncertainties in upper air input data after applying these constraints, especially the heat and moisture budget constraints [*Zhang et al.*, 2001]. This encouraging feature has led us to constrain the air profile data from NWP analysis by using observed surface and TOA measurements to derive the long-term forcing data set. Note that, in this approach, the state variables from NWP analyses are adjusted to balance the observed column budgets rather than the NWP model-predicted column budgets. This is an important difference between the forcing directly derived from NWP analyses and that derived from the proposed approach. On the other hand, the quality of derived forcing fields from this approach will partially depend on the accuracy of these observed flux data, especially precipitation [*Zhang et al.*, 2001]. The quality of surface measurements has been improved in the ARM program, which uses a dense network of surface observations to derive the surface flux data, and a rigorous quality control procedure to process collect data. Furthermore, the flux data have been merged if they are available from different instruments, and have been spatially and temporarily averaged. These procedures reduce the sensitivity of the derived domain-averaged flux data to errors in the individual station measurements. More details about the implementation of the constrained variational analysis approach to analyze the ARM data can be seen in *Zhang et al.* [2001].

[7] In section 2, we will give a detailed description of the proposed approach. In section 3, we will briefly describe the National Oceanic and Atmospheric Administration (NOAA) rapid update cycle (RUC) analyses that are used to provide the required upper air input data. Evaluation of the forcing data derived from this method against those derived from the objective variational analysis during some selected ARM IOPs will be discussed in section 4. Section 5 will show the impact of the derived forcing data on SCM

simulations. A summary and discussions of future work will be given in section 6.

2. Approach

[8] The main element of our approach used to derive the long-term forcing data set is the constrained variational analysis of *Zhang and Lin* [1997]. It was developed for deriving large-scale vertical velocity and advective tendencies from sounding measurements over a network of a small number of stations. The basic idea in this objective analysis approach is to adjust atmospheric state variables either from soundings or from NWP products, given their inevitable uncertainties, by the smallest possible amount to conserve column-integrated mass, moisture, static energy, and momentum. Here we briefly review this approach and describe its adaptation in the present study. More details of the implementation are given by *Zhang et al.* [2001].

[9] As shown in *Zhang and Lin* [1997], the column constraints can be obtained by vertically integrating the governing equations of the large-scale atmospheric fields:

$$\langle \nabla \cdot \mathbf{V} \rangle = -\frac{1}{g} \frac{dp_s}{dt}, \quad (1)$$

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

$$\frac{\partial \langle s \rangle}{\partial t} + \langle \nabla \cdot \mathbf{V} s \rangle = R_{TOA} - R_{SRF} + LP_{rec} + SH + L \frac{\partial \langle q_l \rangle}{\partial t}, \quad (3)$$

$$\frac{\partial \langle \mathbf{V} \rangle}{\partial t} + \langle \nabla \cdot \mathbf{V} \mathbf{V} \rangle + f \mathbf{K} \times \langle \mathbf{V} \rangle + \nabla \langle \phi \rangle = \boldsymbol{\tau}_s, \quad (4)$$

where

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

In the above, \mathbf{V} is the wind, $s = C_p T + gz$ is the dry static energy, q is the mixing ratio of water vapor, and p_s is the surface pressure. q_l is the cloud liquid water content. R is the net downward radiative flux at TOA and at the surface (SRF), $\boldsymbol{\tau}_s$ is the surface wind stress, P_{rec} is precipitation, L is the latent heat of vaporization, SH is the sensible heat flux, E_s is the surface evaporation, and p_t is the TOA pressure. Phase changes associated with ice are neglected for simplicity.

[10] The final analysis product is derived by minimizing the cost function:

$$I(t) = \iint \int_{p, x, y} \left[\alpha_u (u^* - u_0)^2 + \alpha_v (v^* - v_0)^2 + \alpha_s (s^* - s_0)^2 + \alpha_q (q^* - q_0)^2 \right] dx dy dp, \quad (5)$$

with equations (1)–(4) as strong constraints, where the asterisk superscript denotes the analyzed data and subscript

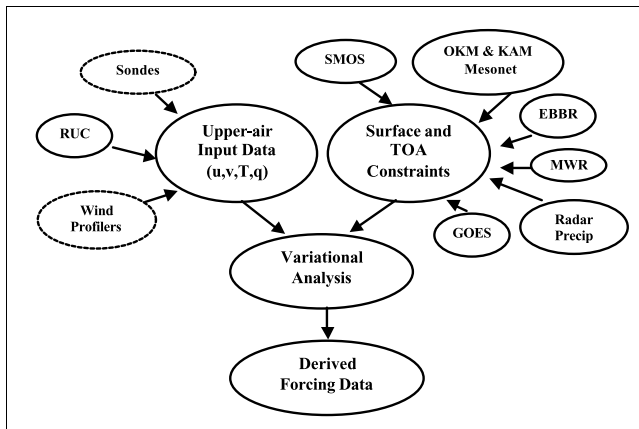


Figure 2. A diagram that illustrates the approach used to derive the long-term forcing using the RUC analyses constrained by the ARM observations. Circles enclosed by dashed lines are for the data that are not available or not used in developing the constrained NWP forcing data. Abbreviations are as follows: SMOS, Surface Meteorological Observation Stations; OKM and KAM, Oklahoma and Kansas mesonet stations, respectively; EBBR, Energy Budget Bowen Ratio (EBBR) stations; MWR, MicroWave Radiometer stations; and GOES, Geostationary Operational Environment Satellite. Details are given by *Zhang et al. [2001]*.

zero denotes original input of state variables. Here α is the weighting function related with error estimates in the initial input. Three practical aspects of the approach are relevant to the present adaptation. One is the input data of the atmospheric state variables. The second is the constraining physical quantities that appear on the right-hand side of equations (1) to (4). The third is the specification of the weighting functions in equation (5).

[11] In the original implementation of *Zhang et al. [2001]* (Figure 2), the upper air input data consisted of five ARM balloon sondes and seven NOAA wind profilers. RUC analyses were used as background fields. Figure 3 shows locations of the ARM sounding stations, the NOAA wind profilers, the RUC analysis grids, and the 13 final analysis grids designed by *Zhang et al. [2001]*. The required constraints were obtained from an intensive array of the ARM surface and TOA measurements, which were described by *Zhang et al. [2001]*. *Zhang et al. [2001]* used the manufacturer error specifications as the instrument uncertainty and 20% of the time standard deviation as aliasing uncertainty to determine the α in equation (5). It was shown that the adjustments made to the input data to conserve the column integral budget of mass, moisture, energy, and momentum are comparable to uncertainties in the original measurements. Compared to the conventional analysis approach in which only a mass constraint was applied, the constrained variational analysis with additional constraints typically exhibits considerably smaller errors [*Waliser et al., 2002*]. The forcing data derived from the variational analysis have been used in recent ARM SCM and CSRM studies.

[12] Long-term continuously observed upper air input data (e.g., sondes) for deriving the large-scale forcing are not available in the observations. However, the ARM

program has collected long-term continuous surface and TOA fluxes near its SGP site for the past several years. This provides the required column constrained variables for the variational analysis approach. Furthermore, balloon soundings are launched four times per day at the ARM SGP facility over the past several years, which provide the necessary information to characterize uncertainties in the NWP RUC products. To develop the multiyear long-term forcing data, we used the NWP (e.g., RUC) analyses as the upper air input data and the ARM surface and TOA flux measurements as the constraints in the constrained variational analysis. We further added the RMS difference between the RUC and balloon soundings at the ARM SGP site in the specification of the α in equation (5). It should be noted that the forcing directly derived from NWP analyses is also dynamically and thermodynamically consistent. However, this consistency is with the model-calculated surface and TOA constraints, which are not the same as observations. In our approach, the atmospheric state variables from NWP analyses are adjusted to balance the observed column budgets. In the following sections, we will show that using the surface and TOA observations to constrain the NWP state variables through the variational analysis approach improves the quality of derived large-scale forcing data.

3. RUC Analyses

[13] The atmospheric state variables from the NOAA RUC-2 analyses have been used as the first input to develop the ARM long-term forcing data set. The RUC is an operational atmospheric prediction system, including a numerical forecast model and an analysis system to initialize that model. It is used to assimilate recent observations aloft and at the surface to provide very high frequency updates of current conditions and short-range forecasts using a mesoscale model. The RUC provides hourly three-dimensional objective analyses over the contiguous

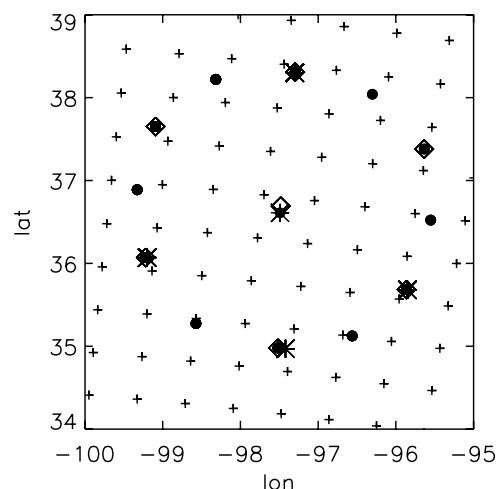


Figure 3. The locations of the ARM five sounding balloons (asterisks), the seven NOAA wind profilers (diamonds), the RUC analysis grids (plus signs), and the SCM variational analysis domain (circles) at the ARM SGP site.

RMS Errors

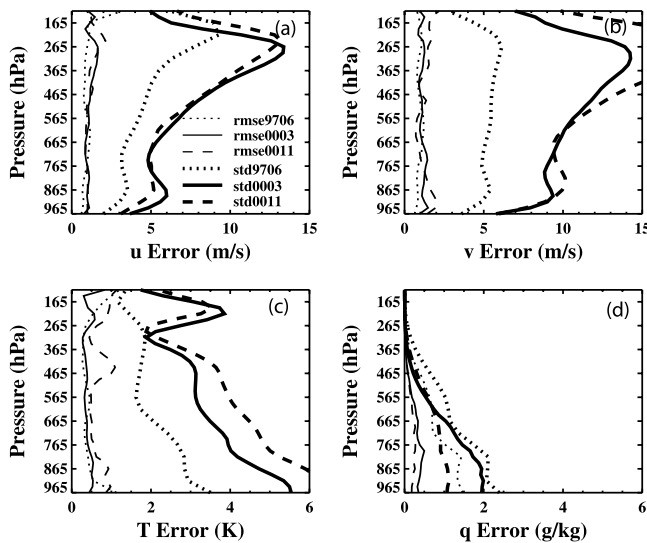


Figure 4. The errors in the RUC analyses of (a) horizontal wind u component, (b) v component, (c) temperature, and (d) moisture during the ARM summer 1997 IOP, spring 2000 IOP, and fall 2000 IOP. Thin lines show the RMS difference (RMSE) between the RUC analyses and the ARM observations. Thick lines represent standard deviations (std) of the observed values.

United States with a 40 km horizontal resolution and 40 vertical sigma levels. The RUC analysis grids around the ARM SGP site are displayed in Figure 3. More detailed description of the RUC analysis is given by Benjamin *et al.* [1998].

[14] To examine the quality of the RUC analyses, Figure 4 shows the root-mean-square (RMS) error of the variational analysis domain-averaged atmospheric state variables from the RUC analyses (thin lines) during the ARM summer 1997 IOP, the spring 2000 IOP, and the late fall 2000 IOP. These three IOPs are from 18 June (2330 UTC) to 18 July (2330 UTC) 1997; 1 March (1730 UTC) to 22 March (0830 UTC) 2000; and 27 November (1730 UTC) to 22 December (0830 UTC) 2000, respectively. Also shown are the observed standard deviations of the observed fields (thick lines) during the three IOPs. RUC analyses typically capture well the temporal evolutions of the observed large-scale atmospheric state variables (not shown) with the RMS error of typically less than 1 m s^{-1} in the horizontal wind components, 0.5 K in the temperature, and 0.5 g kg^{-1} in the moisture within most of the troposphere (Figures 4a–4d). Even though the magnitude of these RMS errors for the domain-averaged variables are small, the differences in these state variables between the RUC analyses and the observed data at individual stations, however, may result in large differences in the derived large-scale forcing fields (e.g., vertical velocity and the large-scale advection terms) [Zhang and Lin, 1997]. It is noted that relatively larger errors are seen in the moisture field during the summer IOP and in the temperature field during the late fall IOP. In comparison with the observed variability itself, these errors in the RUC analyses are generally smaller except for the

moisture field during the summer IOP, which shows errors that are about 60% of the magnitude of the observed variability. This suggests improvement of the moisture analyses during summertime is needed.

[15] The RMS errors shown here are used as RUC uncertainties. They are added to the instrument errors of the balloon soundings and measurement aliasing errors given by Zhang *et al.* [2001] to specify the weighting functions in the cost function for variational minimization. As pointed out by Zhang and Lin [1997], the absolute magnitudes of these weighting functions do not affect the final analysis. Yet their relative distributions among different heights, stations, and different variables do matter in the final product.

4. Evaluation of the ARM Constrained NWP Forcing Data With IOP Observations

[16] Similar to Xie *et al.* [2003], the ARM constrained NWP forcing data developed from the NOAA RUC analyses using the approach described in section 2 are evaluated by comparing to those derived from the objective variational analysis [Zhang and Lin, 1997; Zhang *et al.*, 2001] under three selected cases: A strong convective period from 23 June (2330 UTC) to 29 June (2330 UTC) during the summer 1997 IOP (i.e., the same period as used in Figure 1), 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 late fall 2000 IOP. Similar results are obtained for the entire IOPs. In this study, the forcing derived from the objective variational analysis is considered as the “observed” forcing. Uncertainties in the observed forcing data were discussed by Xie *et al.* [2003] for the summer case. To address the impact of the ARM observed constraints on the derived forcing fields, we will also discuss the forcing derived from the RUC analyses with mass constraint only. The latter can be roughly considered as the forcing data that are derived directly from the RUC analyses, which are not available in the current RUC analysis archive. In the following discussions, we use “RUC-C” to represent the experiment with the RUC analysis input data and all column budget constraints from the ARM observations, “RUC” to represent the experiment with RUC analyses constrained by the column mass only, and “OBS” to represent the experiment with the objective variational analysis.

4.1. Summer Strong Convective Case

[17] Figures 5a and 5b shows the derived omega fields from RUC-C and RUC, respectively, for the summer case. The omega derived from OBS and the observed surface precipitation rates are shown in Figure 1a. It is seen that even though both RUC-C and RUC typically capture the bulk characteristics of the observed omega field, i.e., upward motions during the precipitation events and downward motions during the nonprecipitation periods, RUC-C exhibits considerably better agreement with OBS compared to RUC. The biggest improvements are seen during two strong precipitation periods where RUC produces the omega that is much weaker than OBS. Moderate improvements made by RUC-C are also seen in other periods compared to

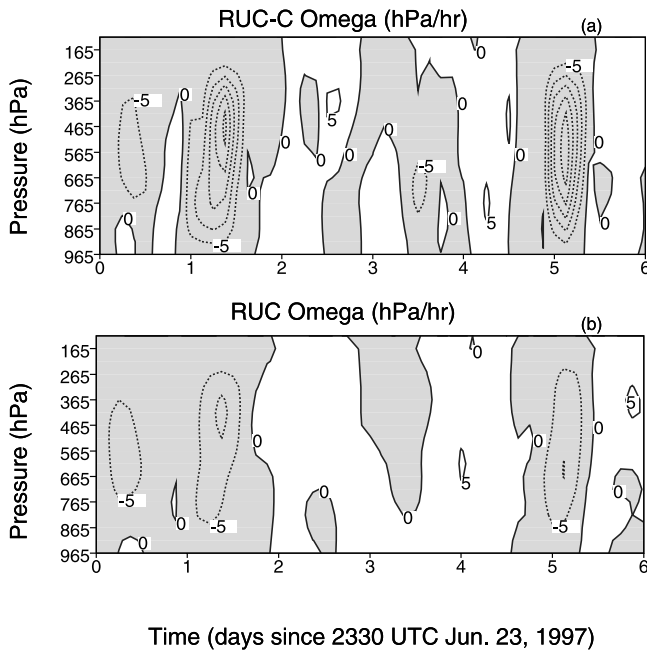


Figure 5. Same as Figure 1a except for the derived omega from (a) the RUC analyses with the ARM constraints (RUC-C) and (b) the RUC analyses with the mass constraint only (RUC).

RUC; for example, RUC-C correctly reproduces the observed downward motions at the middle of day 1 (although a little weak), while RUC generally shows upward motions during the time.

[18] Significant improvements can also be seen in the derived large-scale temperature and moisture forcing fields, i.e., the total advective tendencies of temperature and moisture, from the RUC analyses with the ARM constraints. The adiabatic compression/expansion term is included in the temperature forcing. Figure 6 shows a scatterplot of these derived forcing fields at all vertical levels between RUC-C and OBS and between RUC and OBS. With the ARM constraints, RUC-C captures well the observed temperature and moisture forcing fields in both the magnitude and the temporal variability. It just slightly underestimates the observed advective cooling (Figure 6a) when the cooling is strong ($>20 \text{ K d}^{-1}$). In contrast, the underestimation is significantly larger in RUC (Figure 6c). Similar results for advective moistening can be seen in the moisture forcing (Figures 6b and 6d). In addition, RUC-C shows much higher correlation with OBS in these derived forcing fields than RUC. Examination of the correlation as a function of height shows that the correlation for the temperature forcing between RUC-C and OBS is larger than 0.8 at all vertical levels (larger than 0.95 at the levels above 615 hPa) and the correlation for the moisture forcing is typically larger than 0.7 for the levels below 465 hPa. These are significantly larger than the correlations between RUC and OBS. The weaker correlation in the moisture forcing between RUC-C and OBS compared to that in the temperature forcing could be due to the relatively larger errors in the moisture from the RUC analyses as shown in Figure 4d.

[19] To quantify the error from using the RUC analyses, Figure 7 shows the RMS error of the derived forcing fields

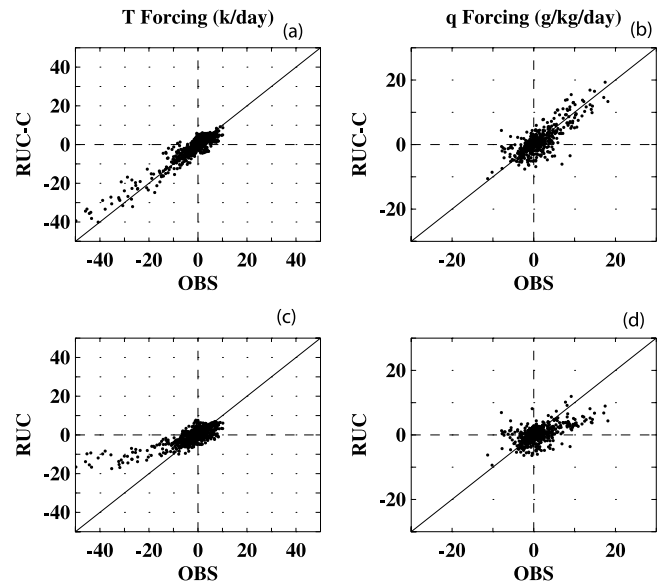


Figure 6. Scatterplots of the derived temperature and moisture forcing data at all vertical levels from (a and b) RUC-C and (c and d) RUC for the selected precipitation period during the summer 1997 IOP.

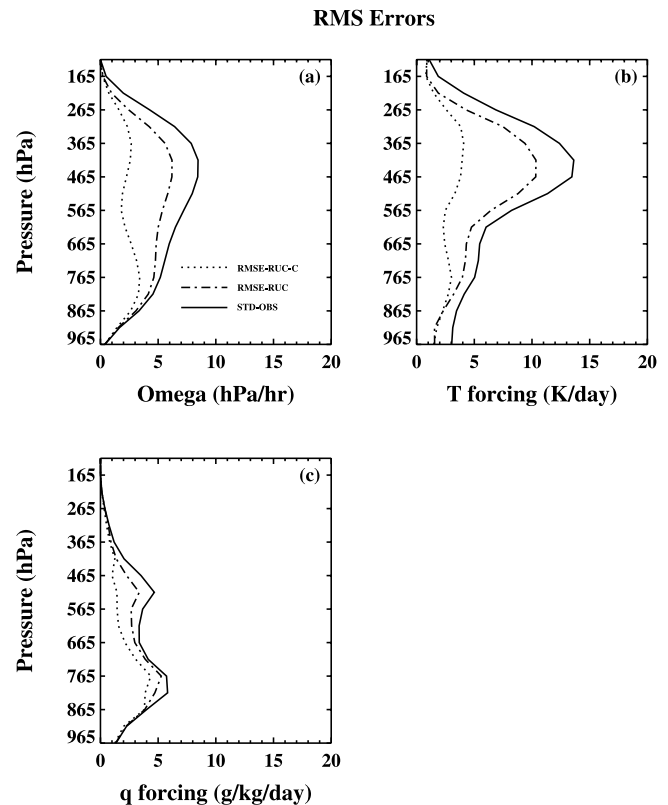


Figure 7. The RMS difference between the derived forcing data from RUC-C and OBS (dotted lines) and from RUC and OBS (dash-dotted lines) for the selected precipitation period during the summer 1997 IOP. Solid lines show the standard deviations of the observed forcing data. (a) Vertical velocity. (b) Temperature forcing. (c) Moisture forcing.

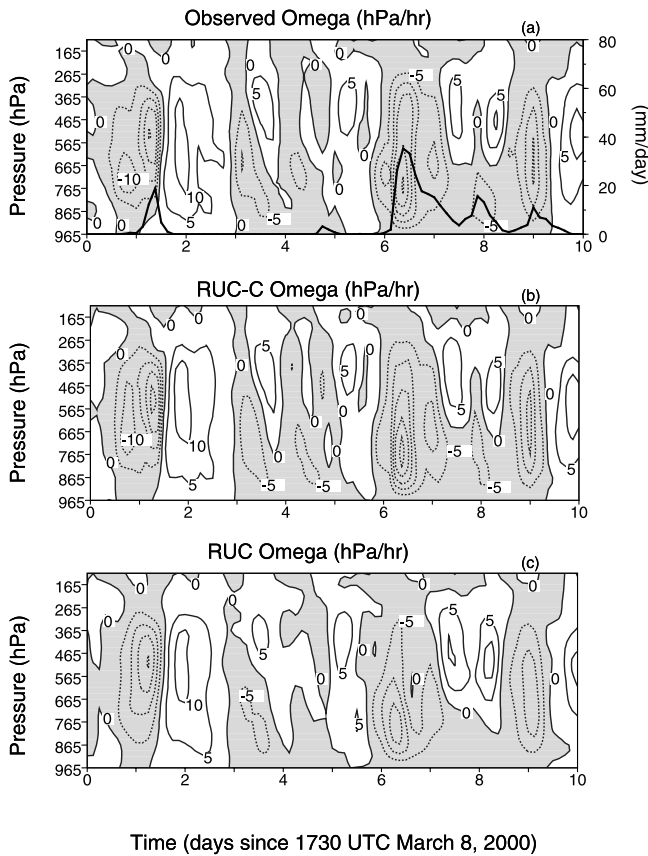


Figure 8. The derived omega from (a) OBS, (b) RUC-C, and (c) RUC for the selected moderate precipitation period in the spring 2000 IOP. Contour interval is 5. Contours less than 0 are shaded. In Figure 8, solid lines are for contours greater than or equal to zero, and dotted lines are for contours less than zero. Thick solid line in Figure 8a is the observed surface precipitation rates (mm d^{-1}).

from RUC-C and RUC. The RMS error of the RUC-C derived forcing fields is within 3 hPa h^{-1} for the omega field, 3 K d^{-1} for the temperature forcing, and $4 \text{ g kg}^{-1} \text{ d}^{-1}$ for the moisture forcing, for this summer case. These are larger than the uncertainty in the observed forcing data due to instrument and measurement errors in the upper air data (sondes), which is within 1 hPa h^{-1} for the omega field, 2 K d^{-1} for the temperature forcing, and $1 \text{ g kg}^{-1} \text{ d}^{-1}$ for the moisture forcing as shown by Xie *et al.* [2003]. However, these errors are much less than those in the RUC derived forcing data. It is also noted that the forcing data from RUC-C have the RMS errors that are considerably smaller than the observed temporal variability itself (STD-OBS) except for the moisture forcing, which shows relatively large errors in the lower troposphere. In contrast, the forcing data derived from RUC exhibit much larger RMS errors, which are of nearly the same magnitude as the observed variability.

[20] The above results clearly show that with the observed column constraints, especially the heat and moisture constraints, the variational analysis can improve the large-scale forcing fields that are derived from the RUC analyses, especially during strong precipitation periods where the constraints are strong. These results also suggest

that improvement of the quality of the RUC analyses is needed, particularly for the moisture field, in order to further improve the constrained NWP forcing data set during the summer convective season. The magnitudes of the errors suggest that the forcing derived from the NWP products cannot completely replace those derived from soundings, yet they may be useful for statistical studies.

4.2. Spring Moderate Precipitation Case

[21] Figure 8 compares the derived omega from RUC-C and RUC with OBS for the spring moderate precipitation case. The thick line in Figure 8a is the surface precipitation rate observed during this period. It is seen that the spring case contained one precipitation event on days 1–2 and one multiday precipitation event on days 7–10. These precipitation events are mainly associated with large-scale frontal systems. Associated with these precipitation events, strong upward motions are seen in the observed omega field (Figure 8a). Similar to the summer case, without the observed column heat and moisture constraints, RUC (Figure 8c) produces weaker and smoother upward motions than the observed. It also shows weaker temporal variability compared to the observations. Corresponding to the first precipitation event on days 1–2, for example, the observed omega shows two peaks with one around 715 hPa and the other one around 515 hPa. The first peak is related to the very weak precipitation that occurred at the end of day 1, and the second peak is associated with the precipitation peak in the middle of day 2. This observed structure is smoothed by the omega derived from RUC, which only shows a single peak in its produced omega field. These problems are significantly reduced in RUC-C when the observed column heat and moisture constraints are used (Figure 8b). The magnitude and temporal variations of the observed omega field are well reproduced by RUC-C.

[22] Figure 9 shows the same information for temperature and moisture as Figure 6 except for the spring case. As with the omega, RUC-C produces the temperature and moisture

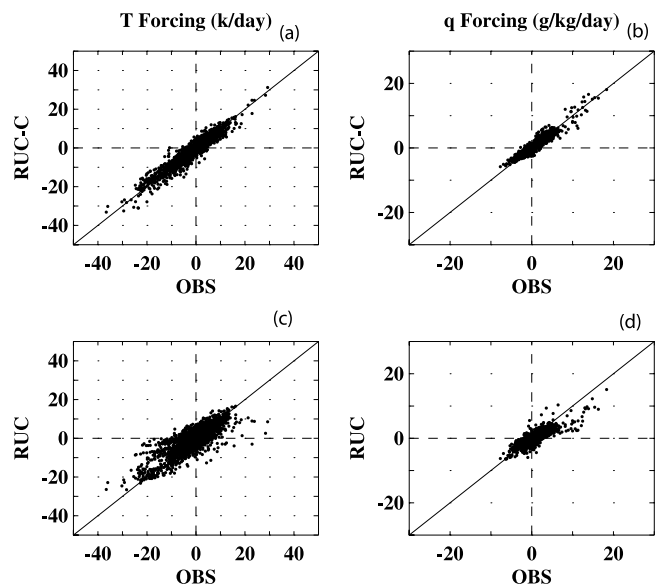


Figure 9. Same as Figure 6 except for the selected moderate precipitation period in the spring 2000 IOP.

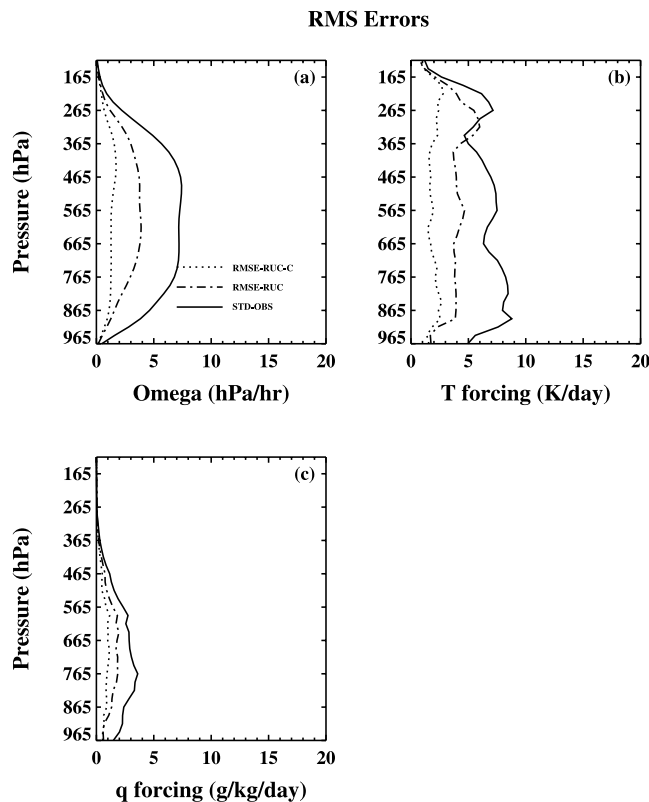


Figure 10. Same as Figure 7 except for the selected moderate precipitation period in the spring 2000 IOP.

forcing data that have better agreement with OBS than RUC. The correlation between the RUC-C and the OBS derived forcing fields is dramatically high. For the temperature forcing, the correlation is larger than 0.95 for the most of the troposphere except for the levels above 265 hPa where the forcing itself is weak. For the moisture forcing, the correlation is larger than 0.9 in the levels below 465 hPa. Above that level, the moisture forcing is very small. It should be noted that, when the forcing itself is very small, any small difference between the two types of forcing data could have significant impacts on the correlation.

[23] The RMS error of these derived forcing data for the spring case is shown in Figure 10. It is seen that the analysis errors from RUC-C are about 1 hPa h^{-1} in the omega, 2.5 K d^{-1} for the temperature forcing, and $1 \text{ g kg}^{-1} \text{ d}^{-1}$ for the moisture forcing, respectively. These are comparable to the uncertainties in the observed forcing data due to instrument and measurement errors in the sounding input data as discussed by Xie *et al.* [2003] and significantly smaller than the observed variability. The better results for the spring case compared to the summer case is partly because the RUC analyses show smaller errors (relatively to the observed variability) in the spring case than the summer case (Figure 4), especially for the moisture analyses. The smaller errors in the RUC analyses during the spring case might also be due to a larger fraction of the spatial variability resolved by the model during the frontal episodes than during the summer convective periods. It is interesting to see that the RMS errors of the forcing data derived from RUC in the spring case are also smaller than those in the

summer case. The magnitude of the RMS error of the RUC derived forcing is just around 50% of the magnitude of the observed variability. This is consistent with the ECMWF model-derived forcing data. Xie *et al.* [2003] found that the ECMWF forcing and the observed forcing show much better agreement in the spring case than in the summer case.

4.3. Fall Nonprecipitation Case

[24] In the above discussions, we have shown that the largest improvements when using our approach are over periods where precipitation events were observed. In this subsection, we will examine whether improvements can still be obtained in nonprecipitation periods where the column heat and moisture constraints are relatively weak, in absence of precipitation. As shown in Figure 11a, the selected fall case was a dry period, and it was dominated by large-scale downward motions. Compared to the summer and the spring cases, the derived omega from OBS during the nonprecipitation period is weaker. In general, both RUC-C and RUC capture the observed omega well although they both tend to underestimate the observed downward motions (Figures 11b and 11c). An encouraging result is that RUC-C still displays noticeably better agreement in the derived omega with OBS than RUC during the quiet period. The problem of underestimation of the observed downward motions on day 2 and days 5–6 and the upward motions

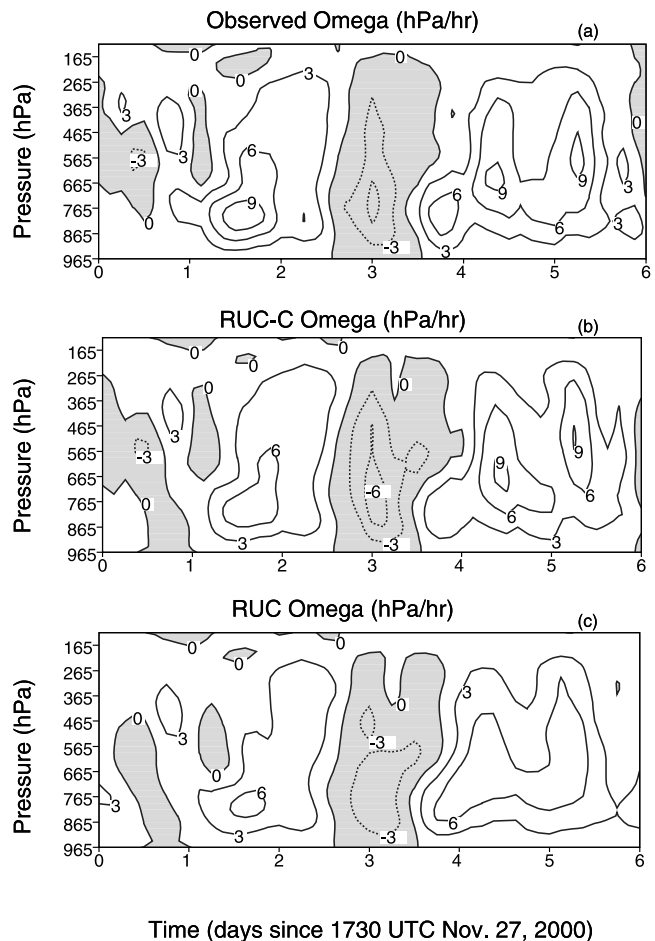


Figure 11. Same as Figure 8 except for the selected nonprecipitation period in the fall 2000 IOP.

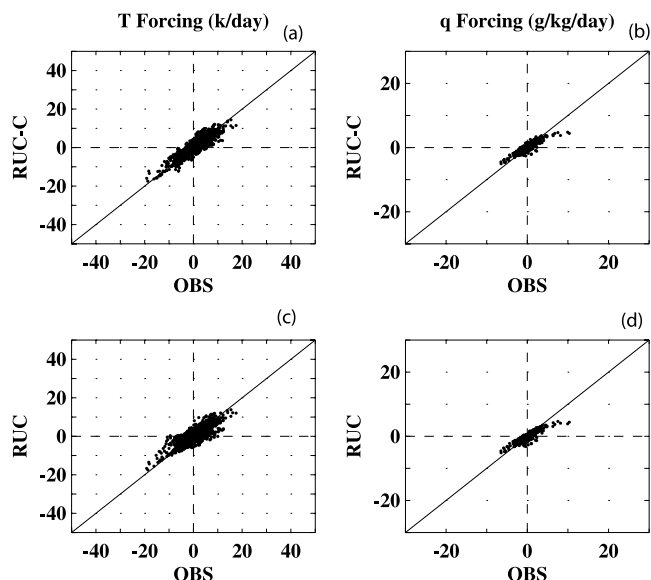


Figure 12. Same as Figure 6 except for the selected nonprecipitation period in the fall 2000 IOP.

on day 3 in RUC is clearly reduced in RUC-C, which produces the vertical motions that are of nearly the same magnitude as the observed values.

[25] As in Figures 6 and 9, Figure 12 shows the correlation between RUC-C and OBS and between RUC and OBS in the derived temperature and moisture forcings. While both RUC-C and RUC reproduced well the observed forcing, RUC-C shows slight improvements in the temperature forcing, consistent with the improvements in the omega field. There is no clear improvement seen in the derived moisture forcing.

[26] Similar results can be seen in the RMS error plots (Figures 13a–13c). With the ARM observed constraints, RUC-C shows slight improvements in the omega field (Figure 13a), slight or some improvements in the derived temperature forcing (Figure 13b), and almost no improvement in the moisture forcing (Figure 13c), compared to RUC.

5. Impacts on SCM Simulations

[27] In this section, we use the NCAR CCM3 SCM (J. J. Hack et al., SCCM user's guide, National Center for Atmospheric Research, Boulder, Colorado, 1998, available at <http://www.cgd.ucar.edu/cms/scm/scm.html>) to investigate how the difference between the two types of forcing data affects SCM simulations. A modified cumulus convection scheme [Xie and Zhang, 2000] is used to avoid apparent model biases associated with poor timing of convection. We will focus on whether the SCM driven by the constrained NWP forcing (RUC-C) can produce simulations that are similar to those from the SCM driven by the observed forcing (OBS) so that the same model deficiencies can be detected.

[28] In the SCM runs, the large-scale total temperature and moisture forcings are specified from the two different forcing data sets and the surface forcing is calculated by the model surface parameterizations. A series of a 36-hour

forecast runs is initiated every day to avoid serious drift of SCM simulations [e.g., Ghan et al., 2000]. For each forecast, the temperature and moisture are initialized with the observations. A composite of 12–36 hour forecasts from the series of 36-hour runs is analyzed. For the convenience of discussion, we use SCM-OBS to represent the SCM run with the observed forcing and SCM-RUC-C to represent the run with the constrained NWP forcing in the following discussions. We will not show results from the SCM driven by the forcing derived from RUC, which generally shows much larger simulation errors than those using the constrained NWP forcing (RUC-C).

[29] Figures 14 and 15 show the SCM biases of simulated temperature and moisture, respectively, driven by the two forcing data sets during the three selected periods. Note that the SCM-produced moisture errors are normalized respectively by the mean specific humidity during the three selected cases in order to have a better comparison for the different seasons since there is much less moisture in the air during the spring/fall cases than the summer case. The mean values of the specific humidity are 4.2 g kg^{-1} , 1.41 g kg^{-1} , and 0.82 g kg^{-1} for the summer, spring, and fall cases, respectively. In general, both runs produce comparable simulation results for most of the periods and the model levels. The magnitudes of the temperature and the normalized moisture biases are generally comparable for the three cases. It is interesting to see that both runs show moist biases near the surface and dry biases in the lower and middle troposphere for all the three cases (Figure 15). This may reflect model deficiencies in vertically transporting

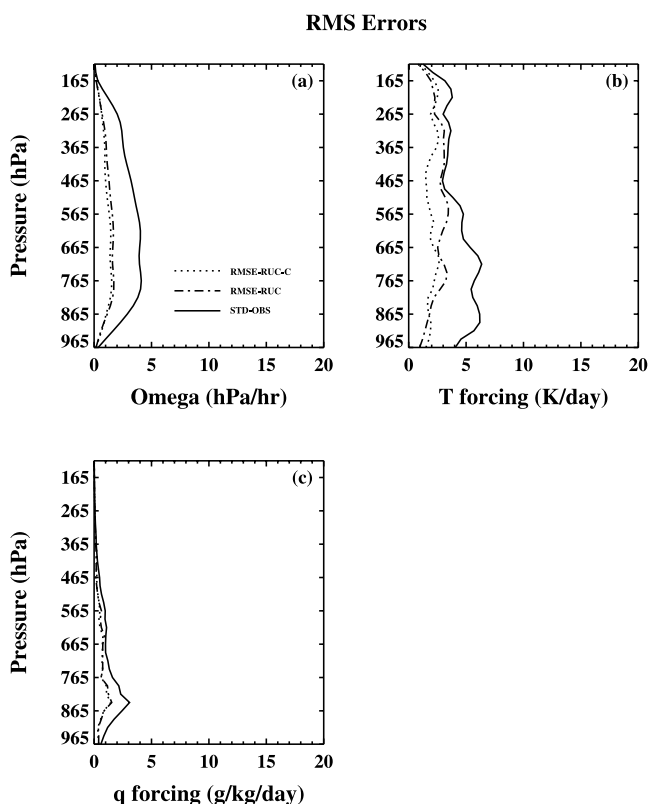


Figure 13. Same as Figure 7 except for the selected nonprecipitation period in the fall 2000 IOP.

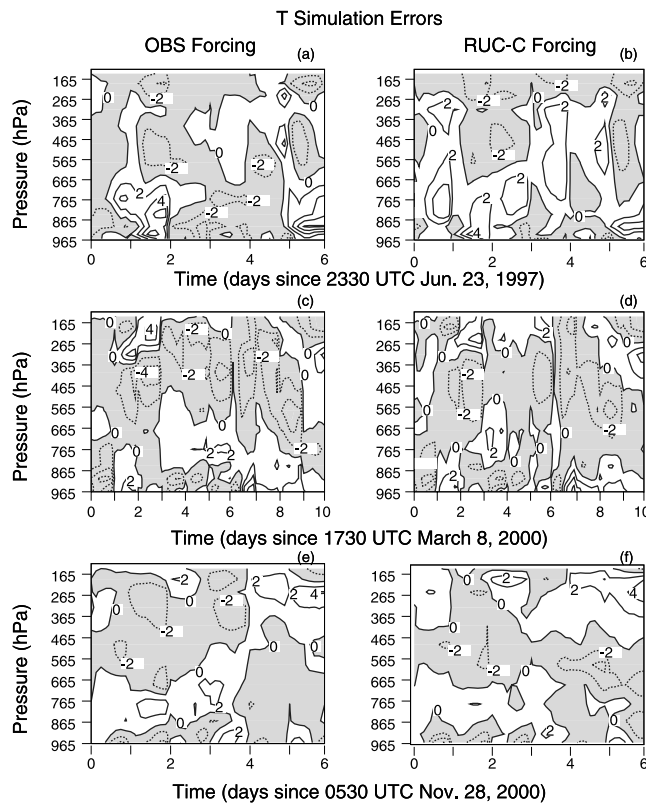


Figure 14. The time-height distributions of differences between the simulated temperature and the observations. (a, c, and e) (left) Biases produced by the SCM with the OBS-derived forcing (SCM-OBS). (b, d, and f) (right) Biases produced by the SCM with the RUC-C-derived forcing (SCM-RUC-C), respectively. Contour interval is 2. Contours less than zero are shaded. In Figure 14, solid lines are for contours greater than or equal to zero, and dotted lines for contours less than zero.

moisture from surface to troposphere. Another important feature shown in Figures 14 and 15 is that the distribution of the model simulation errors is quite similar between these two runs, especially for the fall and spring cases. The biases produced by these two forcing data are well correlated for most of the model levels.

[30] Figures 16a and 16b compare the simulated surface precipitation rates of the two runs during the summer and spring periods, respectively. It is encouraging to see that the simulated surface precipitation rates produced by SCM-RUC-C and SCM-OBS agree quite well with each other for both cases. The temporal correlations between the two simulated precipitation rates are 0.91 for the summer case and 0.97 for the spring case, respectively. For the summer case, both SCM-OBS and SCM-RUC-C produce spurious precipitation on day 1 and overestimate the observed precipitation at the end of day 5 and the end of day 6, while they both underestimate the observed peak at the beginning of day 6. For the spring case, both runs reproduced the observed precipitation events very well. These two runs also generated similar simulations in other important atmospheric fields, such as the column precipitable water and the surface latent and sensible heat fluxes. These fields typically show rather high temporal correlation

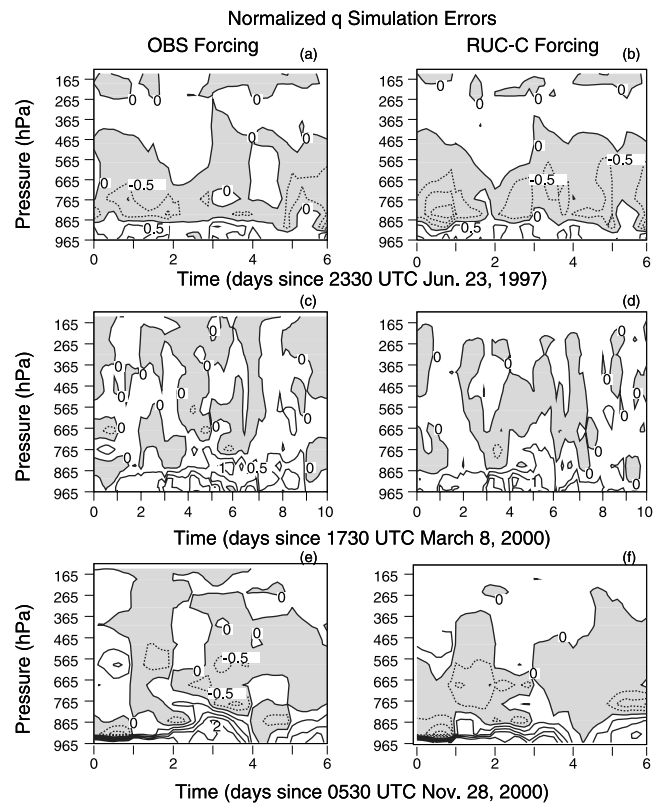


Figure 15. Same as Figure 14 except for the moisture biases. Contour interval is 1. The moisture bias is normalized by the mean specific humidity for the period.

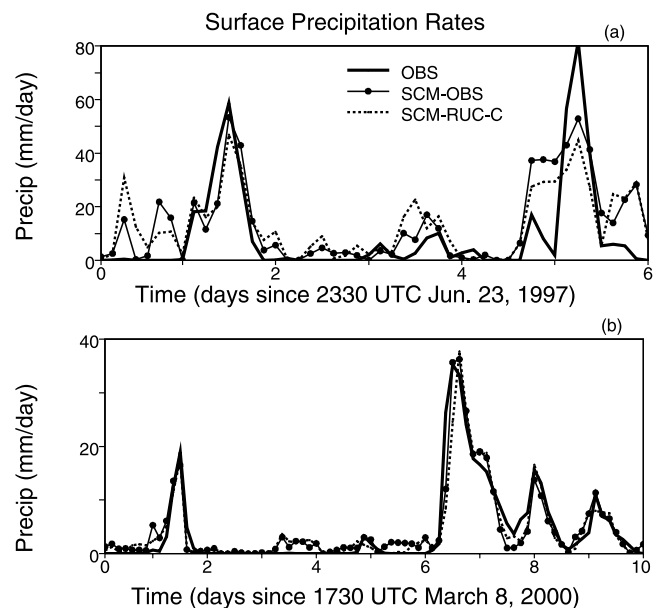


Figure 16. The time series of the observed (thick solid line) and the SCM simulated surface precipitation rates (mm d^{-1}): (a) over the strong convective period in 1997 summer IOP and (b) over the moderate precipitation period in 2000 spring IOP. Thin solid line with circle shows SCM-OBS, and dotted line shows SCM-RUC-C.

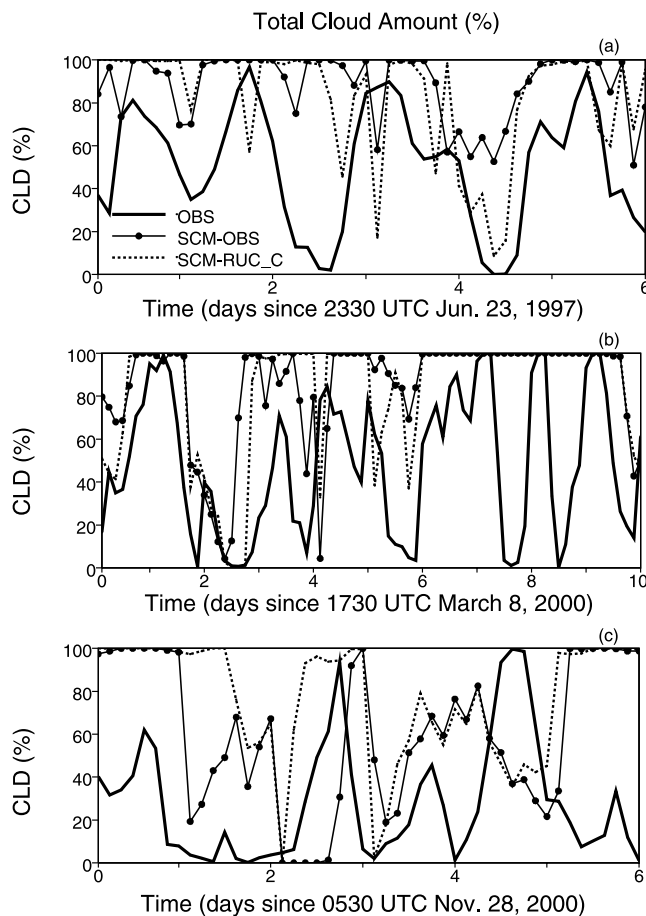


Figure 17. The time series of the observed and the SCM simulated total cloud amounts (in percent): (a) over the strong convective period in 1997 summer IOP, (b) over the moderate precipitation period in 2000 spring IOP and (c) over the nonprecipitation period in 2000 fall IOP. Thick solid line shows the observations. Thin solid line with circles shows SCM-OBS, and dotted line shows SCM-RUC-C.

(>0.90) between SCM-OBS and SCM-RUC-C and similar model errors, particularly for the spring and fall cases.

[31] Compared to these fields, relatively larger differences can be seen in the cloud-related fields (e.g., the cloud amount, cloud liquid water, and radiative fluxes) simulated by the SCM with the two different sets of forcing data. As an example, Figure 17 compares the simulated total cloud amounts with the observed NOAA GOES satellite clouds. Considerable differences are seen between the simulations and the observations, and between the two runs. This illustrates the sensitive nature of the model clouds. It should be noted that the cloud amount is a diagnosed variable in the CCM3 [Kiehl *et al.*, 1996]. It is sensitive to small differences in the model-predicted relative humidity, atmospheric stability, convective precipitation rate, and the prescribed vertical velocity from the forcing data sets. Nevertheless, the simulated clouds from SCM-RUC-C still show a fairly good correlation with those produced from SCM-OBS. The coefficients are 0.63, 0.76, and 0.51 for the summer case, the spring case, and the fall case, respectively. In addition,

both runs show similar model errors in simulating the observed cloud field, i.e., both runs tend to overestimate the observed cloud amounts for all the three cases.

[32] These SCM results obtained with the constrained NWP forcing and the observed forcing provide confidence in using the constrained NWP forcing data for statistical studies of SCM simulations over a much longer time period. The statistics of SCM simulations could further reduce the sensitivity of SCMs to the difference between the constrained NWP forcing and the observed forcing.

6. Summary and Discussion

[33] In this study, we have proposed another approach to develop a multiyear long-term SCM/CSRM forcing data from NWP analyses by using a variational analysis approach constrained with observed surface and TOA measurements. The derived forcing data from the RUC analyses using this approach have been evaluated against those derived from sondes and wind profilers using an objective variational analysis for three selected ARM IOP cases. We have shown that the forcing data derived from the RUC analyses constrained with the ARM surface and TOA observations exhibit considerably better agreement with the observed forcing than those derived from the RUC analyses without these constraints. The improvements tend to be the greatest during the precipitation periods since the observed precipitation provides the strongest constraint on the derived forcing data in the proposed approach. During the nonprecipitation periods, the improvements are moderate due to relatively weak constraints in absence of precipitation. In general, the agreement between the constrained NWP forcing and the observed forcing is better for the spring and fall cases than the summer strong convective case. The agreement in the derived omega and temperature forcing fields between these two data sets is better than that in the derived moisture forcing, especially during the summer case. This is partly related to the relatively larger errors in the RUC moisture analyses. This suggests an improvement of the RUC moisture analyses is needed in the future. It should be noted that the constrained NWP forcing derived the RUC analyses using the same surface and TOA constraints as those used in deriving the “observed” forcing. The improvement can be therefore expected. This is exactly the merit and purpose of the constraining process since the end products are the advective tendency terms and vertical velocity.

[34] It has been shown that the SCM driven by the constrained NWP forcing is able to produce similar simulations in most atmospheric fields as those obtained in the SCM driven by the observed forcing for the three cases. Relatively larger sensitivity to the difference between these two different forcing data sets is seen in those fields that are associated with clouds. Nevertheless, some SCM errors related to clouds can still be detected by using the constrained NWP forcing data. For example, the problem of the SCM overestimating the observed cloud amounts when using the observed forcing is also shown in the SCM using the constrained NWP forcing.

[35] Note that the forcing derived from the RUC analyses is not expected to be exactly the same as those derived from the observations due to errors in the RUC analyses and

uncertainties in the observations. However, this study has shown that adjusting the RUC analyses, based on its error characteristics, to conserve the observed column budgets of mass, heat, moisture, and momentum can significantly improve the quality of the derived forcing data from the RUC analyses. The encouraging results from the SCM tests further provide confidence in using the developed constrained NWP forcing data for statistical studies of SCM simulations over long-time periods.

[36] The results presented in this study can be viewed as a base case in our efforts to develop the ARM long-term forcing data. Currently a one-year forcing data set using the proposed approach has been developed for the year 2000 and has been used in some SCM/CSRM studies. For example, *Del Genio and Wolf* [2003] used the constrained NWP forcing data to study the response of different cloud types to changes in large-scale temperature and moisture forcings in different dynamical regimes. Further improvements to the constrained NWP forcing data are currently under consideration. For example, we are considering the use of more remote sensing retrievals that are available at the ARM SGP site over long periods of time in the analysis of the constrained NWP forcing. These could include (1) the NOAA wind profiler data that could be used to improve the RUC horizontal winds; (2) the high-frequency temperature and moisture retrievals from the atmospheric emitted radiance interferometer (AERI) [Feltz et al., 2003] that could be used to improve the RUC temperature and moisture analyses within the boundary layer; (3) the 6-hourly sondes at the central facility of the ARM SGP site; and (4) the moisture retrievals from the Raman Lidar at the central facility of the ARM SGP site. Studies and experiments to incorporate these data are being conducted and results will be reported separately.

[37] **Acknowledgments.** We greatly appreciate the efforts and dedication of John Yio in the development of the data processing and numerical analysis aspects of this work. We thank the three anonymous reviewers whose valuable comments helped to clarify 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. The developing constrained NWP forcing effort has been strongly supported by the ARM Cloud Parameterization and Modeling Working Group.

References

- Barnes, S. L. (1964), A technique for maximizing details in numerical map analysis, *J. Appl. Meteorol.*, **3**, 396–409.
- Bechtold, P., et al. (2000), A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE, II: Intercomparison of single-column models and a cloud-resolving model, *Q. J. R. Meteorol. Soc.*, **126**, 865–888.
- Benjamin, S. G., et al. (1998), RUC-2—The rapid update cycle version 2, *NWS Tech. Proc. Bull.* **448**, 18 pp., Off. of Meteorol., Natl. Weather Serv., Silver Spring, Md. (Available at <http://www.nws.noaa.gov/om/tpb/448.pdf>).
- Cripe, D. (1998), Single-column modeling: Sensitivity to initial conditions and divergence forcing, paper presented at Seventh ARM Science Team Meeting, U.S. Dep. of Energy, Washington, D. C.
- Del Genio, A. D., and A. B. Wolf (2003), Single-column model simulation of cloud sensitivity to forcing, paper presented at Thirteenth ARM Science Team Meeting, U.S. Dep. of Energy, Broomfield, Colo.
- Feltz, W. F., et al. (2003), Near continuous profiling of temperature, moisture, and atmospheric stability using the atmospheric emitted radiance interferometer (AERI), *J. Appl. Meteorol.*, **42**, 584–597.
- Ghan, S. J., et al. (2000), An intercomparison of single-column model simulations of summertime midlatitude continental convection, *J. Geophys. Res.*, **105**, 2091–2124.
- Hack, J. J., and J. A. Pedretti (2000), Assessment of solution uncertainties in single-column modeling frameworks, *J. Clim.*, **13**, 352–356.
- Kiehl, J. T., et al. (1996), Description of the NCAR Community Climate Model (CCM3), *NCAR Tech. Note NCAR/TN-420+STR*, Natl. Cent. for Atmos. Res., Boulder, Colo.
- Lin, X., and R. H. Johnson (1996), Kinematic and thermodynamic characteristics of the flow over the western Pacific warm pool during TOGA COARE, *J. Atmos. Sci.*, **53**, 695–715.
- Randall, D. A., K.-M. Xu, R. J. C. Somerville, and S. Iacobellis (1996), Single-column models and cloud ensemble models as links between observations and climate models, *J. Clim.*, **9**, 1683–1697.
- Redelsperger, J.-L., et al. (2000), A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE. part I: CRM results, *Q. J. R. Meteorol. Soc.*, **126**, 823–863.
- Waliser, D. E., J. A. Ridout, S. Xie, and M. Zhang (2002), Variational objective analysis for atmospheric field programs: A model assessment, *J. Atmos. Sci.*, **59**, 3436–3456.
- Xie, S. C., and M. H. Zhang (2000), Impact of the convective triggering function on single-column model simulations, *J. Geophys. Res.*, **105**, 14,983–14,996.
- Xie, S. C., et al. (2002), Intercomparison and evaluation of cumulus parameterizations under summertime midlatitude continental conditions, *Q. J. R. Meteorol. Soc.*, **128**, 1095–1135.
- Xie, S. C., R. T. Cederwall, M. Zhang, and J. Yio (2003), Comparison of SCM and CSRM forcing data derived from the ECMWF model and from objective analysis at the ARM SGP, *J. Geophys. Res.*, **108**(D16), 4499, doi:10.1029/2003JD003541.
- Xu, K.-M., et al. (2002), 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.
- Zhang, M. H., and J. L. Lin (1997), 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.
- Zhang, M. H., J. L. Lin, R. T. Cederwall, J. J. Yio, and S. C. Xie (2001), Objective analysis of ARM IOP Data: Method and sensitivity, *Mon. Weather Rev.*, **129**, 295–311.
- R. T. Cederwall and S. Xie, L-103, ASD, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. (xie2@llnl.gov)
- M. Zhang, Marine Sciences Research Center, State University of New York at Stony Brook, Stony Brook, NY 11794-5000, USA.